Novas Tecnologias de Redes Ethernet

Antonio José Figueiredo Enne
Bruno Lima Wanderley
Cristiano Henrique Ferraz

Índice

Capítulo 1: INTRODUÇÃO

1.1. PREÂMBULO
1.2. ETHERNET: DA LAN À WAN
1.3. VISÃO EVOLUTIVA DAS REDES ETHERNET
1.4. ESTRUTURA DO LIVRO
1.4.1. Capítulo 2
1.4.2. Capítulo 3
1.4.3. Capítulo 4
1.4.4. Capítulo 5
1.4.5. Capítulo 6
1.4.6. Capítulo 7
1.4.7. Capítulo 8

Capítulo 2: REDES ETHERNET

2.1. PREÂMBULO
2.2. PADRÃO IEEE 802 (OVERVIEW AND MANAGEMENT)
  2.2.1. Arquitetura IEEE 802
  2.2.2. Métodos de Acesso ao Meio Físico
  2.2.3. Formas de Interconexão de Segmentos de Rede
  2.2.4. Endereçamento Universal
2.3. RESOLUÇÃO DE ENDEREÇOS MAC
  2.3.1. Funcionamento Básico do ARP
  2.3.2. Formato de Quadros ARP
2.4. INDICAÇÃO DE PROTOCOLOS SUPERIORES
2.5. LOGICAL LINK CONTROL (LLC)
  2.5.1. Formato do Quadro LLC
  2.5.2. Indicação de Protocolos Superiores via LLC
    2.5.2.1. Indicação Direta pelo LSAP
    2.5.2.2. Indicação via SNAP
  2.5.3. Tipos de Serviço LLC
2.6. LAN BRIDGING
  2.6.1. Funções Relaying e Filtragem
  2.6.2. Transparent Bridging
2.7. REDES ETHERNET
  2.7.1. Formato dos Quadros Ethernet
  2.7.2. Tipos de Ethernet LANs
  2.7.3. Subcamada MAC
    2.7.3.1. Modo Half-Duplex
    2.7.3.2. Modo Full-Duplex
    2.7.3.3. InterPacketGap
  2.7.4. A Subcamada MAC Control
  2.7.5. Controle de Fluxo em Redes Ethernet
  2.7.6. Simplified Full-Duplex MAC
  2.7.7. Slow Protocols

2.8. A CAMADA FÍSICA DE REDES ETHERNET 802.3
2.8.1. Arquitetura da Camada Física Ethernet 802.3
2.8.2. Redes Operando de 1 Mbit/s a 10 Mbit/s
2.8.3. Redes Operando a 100 Mbit/s (Fast Ethernet)
2.8.4. Redes Operando a 1000 Mbit/s (Gigabit Ethernet)
2.8.5. Redes Operando a 10 Gbit/s
2.8.6. Redes Operando a 40 e a 100 Gbit/s
2.9. FORWARD ERROR CORRECTION (FEC) PARA BASE-R PHYs
2.10. ETHERNET IN THE FIRST MILE (EFM)
2.10.1. Alternativas de Redes de Acesso a Usuários
2.10.2. Tecnologias e Sistemas de Sinais de Camada Física em EFM
2.10.3. Sentidos de Transmissão em Redes de Acesso a Usuários
2.10.4. Tecnologias Ópticas Ponto a Ponto (P2P)
2.10.4.1. Sistemas de Sinais LX10
2.10.4.2. Sistemas de Sinais BX10
2.10.4.3. LX10 e BX10
2.10.4.4. Arquitetura de Redes de Acesso P2P
2.10.5. Tecnologias Ópticas Ponto a Multiponto (P2MP)
2.10.5.1. Passive Optical Networks (PONs)
2.10.5.2. EPONs
2.10.5.3. Forward Error Correction (FEC) para 1000BASE-PX
2.10.5.4. PMDs e Meios Físicos nos Sistemas de Sinais 1000BASE-PX10 e 1000BASE-PX20
2.10.5.5. PMDs e Meios Físicos nos Sistemas de Sinais 10GBASE-PR e 10/1GBASE-PRX
2.10.6. Tecnologias P2P com Cabos de Cobre
2.10.6.1. Considerações Iniciais
2.10.6.2. Padrões 10PASS-TS e 2BASE-TL e os Padrões DSL
2.10.7. Operação, Administração e Manutenção (OAM)

Capítulo 3: LAN BRIDGING COM SPANNING TREES

3.1. PREÂMBULO
3.2. TRANSPRENT BRIDGING
3.3. IEEE 802.1D BRIDGING
3.3.1. Arquitetura de Bridges IEEE 802.1D
3.4. FAMÍLIA DE PROTOCOLOS SPANNING TREE (STP)
3.5. RAPID SPANNING TREE PROTOCOL
3.5.1 Conceitos Básicos
3.5.1.1. Identificadores de Bridges e de Portas de Bridges
3.5.1.2. Root Bridges, Root Ports e Root Path Costs
3.5.1.3. Designated Bridges e Designated Ports
3.5.1.4. Port Roles e Port States
3.5.1.5. Spanning Tree Priority Vector
3.5.2. Formatação de RST BPDUs
3.5.3. Procedimentos de Configuração
3.5.4. Exemplo de Configuração
3.5.5. Procedimentos de Configuração
3.6. FAMÍLIA DE PROTOCOLOS MULTIPLE REGISTRATION (MRP)
3.6.1. Arquitetura do MRP

3.6.2. Operação do MRP
3.6.3. Endereços e Valores EtherType no MRP
3.6.4. Formato das MRPDUs
3.7. MMRP (MULTIPLE MAC REGISTRATION PROTOCOL)
3.7.1. Registro de Informações
3.7.2. Arquitetura do MMRP
3.7.3. Exemplo de Aplicação do MMRP
3.7.4. Utilização do IGMP
3.8. MODOS DE FILTRAGEM POR BRIDGES IEEE 802.1D
3.9. DEFAULT GROUP FILTERING BEHAVIOR
3.10. APRENDIZAGEM DE ENDEREÇOS MAC POR IEEE 802.1D BRIDGES
3.11. TRANSMISSÂO DE QUADROS DE DADOS POR IEEE 802.1D BRIDGES
3.11.1. Quadros de Dados com Endereços Broadcast e com Endereços Unicast Não Registrados
3.11.2. Quadros de Dados com Endereços de Destino Unicast Registrados
3.11.3. Quadros de Dados com Endereços Multicast
3.11.3.1. Bridges Operando no Modo Básico de Filtragem
3.11.3.2. Bridges Operando no Modo Ampliado de Filtragem
3.12. VLAN BRIDGING (PADRÂO IEEE 802.1Q-2014)
3.12.1. Formato de Quadros com VLAN Tags
3.12.2. Tipos de Ethernet Bridges
3.12.3. Tipos de Quadros
3.12.4. Tipos de Estações Finais e de Portas de Bridges
3.12.5. VIDs Especiais
3.12.6. Classificação de VLANs
3.12.7. Endereços MAC Reservados
3.12.8. Valores EtherType Atribuidos pelo IEEE
3.12.9. Prioridade e Precedência para Descarte
3.13. MULTIPLE SPANNING TREE PROTOCOL (MSTP)
3.13.1. Constituição da CIST
3.13.2. Constituição de Uma MSTI
3.13.3. Exemplo de Configuração
3.13.3.1. Configuração da CIST (SST e ISTs)
3.13.3.2. Configuração de uma MSTI na RG2
3.13.4. Formato de MSTP BPDUs
3.13.4.1. Campos Protocol e Protocol Version Identifier
3.13.4.2. Campo BPDU Type
3.13.4.3. Campo CIST Flags
3.13.4.4. Campos Seguintes até o Campo Message Age
3.13.4.5. Campos Message Age até Forward Delay
3.13.4.6. Campos Version 1 Length e Version 3 Length
3.13.4.7. Campos MST Configuration Identifier até campo CIST Remaining Hops
3.13.4.8. Campo MSTI Configuration Messages
3.14. MULTIPLE VLAN REGISTRATION PROTOCOL (MVRP)
3.14.1. Visão Geral do MVRP
3.14.1.1. Operação do MVRP
3.14.1.2. Comportamento de Estações Finais e de Bridges
3.14.1.3. Formato e Codificação de MVRPDUs
3.14.2. Serviço VLAN Registration
3.15. APRENDIZAGEM DE ENFEREÇOS AC POR VLAN-AWARE BRIDGES

3.16. TRANSMISSÂO DE QUADROS DE DADOS EM VLAN-AWARE BRIDGES

Capítulo 4: LAN BRIDGING COM IS-IS

4.1. PREÃMBULO
4.2. UTILIZAÇÂO DO IS-IS PARA ROTEAMENTO DE CAMADA 2
  4.2.1. Protocolo IS-IS
  4.2.2. Extensões do IS-IS para Roteamento de Camada 2
4.3. SHORTEST PATH BRIDGING (SPB)
  4.3.1. Roteamento no SPB
  4.3.2. Constituição de SPTs (Shortest Path Trees)
  4.3.3. Uso de ECTs
  4.3.4. Mitigação de Loops
  4.3.5. Shortest Path Bridging – VID (SPBV)
    4.3.5.1. Identificação de SPTs (ECTs)
    4.3.5.2. Transmissão de Quadros de Dados
  4.3.6. Shortest Path Bridging – MAC SPBM)
    4.3.6.1. Identificação de SPTs
    4.3.6.2. Filtragem de Tráfego
    4.3.6.3. Transmissão de Quadros Unicast e Broadcast
    4.3.6.4. Transmissão de Quadros Multicast
  4.3.7. Distribuição de Tráfego por SPBM ECMP
4.4. TRANSPARENT INTERCONNECTION OF LOTS OF LINKS (TRILL)
  4.4.1. Tipos de Quadros do TRILL
  4.4.2. Encapsulamento de Quadros de Dados do TRILL
    4.4.2.1. Outer Header (Cabeçalho Externo)
    4.4.2.2. TRILL Header
    4.4.2.3. Inner Ethernet Header
  4.4.3. Fine – Grained Labeling (FGL) TRILL
    4.4.3.1. Formato do Quadro de Dados do FGL TRILL
    4.4.3.2. Aspectos Operacionais do FGL TRILL
  4.4.4. Outras Alternativas de Cabeçalho Externo
    4.4.4.1. Conexão de RBridges via PPP (RFC 6361)
    4.4.4.2. Conexão de RBridges via PWs
  4.4.5. Aprendizagem da Localização de Nós Finais
  4.4.6. Protocolo ESADI
  4.4.7. Constituição de Múltiplas Árvores de Distribuição no TRILL
  4.4.8. Transmissão Unicast
  4.4.9. Filtragem de Tráfego
  4.4.10. DRB, Appointed Forwarders e Pseudo Nodes em Links de Acesso
    4.4.10.1. DRB (Designated RBridge)
    4.4.10.2. Appointed Forwarders
    4.4.10.3. Pseudo Node
  4.4.11. Acesso Active – Active ao TRILL
  4.4.12. RBridge Channel
    4.4.12.1. Estrutura de um Quadro do RBridge Channel
    4.4.12.2. RBridge Channel Header
  4.4.13. TRILL OAM e TRILL BFD
    4.4.13.1. Estrutura do TRILL OAM (RFC 7174)
    4.4.13.2. Gerenciamento de Falhas (RFC 7455)

4.4.13.3. Medições de Perdas e de Delays no TRILL (RFC 7456)
4.4.13.4. TRILL BFD (RFC 7175)
4.4.14. Plano de Controle do TRILL
4.4.14.1. Protocolo TRILL Hello
4.4.14.2. Protocolo IS-IS for TRILL
4.4.15. Módulo RBridge MIB (RFC 6850)

Capítulo 5: LAN BRIDGING COM REDES DE PROVEDORES

5.1. PREÃMBULO
5.2. VLANs C-TAGGED
5.3. PROVIDER BRIDGES (PB)
5.3.1. Formato Básico de Quadros MAC PB
5.3.2. Formato do S tag
5.3.3. Conceitos Básicos no PB
5.3.3.1. Componentes VLAN-Aware Bridges
5.3.3.2. Tipos de VLAN-Aware Bridges
5.3.3.3. Tipos de Portas
5.3.4. Interfaces de Serviço para Clientes
5.3.4.1. Interface de Serviço Baseada em Porta (Port-Based)
5.3.4.2. Interface de Serviço C-tagged
5.3.4.3. Interface de Serviço S-tagged
5.3.4.4. Visão Global de Interfaces de Serviço em PB
5.3.4.5. Prioridade de Serviço e Elegibilidade para Descarte
5.3.5. Protocolos de Controle no PB
5.3.5.1. Reserva e Bloqueio de Endereços MAC
5.3.5.2. Operação do MSTP em PB
5.3.6. Aprendizagem de Endereços MAC de Cliente
5.4. PROVIDER BACKBONE BRIDGES (PBB)
5.4.1. Formatos de Quadros em PBB
5.4.1.1. PDU de PBB
5.4.1.2. Campo TCI do I-Tag
5.4.2. Componentes de PBBNs
5.4.2.1. Tipos de Bridges em PBBNs
5.4.2.2. Tipos de Portas de Bridges em PBB
5.4.3. Características do PBB
5.4.4. Conectividade em Backbone VLANs
5.4.5. Interfaces de Serviço para Clientes
5.4.5.1. Interface de Serviço Baseada em Porta
5.4.5.2. Interface de Serviço S-tagged
5.4.5.3. Interface de Serviço C-tagged
5.4.5.4. Interface de Serviço I-tagged
5.4.5.5. Interface de Serviço Transparente
5.4.6. Exemplo de PBBN
5.4.7. Interconexão de PBBNs
5.4.8. Prioridade de Serviço e Elegibilidade para Descarte
5.4.9. Endereços MAC Backbone (B-MAC Addresses)
5.4.10. Proteção do Acesso ao Serviço
5.4.11. Gerenciamento de PBB Bridges
5.5. PROVIDER BACKBONE BRIDGES – TRAFFIC ENGINEERING (PBB-TE)

5.5.1. O PBB-TE como PTN
5.5.2. Funcionamento do PBB-TE
5.5.3. Suporte de Serviço MAC por uma Região PBB-TE
5.5.4. ESPs, TESIs e Shared Forwarding
5.5.4.1. P2P ESPs e P2P TESIs (P2P Eth – LSPs)
5.5.4.2. P2MP ESPs e P2MP TESIs
5.5.5. Constituição e Utilização de ESPs
5.5.6. Gerência de Falhas de Conectividade (CFM)
5.5.7. Comutação de Proteção para P2P TESIs
5.5.8. Plano de Controle do PBB-TE
5.5.9. Generalized Multiprotocol Label Switching (GMPLS)
5.5.9.1. Gerenciamento de Links no GMPLS com LMP
5.5.10. Roteamento no GMPLS
5.5.10.1. Protocolo OSPF
5.5.10.2. Protocolo OSPF – TE
5.5.10.3. Protocolo GMPLS OSPF – TE
5.5.11. Sinalização no GMPLS
5.5.11.1. Protocolo RSVP Versão 1
5.5.11.2. Protocolo RSVP – TE
5.5.11.3. Protocolo GMPLS RSVP – TE
5.5.11.4. Constituição de P2MP TE LSPs (RFC 4875)
5.5.11.5. P2P GMPLS TE LSPs Assimétricos
5.5.12. GMPLS em Redes Ethernet
5.5.13. GMPLS no PBB – TE (RFC 6060)
5.5.13.1. Procedimentos Básicos
5.5.13.2. Encaminhamento Compartilhado (Shared Forwarding)
5.5.13.3. PBB – TE Ethernet Label
5.5.13.4. Identificação de Instâncias de Serviço

Capítulo 6: ETHERNET SOBRE SDH E OTN

6.1. PREÂMBULO
6.2. HIERARQUIA DIGITAL SÍNCRONA
6.2.1. O Quadro STM-N
6.2.2. Elementos de Sinal do SDH
6.2.3. Mapeamento de Tributários PDH
6.2.4. Resumo de Velocidades
6.2.5. Next Generation SDH (NG-SDH)
6.2.6. Concatenação de VCs SDH
6.3. OPTICAL TRANSPORT NETWORKS (OTNs)
6.4. ETHERNET SOBRE SDH E OTN
6.4.1. Mapeamento Nativo de Ethernet em SDH
6.4.2. Transporte de Ethernet sobre OTN
6.4.3. EoSDH pelo LAPS
6.4.4. EoSDH e EoOTN pelo GFP

Capítulo 7: ETHERNET SOBRE MPLS

7.1. PREÂMBULO
7.2. MPLS (MULTIPROTOCOL LABEL SWITCHING)

7.2.1. Tipos de MPLS
7.2.2. Aplicações do MPLS
7.3. ARQUITETURA PWE3
7.3.1. Aspectos Iniciais
7.3.2. Modelo de Referência para a Arquitetura PWE3
7.3.3. Camadas Constituintes da Arquitetura PWE3
7.4. VIRTUAL PRIVATE WIRE SERVICE (VPWS)
7.4.1. VPWS PWs
7.4.2. Transmissão de PDUs de Camada 2
7.4.3. Subcamada Control Word
7.5. ETHERNET SOBRE VPWS
7.5.1. Modos de Operação de PWs
7.5.2. Tipos de VLAN Tag
7.5.3. Procedimentos Gerais
7.6. VIRTUAL PRIVATE LAN SERVICE (VPLS)
7.6.1. Configuração de VPLS VPNs
7.6.2. Aspectos Gerais do VPLS
7.6.3. Procedimentos no VPLS
7.6.4. Topologias em VPLS VPNs
7.6.5. VPLS Hierarquizado (HVPLS)

Capítulo 8: OAM, OPERAÇÂO E GERÊNCIA DE BRIDGES

8.1. PREÂMBULO
8.2. OAM EM REDES ETHERNET (EOAM)
8.2.1. SOAM CFM: Padrão IEEE 802.1Q-2014
8.2.2. DDCFM (Data-Driven and Data-Dependent CFM)
8.2.3. SOAM: Recomendação ITU-T G.8013/Y.1731
8.2.4. SOAM CFM: Recomendação ITU-T G.8013/Y.1731
8.2.5. SOAM PM: Recomendação ITU-T G.8013/Y.1731
8.2.6. OAM de Link: Cláusula 57 do Padrão IEEE 802.3
8.3. MECANISMOS DE PROTEÇÂO EM ETHERNET
8.3.1. Recomendação ITU-T G.8031/Y1342 (Ethernet Protection Switching)
8.3.2. Recomendação ITU-T G.8032/Y.1344 (Ethernet Ring Protection Switching)
8.3.3. RFC 3619 (EAPS)
8.3.4. Padrão IEC 62439
8.3.4.1. PRP (Paralell Redundancy Protocol)
8.3.4.2. HSR (High-Availability Seamless Redundancy)
8.3.4.3. MRP (Media Redundancy Protocol)
8.4. PADRÃO IEEE 802.1AB-2009
8.4.1.1. Princípios de Operação
8.4.1.2. Endereçamento de LLDPDUs
8.4.1.3. Formatação de LLDPDUs e Tipos de TLV
8.4.2. LLDP- Media Endpoint Discovery (LLDP-MED)
8.5. PADRÂO IEEE 802.1AX-2008 (LINK AGGREGATION)
8.5.1. Considerações Gerais
8.5.2. LACP (Link Aggregation Control Protocol)
8.5.3. Marker Protocol
8.5.4. Função Frame Collection e Função Frame Distribution
8.6. GERÊNCIA DE BRIDGES

8.6.1. Bridge Management
  8.6.1.1. Entidade de Gerência de Bridge
  8.6.1.2. FDB da Entidade MAC Relay
8.6.2. Management Information Base – MIB
  8.6.2.1. Conceitos e Funcionamento de MIBs
  8.6.2.2. IEEE 802.1Q MIB

++++++++++++++++++

CAPÍTULO 1

INTRODUÇÃO

1.1. PREÂMBULO

O objetivo básico deste livro, intitulado **Novas Tecnologias de Redes Ethernet**, é o de oferecer aos leitores uma abordagem abrangente e atualizada da tecnologia *Ethernet* e de sua evolução no sentido da oferta de serviços de redes de telecomunicações por parte de provedores.

O livro é endereçado a uma variada audiência, englobando profissionais das áreas de telecomunicações e de TI (tecnologia da informação) que atuam nas mais diversas atividades, tais como planejamento, operação, gerenciamento, manutenção, etc..

A opção por uma abordagem abrangente e didaticamente sequenciada neste livro visa a possibilitar a preparação progressiva do leitor para o aprofundamento dos temas abordados.

Esta obra complementa e expande os conceitos apresentados no livro "TCP/IP sobre MPLS", publicado pela Editora Ciência Moderna e escrito por Antonio José Figueiredo Enne, também um dos autores do presente livro.

1.2. ETHERNET: DA LAN À WAN

A evolução da tecnologia *Ethernet* vem ocorrendo em ritmo considerável.

A expressão "*Ethernet*: *from* LAN *to* WAN", amplamente utilizada por profissionais da área de telecomunicações, resume adequadamente a evolução da *Ethernet*.

A *Ethernet* surgiu como uma tecnologia de LAN no *Xerox Palo Alto Research Center*, em 1973. O seu inventor, o célebre Dr. Robert Metcalfe, declarou recentemente a sua surpresa com o elevado estado atual da evolução de seu invento.

De início, as redes *Ethernet* consistiam em simples barramentos, que posteriormente evoluíram para *hubs* passivos, aos quais as estações finais disputavam o acesso mediante a aplicação do algoritmo CSMA-CD (*Carrier Sensing Multiple Access - Collision Detection*), onde a transmissão ocorria no modo semi-duplex.

A partir da introdução das LAN *bridges* e, posteriormente, dos LAN *switches*, foi rompida a limitação geográfica das LANs, particularmente das *Ethernet* LANs.

As *bridged Ethernet* LANs, constituídas pela interconexão de LANs por meio de *Ethernet bridges* ou de *Ethernet switches*, extrapolaram os limites físicos das *Ethernet* LANs.

Dessa forma, as *bridged Ethernet* LANs passaram a constituir-se em MANs e, mediante a progressiva emissão de padrões mais rigorosos, também a constituir-se em WANs.

A tecnologia *Ethernet* assumiu a predominância quase absoluta no domínio das redes locais, eliminando praticamente outras tecnologias antes promissoras, tais como *Token Ring*, FDDI e ATM LANs. Esse sucesso pode ser atribuído, em suma, à simplicidade e à vantajosa relação custo/desempenho da *Ethernet*.

A técnica de LAN *Bridging* assumiu maior relevância a partir da emissão do padrão IEEE 802.1Q (*Bridges and Bridged Networks*), cuja última edição ocorreu em 2014, que introduziu o conceito de VLANs (*Virtual* LANs), que representa uma importante ferramenta para a constituição de L2VPNs (*Layer 2 Virtual Private Networks*).

1.3. VISÃO EVOLUTIVA DAS REDES ETHERNET

As *bridged* LANs foram inicialmente especificadas no padrão IEEE 802.1D-2004.

O padrão IEEE 802.1Q, hoje na versão 2014, representou um passo significativo na evolução das *bridged LAN*. Nesse padrão foi introduzido o conceito de VLAN, o que serviu de base para outros desenvolvimentos como a criação de redes *Carrier Ethernet*, cujo conceito básico, o de EVCs (*Ethernet Virtual Connections*), é atrelado ao conceito de VLAN ID (*VLAN Identifier*).

O IEEE, além dos abrangentes padrões IEEE 802.1D-2004 e IEEE 802.1Q-2014, emitiu também outros padrões especificamente voltados para determinados aspectos de *LAN Bridging*. São exemplos os padrões IEEE 802.1AB-2009 (*Station and Media Access Control Connectivity Discovery*) e o padrão IEEE 802.1AX-2008 (*Link Aggregation*).

Inicialmente baseadas exclusivamente em protocolos *spanning tree* (STPs), que apresentam algumas deficiências quanto ao uso dos recursos das redes, as *bridged Ethernet* LANs encontram-se em fase de início de utilização de novas concepções funcionais.

Pode ser citada, como principal exemplo, a definição de novas tecnologias *Ethernet* com suporte em roteamento com a métrica *link state*, mediante a aplicação de extensões dos protocolos de roteamento IP, particularmente do protocolo IS-IS (*Intermediate System-to-Intermediate System*). Os exemplos atuais dessa nova concepção são as redes SPB (*Shortest Path Bridging*), hoje definidas no padrão IEEE 802.1Q-2014, e as redes TRILL (*Transparent Interconnection of Lots of Links*), definidas em diferentes RFCs (*Request for Comments*) emitidas pelo IETF (*Internet Engineering Task Force*).

Ao mesmo tempo, passou-se a perseguir a especificação de redes de transporte orientadas a conexão, referidas como PTNs (*Packet Transport Networks*). Embora o termo transporte tenha um significado mais amplo, as PTNs são redes

modo pacote cujo desempenho se aproxima do desempenho de redes de transporte modo circuito, a exemplo das redes SDH e das redes OTN.

No domínio da *Ethernet*, foi definido o conceito de redes CO-*Ethernet* (*Connection-Oriented Ethernet*), para possibilitar a especificação de redes de transporte com suporte em *Ethernet*.

Ficou estabelecido o uso do GMPLS (*Generalized Multiprotocol Label Switching*) como suporte ao plano de controle das redes CO-*Ethernet*. Essas redes, quando controladas pelo GMPLS, são referidas como redes GELS (GMPLS *Ethernet Label Switching*), conforme a RFC 5828 (*Generalized Multiprotocol Label Switching (GMPLS) Ethernet Label Switching Architecture and Framework*).

Como um exempo de aplicação do GELS, foi emitido, inicialmente, o padrão IEEE 802.1Qay, definindo o plano de dados e o gerenciamento do PBB-TE (*Provider Backbone Bridging-Traffic Engineering*), também referido como PBT (*Provider Backbone Transport*). O padrão IEEE 802.1Qay foi obsoletado, sendo seus termos incorporados ao padrão IEEE 802.1Q-2014. O uso do GMPLS para o PBB-TE foi especificado na RFC 6060 (*Generalized Multiprotocol Label Switching* (*GMPLS*) *Control of Ethernet Provider Backbone Traffic Engineering - PBB-TE*).

O PBB-TE é visto como uma PTN, concorrendo com as redes MPLS-TP (MPLS *Transport Profile*), que representam um outro importante exemplo atual de PTNs.

Como redes de provedores de serviço baseadas em *Spanning Trees*, foram especificadas as PBNs (*Provider Bridging Networks*) e as PBBNs (*Provider Backbone Bridging Networks*). O PB (*Provider Bridging*) e o PBB (*Provider Backbone Bridging*) são tecnologias de rede sem conexão, também especificadas no padrão IEEE 802.1Q-2014.

Um grande desafio para as redes *Ethernet* é o de apresentar especificações *Carrier-Class* nos diversos aspectos operacionais, em termos, por exemplo, de gerenciamento, de QoS (*Quality of Service*) e de OAM (*Operation, Administration and Maintenance*).

Diversos padrões são aplicáveis nesse sentido. Para OAM, pode ser citado o padrão IEEE 802.1ag (hoje obsoletado, tendo sido os seus termos incorporados ao padrão 802.1Q-2014), que define os procedimentos para a Gerência de Falhas de Continuidade (CFM). A recomendação ITU-T G.8013/Y.1731 (*OAM Functions and Mechanisms for Ethernet based Networks*) é dedicada a OAM de serviço (SOAM) em redes Ethernet de um modo abrangente, incorporando, consequentemente, também as funções de *Perfomance Monitoring* (PM).

1.4. ESTRUTURA DO LIVRO

Este livro livro encontra-se estruturado nos seguintes oito capítulos:

— Capítulo 1: **Introdução**;
— Capítulo 2: **Redes Ethernet**;

— Capítulo 3: **LAN Bridging com Spanning Trees**;
— Capítulo 4: **LAN Bridging com IS-IS;**
— Capítulo 5: **LAN Bridging em Redes de Provedores;**
— Capítulo 6: **Ethernet sobre SDH e OTN;**
— Capítulo 7: **Ethernet sobre MPLS;**
— Capítulo 8: **OAM, Operação e Gerência de Bridges.**

1.4.1. **Capítulo 2**

O Capítulo 2, que se fundamenta principalmente no padrão IEEE 802.3, destina-se a oferecer aos leitores um embasamento conceitual com relação a LANs, particularmente Ethernet LANs. O propósito é permitir aos leitores um alicerce conceitual sobre o qual fundamentar o conteúdo mais específico dos demais capítulos.

Foram abordados nesse capítulo os seguintes padrões IEEE 802:

— 802-2001 (*Overview and Management*);
— 802.2-1998 (*Logical Link Control* - LLC);
— 802.3-2012 (CSMA/CD *Access Method and Physical Layer Specifications*).

Além de apresentar uma série de aspectos e de conceitos fundamentais para o entendimento das redes Ethernet, o Capítulo 2 e dedica o seu item 2.8 à apresentação da camada física de redes Ethernet IEEE 802.3. Nesse item são abordados os diferentes sistemas de sinais.

O item 2.10 desse capítulo aborda, de forma abrangente, o importante tema *Ethernet in the First Mile* (EFM). Esse tema foi inicialmente especificado no padrão IEEE 802.3ah, que foi obsoletado e incorporado ao padrão IEEE 802.3-2012, na Seção 5 de sua Cláusula 57.

O padrão IEEE 802.3-2012 especifica o conceito de EPON (*Ethernet Passive Optical Networks*), pelo qual são implementadas topologias ponto a multiponto (P2MP) com *splitters* ópticos passivos. As EPONs operam com base no protocolo MPCP (*Multi-Point Control Protocol*), que administra a comunicação entre um MAC virtual em uma estação raiz, denominada OLT (*Optical Line Terminal*), e os diversos MAC localizados nas estações folha da rede P2MP, referidas como ONUs (*Optical Network Units*).

Foram também apresentados no Capítulo 2 os diferentes sistemas de sinais específicos para EFM, tanto utilizando fibras ópticas quanto cabos de cobre.

1.4.2. **Capítulo 3**

O tema *LAN Bridging* pode ser dividido em dois segmentos:

— *LAN Bridging* em redes de usuários;
— *LAN Bridging* em redes de provedores.

Uma outra forma de classificação de *LAN Bridging* é considerar a base de sua concepção. A concepção de *LAN Bridging* com base em *spanning trees* reinou absoluta por muitos anos. Como as *spanning trees* apresentam problemas, particularmente, no que diz respeito à eficiência de uso de recursos de rede, encontram-se em fase de implementação novas concepções de rede Ethernet baseadas em protocolos de roteamento *link state*, particularmente no protocolo IS-IS.

O Capítulo 3 aborda *LAN Bridging* em redes de usuários com base em *spanning trees*. Esse capítulo fundamenta-se nos padrões IEEE 802.1D-2004 e IEEE 802.1Q-2014, e apresenta os seguintes principais aspectos:

— RSTP (*Rapid Spanning Tree Protocol*);
— MSTP (*Multiple Spanning Tree Protocol*);
— MMRP (*Multiple MAC Registration Protocol*);
— MVRP (*Multiple VLAN Registration Protocol*);
— Mecanismo de Aprendizagem de Endereços MAC.

1.4.3. **Capítulo 4**

O Capítulo 4 introduz as tecnologias de redes Ethernet com base no protocolo IS-IS. São cobertas as seguintes tecnologias:

— SPB (*Shortest Path Bridging*), especificado no padrão IEEE 802.1Q-2014;
— TRILL (*Transparent Interconection of Lots of Links*), especificado em diferentes RFCs do IETF.

O SPB, definido inicialmente no padrão IEEE 802.1aq, hoje obsoletado e incorporado ao padrão IEEE802.1Q-2014, opera com roteamento utilizando a métrica *link state*, com suporte em extensões do protocolo IS-IS referidas como ISIS-SPB. Ao não usar o SPT, o SPB representa uma alternativa ao PB e ao PBBB, que, embora com a limitação de também operar sem conexão, apresenta maior eficiência no uso dos recursos da rede que o suporta.

O TRILL, por sua vez, é uma alternativa ao SPB especificado inteiramente pelo IETF. Da mesma forma que o SPB, o TRILL também utiliza uma métrica *link state*, com suporte em extensões do protocolo IS-IS referidas como *IS-IS for TRILL*.

De acordo com o IETF, o TRILL apresenta vantagens sobre o SPB, tais como a possibilidade de uso transparente de diferentes tecnologias para interconexão de seus nós, e a possibilidade de operação sem a necessidade de inundação de rede, o que ocorre em decorrência do uso do protocolo ESADI (*End Station Address Distribution Information*).

Entretanto, como diferentes fornecedores optaram por ambas essas alternativas, aguarda-se um panorama mais claro quanto à prevalência de uma delas.

Embora ainda carecendo de uma especificação mais clara quanto a esse propósito, o SPB e o TRILL destinam-se a operar como redes de provedores.

1.4.4. **Capítulo 5**

No Capítulo 5 são apresentadas as seguintes tecnologias de rede de provedores:

— Extensão de redes *C-tagged*;
— PB (*Provider Bridges*);
— PBB (*Provider Backbone Bridges*);
— PBB-TE (*PBB-Traffic Engineering*).

A simples extensão das redes *C-tagged* para suporte a redes de provedores é considerada inadequada, devido à superposição funcional de endereços MAC e de VLAN-IDs (VIDs) entre o provedor e os usuários. Além disso, o número máximo de instâncias de serviço (até 4094) que podem ser criadas é considerado limitado.

O PB, definido inicialmente no padrão IEEE 802.1ad, hoje obsoletado e incorporado ao padrão IEEE 802.1Q-2014, baseia-se na inserção de um VLAN *tag* adicional pelo provedor de serviço, *tag* esse referido como *S-tag* (*service tag*). Como o PB se resume ao envelopamento do *C-tag* (*Customer tag*) pelo *S-tag* na rede PB (PBN), o PB é também referido como Q-*in*-Q.

O PB representa um avanço com relação à extensão das rede *C-tagged*, mas persistem no PB a superposição de endereços MAC e a questão da escalabilidade, uma vez que o número de valores de S-VID contidos nos *S-tags* (4094) é considerado insuficiente para grandes redes, embora em cada S-VID possam ser mapeados até 4096 (4094 na prática) C-VIDs.

O PBB foi definido inicialmente no padrão IEEE 802.1ah, hoje obsoletado e incorporado ao padrão IEEE 802.1Q-2014, com o propósito de eliminar as limitações do PB. O PBB tem como fundamento a introdução adicional, pelo provedor de serviço, de uma camada MAC (referida como B-MAC, com o B significando *backbone*), além de mais um VLAN *tag* intermediário (referido como I-*tag*, com o I significando *instance*).

Assim sendo, o PBB é também referido como MAC-*in*-MAC, e elimina a superposição de endereços MAC.

Como os I-SIDs nos I-*tags* possuem 24 bits, é solucionada também a questão de escalabilidade existente no PB.

As redes PBB (PBBNs) podem dar acesso a redes PB (quando são utilizados também os *S-tags*) ou podem permitir a conexão direta de redes *C-tagged* (ou seja, redes com C-VLANs) de usuários ou mesmo de redes IEEE 802.1D (sem VLANs) de usuários.

Apesar dessas vantagens, o PBB apresenta ainda a característica limitadora de operar sem conexão, o que impede a implementação de importantes funções operacionais, particularmente a de *Traffic Engineering*. Além disso, o PBB adota os mecanismos convencionais de LAN *bridging* baseados no STP, do que resulta o uso ineficiente dos recursos das PBBNs.

O PBB-TE, definido inicialmente no padrão IEEE 802.1Qay, hoje obsoletado e incorporado ao padrão 802.1Q-2014, que é também referido como PBT (*Provider Backbone Transport*). O PBB-TE opera no modo orientado a conexão, possibilitando a constituição de circuitos virtuais independentes das redes externas, viabilizando a função *Traffic Engineering* e a adoção de mecanismos mais robustos de OAM e de comutação automática de proteção.

O plano de controle do PBB-TE tem como suporte o GMPLS, tanto em termos de roteamento (protocolo GMPLS OSPF-TE ou protocolo GMPLS ISIS-TE) quanto de sinalização (extensão do protocolo GMPLS RSVP-TE).

O PBB-TE representa, em concorrência com o MPLS-TP (MPLS-*Transport Profile*), uma tecnologia PTN (*Packet Transport Network*) emergente.

As redes Ethernet de provedores são, na realidade, formas de transporte de tráfego entre *sites* distantes de redes Ethernet de usuários.

Assim, o Capítulo 4 e o Capítulo 5 constituem as formas de transporte de Ethernet sobre Ethernet, o que pode ser referido como ETHoETH (*Ethernet over Ethernet*).

Diversas outras tecnologias de rede de telecomunicações podem ser utilizadas para transportar o tráfego de redes Ethernet de usuários. Podem ser citados o *Frame Relay*, o ATM, o MPLS e as redes de transporte modo circuito, tais como SDH e OTN, ou mesmo redes ópticas WDM (*Wavelength Division Multiplexing*) ou até as próprias fibras ópticas diretamente utilizadas.

Foram apresentados neste livro, pela sua importância, o transporte de Ethernet sobre SDH e OTN (Capítulo 6) e o transporte de Ethernet sobre MPLS (Capítulo 7).

1.4.5. **Capítulo 6**

O Capítulo 6 dedica-se às formas de transporte de Ethernet sobre SDH (*EoSDH*) e de transporte de Ethernet sobre OTN (*EoOTN*).

Especificamente quanto ao *EoSDH*, foram apresentados, além de uma introdução ao SDH e ao NG-SDH (*Next Generation SDH*), as seguintes tecnologias:

- VCAT (*Virtual Concatenation*), que é parte constituinte do NG-SDH, com base na recomendação ITU-T G.707/Y.1322 (*Network Node Interface for the Synchronous Digital Hierarchy* – SDH);
- LCAS (*Link Capacity Adjustment Scheme*), que é também parte constituinte do NG-SDH, com base na recomendação G.7042/Y.1305 (*Link Capacity Adjustment Scheme* (*LCAS*) *for Virtual Concatenated Signals*);
- *EoSDH* pelo LAPS (*Link Access Procedure-SDH*) com base na recomendação ITU-T X.86/Y.1323 (*Ethernet over LAPS*).

O VCAT e o LCAS são também aplicáveis em OTN.

Finalmente foram apresentados o *EoSDH* e o *EoOTN* com a intermediação do GFP (*Generic Frame Procedure*), com base na recomendação ITU-T G.7041/Y.1303 (*Generic Framing Procedure*).

1.4.6. **Capítulo 7**

O Capítulo 7 aborda o transporte de Ethernet sobre MPLS (*EoMPLS*). Além de uma apresentação genérica das redes MPLS e da arquitetura PWE3 (*Pseudo Wire Emulation Edge-to-Edge*) definida na RFC 3985 (*Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture*), são apresentadas as seguintes aplicações do MPLS:

— VPWS (*Virtual Private Wire Service*);
— VPLS (*Virtual Private LAN Service*).

Como o VPWS é uma aplicação que suporta o transporte de diferentes tecnologias sobre MPLS (*Frame Relay* e ATM, por exemplo), o Capítulo 7, além de apresentar uma visão conceitual do VPWS, concentrou-se no transporte de Ethernet sobre PWs (*pseudowires*) constituídos sobre o VPWS, com base na RFC 4448 (*Encapsulation Methods for Transport of Ethernet over MPLS Networks*).

O transporte de quadros *Ethernet* sobre o VPWS ocorre de forma transparente aos endereços MAC desses quadros. Esse transporte ocorre nos moldes dos serviçoes *E-Line* (*Ethernet Line*) definidos pelo MEF (*Metro Ethernet Forum*).

O VPLS, por sua vez, suporta exclusivamente o transporte de quadros *Ethernet*. Instâncias de serviço VPLS contidas nos PEs de redes MPLS, denominadas VSIs (*Virtual Switching Instances*), emulam *Ethernet switches* associadas a cada VPLS VPN constituída.

Dessa forma, o VPLS opera no padrão convencional das *bridged Ethernet* LANs, baseado na aprendizagem de endereços MAC. Essa forma de operação é característica dos serviços E-LAN (*Ethernet LAN*) definidos pelo MEF.

1.4.7. **Capítulo 8**

Finalmente, o Capítulo 8 é dedicado aos mecanismos necessários à plena operacionalização de redes Ethernet.

Assim, esse capítulo contempla OAM (*Operation, Administration and Maintenance*), assim como alguns protocolos de operação de redes Ethernet. Aborda também, de uma forma geral, o gerenciamento de *bridges*, conforme o padrão IEEE 802.1Q-2014.

No que concerne OAM, são apresentados os seguintes tópicos:

— SOAM CFM (*Service OAM Connectivity Fault Management*), com base nas claúsulas 18 a 22 do padrão IEEE 802.1Q-2014;
— DDCFM (*Data-Driven and Data-Dependent CFM*), com base na Cláusula 27 do padrão IEEE 802.1Q-2014;

— SOAM CFM, com base na recomendação ITU-T G.8013/Y.1731 (*OAM Functions and Mechanisms for Ethernet Networks*);
— SOAM PM (*SOAM*
— *Performance Monitoring*), com base na recomendação ITU-T G.8013/Y.1731;
— OAM de *Link* (*Link OAM*), com base na cláusula 57 do padrão IEEE 802.3-2012 (CSMA/CD *Access Method and Physical Layer Specifications*).

Quanto a padrões operacionais em redes Ethernet, foram contemplados os seguintes tópicos:

— Comutação Automática de Proteção;
— Protocolo LLDP (*Link Layer Discovery Protocol*), com base no padrão IEEE 802.1AB-2009;
— Agregação de *Links*, com base no padrão IEEE 802.1AX-2008.

Para o tópico Comutação Automática de Proteção foram abordados os seguintes padrões:

— Recomendação ITU-T G.8031/Y.1342 (*Ethernet Protection Switching*);
— Recomendação ITU-T G.8032/Y.1344 (*Ethernet Ring Protection Switching*), que define a tecnologia ERP;
— RFC 3619 (*Extreme Networks' Ethernet Automatic Protection Switching*), que define a tecnologia EAPS;
— Padrão IEC 62439 (*Industrial Communication Networks – High Availability Automation Networks*).

A recomendação ITU-T G.8031/Y.1342 define o protocolo *Linear APS* (*Automatic Protection Switching*), destinado à proteção automática de conexões Ethernet ponto a ponto em redes Ethernet com topologias lineares. São suportados mecanismos de proteção 1+1 e mecanismos de proteção I:1, tanto no modo reversivo quanto no modo não reversivo.

A recomendação G.8032/Y.1344, referente ao mecanismo de comutação automática denominado ERP (*Ethernet Ring Protection*), destina-se ao atendimento de redes Ethernet com topologias em anel (*ring*). O protocolo especificado nessa recomendação é o R-APS (*Ring Automatic Protection Switching*).

Da mesma forma que no caso do protocolo R-APS, a RFC 3619 define a tecnologia denominada EAPS (*Extreme Networks' Ethernet Automatic Protection Switching*) com o propósito de oferecer proteção automática em redes Ethernet com topologias em anel.

Finalmente, o Capítulo 8 contempla os diferentes mecanismos de proteção automática em redes Ethernet industriais, com base no padrão IEC 6243, particularmente em redes utilizadas em subestações elétricas.

Dentre tais mecanismos, foram apresentados com maiores detalhes o PRP (*Parallel Redundancy Protocol*) e o HSR (*High Availability Seamless Redundancy*),

ambos especificados no sub-padrão IEC 62439-3, e o MRP (*Media Redundancy Protocol*), especificado no sub-padrão IEC 62439-2.

CAPITULO 2

REDES ETHERNET

2.1. PREÂMBULO

O IEEE 802 é composto por um *Executive Commitee* (EC), pelos *Working Groups* (WGs) e pelos *Technical Advisory Groups* (TAGs) e SGs (*Study Groups*), sendo de competência dos WGs a emissão dos padrões normativos emitidos pelo IEEE 802.

A Figura 2-1 apresenta a relação de WGs, TAGs e SGs ativos.

802.1	Higher Layer LAN Protocols Working Group
802.3	Ethernet Working Group
802.11	Wireless LAN Working Group
802.15	Wireless Personal Area Network (WPAN) Working Group
802.16	Broadband Wireless Access Working Group
802.18	Radio Regulatory TAG
802.19	Wireless Coexistence Working Group
802.21	Media Independent Handover Services Working Group
802.22	Wireless Regional Area Networks
802.24	Smart Grid TAG
OmniRAN EC Study Group	Open Mobile Network Interface for Omni-Range Area Networks based on IEEE 802 access technologies

Figura 2-1 – Relação de IEEE 802 WGs, TAGs e SGs ativos.

A estrutura dos padrões da família IEEE 802 encontra-se ilustrada na Figura 2-2.

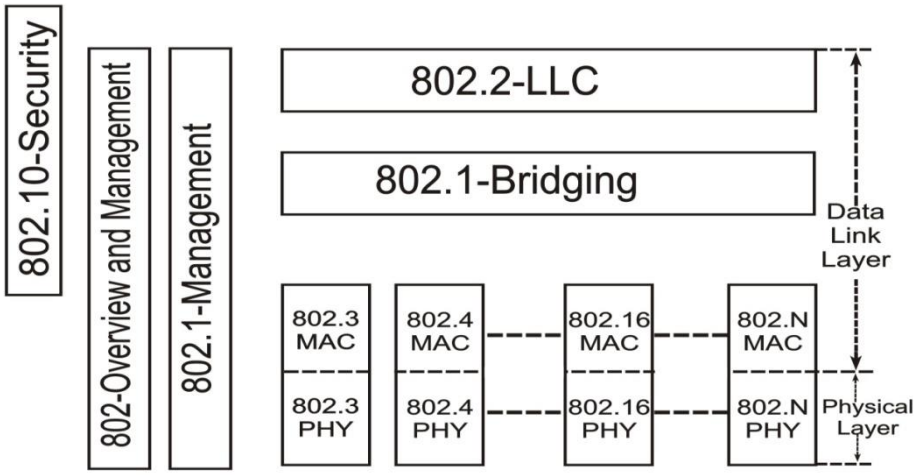


Figura 2-2 – Estrutura das famílias de padrões IEEE 802.

Serão abordados neste capítulo os seguintes padrões IEEE 802:

— 802-2001 (*Overview and Management*);
— 802.2-1998 (*Logical Link Control* - LLC);
— 802.3-2012 (CSMA/CD *Access Method and Physical Layer Specifications*).

2.2. PADRÃO IEEE 802 (OVERVIEW AND MANAGEMENT)

O padrão IEEE 802 apresenta uma visão geral do funcionamento do conjunto de padrões 802 e uma apresentação básica do processo de gerenciamento das respectivas LANs e MANs.

Serão abordados os seguintes aspectos do padrão IEEE 802:

— Arquitetura IEEE 802;
— Métodos de acesso ao meio físico;
— Formas de interconexão de segmentos de rede;
— Gerência de LANs e MANs.

2.2.1. Arquitetura IEEE 802

A arquitetura IEEE 802 consiste de duas camadas:

— Camada física;
— Camada de enlace de dados.

A camada de enlace de dados é composta por uma subcamada MAC (*Media Access Control*) e por uma subcamada LLC (*Logical Link Control*).

A Figura 2-3 apresenta a arquitetura IEEE 802.

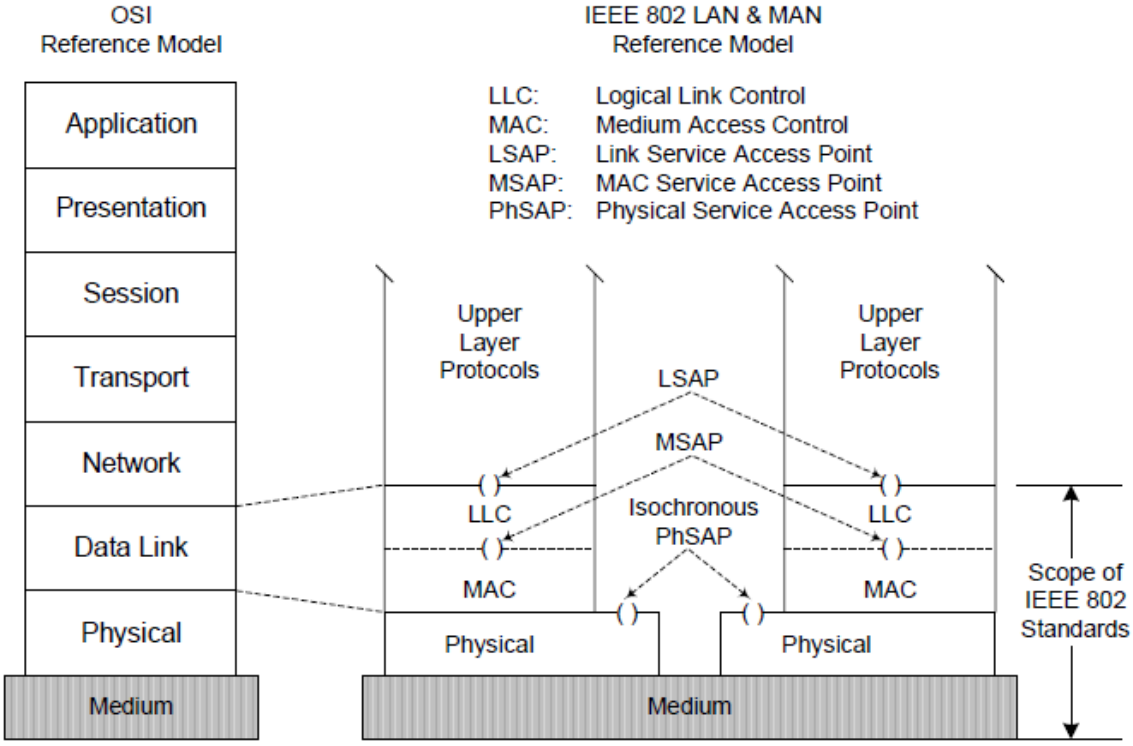


Figura 2-3 – Arquitetura IEEE 802.

A camadas e subcamadas interagem através dos respectivos SAPs (*Service Access Points*).

A camada física provê a capacidade de transmitir e receber bits entre entidades de camada física. Um par de entidades de camada física possibilita o intercâmbio de bits entre dois usuários, e, em alguns casos específicos pode, opcionalmente, dar suporte ao intercâmbio de bits com a preservação da temporização fim a fim entre *PhSAPs* (*Physical SAPs*) de serviço *Ethernet* isócrono.

Embora o serviço provido pela camada física represente uma transferência de bits entre entidades da subcamada MAC, os símbolos codificados nem sempre representam bits individuais. A transmissão na camada física pode utilizar modulações e codificações complexas, podendo inclusive representar mais de um bit por símbolo transmitido.

A subcamada MAC realiza as funções necessárias para o provimento de transferência de dados no modo pacote sem conexão entre diferentes estações. As principais funções do MAC incluem as seguintes:

— Delimitação e reconhecimento de quadros MAC;
— Endereçamento de estações de destino (endereços individuais e endereços de grupo);
— Transporte do endereço da estação de origem;
— Transferência transparente de LLCPDUs;
— Proteção contra erros, por meio de um campo FCS (*Frame Check Sequence*);
— Controle de acesso ao meio físico;

Outras funções da subcamada MAC, aplicadas particularmente quando a implementação inclui dispositivos de interconexão tais como *hubs* ou *bridges*, incluem o controle de fluxo entre uma estação terminal e o dispositivo de interconexão para o tipo de LAN ou MAN utilizado e a filtragem de quadros.

A subcamada LLC desempenha funções que lhe são específicas, de acordo com o tipo de serviço LLC considerado, o que será visto adiante neste capítulo.

2.2.2. **Métodos de Acesso ao Meio físico**

As LANs e MANS são redes modo pacote não orientadas a conexão. As topologias de suporte são variações de barramentos, de anéis e de estrelas. Em consequência, torna-se necessária a definição de métodos de acesso ao meio físico.

O CSMA-CD (*Carrier Sense Multiple Access with Collision Detection*), que será abordado adiante neste capítulo, é o método de acesso utilizado pelas redes *Ethernet*.

2.2.3. **Formas de Interconexão de Segmentos de Rede**

Vários métodos foram desenvolvidos para permitir a interconexão transparente entre estações terminais que não se encontram na mesma LAN física, particularmente no caso do suporte a uma maior extensão geográfica ou a um número maior de estações conectadas.

Os segmentos de LANs e de MANs podem ser interconectados das seguintes formas:

- Interconexão na camada física, utilizando dispositivos geralmente denominados repetidoresou *hubs*.
- Interconexão na subcamada MAC, utilizando dispositivos denominados *bridges*;
- Interconexão na camada de rede, utilizando dispositivos geralmente denominados roteadores.

O propósito dos repetidores é o de simplesmente estender o domínio de alcance de uma rede, constituindo a configuração resultante ainda um único domínio de colisão. Os repetidores multiportas são referidos como *hubs*.

A interconexão na subcamada MAC tem como base as *bridges* e os *LAN switches*, que serão abordados neste capítulo e no capítulo subsequente.

A terceira categoria de interconexão, que se processa em nível de camada de rede (IP, por exemplo), tem como suporte os roteadores, que operam com interfaces suportando as funcionalidades das LANs e MANs que a elas se conectam.

2.2.4. **Endereçamento Universal**

O padrão IEEE 802 define uma formatação de endereçamento universal para os endereços MAC de origem e de destino, válida para todos os tipos de LANs e MANs.

Os endereços MAC são constituídos por 48 bits, divididos em duas partes. A primeira parte corresponde ao valor de OUI (*Organizationally Unique Identifier*), com 24 bits, enquanto a segunda parte, também com 24 bits, é administrada pela entidade à qual foi atribuído o valor de OUI.

A administração universal de endereços LAN MAC iniciou-se com a Xerox *Corporation*, quando os OUIs se denominavam *Block Identifiers* (*Block IDs*). Esse trabalho passou à responsabilidade do IEEE, através da *IEEE Registration Authority*.

Os seis octetos dos endereços MAC são usualmente representados sob a forma hexadecimal. A interpretação dos bits desses octetos dá-se conforme o exemplo ilustrado na Figura 2-4, onde os LSBs (*Least Significant Bits*) de cada octeto estão representados no lado direito.

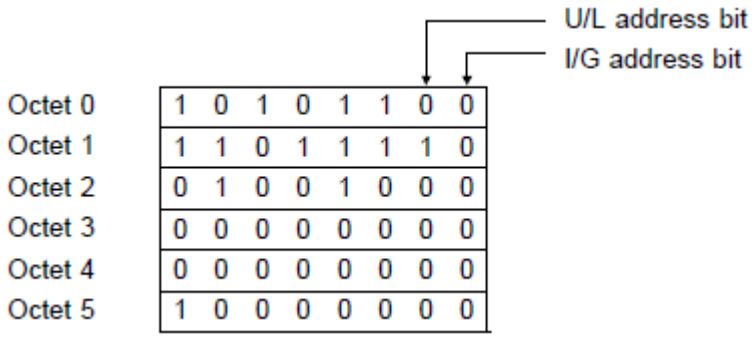


Figura 2-4– Formato de endereços MAC universais.

A representação do endereço universal correspondente a essa figura no formato *bit-stream* encontra-se na Figura 2-5, onde o LSB de cada octeto encontra-se à esquerda, conforme a convenção para a representação da transmissão *bit-serial* da esquerda para a direita adotada para a transmissão dos endereços MAC de LANs Ethernet.

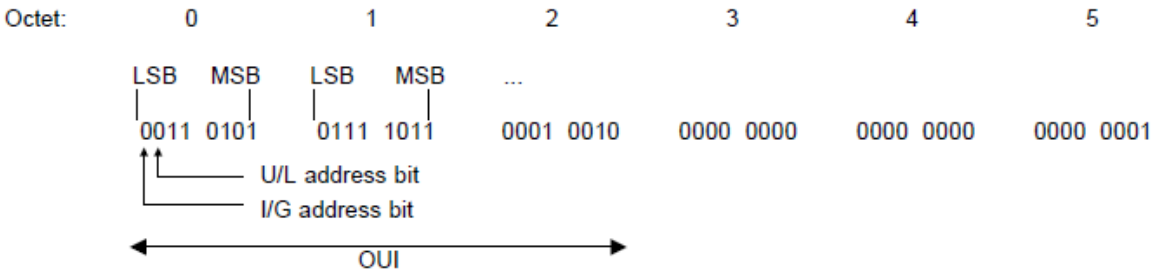


Figura 2-5 – Endereços MAC universais no formato bit-stream.

Como se verifica nessa figura, o bit LSB (*Least Significant Bit*) do primeiro octeto do OUI (I/G *adress bit*) identifica a natureza do endereço MAC contido nos últimos 3 octetos do endereço (que representa os valores atribuídos pelo detentor do valor de OUI). Se o I/G bit for igual a zero, trata-se de um endereço MAC *unicast* (individual). O valor do I/G bit igual a 1 significa que o endereço é *musticast*, referente a um grupo de endereços. Um endereço *broadcast* é um caso particular de endereço de grupo, em que todos os bits são iguais a 1 (endereço FF-FF-FF-FF-FF-FF-FF-FF).

O *U/L bit (Universal/Local Bit*) informa se o endereço MAC foi atribuído por um administrador universal (*U/L bit* igual a zero) ou pela própria entidade local detentora do valor do OUI (*U/L bit* igual a 1).

Como se pode verificar, essa forma de endereçamento permite a um valor de OUI representar aproximadamente 16 milhões (2^{24}) de endereços individuais e 16 milhões de endereços de grupo.

2.3. RESOLUÇÃO DE ENDEREÇOS MAC

O protocolo ARP (*Address Resolution Protocol*) tem como função resolver, ou seja, obter endereços *unicast* de subrede (endereços MAC, por exemplo) a partir de endereços *unicast* de rede (endereços IP, por exemplo).

No caso de IP sobre Ethernet, o ARP resolve endereços MAC *unicast* a partir de endereços IP *unicast* de destino. A rede Ethernet onde se aplica o ARP pode ser um simples segmento *Ethernet LAN* ou uma *Bridged Ethernet LAN*.

2.3.1. Funcionamento Básico do ARP

Uma estação de uma LAN ou MAN envia um quadro ARP *Request* usando o endereço MAC *broadcasting*, mas contendo o endereço IP individual da estação que se deseja mapear (*target IP address*).

Apenas a estação receptora que possui o endereço IP de destino do quadro ARP *Request* responde, informando o seu endereço MAC. A resposta se realiza mediante o envio de um quadro ARP *Reply*, endereçado à estação que enviou o ARP *Request*.

O objetivo dessa troca de quadros ARP é possibilitar às estações a inserção (ou a atualização) de entradas endereço IP → endereço MAC nas tabelas ARP.

A estação à qual foi endereçado o ARP *Request* utiliza as informações contidas nesse ARP *Request* para inserir ou atualizar em sua tabela ARP a entrada referente à estação que gerou o ARP *Request*.

Da mesma forma, a estação receptora do ARP *Reply* insere ou atualiza o registro em sua tabela ARP referente à estação que gerou o ARP *Reply*.

As entradas nas tabelas ARP permanecem válidas por um determinado tempo. Caso se esgote o tempo previsto (*time-out*) sem que tenha ocorrido uma atualização de uma entrada, essa entrada será apagada.

2.3.2. Formato de Quadros ARP

Os quadros ARP, definidos na RFC 826 do IETF, têm o formato apresentado na Figura 2-6.

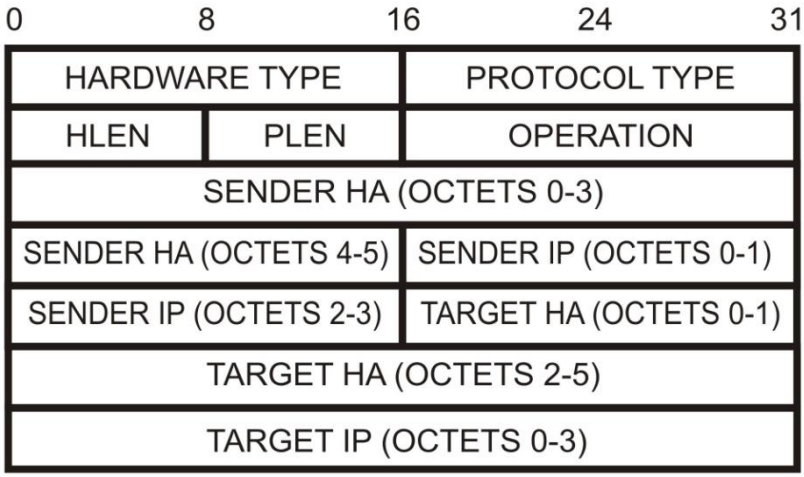


Figura 2-6 – Formato dos quadros ARP.

Os campos *Hardware Type* e *Protocol Type* nessa figura indicam, respectivamente, o tipo de subrede (*Ethernet,* por exemplo) e o protocolo de rede (IP) utilizados.

São exemplos de valores desses códigos:

— Ethernet: *Hardware Type* igual a 1;
— IP: *Protocol Type* igual a 0x0800 .

O campo *Operation* indica o tipo de quadro, com a seguinte codificação:

— ARP *Request*: *Operation* igual a 1;
— ARP *Reply*: *Operation* igual a 2.

Os campos HLEN e PLEN indicam, respectivamente, os comprimentos (*lengths*) dos campos *Hardware Address* (*Sender HA* e *Target HA*) e IP *Address* (*Sender* IP e *Target* IP).

Os demais campos indicam os endereços HA e IP do solicitador (*sender*) e os endereços HA e IP do questionado (*target*) dos quadros ARP enviados. Registra-se que o objetivo do ARP *request* é obter o *target HA* (o campo *target HA* vai vazio no *ARP Request,* e volta preenchido no *ARP response*).

É possível a utilização de *proxies* no processo ARP. Note-se que, nesse caso, os endereços MAC que encapsulam os quadros ARP não coincidem com os endereços MAC (*sender* HA e *target* HA) associados aos endereços IP do solicitador e do questionado (*sender* IP e *target* IP).

2.4. INDICAÇÃO DE PROTOCOLOS SUPERIORES

As redes Ethernet utilizam duas formas para indicar os protocolos superiores envelopados:

— Diretamente pelo MAC, por meio do campo *Length*/*Type*;

— Via subcamada LLC, quando o campo *Length/Type* do MAC indica o tamanho da PDU, ficando para o LLC a incumbência de indicar o protocolo superior.

Como será visto adiante neste capítulo, a forma de indicação depende do tipo de quadro MAC utilizado (MAC DIX 2.0 / Ethernet II ou MAC Ethernet 802.3).

A indicação direta pelo MAC utiliza a codificação *EtherType* (RFC 5342).A codificação *EtherType* destina-se à indicação de um grande número de protocolos superiores, e utiliza um campo com 16 bits.

A título de exemplo, a Figura 2-7 apresenta alguns códigos *EtherType* especificados pelo IETF.

CÓDIGOS	DESTINAÇÕES
0x0800	Internet Protocol Version 4 (IPv4)
0x0806	Address Resolution Protocol (ARP)
0x0808	Frame Relay ARP
0x880B	Point-to-Point Tunneling Protocol (PPTP)
0x880C	General Switch Management Protocol (GSMP)
0x8035	Reverse Address Resolution Protocol (RARP)
0x86DD	Internet Protocol Version 6 (IPv6)
0x8847	MPLS
0x8848	MPLS with upstream-assigned label
0x8861	Multicast Channel Allocation Protocol (MCAP)
0x8863	PPP over Ethernet (PPPoE) Discovery Stage
0x8864	PPP over Ethernet (PPPoE) Session Stage

Figura 2-7 - Exemplos de códigos *EtherType* especificados pelo IETF.

A Figura 2-8, por sua vez, apresenta alguns exemplos de códigos *EtherType* especificados pelo IEEE 802, com as respectivas destinações.

CÓDIGOS	PADRÃO IEEE	DESTINAÇÕES
0x8100	802 .1Q	Customer VLAN Tag Type (C-Tag)
0x8808	802 . 3	Ethernet Passive Optical Network (EPON)
0x888E	802 . 1X	Port-based network access control
0x88A8	802 . 1Q	Service VLAN tag identifier (S-Tag)
0x88B5	802	Local Experimental Ethertype
0x88B6	802	Local Experimental Ethertype
0x88B7	802	OUI Extended Ethertype
0x88C7	802 . 11i	Pré-Authentication
0x88CC	802 . 1AB	Link Layer Discovery Protocol (LLDP)
0x88E5	802 . 1AE	Media Access Control Security
0x88F5	802 . 1ak	Multiple VLAN Registration Protocol (MVRP)
0x88F6	802 . 1Q	Multiple Multicast Registration Protocol (MMRP)
0x890D	802 . 11r	Fast Roaming Remote Request

Figura 2-8 - Exemplos de códigos *EtherType* especificados pelo IEEE 802.

A indicação via LLC é apresentada no item subsequente.

2.5. LOGICAL LINK CONTROL (LLC)

O LLC constitui-se em uma subcamada de enlace de dados de LANs e MANs, e é envelopado pela subcamada MAC dessa camada. O LLC foi definido pelo padrão IEEE 802.2-1998 (*Logical Link Control* – LLC) sendo utilizado como suporte aos padrões de LANs e MANs do IEEE 802, e também do FDDI (*Fiber Distributed Data Interface*). O padrão IEEE 802.2-1998 corresponde ao padrão ISO/IEC 8802-2:1998 (*Logical Link Control* - LLC) .

O protocolo LLC tem como função básica a transferência de informações, de forma controlada ou não controlada, entre um LSAP (*Link Service Access Point*) de origem e um ou mais LSAPs de destino.

Um LSAP, conforme definição OSI, é o ponto da interface entre a subcamada LLC e a camada imediatamente a ela superior, no qual os serviços LLC são providos a essa camada superior. Os LSAPs podem conter a identificação direta dos protocolos superiores, ou a identificação de outra codificação intermediária para essa identificação.

2.5.1. Formato do quadro LLC

O formato do quadro LLC encontra-se representado na Figura 2-9.

DSAP	SSAP	CONTROL	INFORMATION
8 Bits	8 Bits	8 ou 16 Bits	8 x Mbits

Figura 2-9– Formato do quadro LLC.

Os campos dessa figura têm os seguintes significados:

- DSAP (*Destination* SAP): ponto de acesso de serviço no destino;
- SSAP (*Source* SAP): ponto de acesso de serviço na origem;
- *Control*: campo destinado ao controle nos moldes do HDLC controlado (a exemplo do LAP-B);
- *Information*: *payload* (SDU) da LLC PDU, que corresponde à PDU do protocolo imediatamente superior ao LLC (datagrama IP, como importante exemplo).

Os campos DSAP e SSAP (em conjunto denominados LSAP) indicam o protocolo superior diretamente envelopado pelo LLC (IP, IPX, CLNP, ARP,SNAP, etc.). De um modo geral os valores de DSAP e SSAP são iguais, caracterizando assim um valor único de LSAP.

2.5.2. Indicação de Protocolos Superiores via LLC

O LLC, por meio dos LSAPs, provê a multiplexação estatística da camada MAC, como pode ser visualizado na Figura 2-10.

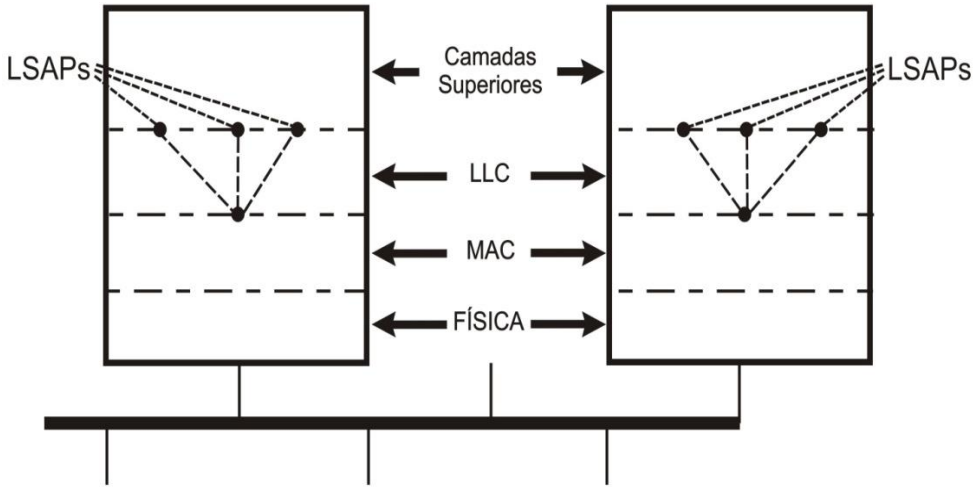


Figura 2-10 – Multiplexação estatística do MAC pelo LLC.

Conforme menção anterior, o LSAP indica o protocolo superior contido na camada de enlace de dados de duas formas:

— Diretamente por codificação própria;
— Pela intermediação de outro protocolo, do qual o principal exemplo é o SNAP (*Subnetwork Access Protocol*).

2.5.2.1. Indicação Direta pelo LSAP

A partir da arquitetura IEEE 802, a indicação dos protocolos superiores em LANs e MANs passou a ser possível pela codificação LSAP contida no quadro do protocolo LLC, liberando assim o protocolo MAC dessa função. O LSAP possibilita a indicação do protocolo superior na origem (campo SSAP) e no destino (campo DSAP).

A Figura 2-11 apresenta alguns exemplos de códigos LSAP.

TIPO	PROTOCOLO	TIPO	PROTOCOLO
04	SNA	80	XNS
05	SNA	81	XNS
06	IP	86	LLC
07	IP	87	LLC
08	SNA	AA	SNAP
09	SNA	AB	SNAP
0C	SNA	BC	VINES IP (VIP)
0D	SNA	BD	VINES IP (VIP)
10	NETWARE (IPX)	E0	XNS
42	BPDU (IEEE 802.1D)	E1	XNS
43	BPDU (IEEE 802.1D)	F0	NETBIOS
7E	ISO 8202 (X.25)	F1	NETBIOS

Figura 2-11 – Exemplos de códigos LSAP.

2.5.2.2. Indicação via SNAP

Em determinadas condições especiais utiliza-se o protocolo SNAP em complemento ao LSAP. Para isso, o LSAP deve indicar a presença do SNAP mediante o uso do código LSAP hexadecimal AA ou AB. A utilização do SNAP pode justificar-se pelas seguintes razões:

— Inexistência de código LSAP para a indicação direta de um protocolo superior;
— Necessidade de unificar o processo de identificação de protocolos superiores em inter-redes contendo *Ethernet* LANs legadas e LANs ou MANs que utilizam a arquitetura IEEE 802 (isto é, contendo a subcamada LLC);
— Outras condições especiais.

Observa-se que, embora exista um LSAP para IPv4, não existe um LSAP para o ARP. Como o ARP está sempre presente em redes IPv4, é necessário utilizar o SNAP caso se utilize o IPv4 sobre LLC.

O SNAP é constituído por duas subcamadas:

— A subcamada OUI;
— A codificação correspondente ao OUI utilizado.

O SNAP aponta para a codificação utilizada por meio de sua subcamada OUI (*Organizationally Unique Identifier*).

Podem ser mencionados os seguintes exemplos de valores de OUI:

— OUI igual 00-00-00 (codificação *EtherType*);
— OUI igual 00-80-C2 (codificação para tipos de MAC).

Para o exemplo do uso do SNAP apontando a codificação *EtherType* em LANs e MANs suportando o IPv4, temos a arquitetura da Figura 2-12.

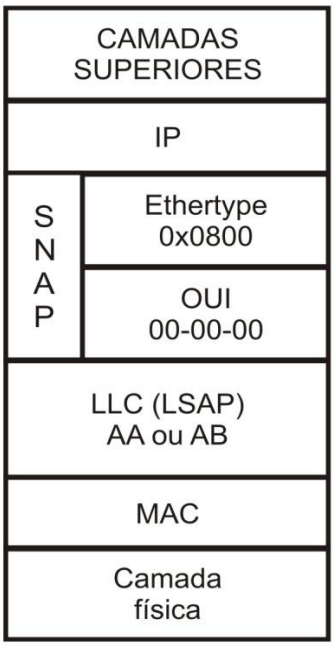


Figura 2-12– Arquitetura de LANs e MANs com SNAP em rede IP.

Observa-se que, no caso de uso do SNAP, o campo de controle do LLC apresenta o valor hexadecimal 03, o que significa tratar-se de um quadro de informação não numerado.

Podem ocorrer condições em que o protocolo SNAP deve apontar para a utilização de uma codificação que não seja a *EtherType*, como, por exemplo, a codificação que indica tipos de camada MAC (802.3 e 802.5, por exemplo).

A Figura 2-13 apresenta alguns valores de PID destinados a apontar para os tipos de MAC.

VALORES DO PID		TIPOS DE QUADROS
COM PRESERVAÇÃO DO FCS	SEM PRESERVAÇÃO DO FCS	
00-01	00-07	B02.3/ETHERNET
00-02	00-08	802.4
00-03	00-09	802.5
00-04	00-0A	FDDI
———	00-0B	802.6(DQDB)
———	00-0D	FRAGMENTOS
———	00-03	BPDUs (802.1d)
———	00-0F	BPDUs DE BRIDGES SOURCE ROUTING

Figura 2-13 – Valores de PID para tipos de MAC.

Um importante exemplo de protocolos superiores indicados pelo LSAP são os protocolos de rede. O protocolo IPv4, por exemplo, pode ser indicado diretamente pelo LSAP pelos códigos 0x06 ou 0x07. Alternativamente, o LSAP pode apontar para o uso de SNAP, que por sua vez aponta para a codificação *EtherType* (pelo OUI 00-00-00), que indica então o protocolo IPv4 pelo código 0x0800.

2.5.3. **Tipos de Serviço LLC**

O LLC oferece três alternativas de serviço:

— Serviço sem conexão e sem confirmação (LLC Tipo 1);
— Serviço com conexão e com confirmação (LLC Tipo 2);
— Serviço sem conexão e com confirmação (LLC Tipo 3).

O LLC Tipo 1 é de uso predominante, partindo-se do princípio de que o controle se realiza por protocolos superiores (TCP, principalmente). Nesse tipo de serviço, a função do LLC é a de indicar o protocolo superior envelopado. Nesse caso, o campo *Control* adota o valor fixo 0x03 (interpretado como UI ou *Unnumbered Information*, ou seja, que representa a ausência de funções de controle no nível de enlace de dados).

O LLC Tipo 2, que opera de modo controlado (de forma análoga, por exemplo, à do protocolo LAP-B), é utilizado em casos especiais onde a função de controle se torna necessária.

O LLC Tipo 3 não foi praticamente utilizado, sendo desconsiderado na literatura técnica sobre LANs e MANs.

2.6. LAN BRIDGING

Como visto anteriormente neste capítulo, as LANs (e também as MANs) podem ser interconectadas com base em sua subcamada MAC, mediante a utilização de *bridges*.

O termo LAN *switch*, ou simplesmente *switch*, é frequentemente utilizado para alguns tipos de *bridge*, embora o padrão IEEE 802 não reconheça qualquer distinção consistente entre os termos *bridge* e *switch*. Esse padrão optou pelo uso exclusivo do termo *bridge*.

2.6.1. Funções Relaying e Filtragem

As funções básicas das *bridges* são o *relaying* (retransmissão) e a filtragem de quadros MAC.

O *relaying* consiste no encaminhamento correto do quadro MAC pelas *bridges*, com base no endereço MAC de destino. A filtragem consiste na possibilidade de bloquear o envio de certos quadros por determinadas portas das *bridges*.

A Figura 2-14 ilustra o posicionamento da função de *bridging* nas subcamada MAC.

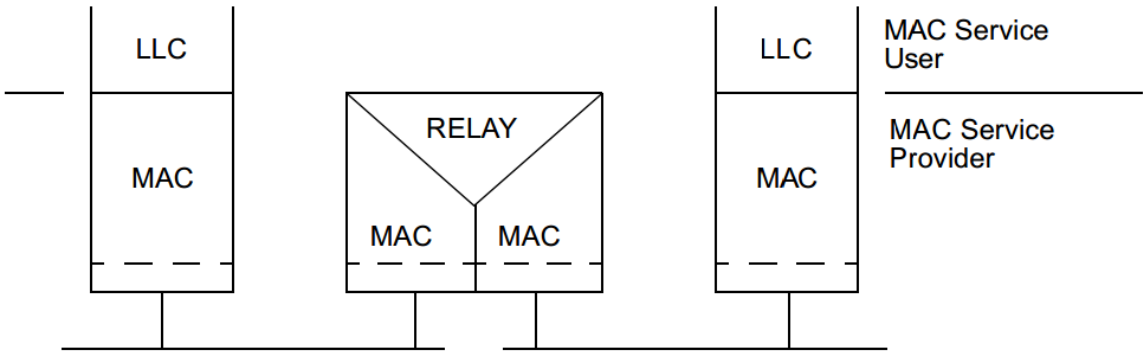


Figura 2-14– Posicionamento da função de *bridging* na subcamada MAC.

As funções de *relaying* e de filtragem estão contidas inteiramente na subcamada MAC.

2.6.2. Transparent Bridging

O padrão ISO/IEC 15802-3 define o modo de operação *Transparent Bridging*, assim denominado porque o seu funcionamento não requer que os quadros MAC de dados transportem qualquer informação adicional. Essa é a opção utilizada em *bridged Ethernet* LANs.

Menciona-se o fato de que, além do modo de operação *Transparent Bridging*, estão definidos também os modos de operação *Source-Routing Bridging* e *Source-Routing Transparent Bridging*, usados em outros tipos de rede IEEE 802 que não Ethernet.

2.7. REDES ETHERNET

As redes *Ethernet* tiveram seu início em 1973, mediante a emissão de um memorando escrito pelo Dr. Robert Metcalfe, no *Xerox Palo Alto Research Center*.

O sucesso da *Ethernet* foi tão significativo que excedeu as expectativas do próprio Dr. Robert Metcalfe, segundo suas recentes palavras. As vendas de *Ethernet switching ports* foram surpreendentes para o próprio inventor da *Ethernet*.

A predominância da *Ethernet* como LAN é hoje praticamente absoluta. Principalmente no ambiente metropolitano, mas também em longa distância, a utilização da *Ethernet* vem recebendo uma grande atenção pelo setor de desenvolvimento e padronização.

Em 1979 foi emitido o primeiro padrão *Ethernet,* sob a denominação DIX, pela associação das empresas *Digital* (DEC), *Intel* e *Xerox*. Uma segunda versão do padrão DIX foi divulgada em 1982. Essa versão é referida como Ethernet DIX V2.0 ou Ethernet II.

O padrão IEEE 802.3 foi inicialmente emitido em 1983, dando início à elaboração de uma considerável série de padrões relativos à *Ethernet.* O padrão IEEE 802.3 recebeu a denominação IEEE 802.3 *Carrier Sense Multiple Access with Collision Detection* (CSMA-CD) *Access Method and Physical Layer Specifications*. Esse padrão foi reconhecido posteriormente pela *International Standards Organization* (ISO), o que lhe concedeu o caráter oficial de padrão mundial.

As redes 802.3 *Ethernet* diferem ligeiramente das redes Ethernet definidas pelo padrão DIX. Para maior clareza será convencionado, no presente livro, que a simples expressão redes Ethernet refere-se às duas opções de redes *Ethernet*.

As redes Ethernet podem hoje ser classificadas em dois tipos:

— Redes Ethernet em barramento ou com *Ethernet hubs*;
— Redes *switched Ethernet* (ou redes *bridged Ethernet*).

As redes Ethernet em barramento representam a forma original de uso da Ethernet em LANs. Esse tipo de rede Ethernet caracteriza-se pelo uso do CSMA-CD para a resolução de contenda no acesso ao meio, podendo os segmentos de LAN ser interconectados por meio de repetidores. Para maior eficiência topológica, os barramentos passaram a ser colapsados com o uso dos *Ethernet hubs*, mantendo-se a mesma lógica operacional dos barramentos.

Em uma escala evolutiva, surgiu a concepção de redes *switched Ethernet*(ou redes *bridged Ethernet*).

Além de uma topologia simplificada contendo um único *Ethernet switch*, as redes *switched Ethernet* podem apresentar múltiplos *Ethernet switches*,

interconectados em diferentes topologias. São exemplos a topologia em estrela, a topologia *mesh* e a topologia *hub-and-spoke*.

As redes *switched Ethernet* apresentam as seguintes características marcantes:

— Possibilitam o acesso ao meio sem disputa;
— Podem operar em modo *full-duplex*;
— Alcançam dimensões metropolitanas ou mesmo de longa distância;
— Operam com múltiplas interfaces físicas, com velocidades de até 100 G*bit/s*.

Os padrões *Ethernet*, que utilizam a arquitetura representada pelas camadas físicas e de enlace de dados (subcamada MAC e subcamada LLC), são representados pela seguinte notação:

— Velocidade/modo de codificação/meio físico ou comprimento.

Assim, por exemplo, o padrão 10BASE5 significa uma rede que opera com 10 Mbit/s, com modulação em banda base (*baseband*) e com extensão de até 500 metros. Os padrões 10GBASE-T e 10GBASE-LRM, como outros exemplos, representam redes Ethernet com 10 Gbit/s com modulação *baseband*, operando, respectivamente, sobre UTP (*Unshielded Twisted Pair*) e fibras ópticas multimodo.

A relação entre o a padrão IEEE 802.3-2012 e o modelo de referência OSI pode ser vista na Figura 2-18 adiante neste capítulo.

2.7.1. **Formato dos Quadros Ethernet**

As redes *Ethernet*, independentemente do padrão e do tipo de rede, utilizam um único formato para os quadros MAC. A subcamada LLC, quando empregada, segue o formato definido pelo padrão IEEE 802.2.

Os quadros Ethernet possuem o formato apresentado na Figura 2-15.

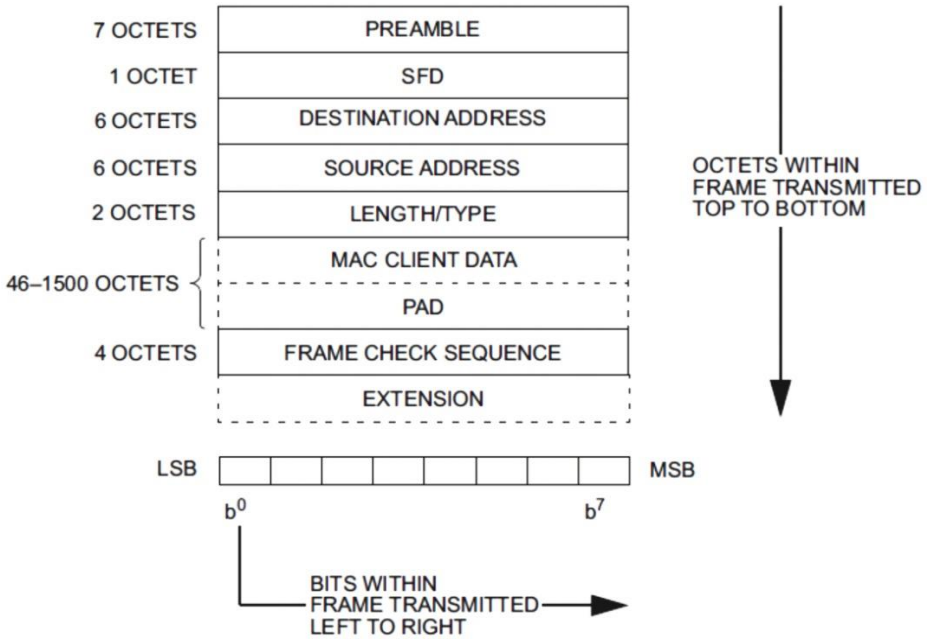


Figura 2-15 – Formato dos quadros Ethernet.

Como se observa nessa figura, os octetos dos quadros são transmitidos no modo usual, isto é, de cima para baixo. Os bits no interior dos octetos, por sua vez, são transmitidos da esquerda para a direita, o que significa que o primeiro bit de um octeto transmitido é o LSB bit.

O campo *Preamble* (preâmbulo), com 7 octetos, é utilizado para possibilitar ao componente PLS da camada física alcançar estabilidade de sincronização com o *timing* do quadro MAC recebido da subcamada MAC.

O campo *Start Frame Delimiter* (SFD) é representado pela sequência 10101011, e indica o início do quadro.

Os campos *Destination Address* (MAC DA) e *Source Address* (MAC SA) são, respectivamente, os endereços MAC de destino e de origem do quadro MAC. Esses endereços são estruturados em 48 bits, da forma descrita anteriormente neste capítulo. Como visto, o LSB bit do primeiro octeto do OUI indica se o endereço é individual ou de grupo.

O campo *Length/Type*, com 2 octetos, pode ter dois diferentes significados, dependendo de seu valor numérico, o que será visto adiante neste capítulo.

O campo *MAC Client Data* representa o *payload*, ou a SDU, do quadro MAC.

Caso o número de octetos de dados do protocolo do cliente MAC seja inferior ao tamanho mínimo desse campo (46 octetos), deve ocorrer o preenchimento (*padding*) desse tamanho mínimo por meio da inserção de octetos com todos os seus bits zerados.Tais octetos assim inseridos representam o subcampo *PAD* existente no formato do quadro MAC.

Constituindo o *trailer* dos quadros MAC, o campo *Frame Check Sequence* (FCS), com 4 octetos, transporta um valor de CRC (*Cyclic Redudancy Check*), que permite a detecção de erros de transmissão no conteúdo desses quadros.

O valor de FCS é definido a partir da geração de um polinômio na origem, e protege os campos de endereço, de comprimento/tipo e de dados (incluindo o PAD,se existir).

2.7.2. **Tipos de Ethernet LANs**

Existem atualmente dois tipos de Ethernet LANs:

— LAN Ethernet DIX V2.0 ou Ethernet II;
— LAN Ethernet 802.3.

As redes Ethernet II efetuam a indicação de protocolos superiores por meio de um campo com esse propósito existente na subcamada MAC. Essa indicação ocorre pela codificação *EtherType*, nesse caso diretamente, ou seja, sem necessidade do LLC e do SNAP.

O tipo de rede é identificado com base no campo *Type/Length*. Se o valor desse campo for igual ou superior a 1536, trata-se de um quadro do tipo Ethernet II. Nesse caso, o valor contido nesse campo corresponde ao *EtherType*. Para isso, os valores de *EtherType* são todos superiores a 1536.

Se o valor do campo *Type/Length* for igual ou inferior a 1500, trata-se de um quadro do tipo Ethernet 802.3, que obrigatoriamente utiliza o LLC (definido no padrão IEEE 802.2). Nesse caso, existem duas hipóteses para a indicação do protocolo de camada superior contido no quadro:

— diretamente através do LSAP;
— através do SNAP, cuja utilização é apontada pelo LSAP.

O exame do cabeçalho LLC permite determinar se o cabeçalho LLC é seguido de um cabeçalho SNAP. O cabeçalho LLC inclui o LSAP, composto de dois endereços, chamados SAPs (DSAP e SSAP, respectivamente *Destination Service Access Point* e *Source Service Access Point*, de um octeto cada). No caso do uso de SNAP, ambos os campos assumem o valor 0xAA.

O cabeçalho SNAP, mediante a utilização do OUI 0x00-00-00, permite que a indicação dos protocolos superiores ocorra pelo *EtherType*.

Quando a indicação do protocolo superior for feita diretamente pelo LSAP, os valores dos campos DSAP e SSAP corresponde ao código do respectivo protocolo.

Quando se utiliza o IPv4 sobre Ethernet 802.3, é necessário utilizar o SNAP, uma vez que o ARP (*Address Resolution Protocol*) não possui código LSAP.

Conforme o item 3.2.7 da Seção 1 do padrão IEEE 802.3-2012, o campo de dados da PDU MAC (*MAC Client Data*) deve suportar pelo menos um dos seguintes comprimentos:

— 1500 octetos: quadros básicos;
— 1504 octetos: quadros *Q-tagged*;
— 1982 octetos: quadros envelopados.

Os quadros envelopados permitem a inclusão de prefixos e sufixos adicionais exigidos por protocolos de encapsulamento de camadas superiores, tais como PB (*Provider Bridges*) e PBB (*Provider Backbone Bridges*), os quais serão tratados mais adiante neste livro. O campo original de dados de cliente continua a ser de 1500 octetos.

Embora o padrão IEEE 802.3 tenha um limite máximo de 1500 octetos para o campo de dados de cliente do quadro, ressalva-se que quadros maiores, geralmente designados por *jumbo frames*, são suportados por fabricantes, particularmente para interfaces Ethernet em altas velocidades.

2.7.3. **Subcamada MAC**

A Cláusula 4 da Seção 1 do padrão IEEE 802.3-2012 especifica a subcamada MAC para redes Ethernet.

A subcamada MAC define um conjunto functional independente, construído sobre o conjunto functional dependente do meio físico definido pela camada PHY. A subcamada MAC está localizada abaixo da subcamada LLC (ou outro cliente MAC). Em conjunto, as subcamadas LLC e MAC representam as funções descritas no modelo OSI para a camada de enlace de dados (*Data Link Layer*).

A subcamada MAC desempenha as seguintes funções:

— Encapsulamento dos dados (TX e RX);
— Demarcação dos quadros;
— Endereçamento;
— Detecção de erros de transmissão da camada física;
— Gerenciamento do acesso ao meio;
— Alocação do meio (*collision avoidance*);
— Resolução de disputa no acesso ao meio (*collision handling*).

O modelo contempla uma subcamada opcional, denominada *MAC control*, localizada entre a subcamada LLC (ou outro cliente MAC) e a subcamada MAC. A especificação dessa subcamada encontra-se na Cláusula 31 da Seção 2 do padrão IEEE 802.3-2012. Esse assunto será tratado adiante neste capítulo.

O padrão define duas formas de operação da subcamada MAC:

— Modo *half-duplex*, no qual as estações alternam-se no acesso ao meio, utilizando os algoritmos CSMA/CD especificados;
— Modo *full-duplex*.

Quando um quadro é apresentado pelo cliente MAC para ser transmitido, a transmissão tem início assim que possível, mas em conformidade com algumas regras de adiamento. Existem regras para o modo *half-duplex* e regras para o modo *full-duplex*.

2.7.3.1. Modo Half-Duplex

Mesmo quando não tem nada a transmitir, a subcamada CSMA/CD MAC observa o meio físico para verificar se há tráfego presente, usando para isso o sinal *carrierSense* provido pela subcamada PLS (*Physical Layer Signalling*). Sempre que o meio encontra-se ocupado, essa subcamada dá preferência ao pacote em curso, atrasando para tanto qualquer transmissão própria pendente.

Após a passagem do último bit do pacote em curso (ou seja, no momento em que o *carrierSense* muda de TRUE para FALSE), o CSMA/CD continua a adiar a transmissão por um *interPacketGap* apropriado.

Caso ao final do *interPacketGap* haja um quadro à espera para ser transmitido, a transmissão é iniciada, sem importar o valor do *carrierSense*. Ao completar-se a transmissão (ou imediatamente, caso não houvesse nada a transmitir), a subcamada CSMA/CD MAC retorna a seu estado original de observação do sinal *carrierSense*.

Uma vez que a subcamada CSMA/CD tenha finalizado o adiamento e iniciado a transmissão, ainda permanence a possibilidade de haver uma disputa pelo meio. Podem ocorrer colisões até que a obtenção do meio tenha sido alcançada mediante o adiamento efetuado pelas subcamadas CSMA/CD de todas as demais estações.

As colisões são detectadas mediante a observação do sinal *collisionDetect* provido pela camada fisica. Quando se dectecta uma colisão durante a transmissão de um quadro, a transmissão não é interrompida de imediato. A transmissão continua até que termine a transmissão de bits adicionais especificados pelo parâmetro *jamSize* (contados a partir do *collisionDetect*).

Esse asseguramento da colisão, ou *jam*, garante que a duração da colisão seja suficiente para que todas as estações que estão transmitindo na rede detectem a colisão. O valor indicado para o *jamSize* para todas as camadas físicas operando no modo half-duplex (até inclusive 1Gbit/s) corresponde a 32 tempos de bit. Assim, por exemplo, para 10 Mbit/s, o *jamSize* é igual a 3,2 μs.

Quando uma tentativa de transmissão for interrompida devido a uma colisão, a subcamada CSMA/CD volta a tentar transmitir até que obtenha sucesso ou que tenha ocorrido um número de tentativas (*attemptLimit*) interrompidas devido a colisões.

O agendamento das retransmissões é determinado por um processo de aleatorização controlada, referida como "*truncated binary exponential backoff*."

Após uma colisão, cada estação que colidiu gera um número aleatório próprio (*r*), que se encontra em uma faixa especificada de valores.

Cada uma dessas estações aguarda um número de períodos igual a *r* que lhe corresponde, períodos esses referidos como *slot times*, até que possa tentar uma nova retransmissão.

Ao final de cada *slot time*, as estações que vão tentar retransmitir (ou transmitir no caso de um quadro novo na rede) devem acrescentar um tempo de espera igual um *interPacketGap*.

Registra-se que essas estações devem também observar o valor do *carrierSense* para verificar se existe ou não uma transmissão bem sucedida em curso, imediatamente após do final do *slot time*. Se esse valor for TRUE, as estações devem suspender as tentativas de transmissão/retransmissão, postergando-as para o final da transmissão do quadro em curso com o acréscimo de um *interPacketGap*.

O valor *r* para cada estação que colidiu é definido aleatoriamente dentro da faixa $0 \leq r \leq 2^k$, onde *k* é o número seqüencial da tentativa de retransmissão, cujo valor máximo é igual a 10.

Assim, se *k*=1 (primeira tentativa de retransmissão do quadro da estação considerada), o valor de *r* se encontra na faixa de {0,1}. Se *k*=2, o valor de *r* se encontra na faixa {0,1,2,3}, e assim sucessivamente. A partir da décima tentativa de retransmissão (*k*=10), o valor de *r* permanece na faixa {0,1,2,3,...1023}.

Observa-se que, como as faixas de valores de *r* aumentam com o número de tentativas de retransmissão e a escolha do valor de *r* nessas faixas ocorrem aleatoriamente, a probabilidade de sucesso nas retransmissões também aumenta.

Cada estação transmissora também limita o número de tentativas de retransmissão de um quadro para *k*=15 (o que representa a 16ª tentativa, incluindo aquela da primeira colisão). Após esse número de tentativas, a estação transmissora dá por encerrada a transmissão e descarta o quadro, registrando a ocorrência de um erro.

O valor de um *slot time* corresponde ao período mínimo que garanta que a colisão tenha sido detectada por todas as estações da rede.

Para 10 Mbit/s e 100 Mbit/s o valor do *slot time* corresponde a 512 tempos de bit. Assim, para 10 Mbit/s o seu valor é igual a 51,2 μs, enquanto que para 100 Mbit/s é igual 5,12 μs.

Para 1 Gbit/s, em função da alta velocidade, o valor do *slot time* foi ampliado para 4096 tempos de bit, do que resulta que o seu valor passa a ser igual a 4,096 μs.

Foi estabelecido que uma transmissão bem sucedida não deve terminar antes do momento em que ocorra o início das próximas tentativas de retransmissão. Assim, os *slots times*, medidos em tempos de bit, devem corresponder ao tamanho mínimo dos quadros.

Então, para 10 Mbit/s e 100 Mbit/s, em que os *slot times* correspondem a 512 tempos de bit, ou a 64 tempos de octeto, o tamanho mínimo de quadros é igual a 64 octetos. Para 1Gbps, cujo *slot time* corresponde a 4096 tempos de bit (ou a 512 tempos de octecto), o tamanho mínimo de quadros é igual a 512 octetos.

2.7.3.2. Modo Full-Duplex

A Seção 1 do padrão IEEE 802.3-2012 define dois tipos de subcamada MAC:

— A subcamada MAC especificada na Cláusula 4 dessa Seção;
— A subcamada *Simplified Full-Duplex MAC*, especificada no Anexo 4A dessa Seção.

A primeira delas está descrita neste subitem. A subcamada *Simplified Full-duplex MAC* será vista adiante neste capítulo.

No modo full-duplex não há disputa por um meio compartilhado. A camada física pode indicar à camada MAC que estão ocorrendo transmissões simultâneas das duas estações, mas, uma vez que essas transmissões não interferem mutuamente, uma camada MAC que opere em modo *full-duplex* não deve reagir a essas indicações da camada física.

Para a operação no modo *full-duplex*, são necessárias as seguintes condições:

— O meio físico deve ser capaz de suportar a recepção e a transmissão simultâneas;
— Deve haver exatamente duas estações na LAN para que o meio físico seja tratado como um *link* ponto a ponto, e, ao não haver disputa por um meio compartilhado, não há necessidade de utilizar os algoritmos de acesso múltiplo (ou seja, CSMA/CD);
— Ambas as estações na LAN devem estar configuradas para operação *full-duplex*.

A operação em modo *full-duplex* pode requerer um método de controle de fluxo. O controle de fluxo em Ethernet consiste em um mecanismo que permite que uma interface ou uma porta de um *switch* solicite uma pausa curta na transmissão dos quadros. O controle de fluxo será visto adiante neste capítulo.

No modo *full-duplex*, o CSMA/CD MAC não adia a transmissão de pacotes pendentes com base no sinal *carrierSense* da subcamada PLS. Em vez disso, usa a variável interna *transmitting* para manter o estado apropriado do MAC enquanto a transmissão de um quadro esteja em curso. Após a transmissão do ultimo bit desse quadro (quando a variável *transmitting* muda de TRUE para FALSE), o MAC continua a adiar a transmissão durante um *interPacketGap* apropriado.

2.7.3.3. InterPacketGap

O padrão IEEE 802.3-2012 define um período de tempo entre quadros Ethernet com o objetivo de prover um tempo de recuperação inter-quadros para outras subcamadas ethernet e para o meio físico. Esse período de tempo é referido como *Inter-Packet Gap* (IPG). O valor mínimo do IPG é denominado *interPacketGap* (que se trata de um parâmetro MAC).

O valor especificado para o *interPacketGap* para todas as interfaces físicas Ethernet é o que corresponde a 96 tempos de bit. Esse tempo é de 9,6 μs para 10 Mbit/s, de 0,96 μs para 100 Mbit/s, de 96 ns para 1 Gbit/s, de 9,6 ns para 10 Gbit/s de 2,4 ns para 40 Gbit/s e de 0,96 ns para 100 Gbit/s.

Como visto anteriormente neste capítulo, as regras para adiamento da transmissão de quadros nos modos *half-duplex* e *full-duplex* utilizam os valores de *interPacketGap*.

2.7.4. **A Subcamada MAC Control**

A subcamada *MAC Control*, de caráter opcional, especificada na Cláusula 31 da Seção 2 do padrão IEEE 802.3-2012 para uso com o MAC CSMA/CD, define o suporte a algumas funções de controle, sendo possível a incorporação de novas funções ao padrão no futuro.

As funções desempenhadas atualmente pela subcamada *MAC Control* encontram-se relacionadas na Tabela 31A-1 do Anexo 31A da Seção 2 do padrão IEEE 802.3-2012.

Atualmente, encontram-se especificadas para atendimento por meio da subcamada MAC Control, as seguintes funções:

— Função PAUSE;
— Função PFC (*Priority-based Flow Control*);
— Funções do protocolo *GEPON Multipoint MAC Control* (GEPON MPCP);
— Funções do protocolo *10G-EPON Multipoint MAC Control* (10G-EPON MPCP).

Essas funções serão abordadas adiante neste capítulo.

A arquitetura das subcamadas Ethernet, incluindo a subcamada *MAC Control*, dentro do contexto da arquitetura OSI encontra-se representada na Figura 2-16.

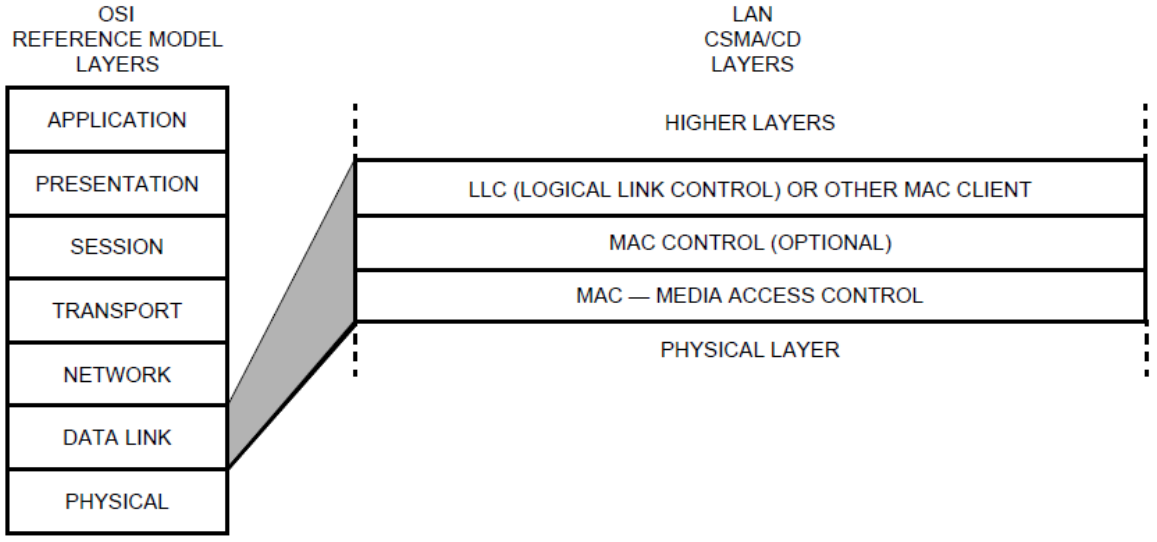


Figura 2-16 – Arquitetura Ethernet incluindo a subcamada *MAC Control*.

O formato do quadro *MAC Control* (ou seja, do quadro MAC englobando a subcamada *MAC Control*) encontra-se representado na Figura 2-17.

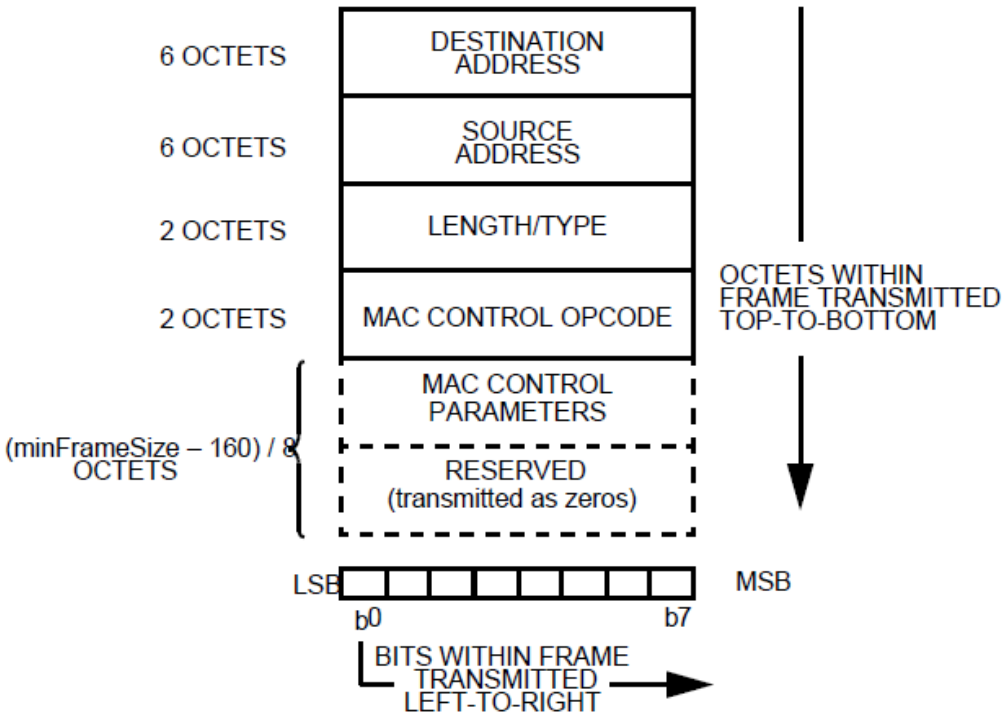


Figura 2-17 – Formato do quadro MAC Control.

O campo *MAC Control Opcode* identifica as funções atendidas pela Subcamada MAC Control.

O campo referente ao endereço de destino pode se constituir em um endereço individual ou em um endereço *multicast* (incluindo o endereço *broadcast*), dependendo do valor de *Opcode*. Para os *Opcodes* referentes à função PAUSE e à função PFC (*Priority-based Flow Control*) do *MAC Control*, por exemplo, é utilizado o endereço MAC *multicast* 01-80-C2-00-00-01.

O campo *Length/Type*, no caso significando *Type*, utiliza o código *EtherType* de valor 88-08, atribuído universalmente para a camada *MAC Control* de LANs CSMA/CD.

O campo *MAC Control Parameters* contém os parâmetros específicos para cada *Opcode*.

2.7.5. **Controle de Fluxo em Redes Ethernet**

Serão abordadas neste subitem as duas funções da subcamada MAC Control que dizem respeito a controle de fluxo em redes Ethernet. Essas funções são as seguintes:

— A função PAUSE, com o *opcode* 00-01;
— A função PFC (*Priority-based Flow Control*), com o *opcode* 01-01.

A função PAUSE, referente à operação *PAUSE MAC Control*, encontra-se especificada no Anexo 31B da Seção 2 do padrão IEEE 802.3-2012.

Essa operação possibilita que seja determinado ao receptor de um quadro PAUSE que pare de enviar quadros, com exceção de quadros de controle, por um período de tempo indicado no campo *MAC Control Parameters* do quadro PAUSE.

Um cliente do *MAC Control* que deseja solicitar essa operação deve gerar uma primitiva de serviço *MA_CONTROL.request* com esse propósito.

O endereço de destino indicado nessa primitiva de serviço é o endereço *multicast* reservado 01-80-C2-00-00-01.

O valor hexadecimal de *Opcode* indicado é igual a 00-01, que corresponde à operação PAUSE.

Essa primitiva de serviço deve indicar, no campo *MAC Control Parameters*, também o período de tempo que o cliente do *MAC Control* definiu para a duração da interrupção do envio de quadros de serviço. Esse tempo é medido em *pause_quanta*, onde um *pause_quantum* equivale a 512 períodos de bit da implementação em questão. A gama de valores possíveis é de 0 a 65535 *pause_quanta*.

A função PFC, referente à operação *PFC MAC Control*, encontra-se especificada no anexo 31D da Seção 2 do padrão IEEE 802.3-2012.

O padrão IEEE 802.1Q, em sua última versão (IEEE 802.1Q-2014), incorporou, em sua Cláusula 36 (*Priority-based Flow Control*), a definição mais detalhada da função PFC.

Essa operação possibilita que seja determinado ao receptor de um quadro PFC que pare de enviar quadros de dados com as prioridades indicadas no campo *MAC Control Parameters* dessa função, por um período de tempo também indicado nesse campo.

Da mesma forma que na função PAUSE, o endereço MAC de destino indicado pela função PFC é o endereço *multicast* 01-80-C2-00-00-01.

O valor hexadecimal de *opcode* indicado deve ser igual a 01-01, que corresponde a operação PFC.

Para o caso particular de transmissão *full-duplex* ponto a ponto, foi definida uma técnica de controle de fluxo referida com *Full-duplex Carrier Deferral,* no anexo 4A da Seção 1 do padrão IEEE 802.3-2012.

Essa técnica baseia-se no fato de que o algoritmo CSMA/CD requer, para a transmissão *half-duplex,* que o MAC monitore os sinal *carrierSense* enviado pela camada física. Se esse sinal indicar que o meio físico está ocupado, a eventual transmissão de quadros deve ser retardada. Essa indicação ocorre pela presença do valor TRUE nesse sinal.

A transmissão só poderá ocorrer após a disponibilização do meio físico, o que será indicado pela presença do valor FALSE no sinal *carrierSense*.

Como na transmissão *full-duplex* ponto a ponto não existe necessidade dessa indicação, a variável *carrierSense* passa a ser utilizada, na técnica *Full-Duplex Carrier Deferral,* como uma indicação que o meio físico não está disponível por uma

variedade de possíveis razões, o que permite o retardamento da transmissão de quadros.

A técnica *Full-Duplex Carrier Deferral* é, portanto, um mecanismo genérico de controle de fluxo e de velocidade, definido no próprio MAC com relação direta com a camada física.

2.7.6. **Simplified Full-Duplex MAC**

O anexo 4A da Seção 1 do padrão IEEE 802.3-2012 especifica uma versão simplificada do MAC completo definido na cláusula 4 da mesma Seção 1 do padrão, versão essa denominada *Simplified Full-Duplex Medium Access Control.* Essa versão se destina às redes que não requerem o modo operacional *half-duplex* (semi-duplex).

O MAC *full-duplex* simplificado provê as mesmas interfaces de subcamada que o MAC completo, preservando inclusive todas as indicações de estado. Dessa forma, não existe o problema da incompatibilidade entre esses dois tipos de MAC.

Ao contrário do MAC completo, no entanto, o MAC *full-duplex* não é sobrecarregado pela necessidade de suportar também o modo de operação semi-duplex, o que simplifica o MAC e reduz os seus custos.

Uma outra vantagem do MAC *full-duplex* simplificado é que ela explicitamente inclui o suporte à técnica *full-duplex Carrier Deferral* para controle de fluxo e de velocidades. Em consequência, torna-se desnecessária a utilização, com o MAC *full-duplex* simplificado, da problemática técnica de ajuste (*stretching*) referida como IPG (*Inter-Packet Gap*).

A técnica IPG, também definida na cláusula 4 da Seção 1 do IEEE 802.3-2012, é um mecanismo de adaptação de velocidades que reduz a taxa de transmissão de quadros MAC de forma tal que, ao longo do tempo, o número de bits transmitido pelo MAC iguala o número que pode ser transmitido pela camada física.

Um exemplo é o ajustamento da interface MAC operando a 10Gbit/s à camada física SONET OC-192 que transmite na velocidade de *payload* de 9,9532 Gbit/s.

O uso do MAC *full-duplex* simplificado é obrigatório nas EPONs, que serão tratadas mais adiante neste capítulo.

As redes P2P EFM (*Ethernet in the First Mile* ponto a ponto), ópticas ou utilizando pares de cobre, podem, opcionalmente, utilizar a subcamada MAC convencional. É recomendável, no entanto, a utilização do MAC *full-duplex* simplificado indiscriminadamente para todos os sistemas de sinais definidos para as redes EFM.

O MAC *full-duplex* simplificado é indicado para aplicação em qualquer padrão com camada física *full-duplex*, em adição à sua aplicabilidade para os sistemas de sinais EFM.

2.7.7. **Slow Protocols**

Conforme o Anexo 57A da Seção 5 do padrão IEEE 802.3-2012, existem duas class

es distintas de protocolos utilizados para controlar diferentes aspectos da operação de dispositivos IEEE 802.3-2012:

— Protocolos tais como o destinado à operação PAUSE na subcamada *MAC Control*, que necessitam processar as PDUs recebidas e respondê-las rapidamente para evitar degradação do desempenho;
— Protocolos tais como o LACP (*Link Aggregation Control Protocol*) e protocolos de OAM (Operação Administração e Manutenção), com requisitos menos rigorosos de frequência e de latência, que podem ser implementados em *software*, referidos como protocolos lentos (*slow protocols*).

Por *default*, o número máximo de quadros que podem ser transmitidos em um segundo em um protocolo lento é igual a 10. É recomendado também que o tamanho máximo de um quadro de um protocolo lento seja limitado a 128 octetos.

Os protocolos lentos são agrupados e tratados separadamente por processos próprios.

Como regra geral, as PDUs de um protocolo lento, contidas em um quadro MAC desse protocolo, devem ser endereçadas ao endereço *multicast* reservado *Slow_Protocols_Multicast address* (01-80-C2-00-00-02). Fica aberta a possibilidade, contudo, da especificação de endereços de destino *unicast* para futuras definições de novos subtipos de protocolos lentos.

O campo *Length/Type* do quadro MAC de um protocolo lento adota o significado *Type*, e o código *EtherType* utilizado, referido como valor *Slow_Protocols_Type*, é 88-09.

O primeiro octeto do campo de informação do cliente MAC, que se localiza logo após o campo *Length/Type*, é o identificador do subtipo de protocolo lento. Assim, por exemplo, o valor 1 identifica o LACP, o valor 2 identifica o *Link Aggregation - Marker Protocol* e o valor 3 identifica o protocolo OAM.

2.8. A CAMADA FÍSICA DE REDES ETHERNET IEEE 802.3

As redes *Ethernet* 802.3 apresentam um grande número de especificações de meios físicos, que variam em termos de velocidade, forma de codificação do sinal e natureza do meio, dentre outros parâmetros.

Alguns desses padrões dizem respeito a redes de acesso distribuído em barramento ou em *Ethernet hubs*, enquanto outros prestam-se a configurações ponto a ponto (caso geral de redes *switched Ethernet*).

As interfaces podem operar em modo *half-duplex* ou *full-duplex*. Em modo *half-duplex*, o CSMA-CD, como vimos, disciplina o acesso ao meio. O modo *full-duplex*, por outro lado, exige que as seguintes condições sejam preenchidas:

— O meio físico suporta operação *full-duplex*;
— Somente há duas estações na LAN (configuração ponto a ponto);
— Ambas estações são configuradas para operação *full-duplex*.

Estas condições são atendidas, por exemplo, em redes *switched Ethernet*, desde que as interfaces se deixem programar para operação *full-duplex*.

Na operação *half-duplex*, mesmo quando não há nada a transmitir, a subcamada CSMA/CD monitora o meio físico pela observação da variável *carrierSense*.

Na operação *full-duplex* o CSMA/CD MAC não leva em consideração a variável *carrierSense*. Em vez disso, ele usa a variável interna *transmitting* para manter o estado MAC apropriado enquanto dura a transmissão.

2.8.1. Arquitetura da Camada Física Ethernet 802.3

A Figura 2-18 apresenta a arquitetura do padrão IEEE 802.3-2012 e sua relação com o modelo de referência OSI. Nessa figura, a camada física correspondente a cada agrupamento de velocidades pode ser visualizada.

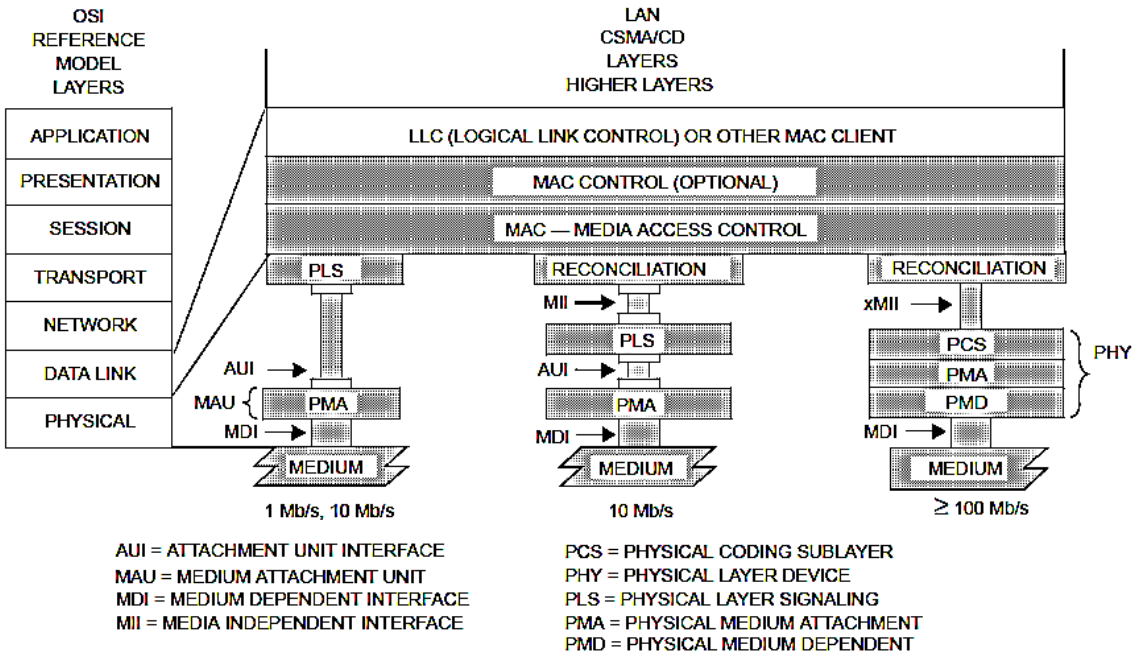


Figura 2-18 – Relação entre o padrão IEEE 802.3-2012 e o modelo de referência OSI.

Nessa figura, o termo xMII é usado como um termo genérico nas implementações das *Media Independent Interfaces* que operam a 100 Mbit/s ou a velocidades superiores. Por exemplo, para 100 Mbit/s a interface denomina-se MII, para 1 Gbit/s denomina-se GMII, para 10 Gbit/s denomina-se XMII, para 40 Gbit/s denomina-se XLGMII e para 100 Gbit/s denomina-se CGMII.

Na medida em que as velocidades vão aumentando, as arquiteturas das respectivas camadas físicas podem sofrer ligeiras alterações. Em algumas camadas físicas, por exemplo, não existe a subcamada PMD. Além disso, algumas camadas físicas utilizam uma subcamada de auto-negociação (AN) em caráter obrigatório ou opcional, imediatamente acima do meio físico.

Um outro exemplo é a utilização da subcamada física opcional WIS (*WAN Interface Sublayer*) para a interface física 10GBASE-W.

Para ilustrar esses outros tipos de interface física, na Figura 2-19 encontra-se representada a relação entre o modelo IEEE 802.3-2012 e o modelo de referência OSI para 40 Gbit/s e 100 Gbit/s.

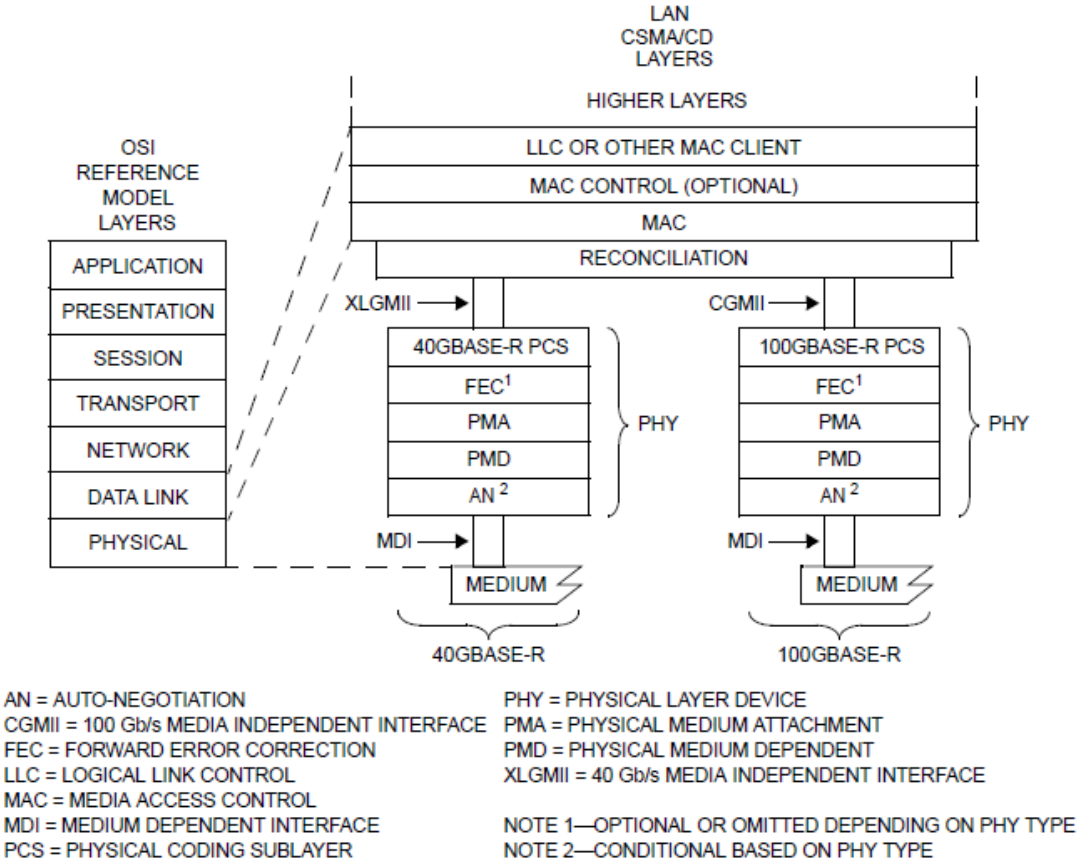


Figura 2-19 – Posicionamento de 40 Gbit/s e 100 Gbit/s na arquitetura OSI.

2.8.2. **Redes Operando de 1 Mbit/s a 10 Mbit/s**

A seção 1 do padrão IEEE 802.3-2012 define um quadro conceitual básico de redes Ethernet IEEE 802.3, particularmente no que concerne à especificação da camada física, e especifica as redes operando de 1 Mbit/s a 10 Mbit/s.

Redes *Ethernet* operando nessa faixa de velocidades suportam os modos de operação *half-duplex* e *full-duplex*. Existem, nessa faixa de velocidades, redes *Ethernet* operando nas velocidades de 1 Mbit/s, 2 Mbit/s e 10 Mbit/s, normalmente em *baseband*.

Foram especificadas as seguintes camadas físicas nessa faixa de velocidades operacionais:

— 1BASE5: camada física para LAN CSMA-CD a 1 Mbit/s, sobre dois pares de fios telefônicos trançados;
— 2BASE-TL: camada física para *link* ponto a ponto de até 5,696 Mbit/s, sobre um único par de fios de cobre;
— 10BASE2: camada física para LAN CSMA-CD a 10 Mbit/s, sobre cabo coaxial RG 58;

— 10BASE5: camada física para LAN CSMA-CD a 10 Mbit/s, sobre cabo coaxial (*thicknet*);
— 10BASE-T: camada física para LAN CSMA-CD a 10 Mbit/s, sobre dois pares de fios telefônicos trançados;
— 10BASE-Te: versão de 10BASE-T com maior eficiência de energia;
— 10BASE-F: camada física para LAN CSMA-CD a 10 Mbit/s, sobre cabo de fibra óptica;
— 10BASE-FB port: porta em um repetidor com uma MAU 10BASE-FB para conexão a outra porta similar em outro repetidor;
— 10BASE-FB segment: segmento de fibra óptica que provê uma conexão ponto a ponto entre duas portas 10BASE-FB em repetidores;
— 10BASE-FL segment: segmento de fibra óptica que provê uma conexão ponto a ponto entre duas MAUs 10BASE-FL;
— 10BASE-FP segment: segmento de junção de fibra óptica, que inclui uma estrela 10BASE-FP Star e todos os pares de fibras a ela conectados;
— 10BASE-FP Star: dispositivo passivo que se usa para acoplar pares de fibra entre si para formar um segmento 10BASE-FP, distribuindo os sinais ópticos recebidos em uma porta qualquer a todas as portas de saída(inclusive a própria);
— 10BROAD36: camada física para LAN CSMA-CD a 10 Mbit/s sobre um cabo *broadband*;
— 10PASS-TS: camada física para para *link* ponto a ponto de até 10 Mbit/s, sobre um par de fios de cobre.

2.8.3. **Redes Operando a 100 Mbit/s (Fast Ethernet)**

As redes *Ethernet* IEEE 802.3 operando a 100 Mbit/s, em *baseband,* foram especificadas na Seção 2 do padrão IEEE 802.3-2012.

Redes *Ethernet* operando a 100 Mbit/s suportam os modos de operação *half-duplex* e *full-duplex.*

Foram especificadas as seguintes camadas físicas para essa velocidade operacional:

— 100BASE-BX10: camada física para *link* ponto a ponto de 100 Mbit/s, sobre uma fibra óptica monomodo (com duas especificações, a 100BASE-BX10-D e a 100BASE-BX10-U);
— 100BASE-FX: camada física para LAN CSMA-CD a 100 Mbit/s, sobre duas fibras ópticas multimodo;
— 100BASE-LX10: camada física para *link* ponto a ponto de 100 Mbit/s, sobre duas fibras ópticas monomodo;
— 100BASE-T: camada física para LAN CSMA-CD a 100 Mbit/s;
— 100BASE-T2: camada física para LAN CSMA-CD a 100 Mbit/s, sobre pares de cabos balanceados de Categoria 3 ou superior;

— 100BASE-T4: camada física para LAN CSMA-CD a 100 Mbit/s, sobre quatro pares de cabos trançados não blindados (UTP) de Categoria 3, 4 ou 5;
— 100BASE-TX: camada física para LAN CSMA-CD a 100 Mbit/s, sobre dois pares de cabos trançados de Categoria 5 não blindados (UTP) ou blindados (STP);
— 100BASE-X: camada física para LAN CSMA-CD a 100 Mbit/s, utilizando a subcamada PMD (*Physical Medium Dependent*) e a interface MDI (*Medium Dependent Interface*) do grupo de padrões ISO/IEC 9314.

2.8.4. **Redes Operando a 1000 Mbit/s (Gigabit Ethernet)**

As redes *Ethernet* IEEE 802.3 operando a 1000 Mbit/s, em *baseband*, foram especificadas na Seção 3 do padrão IEEE 802.3-2012. Tais redes são referidas como redes *Gigabit Ethernet*.

As redes *Ethernet* a 1000 Mbit/s suportam os modos de operação *half-duplex* e *full-duplex*, e representam as redes de maior velocidade que possibilitam a operação do CSMA-CD (ou seja, possibilitam a contenda para o acesso ao meio).

Foram especificadas as seguintes camadas físicas para essa velocidade operacional:

— 1000BASE-BX10: camada física para *link* ponto a ponto a 1000 Mbit/s, sobre uma fibra óptica monomodo;
— 1000BASE-X: especificação para LAN CSMA-CD a 1000 Mbit/s, que utiliza uma camada física derivada do padrão ANSI-X3.230-1994 (*Fibre Channel - Physical and Signaling Interface*: *FC-PH*);
— 1000BASE-CX: camada física 1000BASE-X, sobre montagens de cabos *jumper* especiais de cobre balanceados e blindados;
— 1000BASE-KX: camada física IEEE 802.3 para 1000 Mbit/s, utilizando codificação 1000BASE-X sobre um *backplane* elétrico;
— 1000BASE-LX: camada física 1000BASE-X, usando dispositivos laser com comprimento de onda longo sobre fibras ópticas monomodo e multimodo;
— 1000BASE-SX: camada física 1000BASE-X, utilizando dispositivos *laser* com comprimento de onda curto sobre fibras ópticas multimodo;
— 1000BASE-LX10: camada física para *link* ponto a ponto de 1000 Mbit/s, sobre duas fibras ópticas monomodo ou multimodo;
— 1000BASE-PX10: camada física para *link* ponto a ponto de 1000 Mbit/s, sobre uma fibra óptica monomodo, com alcance de até 10 km;
— 1000BASE-PX20: camada física para *link* ponto a ponto de 1000 Mbit/s, sobre uma fibra óptica monomodo, com alcance de até 20 km;

— 1000BASE-T: camada física para LAN CSMA-CD a 1000 Mbit/s, que utiliza quatro pares de cabos de cobre balanceados e blindados de Categoria 5.

2.8.5. **Redes Operando a 10 Gbit/s**

As redes *Ethernet* IEEE 802.3 operando a 10 G*bit/s* foram especificadas na Seção 4 do padrão IEEE 802.3.

Essas redes operam em *baseband* e suportam exclusivamente o modo de operação *full-duplex*.

Os objetivos dessas redes são os seguintes:

— Suportar PMDs para LANs operando a 10 G*bit/s* e PMDs para WANs operando na velocidade de SONET STS-192c/SDH VC-4-64c.
— Suportar redes construídas com fibras ópticas que atendam à norma ISO/IEC 11801:1995.
— Suportar uma extensão nominal de até 40 km;
— Suportar operação em montagens de cabos *twinaxial* para aplicações em distribuidores de fiação e *data center*;
— Suportar operação em determinados meios de cobre especificados na norma ISO/IEC 11801:2002;
— Suportar taxas de erros de bits (BER) de 10^{-12} ou melhores.

A camada física das redes 10 *Gigabit Ethernet* utiliza as seguintes opções de conjuntos ópticos:

— 10GBASE-E: especificação de PMD IEEE 802.3 para transmissão serial a 10 Gbit/s utilizando comprimento de onda extra longo;
— 10GBASE-L: especificação de PMD IEEE 802.3 para transmissão serial a 10 Gbit/sutilizando comprimento de onda longo;
— 10GBASE-S: especificação de PMD IEEE 802.3 para transmissão serial a 10 Gbit/s utilizando comprimento de onda curto.

Além dos conjuntos ópticos, algumas interfaces elétricas são definidas para as redes 10 *Gigabit Ethernet.*

No que concerne à codificação, as redes 10 *Gigabit Ethernet* utilizam as seguintes opções:

— 10GBASE-R: subcamada de codificação da camada física IEEE 802.3 para operação serial a 10 Gbits/s;
— 10GBASE-W: subcamada de codificação da camada física IEEE 802.3 para operação serial a 10 Gbit/s compatível em velocidade binária e em formato com a SONET STS-192c/SDH VC-4-64c;
— 10GBASE-X: subcamada de codificação da camada física IEEE 802.3 para operação serial a 10 Gbit/s com XAUI e PMDs de quatro pistas.

Foram especificadas as seguintes camadas físicas baseadas em interfaces ópticas para essa velocidade operacional:

— 10GBASE-ER: camada física para 10 Gbit/s, utilizando codificação 10GBASE-R e conjunto óptico 10GBASE-E;
— 10GBASE-EW: camada física para 10 Gbit/s, utilizando codificação 10GBASE-W e conjunto óptico 10GBASE-E;
— 10GBASE-LR: camada física para 10 Gbit/s, utilizando codificação 10GBASE-R e conjunto óptico 10GBASE-L;
— 10GBASE-LRM: camada física para 10 Gbit/s, utilizando codificação 10GBASE-R e conjunto óptico de comprimento de onda longo para fibra multímodo;
— 10GBASE-LW: camada física para 10 Gbit/s, utilizando codificação 10GBASE-W e conjunto óptico 10GBASE-L;
— 10GBASE-LX4: camada física para 10 Gbit/s, utilizando codificação 10GBASE-X sobre quatro pistas WWDM sobre fibra óptica multimodo;
— 10GBASE-SR: camada física para 10 Gbit/s, utilizando codificação 10GBASE-R e conjunto óptico 10GBASE-S;
— 10GBASE-SW: camada física para 10 Gbit/s, utilizando codificação 10GBASE-W e conjunto óptico 10GBASE-S.

Além disso, foram especificadas as seguintes camadas físicas baseadas em interfaces elétricas para essa velocidade operacional:

— 10GBASE-T: camada física de 10 Gbit/s, utilizando quatro pares de cabo balanceado de cobre de Classe E ou Classe F;
— 10GBASE-CX4: camada física para 10 Gbit/s, utilizando codificação 10GBASE-X em quatro pistas sobre cabeamento de cobre balanceado blindado;
— 10GBASE-KX4: camada física para 10 Gbit/s, utilizando codificação 10GBASE-X, sobre um *backplane* elétrico;
— 10GBASE-KR: camada física para 10 Gbit/s, utilizando codificação 10GBASE-R, sobre um *backplane* elétrico.

Foram também especificadas, na Seção 5 do padrão IEEE 802.3-2012, as seguintes camadas físicas para operação ponto a multiponto (10G-EPON):

— 10/1GBASE-PRX: camada física de 10 Gbit/s *downstream* e 1 Gbit/s *upstream*(10/1G-EPON) para um enlace ponto a multiponto sobre uma fibra óptica monomodo;
— 10GBASE-PR: camada física de 10 Gbit/s (10/10G-EPON) para um enlace ponto a multiponto sobre uma fibra óptica monomodo;

2.8.6. **Redes Operando a 40 e a 100 Gbit/s**

As redes *Ethernet* IEEE 802.3 operando a 40 e a 100 G*bit/s* foram especificadas na Seção 6 do padrão IEEE 802.3-2012.

Os objetivos dessas redes são os seguintes:

— Suportar a MAC Ethernet full-duplex;
— Preservar o formato de quadro e os tamanhos máximo e mínimo do quadro MAC definido no padrão IEEE 802.3 MAC;
— Suportar taxas de erros de bits (BER) de 10^{-12} ou melhores;
— Suportar OTN (*Optical Transport Network*).
— Suportar camadas físicas a 40 Gbit/s sobre os seguintes meios:
  - Pelo menos 10 km sobre fibra monomodo (SMF);
  - Pelo menos 2 km sobre fibra monomodo (SMF);
  - Pelo menos 100 m sobre fibra multimodo (MMF) OM3;
  - Pelo menos 7 m sobre um conjunto de cabos de cobre;
  - Pelo menos 1 m sobre um *backplane*;
— Suportar camadas físicas a 100 Gbit/s sobre os seguintes meios:
  - Pelo menos 40 km sobre fibra monomodo (SMF);
  - Pelo menos 10 km sobre fibra monomodo (SMF);
  - Pelo menos 100 m sobre fibra multimodo (MMF) OM3;
  - Pelo menos 7 m sobre um conjunto de cabos de cobre;

No que concerne à codificação, as redes 40 e 100 *Gigabit Ethernet* utilizam as subcamadas de codificação da camada física 40GBASE-R e 100GBASE-R, baseadas em codificação 64B/66B, respectivamente, e contêm uma subcamada de correção de erros (FEC) opcional.

As camadas físicas definidas para 40 Gbit/s são as seguintes:

— 40GBASE-CR4: Codificação 40GBASE-R em quatro pistas em cabo de cobre STP, com até 7 m;
— 40GBASE-FR: Codificação 40GBASE-R em um comprimento de onda em fibra óptica monomodo, com alcance de pelo menos 2 km;
— 40GBASE-KR4: Codificação 40GBASE-R em quatro pistas em um *backplane* elétrico, com alcance de pelo menos 1 m;
— 40GBASE-LR4: Codificação 40GBASE-R em quatro pistas WDM em fibra óptica monomodo, com alcance de pelo menos 10 km;
— 40GBASE-SR4 Codificação 40GBASE-R em quatro pistas de fibra multimodo, com alcance de pelo menos 100 m.

As camadas físicas definidas para 100 Gbit/s são as seguintes:

— 100GBASE-CR10: Codificação 100GBASE-R em dez pistas em cabo de cobre STP, com até 7 m;

— 100GBASE-ER4: Codificação 100GBASE-R em dez pistas WDM em fibra óptica monomodo, com alcance de pelo menos 40 km;
— 100GBASE-LR4: Codificação 100GBASE-R em dez pistas WDM em fibra óptica monomodo, com alcance de pelo menos 10 km.
— 100GBASE-SR10: Codificação 100GBASE-R em dez pistas de fibra multimodo, com alcance de pelo menos 100 m.

2.9. FORWARD ERROR CORRECTION (FEC) PARA BASE-R PHYS

Anteriormente à utilização do mecanismo FEC (*Forward Error Correction*) em camadas físicas óticas Ethernet, algumas camadas físicas Ethernet baseadas em cabos de cobre já utilizavam esse mecanismo. São exemplos a PHY 10PASS-TS e a PHY 1000BASE-T.

A Seção 5 do padrão IEEE 802.3-2012 especifica a utilização opcional de uma subcamada FEC para dois grupos de PHYs ópticas:

— 1000BASE-PX PHYs (P2MP PHYs), na Cláusula 65;
— BASE-R PHYs, na Cláusula 74.

O uso de FEC para as 1000BASE-PX PHYs será abordado adiante neste capítulo, na parte referente às tecnologias ópticas ponto a multiponto em redes EFM (*Ethernet in The First Mile*).

O presente subitem objetiva a abordagem da utilização opcional de FEC em BASE-R PHYs, o que foi especificado na cláusula 74 do padrão acima mencionada.

Essa especificação engloba a 10BASE-R PHYs e outras BASE-R PHYs (40GBASE-R e 100GBASE-R, na presente versão do padrão). Foram citadas no padrão, por exemplo, as PHYs P2P 10GBASE-KR, 40GBASE-KR4, 40GBASE-CR4 e 100GBASE-CR10 como usuárias opcionais de FEC.

A Cláusula 76 da Seção 5 do padrão IEEE 802.3-2012, por sua vez, especifica, em complemento à Cláusula 74, o uso da codificação FEC para as PHYs 10GBASE-PR e 10/1GBASE-PRX. Tais PHYs, como será visto adiante, são P2MP PHYs utilizadas em 10G-EPONS.

Os principais objetivos para o uso de FEC em BASE-R PHYs são os seguintes:

— Suportar o modo de operação full-duplex do MAC;
— Suportar as subcamadas PCS, PMA e PMD definidas para 10GBASE-R;
— Suportar um objetivo de BER (*Bit Error Rate*) de 10^{-12} ou melhor.

A subcamada FEC BASE-R opcional é inserida na camada física entre as subcamadas PCS e PMA. Essa subcamada provê um ganho de codificação que possibilita aumentar o *link budget* óptico e melhorar o BER para acima de 10^{-12}.

2.10. ETHERNET IN THE FIRST MILE (EFM)

A EFM (*Ethernet in the First Mile*), isto é, o uso redes Ethernet no acesso entre usuários e provedores,foi inicialmente especificada pelo padrão IEEE 802.3ah, emitido em 2004, e posteriormente incorporada na Seção 5 do padrão IEEE 802.3-2012.

O processo de definição da EFM teve início em 2000, através de uma iniciativa do IEEE 802 LAN/MAN *Standards Commitee* (LMSC).

A EFM combina um conjunto mínimo de extensões às subcamadas MAC e Controle MAC do padrão IEEE 802.3-2012, com uma família adicional de camadas físicas. Tais camadas físicas dizem respeito à utilização, como meio físico, de fibras ópticas e cabos de cobre do padrão de voz.

Nesse contexto foram definidas subcamadas PMD(*Physical Medium Dependent*) para conexões ponto a ponto (P2P) em redes de acesso a usuários, tanto em termos de fibras ópticas quando de cabos de cobre.

Foi também introduzido o conceito de EPONs (*Ethernet Passive Optical Networks*), no qual se pressupõe a implementação de topologias ponto a multiponto (P2MP) com *splitters* ópticos passivos. Tais topologias P2MP utilizam extensões às subcamadas Controle MAC e Reconciliação (RS) da Seção 5 do padrão IEEE 802.3-2012, além PMDs para fibras ópticas compatíveis com essa tecnologia.

Finalmente, o padrão incorpora a definição de um mecanismo de OAM específico para as redes EFM.

A EFM é também apropriada para a conexão de usuários, corporativos ou residenciais, a ISPs (*Internet Service Providers*), constituindo assim uma solução alternativa para acesso de banda larga à *Internet*.

Ressalta-se também a possibilidade de utilização dos padrões EFM para acesso a redes *Carrier Ethernet* e, na medida em que isso se viabilize,na constituição interna dessas redes.

Considerando-se o contexto de utilização da EFM, pressupõe-se que as interfaces físicas definidas pelo padrão se destinem à transmissão *Ethernet full-duplex*. Um outro pressuposto é o de que essas interfaces físicas possibilitem a transmissão *Ethernet* no modo nativo, como rede de transporte, ou seja, diretamente aplicada sobre os meios físicos.

2.10.1. **Alternativas de Redes de Acesso a Usuários**

Uma rede de acesso a usuários destina-se à interconexão de equipamentos de usuários a um ponto de presença (POP) de um provedor. O POP pode representar a terminação de um circuito ou de uma rede comutada de um provedor de serviços de rede, assim como de uma rede de um ISP.

O objetivo da EFM é a definição de padrões para a utilização de quadros Ethernet contendo datagramas IP nas redes de acesso a usuários, tendo como meio de transmissão fibras ópticas ou cabos de cobre.

Uma alternativa ao uso de fibras ópticas ou de cabos de cobre para a transmissão de quadros Ethernet em redes de acesso a usuários, é a utilização de tecnologias *wireless* de banda larga.

A primeira grande experiência de transmissão *wireless* de quadros MAC Ethernet teve como suporte o padrão IEEE 802.11b (*Wireless LAN Medium Access Control and Physical Layer Specifications*), que pode possibilitar transmissões nominais de até 11 Mbit/s.

Na esteira da grande aceitação do padrão IEEE 802.11b para o acesso *wireless* à *Internet*, foram emitidos os padrões IEEE 802.15 e IEEE 802.16.

O padrão IEEE 802.15 destina-se à especificação das WPANs (*Wireless Personal Area Networks*), enquanto o padrão IEEE 802.16 é endereçado às WMANs (*Wireless Metropolitan Area Networks*).

Ocorre hoje uma intensa atividade no IEEE 802 LMSC, e em outros organismos de padronização, no sentido da definição de outros padrões para redes de acesso *wireless*.

2.10.2. **Tecnologias e Sistemas de Sinais de camada física em EFM**

A Seção 5 do padrão IEEE 802.3-2012 define os seguintes três tipos de tecnologia para uso na EFM:

- Tecnologias ópticas ponto a ponto (P2P);
- Tecnologias ópticas ponto a multiponto (P2MP);
- Tecnologias com cabos de cobre ponto a ponto.

Para cada um desses tipos de tecnologia foram especificados diferentes sistemas de sinais de camada física, conforme a Figura 2-20.

Name	Location	Rate[a]	Nominal reach (km)	Medium	Clause
100BASE-LX10	ONU/OLT[b]	100 Mb/s	10	Two single-mode fibers	58
100BASE-BX10-D	OLT	100 Mb/s	10	One single-mode fiber	58
100BASE-BX10-U	ONU				
1000BASE-LX10	ONU/OLT[b]	1000 Mb/s	10 0.55	Two single-mode fibers Two multimode fibers	59
1000BASE-BX10-D	OLT	1000 Mb/s	10	One single-mode fiber	59
1000BASE-BX10-U	ONU				
1000BASE-PX10-D	OLT	1000 Mb/s	10	One single-mode fiber PON	60
1000BASE-PX10-U	ONU				
1000BASE-PX20-D	OLT	1000 Mb/s	20	One single-mode fiber PON	60
1000BASE-PX20-U	ONU				
10/1GBASE-PRX-D1	OLT	10 Gb/s (tx) 1000 Mb/s (rx)	10	One single-mode fiber PON	75
10/1GBASE-PRX-U1	ONU	1000 Mb/s (tx) 10 Gb/s (rx)			
10/1GBASE-PRX-D2	OLT	10 Gb/s (tx) 1000 Mb/s (rx)	20	One single-mode fiber PON	75
10/1GBASE-PRX-U2	ONU	1000 Mb/s (tx) 10 Gb/s (rx)			
10/1GBASE-PRX-D3	OLT	10 Gb/s (tx) 1000 Mb/s (rx)	20	One single-mode fiber PON	75
10/1GBASE-PRX-U3	ONU	1000 Mb/s (tx) 10 Gb/s (rx)			
10GBASE-PR-D1	OLT	10 Gb/s	10	One single-mode fiber PON	75
10GBASE-PR-U1	ONU				
10GBASE-PR-D2	OLT	10 Gb/s	20	One single-mode fiber PON	75
10GBASE-PR-U1	ONU				
10GBASE-PR-D3	OLT	10 Gb/s	20	One single-mode fiber PON	75
10GBASE-PR-U3	ONU				
10PASS-TS-O	CO[c]	10 Mb/s[d]	0.75[e]	One or more pairs of voice grade copper cable	62
10PASS-TS-R	Subscriber[c]				
2BASE-TL-O	CO[c]	2 Mb/s[f]	2.7[g]	One or more pairs of voice grade copper cable	63
2BASE-TL-R	Subscriber[c]				

Figura 2-20– Sistemas de sinais da camada física na EFM.

Os componentes dessa figura serão abordados em itens posteriores deste capítulo.

2.10.3. **Sentidos de Transmissão em Redes de Acesso a Usuários**

Uma rede de acesso a usuários envolve, de um lado, o provedor de serviços, e, do outro lado, o usuário. A transmissão no sentido do provedor para o usuário é referida como transmissão *downstream*, enquanto a transmissão no sentido contrário é referida como transmissão *upstream*.

No caso de redes de acesso por fibras ópticas, com topologia P2P ou P2MP, o componente óptico que se localiza nas instalações do provedor é denominado *Optical Line Terminal* (OLT). O componente óptico localizado nas dependências do usuário denomina-se *Optical Network Unit* (ONU).

2.10.4. **Tecnologias Ópticas Ponto a Ponto (P2P)**

Conforme a Figura 2-20 anterior, foram definidos, pelo padrão IEEE 802.3-2012, seis sistemas de sinais de camada física destinados à comunicação óptica ponto a ponto (P2P) em redes EFM, que constituem as famílias LX10 e BX10.

Tais sistemas de sinais apresentam os seguintes aspectos:

- Todas as opções têm o alcance máximo de 10 km quando suportadas por fibras ópticas monomodo;
- Existem opções com velocidades de transmissão de 100 Mbit/s e de 1000 Mbit/s;
- Existem opções com suporte em duas fibras ópticas (família LX10) e em uma única fibra óptica (família BX10).

2.10.4.1. Sistemas de Sinais LX10

Os sistemas de sinais de camada física LX10, independentemente da velocidade de transmissão, têm o alcance máximo de 10 km e utilizam duas fibras ópticas monomodo, cada uma para transmissão em um sentido.

A letra L (*long wavelength*) indica a utilização de fibras ópticas monomodo usando *lasers* operando com comprimentos de onda longos.

Ressalva-se que o sistema de sinais 1000BASE-LX10 pode opcionalmente utilizar fibras ópticas multimodo, a exemplo do que ocorre com o sistema de sinais 1000BASE-LX, anterior ao padrão IEEE 802.3ah, no qual se baseou a definição do sistema de sinais 1000BASE-LX10. Nesse caso, o alcance do enlace de acesso é de apenas 0,55 km.

2.10.4.2. Sistemas de Sinais BX10

Os sistemas de sinais BX10 utilizam uma única fibra óptica monomodo, com transmissão *full-duplex*. A letra B (*bidirecional long wavelength*) indica transmissão *full-duplex* em uma única fibra óptica monomodo utilizando *lasers* com comprimentos de onda longos.

Existem duas opções tecnológicas para a transmissão *full-duplex* em uma única fibra óptica:

— Com utilização de dois comprimentos de onda (tecnologia *Dual Wavelength*);
— Com utilização de um único comprimento de onda (tecnologia *Single Wavelength*).

O padrão adotou a tecnologia *Dual Wavelength* para os sistemas de sinais BX10.

As opções de tecnologia serão abordadas a seguir.

- **Tecnologia Dual Wavelength**

A tecnologia *Dual Wavelength* representa uma forma de CWDM (*Coarse Wavelength Division Multiplex*). Cada extremidade da fibra óptica utiliza um dos dois comprimentos de onda como portador dos sinais ópticos.

Nessa tecnologia, os transceptores ópticos utilizados no OLT são diferentes dos transceptores ópticos utilizados na ONU. Os custos relativos aos transceptores ópticos do OLT são mais elevados, o que não foi, contudo, um fator determinante para a escolha da tecnologia *Dual Wavelength* pelo IEEE.

Para essa escolha foram considerados diversos parâmetros além dos custos, tais como confiabilidade e complexidade.

Na tecnologia *Dual Wavelength*, a denominação dos sistemas de sinais que se encontram nos OLTs utiliza a letra D (de *downstream*), como nos exemplos dos sistemas de sinais 100BASE-BX10-D e 1000BASE-BX10-D.

Do lado do usuário, a denominação dos sistemas de sinais que se encontram nas ONUs utiliza a letra U (de *upstream*), como nos exemplos dos sistemas de sinais 100BASE-BX10-U e 1000BASE-BX10-U.

Os comprimentos de onda adotados para transmissão no sentido *downstream* para os sistemas de sinais 100BASE-BX10 encontram-se na faixa entre 1.480 nm e 1.580 nm (valor nominal igual a 1.550 nm). Os comprimentos de onda adotados para a transmissão no sentido *upstream* encontram-se na faixa entre 1.260 nm e 1.360 nm (valor nominal igual a 1.310 nm).

Para os sistemas de sinais 1000BASE-BX10 essas faixas encontram-se, respectivamente, entre 1.480 nm e 1.500 nm (valor nominal igual a 1.490 nm), e entre 1.260 nm e 1.360 nm (valor nominal igual a 1.310 nm).

Os transceptores utilizados nos sistemas de sinais BX10, que utilizam a tecnologia *Dual Wavelength*, possuem em seus conjuntos ópticos, além das submontagens de transmissão e de recepção, submontagens de multiplexação.

Na submontagem de multiplexação de um transceptor ocorre a separação das subfaixas correspondentes aos comprimentos de onda de transmissão e de recepção.

- **Tecnologia Single Wavelength**

A tecnologia *Single Wavelength* para a transmissão *duplex* em uma única fibra óptica utiliza um único comprimento de onda para a transmissão dos sinais ópticos nos dois sentidos.

A distinção entre os sinais ópticos nos dois sentidos tem como base o posicionamento das submontagens de transmissão e de recepção em um transceptor, e a própria direcionalidade desses sinais.

Diferentemente dos transceptores que operam na tecnologia *Dual Wavelength* e utilizam dispositivos de acoplamento de sinais, os transceptores da tecnologia *Single Wavelength* utilizam dispositivos de acoplamento que não operam com base em comprimentos de onda.

O comprimento de onda nominal recomendável para a tecnologia *Single Wavelength* é o de 1.310 nm.

Os PMDs em cada extremidade do *link* óptico são os mesmos, o que se reflete em uma vantagem logística e de custos sobre a tecnologia *Dual Wavelength.*

Por outro lado, a operação dos acopladores, pela sua complexidade e menor confiabilidade (robustez), representa uma desvantagem da tecnologia *Single Wavelength.*

Conforme menção anterior neste capítulo, o IEEE optou pela tecnologia *Dual Wavelength* na definição dos sistemas de sinais BX10.

2.10.4.3. LX10 e BX10

Conforme menção anterior, os sistemas de sinais LX10 utilizam duas fibras ópticas, enquanto os sistemas de sinais BX10 utilizam apenas uma fibra óptica.

A escolha entre o uso desses sistemas de sinais por parte de provedores de serviço, deve considerar as vantagens e desvantagens das duas opções.

Por um lado, os sistemas de sinais BX10 têm a vantagem de utilizar apenas uma fibra óptica, enquanto os sistemas de sinais LX10 utilizam duas fibras ópticas.

Em contrapatida, na tecnologia para transmissão *duplex* adotada para os sistemas de sinais BX10 (com tecnologia *Dual Wavelength*), os custos associados as componentes são mais elevados, em decorrência principalmente da operação com dois comprimentos de onda em uma única fibra óptica.

2.10.4.4. Arquitetura de Redes de Acesso P2P

A Figura 2-21 apresenta a arquitetura de redes de acesso P2P.

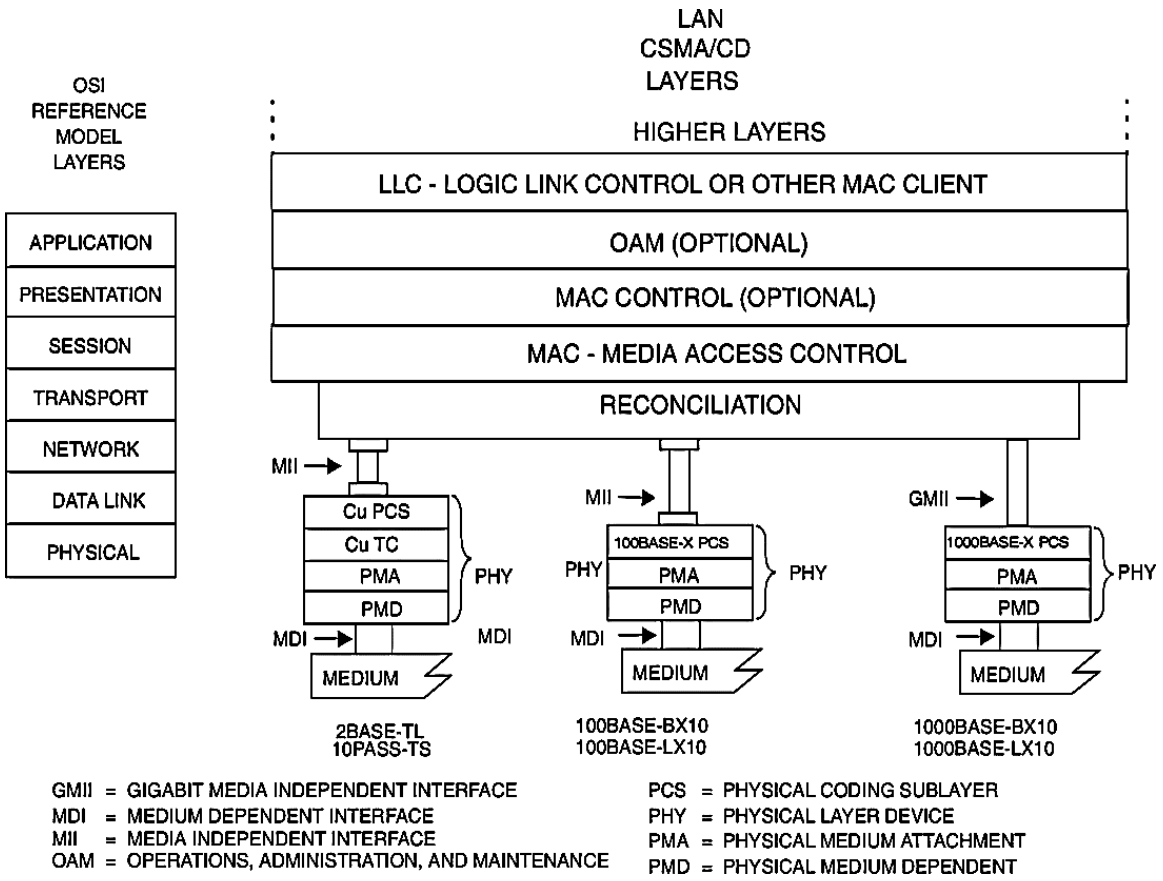


Figura 2-21 – Arquitetura de redes de acesso P2P.

Observa-se que essa figura abrange também o caso de redes de acesso P2P baseadas em cabos de cobre, que serão abordadas adiante neste capítulo.

O anexo 4A, de caráter normativo, que se encontra na Seção 1 do padrão IEEE 802.3-2012, especifica um método de controle de acesso ao meio (MAC) simplificado para uso em redes EFM. Tal método denomina-se *Simplified Full-Duplex Media Access Control*, e será abordado adiante neste capítulo.

As redes EFM ópticas suportam exclusivamente *links full-duplex*. As redes EFM ópticas P2P podem utilizar o MAC completo no modo *full-duplex* definido pela cláusula 4 da Seção 1 do padrão IEEE 802.3-2012, ou podem usar o MAC simplificado do anexo 4A desse padrão.

As redes EFM ópticas P2MP utilizam exclusivamente a versão simplificada do MAC, enquanto as redes EFM utilizando cabos de cobre adotam ou essa versão simplificada do MAC ou o MAC completo definido pela Cláusula 4 da Seção 1 do padrão.

2.10.5. **Tecnologias Ópticas Ponto a Multiponto (P2MP)**

Para as topologias P2MP, a EFM introduz duas famílias de sistemas de sinais de camada física baseados na 1000BASE-X,com adições às subcamadas de reconciliação (RS), de codificação física (PCS) e de ligação à camada física (PMA),

e com correção de erros (FEC) opcional, que são as famílias 1000BASE-PX10 e 1000BASE-PX20.

Conforme a Figura 2-20 anterior, foram definidos, pela Seção 5 do padrão IEEE 802.3-2012, quatro sistemas de sinais de camada física em 1000 Mbit/s destinados à comunicação óptica ponto a multiponto (P2MP) em redes ópticas passivas (PONs, *Passive Optical Networks*) para uso em EFM, e que constituem as famílias 1000BASE-PX10 e 1000BASE-PX20.

Tais sistemas de sinais apresentam os seguintes aspectos:

— Existem opções com o alcance máximo de 10 km e opções com o alcance máximo de 20 km, todas operando com suporte em fibras ópticas monomodo;
— Todas as opções operam na velocidade de sinais de 1250 Mbit/s (velocidade líquida de 1000 Mbit/s);
— Todas as opções operam no modo *full-duplex* sobre uma árvore cujos ramos constituem-se em uma única fibra óptica monomodo;
— Em todas as opções as fibras ópticas monomodo operam com suporte na tecnologia EPON (*Ethernet Passive Optical Network*).

Os quatro sistemas de sinais acima referidos para comunicação óptica P2MP em EFM são os seguintes:

— 1000BASE-PX10-D (*Passive Optical Network Downstream* 10 km);
— 1000BASE-PX10-U(*Passive Optical Network Upstream* 10 km);
— 1000BASE-PX20-D(*Passive Optical Network Dowstream* 20 km);
— 1000BASE-PX20-U(*Passive Optical Network Upstream* 20 km).

Também para a comunicação óptica ponto a multiponto (P2MP) em PONs, a EFM introduz uma família de sistemas de sinais de camada física baseados na 10GBASE-R, com adições às subcamadas de reconciliação (RS), de codificação física (PCS) e de ligação à camada física (PMA). Para essa família, denominada 10GBASE-PR, a correção de erros (FEC) é obrigatória.

Os seis sistemas de sinaisdessa família são os seguintes:

— 10GBASE-PR-D1;
— 10GBASE-PR-U1;
— 10GBASE-PR-D2;
— 10GBASE-PR-U2;
— 10GBASE-PR-D3;
— 10GBASE-PR-U3.

Essa família suporta sinais 10GBASE-R no sentido *downstream* e no sentido *upstream*.

As seguintes combinações podem ser formadas:

— 10GBASE–PR–D1 e 10GBASE–PR–U1, com balanço de potência PR10, alcance de pelo menos 10 km e um *split ratio* de pelo menos 1:16;

— 10GBASE–PR–D2 e 10GBASE–PR–U1, com balanço de potência PR20, com alcance de pelo menos 20 km e um *split ratio* de pelo menos 1:16, ou com alcance de pelo menos 10 km e um *split ratio* de pelo menos 1:32;
— 10GBASE–PR–D3 e 10GBASE–PR–U3, com balanço de potência PR30, com alcance de pelo menos 20 km e um *split ratio* de pelo menos 1:32.

Adicionalmente, a EFM introduz a família de sistemas de sinais 10/1GBASE-PRX, que suportam sinais 10GBASE-R no sentido *downstream* e sinais 1000BASE-X no sentido *upstream*.

Os seis sistemas de sinais dessa família são os seguintes:

— 10/1GBASE-PRX-D1;
— 10/1GBASE-PRX-U1;
— 10/1GBASE-PRX-D2;
— 10/1BASE-PRX-U2;
— 10/1GBASE-PRX-D3;
— 10/1GBASE-PRX-U3.

As seguintes combinações podem ser formadas:

— 10/1GBASE–PRX–D1 e 10/1GBASE–PRX–U1, com balanço de potência PRX10, alcance de pelo menos 10 km e um *split ratio* de pelo menos 1:16;
— 10/1GBASE–PRX–D2 e 10/1GBASE–PRX–U2, com balanço de potência PRX20, com alcance de pelo menos 20 km e um *split ratio* de pelo menos 1:16;
— 10/1GBASE–PRX–D3 e 10/1GBASE–PRX–U3, com balanço de potência PRX30, com alcance de pelo menos 20 km e um *split ratio* de pelo menos 1:32.

2.10.5.1. Passive Optical Networks (PONs)

Uma PON é uma rede óptica P2P ou P2MP, cuja ODN (*Optical Distribution Network*) utiliza exclusivamente componentes eletricamente passivos. Uma ODN é composta, em uma PON, pelas fibras ópticas utilizadas, e por acopladores, *splitters*, conectores e redes.

A idéia da utilização de fibras ópticas para o acesso a residências e a empresas denomina-se FTTx, englobando FTTH **(***Fiber-to-the-Home***)**, FTTC *(Fiber-to-the-Curb)* e FTTP (*Fiber-to-the-Premises*). Os altos custos das soluções propostas inicialmente, decorrentes do uso de muitas fibras ópticas P2P e de componentes ativos, tornaram-se uma barreira para a proliferação dessas soluções.

Em resposta a tal desafio foram definidas as PONs, principalmente as P2MP PONs, que apresentam uma considerável redução de custos.

Embora existam as P2P PONs, as P2MP PONs são mais utilizadas, devido ao uso compartilhado de fibras ópticas e de componentes.

Podem ser citadas as seguintes opções de P2MP PONs:

— ATM PONs (APONs);
— *Broadband* PONs (BPONs);
— *Gigabit-Capable* PONs (GPONs) ;
— *Ethernet* PONs (EPONs).

Observa-se que todas as opções de P2MP PONs acima citadas atendem o que se denomina *triple-play applications*, ou seja, suportam dados, vídeo digital e VoIP (voz sobre IP).

- **APONs e BPONs**

As APONs e as BPONs, definidas na década de 1990, utilizam o ATM como tecnologia de suporte. Foram definidas, pela série de recomendações G.983.x do ITU-T, as subcamadas APON PMD e BPON PMD, que interfaceiam o meio óptico com o ATM.

As BPONs são uma evolução das APONS, no sentido de possibilitar maiores velocidades de transmissão. As BPONs são capazes de operar com velocidades de até 1.250 Mbit/s (sentido *downstream*) e de até 622 Mbit/s (sentido *upstream*).

- **GPONs**

As GPONs, padronizadas pelo ITU-T pela série de recomendações G.984.x (GPON) e G.987.x (XG-PON) objetivam possibilitar maiores velocidades de transmissão e maiores distâncias de alcance que os demais tipos de P2MP PONs. Oferecem também um suporte robusto para OAM e QoS.

Independentemente do tipo de PMD utilizado, as GPONs inserem um cabeçalho PCB (*Physical Control Block*) a cada 125 μs, formando uma sucessão de quadros GTC (GPON *Transmission Convergence*).

Um cabeçalho PCB contém informações relativas a sincronismo e a OAM, além de bits de controle e de bits destinados à indicação do sentido de transmissão (*upstream* e *downstream*).

A recomendação ITU-T G.984.2 (*Gigabit-Capable Passive Optical Networks (GPONs): Physical Media Dependent (PMD) Layer Specification*) especifica, para a GPON, o suporte às velocidades de 2,5 Gbit/s e 1,25 Gbit/s no sentido downstream e de 155 Mbit/s, 622 Mbit/s, 1,25 Gbit/s e 2,5 Gbit/s no sentido upstream. A opção preferencialmente implementada opera com 2,5 Gbit/s no sentido downstream e 1,25 Gbit/s no sentido upstream.

A XG-PON (especificada nas recomendações ITU-T G.987.x) opera com 10 Gbit/s no sentido *downstream* e com 2,5 Gbit/s no sentido *upstream*.

Este livro se absterá de uma abordagem mais profunda das GPONs, por tratar-se de uma tecnologia de PON definida pelo ITU-T, e que não consta das definições de EFM contidas no padrão IEEE 802.3-2012.

2.10.5.2. EPONs

A Cláusula 64 do padrão IEEE 802.3ah, posteriormente incorporada à Cláusula 64 da Seção 5 do padrão IEEE 802.3-2012, especifica o controle MAC multiponto para as EPONs operando a 1000 Mbit/s (também referidas como GEPONs). A Seção 5 do padrão IEEE 802.3-2012 especificou também, na Cláusula 77, o controle MAC multiponto para as 10G-EPONs.

O padrão IEEE 802.3-2012 define o mecanismo de funcionamento e os protocolos requeridos para conciliar a topologia P2MP utilizada nas EPONs com a tecnologia *Ethernet*. As P2MP PONs, como no exemplo das EPONs, apresentam uma topologia assimétrica, e utilizam os conceitos de OLT, ONU, transmissão *downstream* e transmissão *upstream* adotados pelas soluções ópticas P2P já abordadas neste capítulo.

A Figura 2-22 apresenta uma configuração topológica simplificada de uma EPON.

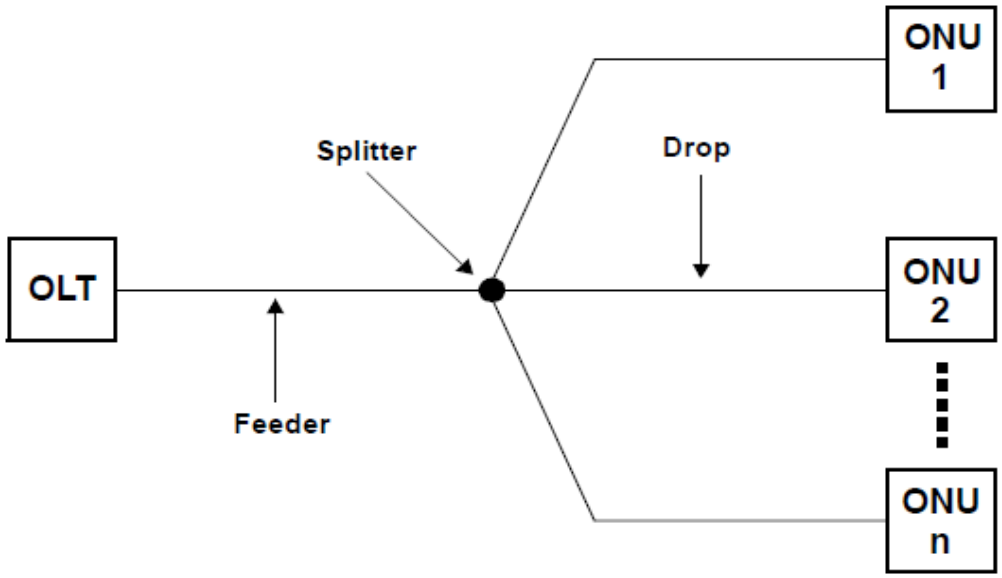


Figura 2-22 – Configuração de uma EPON.

Embora a topologia das EPONs seja assimétrica, a transmissão pode ocorrer de modo simétrico. Este é o caso das redes GEPON,em que para ambos os sentidos de transmissão foi adotada a velocidade de 1 Gbits/s. No caso das redes 10G-EPON, a transmissão pode ser simétrica (10 Gbit/s em ambos sentidos) ou assimétrica (10 Gbit/s *downstream* e 1 Gbit/s *upstream*).

É importante observar, como será visto com maiores detalhes posteriormente neste item, que as EPONs, embora sendo redes P2MP, operam por emulação de circuitos P2P, um para cada ONU contida na rede.

No sentido *downstream*, os sinais transmitidos são simplesmente distribuídos por um distribuidor (*splitter*) passivo, ou por um conjunto de *splitters* em cascata.

No sentido *upstream*, os sinais enviados por cada ONU atingem exclusivamente a OLT, não alcançando, portanto, as demais ONUs. Para evitar colisões, adota-se um processo TDMA de transmissão, onde cada ONU recebe autorização (*grant*) para ocupar janelas de transmissão. Esse processo de arbitragem de janelas de transmissão é controlado pelo protocolo MPCP (*Multi-Point Control Protocol*).

Todos os sistemas de sinais adotados pelas EPONs operam com uma única fibra óptica, no modo *Dual Wavelength*. Em conseqüência, eles se subdividem em dois tipos, conforme a sua localização na rede. Assim, existem os sistemas de sinais-D do lado do OLT (o D denota transmissão *Downstream*), e os sistemas -U do lado das ONUs (transmissão *Upstream*).

O mecanismo de transporte das EPONs é *Ethernet*, sobreposta nativamente às PMDs especificadas para as EPONs.As EPONs utilizam opcionalmente a técnica FEC (*Forward Error Correction*).

Os aspectos básicos referentes à operação das EPONs acima mencionados serão melhor abordados subsequentemente neste item.

- **Topologias das EPONs**

As EPONs podem apresentar diferentes topologias.

A topologia com um único *splitter* encontra-se ilustrada na Figura 2-23.

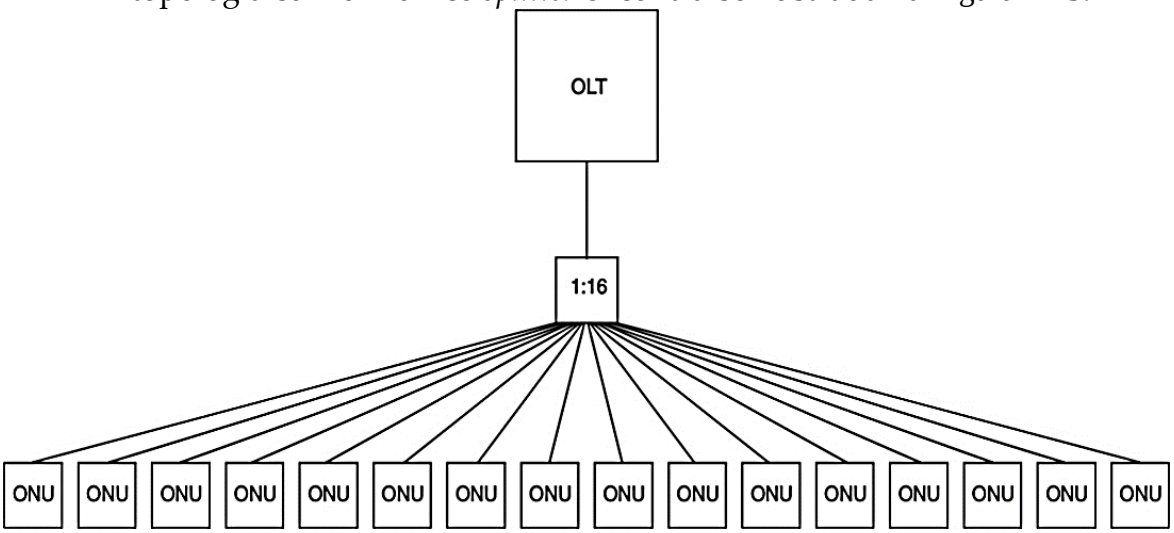


Figura 2-23– Topologia com um único splitter.

Outras topologias podem se basear em uma árvore de *splitters*, o que se encontra ilustrado na Figura 2-24.

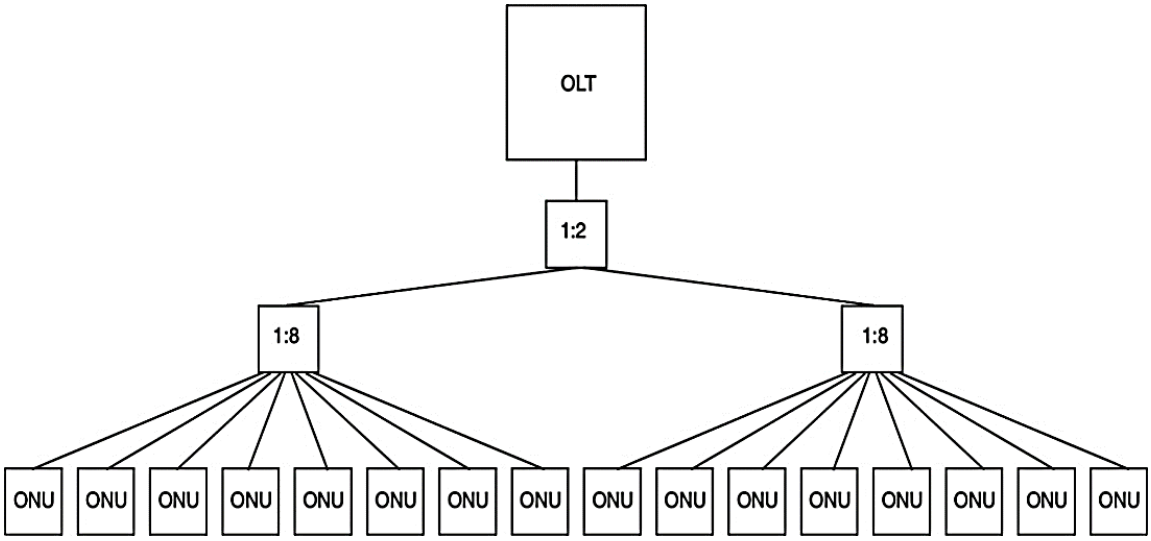


Figura 2-24 – Topologia com árvore de splitters.

- **Arquitetura de EPONs**

Na Figura 2-25 encontra-se representada a arquitetura simplificada de uma aplicação típica de uma EPON, que evidencia a simplicidade desse tipo de rede.

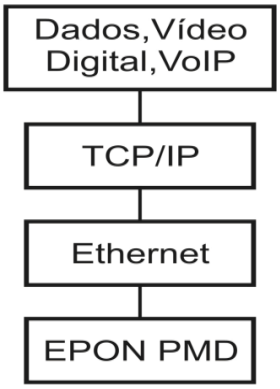


Figura 2-25 – Arquitetura simplificada das EPONs.

A Figura 2-26, por sua vez, apresenta a arquitetura completa das EPONs, para o caso particular das GEPONs, conforme a Cláusula 64 da Seção 5 do padrão IEEE 802.3-2012.

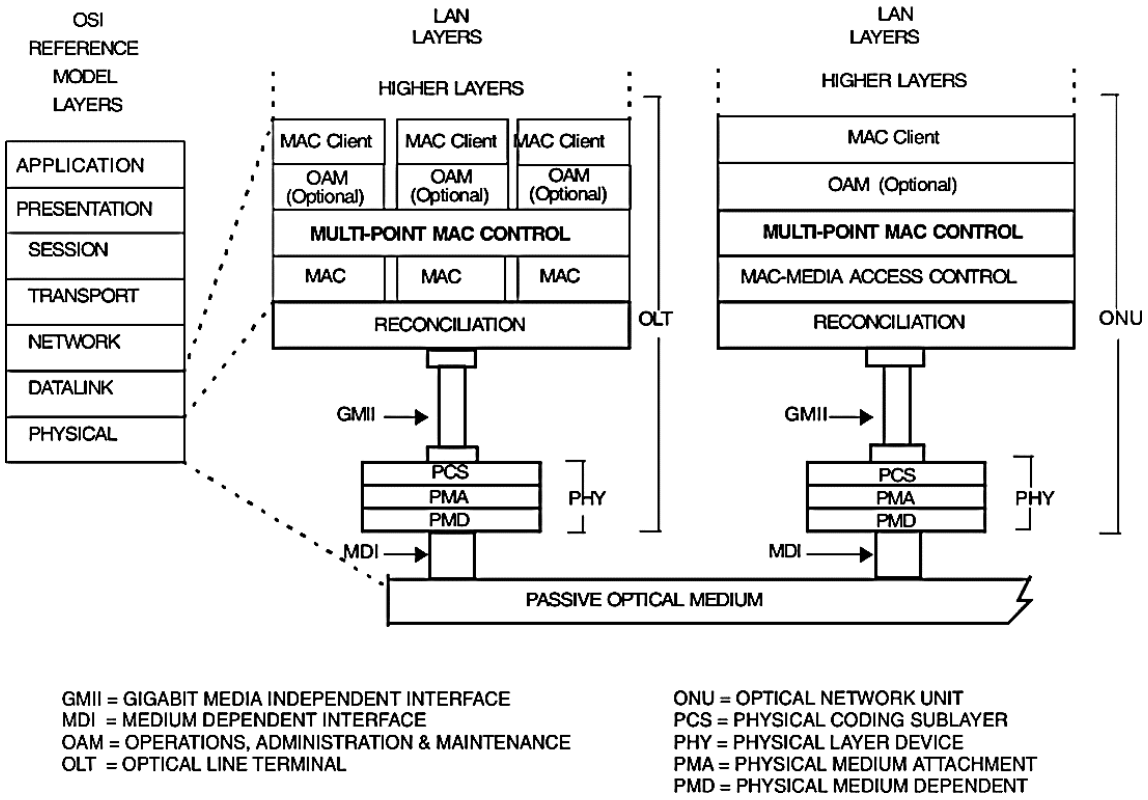


Figura 2-26 – Arquitetura das EPONs.

Como se observa nessa figura, a arquitetura das EPONs, comparativamente à arquitetura das redes EFM P2P, introduz uma nova subcamada, denominada *Multi-Point MAC Control* (MPMC), que se encontra identificada na figura.

A MPMC é uma extensão da subcamada *MAC Control* definida na Cláusula 31 da Seção 2 do padrão IEEE 802.3-2012. Verifica-se que a subcamada MPMC é utilizada apenas no plano de controle das redes *Ethernet*.

É na MPMC que residem as funções relativas ao protocolo MPCP, que inicializa, controla e governa as EPONs.

Observa-se também, nessa figura, a existência de múltiplas instâncias MAC do lado da OLT. A OLT possui pelo menos uma instância MAC associada a cada ONU. Cada uma dessas instâncias individuais MAC proporciona um serviço de emulação ponto a ponto entre a OLT e uma ONU.

A emulação de conexões P2P nas EPONs tem como base a utilização de um campo LLID (*Logical Link Identifier*). Um par constituído por uma instância MAC na OLT e a instância MAC de uma determinada ONU é identificado por um valor de LLID.Tais valores são atribuídos por meio do MPCP, como será visto adiante.

No sentido *downstream*, a rede EPON é um meio de transmissão em difusão (*broadcast*). Além dessas instâncias MAC utilizadas para comunicação ponto a ponto, existe uma instância MAC usada para comunicação simultânea com todas as ONUs. Essa instância aproveita a natureza *broadcast* do canal *downstream* para enviar uma única cópia do quadro, a qual é aceita por todas as ONUs. Essa

instância MAC é denominada *Single Copy Broadcast* (SCB MAC). O LLID associado denomina-se LLID *broadcast*.

Além disso, a OLT pode possuir uma instância MAC *multicast* associada a cada LLID *multicast*, cuja utilização é opcional.

Para as 10G-EPONs (10GBASE-PR e 10/1GBASE-PRX), observa-se a existência das devidas adequações nas respectivas arquiteturas, como, por exemplo, o uso da interface XGMII (10 GIGABIT MEDIA INDEPENDENT INTERFACE) imediatamente abaixo da subcamada de reconciliação.

Tais arquiteturas encontram-se descritas na Clásula 77 da Seção 5 do padrão IEEE 802.3-2012.

- **Multi-Point Control Protocol (MPCP)**

O MPCP opera com base em dois conceitos fundamentais:

— Emulação de conexões P2P nas EPONs;
— Alocação de bandas passantes no sentido *upstream* por TDMA.

A alocação dinâmica de bandas passantes por TDMA será vista em subitens subsequentes neste capítulo.

- **Transmissão de Quadros de Dados**

A Figura 2-27 permite a visualização de transmissão de quadros de dados em EPONs, no sentido *downstream*.

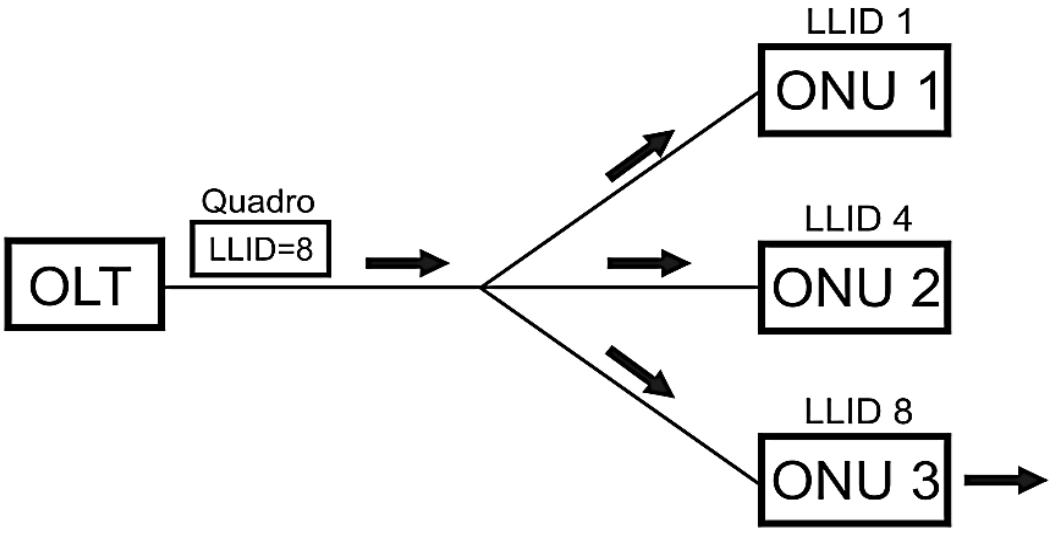


Figura 2-27 – Transmissão no sentido downstream.

Como a transmissão no sentido *downstream* é em *broadcasting*, o quadro de dados será transmitido para todas as ONUs. Como o quadro conduz o valor de LLID igual a 8 em seu preâmbulo nesse exemplo, apenas a ONU 3, associada ao LLID 8, aceitará o quadro.

Por sua vez, a Figura 2-28 representa a transmissão de um quadro de dados em EPONs no sentido *upstream*.

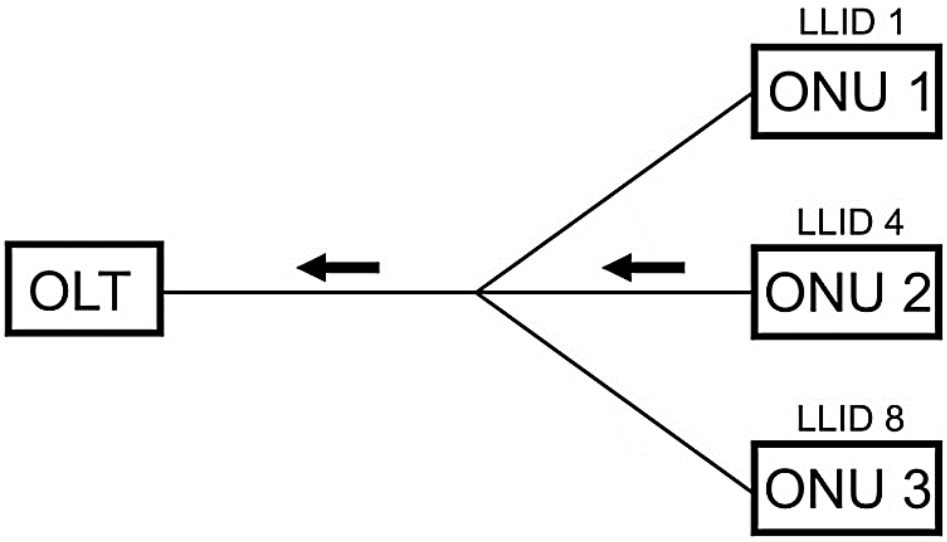


Figura 2-28 – Transmissão no sentido upstream.

Nessa figura, supõe-se que a ONU 2 transmitiu um quadro de dados dentro da janela de transmissão (*time slot*) que lhe foi concedida anteriormente pela OLT. Dentro dessa janela de transmissão apenas a ONU 2 está autorizada a transmitir quadros de dados, para que não ocorra colisão.

O quadro de dados transmitido pela ONU 2 conduz o LLID com o valor 4. Isso indicará à OLT a instância MAC a ser utilizada, associada ao LLID 4.

- **Tipos de Processamento no MPCP**

O MPCP opera com base em três tipos de processamento:

— *Discovery processing*;
— *Report processing*;
— *Gate processing*.

Pelo *discovery processing* cada ONU é descoberta e registrada na OLT.

O *report processing* gerencia a geração e a coleta de mensagens *report*, através das quais as ONUs comunicam à OLT as suas necessidades de banda passante.

O *gate processing* gerencia a geração e a coleta de mensagens *gate*, através das quais é obtida a multiplexação de múltiplas ONUs como transmissoras de quadros.

- **Resumo das Mensagens MPCP**

As MPCPDUs apresentam o formato genérico da Figura 2-29.

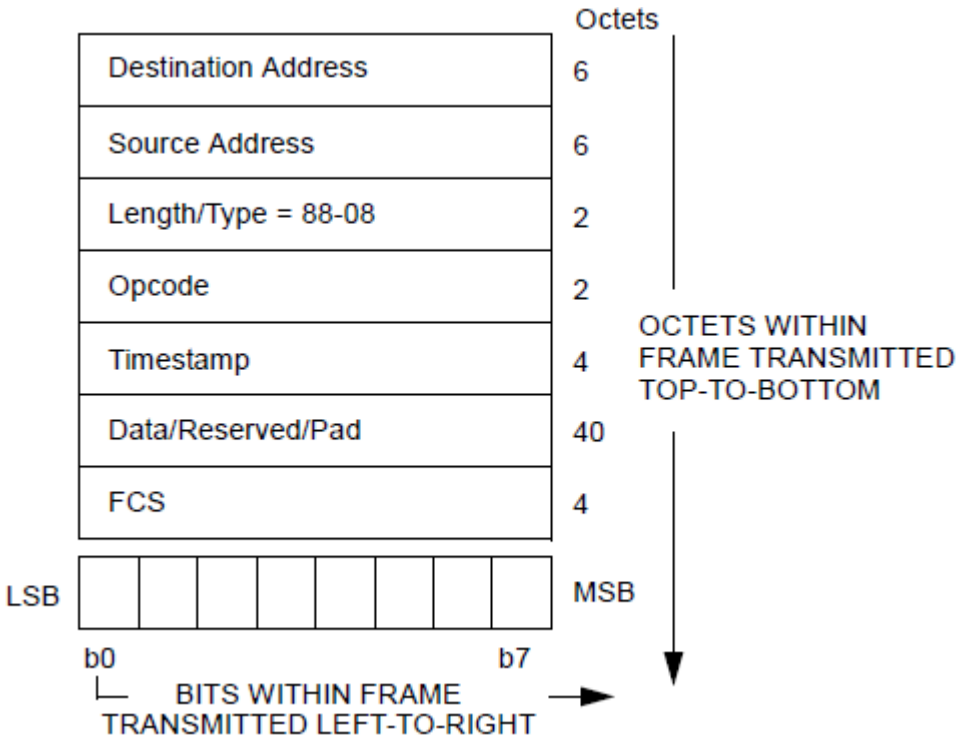


Figura 2-29 – Formato genérico das MPCPDUs.

Nessa figura, o campo *opcode* identifica a MPCPDU específica encapsulada, conforme a Tabela 31A-1 no Anexo 31A da Seção 2 do padrão IEEE 802.3-2012.

A Figura 2-30 apresenta um resumo das mensagens MPCP (MPCPDUs), com os seus principais atributos.

MPCPDU	*Opcode*	ORIGEM	PROCESSO	LLID	ENDEREÇO DE DESTINO
GATE	0X0002	OLT	*DISCOVERY*	SCB	*MULTICAST*
				PRÓPRIO	
			GATE	PRÓPRIO	*UNICAST*
REPORT	0X0003	ONU	*REPORT*	PRÓPRIO	*UNICAST*
REGISTER_REQ	0X0004	ONU	*DISCOVERY*	SCB	*MULTICAST*
REGISTER	0X0005	OLT	*DISCOVERY*	SCB	*UNICAST*
REGISTER_ACK	0X0006	ONU	*DISCOVERY*	PRÓPRIO	*MULTICAST*

Figura 2-30 – Principais atributos das MPCPDUs.

Na Cláusula 64.3 da Seção 5 do padrão IEEE 802.3-2012 encontram-se os formatos e a descrição dos campos para os tipos de MPCPDU listados na figura.

2.10.5.3. Forward Error Correction (FEC) para 1000BASE-PX

A Cláusula 65 da Seção 5 do padrão IEEE 802.3-2012 define o mecanismo de FEC opcional aplicável às famílias de sistemas de sinais P2MP 1000BASE-PX.

Segundo essa cláusula, os principais objetivos da subcamada FEC para 1000BASE-PX são manter o formato dos quadros compatível com a PCS (*Physical Coding Sublayer*) 1000BASE-X, permitir a compatibilidade com dispositivos 1000BASE-X legados e suportar um objetivo de BER de 10^{-12}, ou melhor, na PCS.

O mecanismo FEC é capaz de melhorar o *link budget* óptico ou aumentar a distância alcançável por uma *link* óptico. Por exemplo, um *link* óptico P2MP que alcança a distância de 10 km poderia passar a alcançar até 15 km após a introdução do mecanismo FEC. A utilização do mecanismo FEC nas PHYs 1000BASE-PX possibilita também a amplicação do *split ratio* para cima do valor especificado de 1:16.

Tal mecanismo opera mediante a inserção adicional de informações no quadro Ethernet, informações essas que resultam de um conjunto de funções aritméticas não-binárias. Essas informações adicionais possibilitam a correção de erros automaticamente na recepção do quadro MAC.

A função FEC contempla três blocos funcionais:

— Codificador FEC;
— Decodificador FEC;
— Sincronização FEC.

O codificador FEC opera do lado transmissor do *link*, enquanto o decodificador FEC e a sincronização FEC operam na recepção.

Do lado da transmissão, a subcamada FEC recebe os pacotes da PCS, realiza a codificação FEC, acrescenta os octetos de paridade FEC e envia os dados para a PMA. Na recepção, a subcamada FEC recebe esse dados da PMA, realiza o alinhamento dos octetos, detecta a *Start FEC Framing Sequence*, realiza a decodificação, corrige erros quando necessário e possível, substitui os octetos de paridade FEC por octetos vazios e envia os dados para a PCS.

Embora especificada apenas para as PHYs P2MP 1000BASE-PX aplicáveis às EPONS, a utilização da FEC do presente subitem pode ser estendida para *links* ópticos P2P 1000BASE-X.

2.10.5.4. PMDs e Meios Físicos nos Sistemas de Sinais 1000BASE-PX10 e 1000BASE-PX20

A Cláusula 60 da Seção 5 do padrão IEEE 802.3-2012 especifica as subcamadas *Physical Medium Dependent* (PMDs) para os sistemas de sinais 1000BASE-PX10 e 1000BASE-PX20. Essa cláusula especifica também as características ópticas das fibras ópticas utilizadas em cada um desses sistemas de sinais.

Como visto anteriormente, os sistemas de sinais 1000BASE-PX10 e 1000BASE-PX20 provêem *links* Ethernet P2MP utilizando uma topologia em árvore constituída por fibras ópticas monomodo, com suporte na tecnologia *Dual Wavelength*.

A camada física que suporta, com as alterações necessárias, os sistemas de sinais 1000BASE-PX, é a mesma que suporta os sistemas de sinais 1000BASE-LX e 1000BASE-BX.

Uma PMD 1000BASE-PX20-D pode interoperar com uma PMD 100BASE-PX10-U, o que possibilita maior flexibilidade operacional. Registra-se também que o uso de *links* utilizando FEC pode aumentar a razão de *splitting* (1:N) ou a distância de alcance nesses *links*.

As PMDs de cada um dos sistemas de sinais de que trata este subitem possuem os atributos primários da Figura 2-31.

Description	1000BASE-PX10-U	1000BASE-PX10-D	1000BASE-PX20-U	1000BASE-PX20-D	Unit
Fiber type	B1.1, B1.3 SMF				
Number of fibers	1				
Nominal transmit wavelength	1310	1490	1310	1490	nm
Transmit direction	Upstream	Downstream	Upstream	Downstream	
Minimum range	0.5 m to 10 km		0.5 m to 20 km		
Maximum channel insertion loss	20	19.5	24	23.5	dB
Minimum channel insertion loss	5		10		dB

Figura 2-31 – Atributos primários das famílias 1000BASE-PX.

As faixas de comprimento de onda para as famílias 1000BASE-PX10 e 1000BASE-PX20 são de 1480 nm a 1500 nm *downstream*, e 1260 nm a 1360 nm *upstream*.

2.10.5.5. PMDs e Meios Físicos nos Sistemas de Sinais 10GBASE-PR e 10/1GBASE-PRX

Para atender os objetivo das PMDs nos sistemas de sinais das famílias 10GBASE-PR e 10/1GBASE-PRX, a Cláusula 75 da Seção 5 do padrão IEEE 802.3-2012 define as seguintes três classes de *power budget*:

— Classe *Low Power Budget*;
— Classe *Medium Power Budget*;
— Classe *High Power Budget*.

A Figura 2-32 apresenta os atributos primários desses tipos de *power budget*.

Description	Low Power Budget		Medium Power Budget		High Power Budget		Units
	PRX10	PR10	PRX20	PR20	PRX30	PR30	
Number of fibers	1						–
Nominal downstream line rate	10.3125						GBd
Nominal upstream line rate	1.25	10.3125	1.25	10.3125	1.25	10.3125	GBd
Nominal downstream wavelength	1577						nm
Downstream wavelength tolerance	–2, +3						nm
Nominal upstream wavelength	1310	1270	1310	1270	1310	1270	nm
Upstream wavelength tolerance	± 50	± 10	± 50	± 10	± 50	± 10	nm
Maximum reach	≥10		≥20		≥20		km
Maximum channel insertion loss	20		24		29		dB
Minimum channel insertion loss	5		10		15		dB

Figura 2-32 – Atributos primários dos power budgets das famílias 10GBASE-PR e 10/1GBASE-PRX.

As faixas de comprimentos de onda para a família 10GBASE-PR são de 1575 nm a 1580 nm *downstream*, e 1260 nm a 1280 nm *upstream*.

Para a família 10/1GBASE-PRX, as faixas de comprimentos de onda são de 1575 nm a 1580 nm *downstream*, e 1260 nm a 1360 nm *upstream*.

2.10.6. **Tecnologias P2P com Cabos de Cobre**

A Cláusula 61 da Seção 5 do padrão IEEE 802.3-2012 define duas famílias de sistemas de sinais de camada física para redes EFM P2P com cabos de cobre (cabos de cobre de pares trançados não blindados):

— 10PASS-TS;
— 2BASE-TL.

2.10.6.1. Considerações Iniciais

A família de sistemas de sinais 10PASS-TS opera com a velocidade de transmissão nominal de 10 Mbit/s, com modulação *passband* (ou seja, que permite a coexistência com telefonia tradicional no mesmo par), em pares trançados (T, de *twisted pair*), com alcance curto (S, de *short reach*).

A família de sistemas de sinais 2BASE-TL, por sua vez, opera com a velocidade de transmissão nominal de 2Mbit/s, com modulação *baseband*(ou seja, que não permite a coexistência com telefonia tradicional no mesmo par), em pares trançados (T), com alcance longo (L, de *long reach*).

O acréscimo da letra O (10PASS-TS-O e 2BASE-TL-O) é utilizado para representar a terminação do enlace de cobre no lado do provedor (CO, de *central office)*, enquanto a terminação que se encontra no lado do usuário(CPE, de *customer premises equipment*), é representada pela letra R (10PASS-TS-R e 2BASE-TL-R).

Observa-se que os sistemas de sinais 10PASS-TS e 2BASE-TL, embora especificados para operar nominalmente com as velocidades de 10 Mbit/s e 2 Mbit/s, respectivamente, podem operar com velocidades inferiores ou superiores, em decorrência das características operacionais das *Digital Subscriber Lines* (DSL) que os suportam.

Os membros dessas famílias de sistemas de sinais com cabos de cobre representam as últimas quatro linhas da Figura 2-20 anterior deste livro.

Os sistemas de sinais para redes EFM baseados em cabos de cobre transportam quadros Ethernet sobre PMDs DSL. Utilizam o modo *full-duplex* para a transmissão desses quadros.

O uso de cabos telefônicos de cobre constitui uma alternativa ainda amplamente disponível para redes de acesso, sendo especificado para redes EFM.

As alternativas ópticas P2P ou P2MP operam com alcance geográfico e velocidades de transmissão bem maiores que as alternativas sobre cabos de cobre. Contudo, as tecnologias definidas para as redes EFM baseadas em cabos de cobre representam uma escala intermediária entre as soluções atuais, limitadas,e as tecnologias ópticas.

2.10.6.2. Padrões IEEE 10PASS-TS e 2BASE-TL e os padrões DSL

As especificações da PMA e da PMD 10PASS-TS baseiam-se no transceptor VDSL (*Very High Speed DSL*) especificado no padrão ANSI T1.424, denominado *Interface between Networks and Customer Installations: Very-high bit-rate Digital Subscriber Line (VDSL) Metallic Interface* (*DMT Based*).

As especificações da PMA e da PMD 2BASE-TL baseiam-se no transceptor SHDSL (*Single-pair High-speed DSL*) especificado na recomendação ITU-T G.991.2 (*Single-Pair High-Speed Digital Subscriber Line (SHDSL) Transceivers*).

O ITU-T definiu três famílias de padrões DSL, genericamente referidos como xDSL, para operação em pares de fios de cobre, que são as seguintes:

— HDSL (*High Bit Rate ou High Speed DSL*), especificado pela série de recomendações G.991;
— ADSL (*Asymetric* DSL), especificado pela série de recomendações G.992.
— VDSL (*Very High Speed* DSL), especificado pela série de recomendações G.993.

A série de recomendações ITU-T G.991 define dois tipos de HDSL:

— HDSL, especificado na recomendação ITU-T G.991.1 (*High bit rate digital subscriber line (HDSL) transceivers*) ;

— SHDSL (*Single-Pair High-speed* DSL), especificado pela recomendação G.991.2.

A série de recomendações G.992 define cinco tipos de ADSL:

— ADSL, especificado na recomendação ITU-T G.992.1 (*Asymmetric digital subscriber line (ADSL) transceivers*);
— *Splitterless ADSL*, especificado na recomendação ITU-T G.992.2 (*Splitterless asymmetric digital subscriber line (ADSL) transceivers*);
— ADSL2, especificado na recomendação ITU-T G.992.3 (*Asymmetric digital subscriber line transceivers 2 – ADSL2)*;
— *Splitterless ADSL2*, especificado na recomendação ITU-T G.992.4 (*Splitterless asymmetric digital subscriber line transceivers 2 - splitterless ADSL2*) ;
— ADSL2+ (*Extended Bandwidth* ADSL2), especificado na recomendação ITU-T G. 992.5 (*Asymmetric digital subscriber line 2 transceivers (ADSL2) - Extended bandwidth ADSL2 - ADSL2plus*) .

Finalmente, a série de recomendações ITU-T G.993 define dois tipos de VDSL:

— VDSL, especificado na recomendação ITU-T G.993.1 (*Very high speed digital subscriber line transceivers - VDSL*);
— VDSL2, especificado na recomendação ITU-T G.993.2. (*Very high speed digital subscriber line transceivers 2 – VDSL2*)

Complementarmente, o ITU-T definiu as seguintes recomendações relativas à tecnologia DSL:

— Recomendação ITU-T G.994.1 (*Handshake Procedures for Digital Subscriber Line Transceivers*);
— Recomendação ITU-T G.995.1 (*Overview of Digital Subscriber Line Recommendations*);
— Recomendação ITU-T G.996.1 (*Test Procedures for Digital Subscriber Line Transceivers*);
— Recomendação ITU-T G.997.1 (*Physical Layer Management for Digital Subscriber Line Transceivers*).

- **Single-Pair High-Speed DSL (SHDSL)**

O padrão SHDSL suporta velocidades simétricas. A operação básica ocorre sobre dois fios (pares metálicos trançados), mas existe a opção de operação sobre 4 fios.

A operação sobre dois fios é capaz de suportar velocidades de transmissão na faixa de 192 Kbit/s a 2.312 Kbit/s, com incrementos de 8 Kbit/s.

A operação sobre quatro fios, que permite o uso de *loops* mais extensos, pode suportar velocidades de transmissão na faixa de 384 Kbit/s a 4.624 Kbit/s, com incrementos de 16 Kbit/s.

O padrão SHDSL adota a codificação de linha TC-PAM (*Trellis Coded Pulse Amplitude Modulation*).

- **Very High Speed DSL (VDSL) e VDLS2**

Os padrões VDSL possibilitam a transmissão assimétrica ou simétrica de agregados de dados, a velocidades de até dezenas de Mbit/s, em pares de cobre trançados.

Os padrões VDSL englobam também planos mundiais de frequências, que permitem a transmissão assimétrica e a transmissão simétrica em um mesmo grupo de pares trançados. Os níveis de potência de transmissão foram especificados de forma a minimizar interferências com outros sistemas de transmissão.

A tecnologia adotada pelo VDSL explora a infraestrutura existente de cabos de cobre, originalmente instalada para o atendimento do serviço POTS (*Plain Old Tephone Service*).

O padrão VDSL pode utilizar modulação QAM (*Quadrature Amplitude Modulation*). Pode utilizar, alternativamente, a modulação DMT (*Discrete Multi Tone*). O DMT VDSL foi a técnica de modulação especificada para uso no sistema de sinais 10PASS-TS, conforme definição do padrão IEEE 802.3-2012.

2.10.7. **Operação, Administração e Manutenção (OAM)**

A cláusula 57 da Seção 5 do padrão IEEE 802.3-2012 define uma subcamada OAM, com o propósito de prover mecanismos para a monitoração de operação de *links* EFM, tais como mecanismos para indicação de falhas e para controle de *loopback* remoto. O OAM de *link* pode ser implementado, na realidade, em qualquer *link* Ethernet *full-duplex* ponto a ponto ou ponto a ponto emulado.

A subcamada OAM de *link* posiciona-se na arquitetura EFM conforme a Figura 2-33.

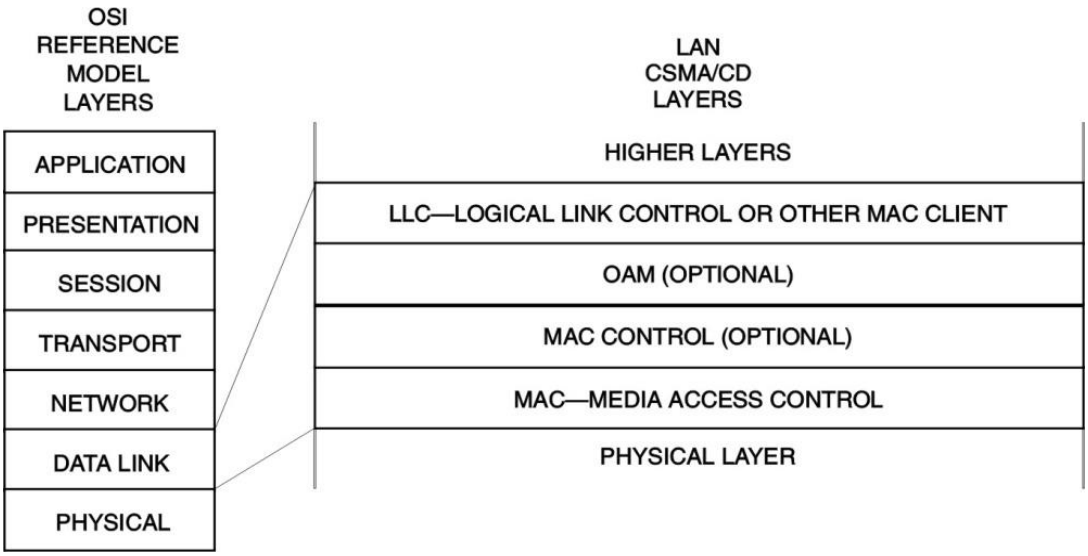


Figura 2-33 – OAM na arquitetura EFM.

As informações de OAM de camada de enlace são transportadas em OAMPDUs em quadros *Slow Protocols*. O conceito de *Slow Protocols* foi apresentado anteriormente neste capítulo.

As OAMPDUs contêm informações apropriadas de controle e de estado usadas para monitorar, testar e localizar falhas em *links* cobertos pelo OAM.

A subcamada OAM definida para EFM tem como propósito atender as funções de OAM em *links* ponto a ponto ou em *links* ponto a ponto emulados (como nas EPONs).

Os objetivos da subcamada OAM são os seguintes:

— Descobrimento (*discovery*);
— Monitoramento de *links*;
— Indicação de falhas remotas;
— *Loopback* remoto;
— Outros (*miscellaneous*).

O descobrimento identifica os dispositivos nas duas pontas do *link*, juntamente com suas propriedades de OAM.

O monitoramento de *links* serve para detectar e indicar erros no enlace em decorrência de várias condições.

A indicação remota de falhas permite à ponta do *link* que tenha detectado falhas de recepção, informar a ponta remota sobre as falhas detectadas, mediante *flags* específicos nas OAMPDUs. Entre as condições de falhas encontram-se a perda do sinal recebido, um erro irrecuperável ou outros eventos críticos.

Através do *loopback* remoto, uma entidade de OAM pode colocar a ponta remota em modo *loopback* mediante o envio de OAMPDUs.

As funções e procedimentos de OAM de *link* serão tratadas com maiores detalhes no capítulo referente a OAM deste livro.

CAPÍTULO 3

LAN BRIDGING com SPANNING TREES

3.1. PREÂMBULO

O tema *Ethernet bridging*, ou *Ethernet switching*, assim como *LAN Bridging* em geral, foi tratado inicialmente pelo padrão IEEE 802.1D-2002 que foi incorporado à versão atual desse padrão emitida em 2004 (padrão IEEE 802.1D-2004).

Foi emitido posteriormente o padrão IEEE 802.1Q, que hoje se encontra na versão emitida em 2014 (padrão IEEE 802.1Q-2014), que ampliou a concepção de *LAN Bridging*, introduzindo o conceito de VLANs (*Virtual LANs*).

O padrão IEEE 802.1D-2004 (*Media Access Control* (*MAC*) *Bridges*) e o padrão IEEE 802.1Q-2014 (*Bridges and Bridged Networks*) abordam *LAN Bridging* com base em *spanning trees*. Uma nova abordagem surgiu mais recentemente, tendo como fundamento a constituição de *bridged LANs* com suporte em extensões do protocolo IS-IS (*Intermediate System-to-Intermediate System*), o que será visto no capítulo 4 deste livro, intitulado *LAN Bridging* com IS-IS.

O presente capítulo restrige-se à abordagem de *bridged LANs* com *spanning trees* em redes de usuário. *LAN Bridging* em redes de provedor, também baseado em *spanning trees*, constitui-se no objeto do capítulo 5 deste livro, denominado *Provider LAN Bridging*.

Ocorreram algumas incorporações de padrões IEEE 802.1, antes existentes isoladamente, ao padrão 802.1Q-2014. Foram incorporados, dentre outros, os seguintes padrões (onde estão mencionadas as respectivas seções no padrão 802.1Q-2014):

- Padrão IEEE 802.1ad (*Provider Bridges*), nas Seções 15 e 16;
- Padrão IEEE 802.1ah (*Provider Backbone Bridges*), nas Seções 25 e 26;
- Padrão IEEE 802.1Qay (*Provider Backbone Bridge – Traffic Engineering*), nas Seções 25 e 26;
- Padrão IEEE 802.1ag (*Conectivity Fault Management*), nas Seções 18 a 22 e 29;
- Padrão IEEE 802.1ak (*Multiple Registration Protocol*), nas Seções 10 e 11;
- Padrão IEEE 802.1s (*Multiple Spanning Trees*), nas Seções 13 e 14.

A versão do padrão IEEE 802.1Q emitida em 2014 passou a incorporar também o padrão IEEE 802.1aq, que especifica o SPB (*Shortest Path Bridging*). O SPB será parte do capítulo 4 deste livro.

O padrão IEEE 802.1G (*Remote Bridging*) foi obsoletado e cancelado pelo IEEE.

3.2. TRANSPARENT BRIDGING

As *bridges* Ethernet operam na concepção *transparent bridging*.

A Figura 3-1 apresenta um exemplo de configuração de uma *bridged* LAN operando em *transparent bridging*.

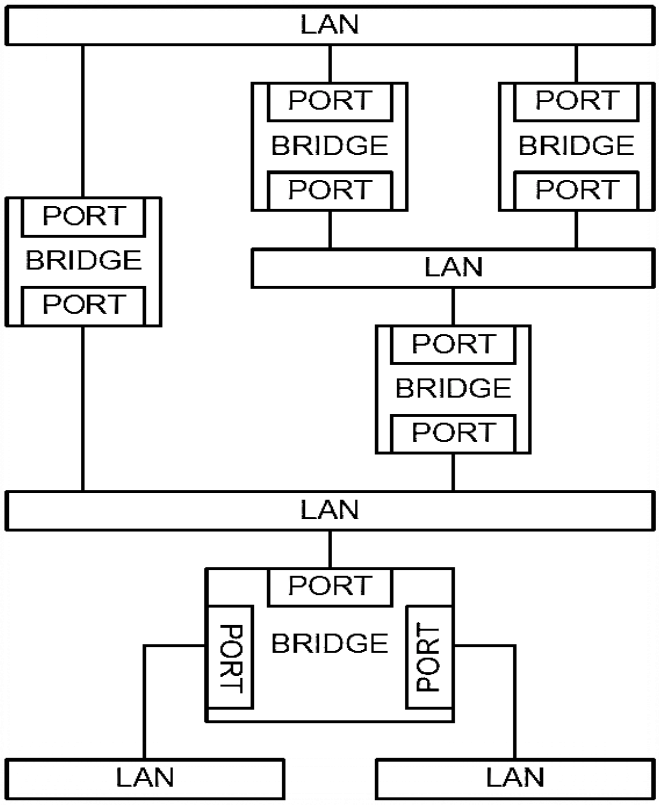


Figura 3-1– Configuração de uma bridged LAN transparente.

As *bridges* transparentes filtram e comutam quadros MAC entrantes em uma de suas portas, encaminhando-os por uma outra porta ou por outras de suas portas, sem a introdução de alterações nesses quadros MAC. A filtragem e a comutação ocorrem com base em tabelas contidas nas *bridges*, que podem ser montadas estática ou dinamicamente.

O roteamento por tabelas estáticas tem a sua utilização limitada a redes de pequeno porte, e não será considerado neste capítulo.

A filtragem e roteamento por tabelas dinâmicas fundamentam-se nos seguintes tipos básicos de processo:

— Montagem de *spanning trees* (STs), por um STP (ST *protocol*);
— Transmissão de quadros por inundação da rede;
— Aprendizagem de endereços MAC (MAC *address learning*);
— Registro de participação em endereços de grupo;
— Registro de participação em VLANs.

As *bridges* transmitem quadros *unicast, multicast* e *broadcast*. A aprendizagem dinâmica de endereços MAC baseia-se em endereços *unicast*, como será visto adiante neste capítulo.

Precedendo o processo de aprendizagem de endereços MAC *unicast*, as *bridged* LANs devem passar por um processo de eliminação de *loops* físicos, por meio de um STP. As STs definidas nas *bridged* LANs aplicam-se também à transmissão de quadros MAC *multicast* e *broadcast*.

As associações, em uma *bridge*, entre endereços MAC de estações finais de destino e as portas de saída de quadros a elas endereçadas, são registradas no *Filtering Database* dessa *bridge*.

Os quadros *broadcast* e os quadros *unicast* cujos endereços não tenham ainda sido aprendidos, são transmitidos por inundação (*flooding*) da rede, obedecida a respectiva *Spanning Tree*, o que proporciona a realização do processo de aprendizagem de endereços MAC de origem.

A transmissão de quadros MAC *multicast* fundamenta-se na aplicação prévia do MMRP (*Multiple Multicast Registration Protocol*).

Como será visto neste capítulo, em *bridged* LANs que suportam VLANs pode ocorrer também um processo de registro, em *bridges* e em estações finais, da participação das demais *bridges* e estações finais em cada uma das VLANs. Tal processo tem como base o MVRP (*Multiple* VLAN *Registration Protocol*). O MVRP processa-se com suporte na MST (*Multiple Spanning Tree*) da respectiva VLAN.

3.3. IEEE 802.1D BRIDGING

A versão atual do padrão IEEE 802.1D foi emitida em 2004 (padrão IEEE 802.1D-2004).

O conceito de LAN Bridging foi definido pelo padrão IEEE 802, sendo posteriormente ampliado pelos padrões IEEE 802.1D-2004 e IEEE 802.1G (*Remote Bridging*, agora cancelado). Será enfatizada neste subitem, assim como em todo este livro, a operação de *bridged* (ou *switched*) *Ethernet* LANs.

O padrão IEEE 802.1D-2004 não considera o conceito de VLAN, o que veio a se efetivar através do padrão IEEE 802.1Q-2014.

Dentro da família de padrões STP (*Spanning Tree Protocol*), utiliza-se atualmente em IEEE 802.1D *bridging* o RSTP (*Rapid Spanning Tree Protocol*), que representa a versão 2 dessa família de protocolos.

O STP básico (versão 0) foi obsoletado no padrão 802.1D-2004, embora existam ainda em operação implementações desse protocolo. A versão 1 do STP, que é uma extensão da versão 0 desse protocolo definida pelo padrão IEEE 802.1G, não teve aceitação e encontra-se hoje obsoletada.

O RSTP define uma única ST para toda uma rede *bridged Ethernet*. No interior de uma RSTP ocorre a aprendizagem dinâmica de endereços MAC *unicast* para toda a rede. Essa aprendizagem tem como base o próprio tráfego de dados, o que será visto adiante neste capítulo.

O padrão IEEE 802.1D-2004 especifica também o suporte ao MMRP (*Multiple Mac Registration Protocol*), pertinente à família de protocolos MRP (*Multiple Registration Protocol)*.

3.3.1. **Arquitetura de Bridges IEEE 802.1D**

A Figura 3-2 ilustra a arquitetura de uma *bridge* IEEE 802.1D com duas portas, que pode ser estendida para *bridges* com múltiplas portas.

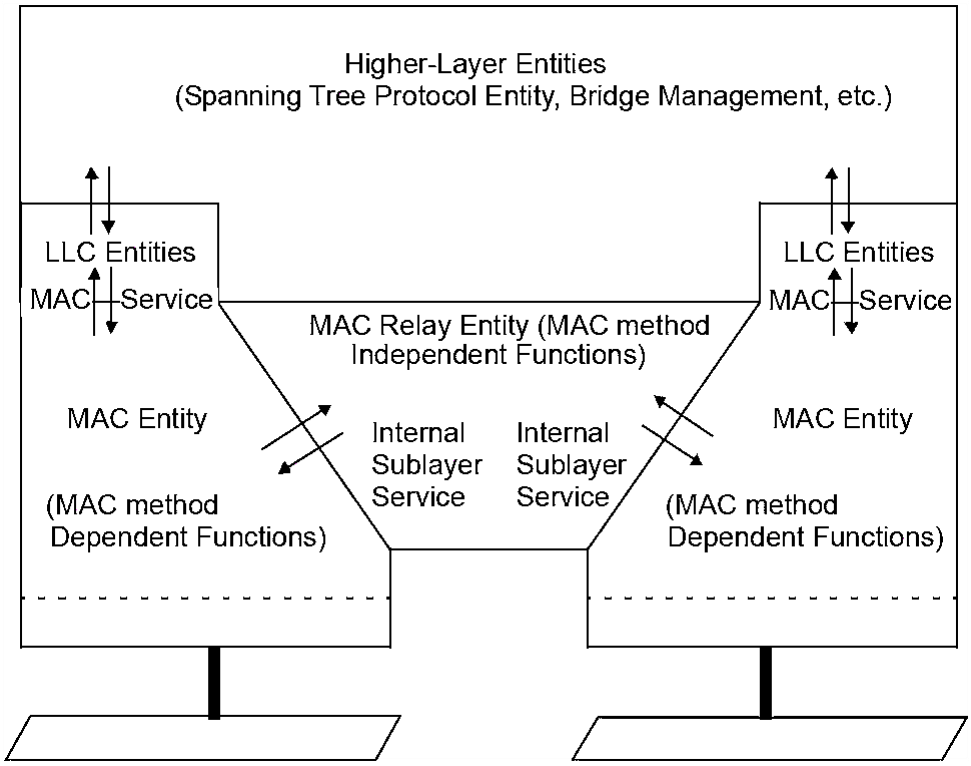


Figura 3-2– Arquitetura de IEEE 802.1D bridges.

A parte superior dessa figura, relativa às entidades de camadas superiores, diz respeito a protocolos de gerenciamento da *bridge* e a protocolos de controle, tais como aqueles destinados à montagem de STs e aqueles destinados a registro de participação em VLANs ou em endereços de grupo de estações finais.

A entidade *MAC Relay* é responsável pelas funções independentes do método MAC, enquanto as entidades MAC das portas encarregam-se das funções dependentes do método MAC.

Como se observa, as entidades MAC das portas da *bridge* prestam serviço de recepção e de transmissão de quadros à entidade *MAC Relay*. Tal serviço é referido como ISS (*Internal Sublayer Service*), representado por setas na figura.

Quando da recepção de um quadro de dados por uma porta, a entidade MAC receptora encaminha esse quadro de dados à entidade *MAC Relay* por uma primitiva de serviço M_UNITDATA_INDICATION, utilizando o ISS, contendo essa primitiva os parâmetros aplicáveis.

Após a devida realização de funções de filtragem e roteamento, a entidade *MAC Relay* envia o quadro de dados para a entidade MAC da porta de destino, utilizando o ISS, por uma primitiva de serviço M_UNIDATA_REQUEST, solicitando a sua transmissão. Essa primitiva de serviço contém um conjunto de parâmetros igual ao da primitiva de serviço M_UNITDATA_INDICATION com o acréscimo do parâmetro *access_priority.*

3.4. FAMÍLIA DE PROTOCOLOS SPANNING TREE (STP)

A família de protocolos STP é composta pelos seguintes membros:

— STP: versão zero;
— Extensão do STP: versão 1;
— *Rapid* STP (RSTP): versão 2;
— *Multiple* STP (MSTP): versão 3.

O STP e o RSTP são englobados em uma categoria referida como SST (*Single Spanning Trees*), enquanto o MSTP representa a categoria MST (*Multiple Spanning Trees*).

Na categoria de protocolos SST, o IEEE não mais recomenda a implementação do ST e de sua extensão, mas apenas do RSTP.

O objetivo básico desses protocolos é definir, com base em um processo dinâmico, uma árvore lógica sem *loops* sobre uma rede *bridged Ethernet* contendo *loops* físicos. Esse resultado é obtido pelo bloqueio lógico de um ou mais *links* físicos, mantendo-os em *stand-by*. A rede lógica constituída é também referida como a topologia ativa da rede (caso do RSTP), e também referida como topologia ativa de uma VLAN ou de um grupo de VLANs.

Quando uma *bridge* detecta uma condição de falha na rede, o protocolo STP (RSTP ou MSTP) é automaticamente reativado, no sentido de definir uma nova topologia lógica, também sem *loops* lógicos, que possibilite o restabelecimento do funcionamento da rede *bridged Ethenet*.

A Figura 3-3 apresenta o funcionamento da família de protocolos STP em uma *bridge* com duas portas, comportamento esse que pode ser estendido para *bridges* com múltiplas portas.

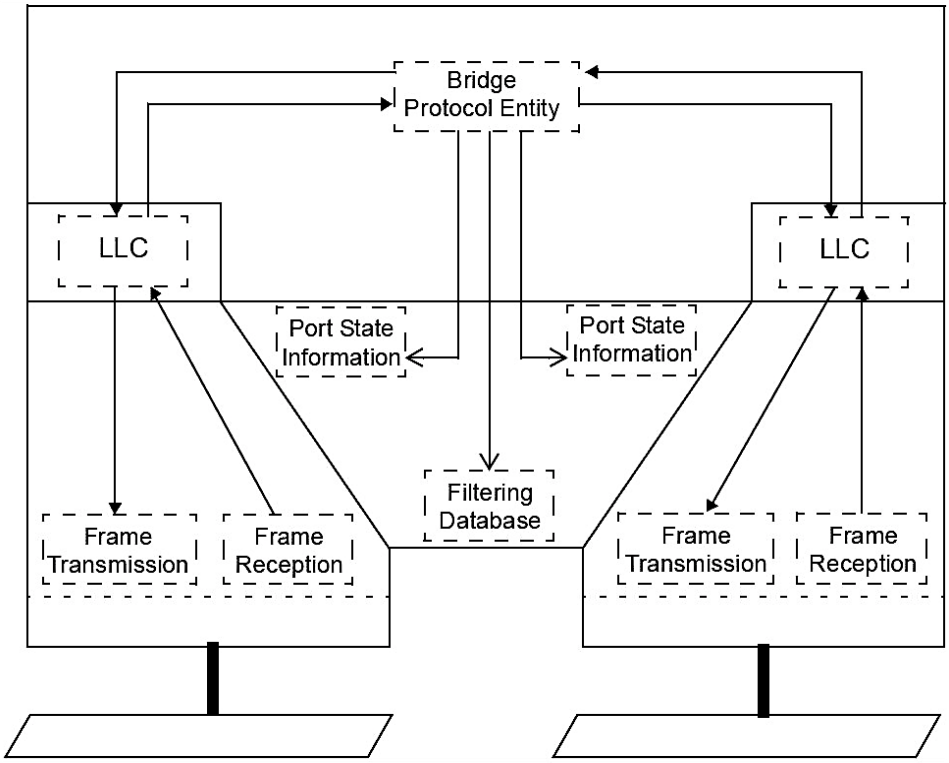


Figura 3-3– Funcionamento do STP em uma bridge.

A entidade *Bridge Protocol* nessa figura é responsável pela geração de BPDUs (*Bridge Protocol Data Units*) e pelo processamento dessas BPDUs, como será visto adiante neste capítulo, quando serão abordados o RSTP e o MSTP.

3.5. RAPID SPANNING TREE PROTOCOL (RSTP)

O padrão IEEE 802.ID-2014 define uma opção de protocolo da família STP, do tipo SST, denominada *Rapid Spanning Tree Protocol* (RSTP), cuja diferença com relação ao STP (versão 0 e versão 1) reside na maior rapidez para recuperação da conectividade em caso de falhas na rede. O RSTP foi inicialmente definido pelo padrão IEEE 802.1w-2001, que foi incorporado na versão de 2004 do padrão IEEE 802.1D.

O RSTP oferece retrocompatibilidade (*backward compatibility*) com as implementações remanescentes do STP, para facilitar a migração das implementações legadas para o RSTP. Observa-se, contudo, que as redes mistas resultantes nem sempre possibilitam a rápida reconfiguração em caso de falhas.

Serão considerados os seguintes quesitos relativamente ao RSTP:

— Conceitos básicos;
— Funções de porta e estados de porta (*Port Roles* e *Port States*);
— Formatação de RSTP BPDUs;
— Procedimentos de configuração;
— Procedimentos de reconfiguração em caso de falhas;

— Exemplo de implementação.

3.5.1. **Conceitos Básicos**

O RSTP fundamenta-se nos seguintes conceitos básicos:

— Identificador de *bridge*;
— Identificador de porta de *bridge*;
— *Root bridge*;
— *Root path cost*;
— *Root port* de uma *bridge*;
— *Designated bridge* para uma LAN ativa;
— *Designated port* em uma *designated bridge* para uma LAN ativa;
— *Edge ports;*
— *Alternate ports*;
— *Backup ports*;
— *Enabled/disabled ports*;
— *Spanning tree priority vector*.

E*dge port*, *alternate port* e *backup port* representam os *ports roles*, ou seja, os papéis que uma porta de *bridge* pode desempenhar.

Uma *enabled port* pode encontrar-se no estado *forwarding* ou no estado *blocking*. As portas pertinentes a uma RSTP ativa encontram-se no estado *forwarding*.

3.5.1.1. Identificadores de Bridges e de Portas de Bridges

Cada *bridge* recebe um identificador, que consiste em um endereço MAC e em um componente de prioridade administrável. Tais identificadores são totalmente ordenados, isto é, seus valores são sempre distintos e um deles possui sempre maior valor que os outros.

Para cada bridge deve ser atribuído um endereço MAC único universalmente administrável, denominado *bridge address*. Pode ser também utilizado como endereço MAC de uma *bridge* o endereço MAC individual de uma de suas portas.

Os valores recomendados de prioridade para *bridges* encontra-se na faixa entre 0 e 61440, em degraus de 4096.

Cada porta de *bridge* recebe um identificador, que consiste também de um endereço MAC e de um componente de prioridade administrável. Os valores recomendados de prioridade para portas de *bridge* encontram-se na faixa de 0 a 240, em degraus de 16.

3.5.1.2. Root Bridges, Root Ports e Root Path Costs

Uma RSTP é centralizada em uma *root bridge*. A *root bridge* é selecionada pelo protocolo STP como sendo a *bridge* que possui o identificador de maior prioridade (isto é, de menor valor de prioridade), e representa a raiz da *spanning tree* configurada para uma rede *bridged Ethernet*. Em caso de igualdade entre identificadores de maior prioridade, é escolhida a *bridge* com o menor valor de identificação.

Cada uma das demais *bridges* possui uma porta cujo custo total para atingir a *root bridge* (*root path cost*) é o menor relativamente aos *root path costs* para as outras portas da *bridge*. Essa porta é selecionada como sendo a *root port* da *bridge*. Em caso de igualdade de *root path costs*, aplica-se um critério de desempate que será visto adiante neste capítulo.

É pela *root port* de uma bridge que ocorre a transmissão de quadros de dados dessa *bridge* no sentido da *root bridge*, e pela qual ocorre a recepção de quadro de dados no sentido oposto.

3.5.1.3. Designated Bridges e Designated Ports

Quando uma LAN, seja ela uma LAN distribuída (*shared* LAN) ou uma LAN ponto a ponto, de uma *bridged Ethernet* LAN, é conectada a uma única *bridge*, essa *bridge* é selecionada para desempenhar o papel de *designated bridge* para essa LAN. Caso uma LAN se conecte à rede por duas ou mais *bridges*, a *designated bridge* para essa LAN será a *bridge* que apresentar o menor *path cost* para alcançar a *root bridge* a partir dessa LAN. Em caso de igualdade desses custos, a *designated bridge* será escolhida pelo mesmo critério que aquele utilizado para a escolha da *root bridge*.

A porta de uma *designated bridge* que se conecta à LAN considerada é a *designated port* dessa *designated bridge* para essa LAN. Caso exista mais de uma porta da *designated bridge* para a referida LAN, a *designated port* para essa LAN será escolhida pelo mesmo critério que o utilizado para as *root ports*.

Cada *root port* conecta-se, via uma LAN distribuída ou uma LAN ponto a ponto, à *designated port* que corresponde a essa LAN. A transmissão e a recepção de quadro de dados por portas designadas ocorre no sentido inverso daquele das *root ports*.

Uma porta de *bridge* conectada a uma LAN que não tenha conexão com outra ou com outras *bridges*, pode ser administrativamente configurada como uma *edge port* (porta de borda). Uma porta de borda é um caso particular de *designated port*.

3.5.1.4. Port Roles e Port States

Os protocolos da família STP atribuem determinadas funções às portas das bridges, funções essas denominadas *port roles* (papéis das portas). Para o caso do RSTP foram definidos os seguintes *port roles*:

— *Root port;*
— *Designated port;*
— *Alternate port;*
— *Backup port;*
— *Enabled port;*
— *Disabled port.*

Nos primeiros quatro *port roles* acima, a porta encontra-se habilitada (*enabled*). Em portas desabilitadas (*disabled*), ou não existem *links* a elas conectados ou elas se encontram administrativamente desabilitadas.

Para o RSTP, uma porta de *bridge* pode encontrar-se em um dos seguintes *port states* (estados das portas):

— *Learning* (aprendendo);
— *Forwarding* (encaminhando);
— *Discarding* (descartando).

A transição do estado *forwarding* para o estado *blocking* é imediata. A transição oposta, no entanto, requer a introdução de um *delay* entre o momento em que se seleciona uma porta bloqueada para passar ao estado *forwarding*, e o momento em que essa porta começa efetivamente a operar no estado *forwarding*. Tal medida visa a evitar a ocorrência de *loops* na rede em decorrência da mudança de estado.

Os conceitos de *root port* e de *designated port* foram já definidos anterioremente nos subitens anteriores deste capítulo. É pela *root port* de uma *bridge* que ocorre a comunicação dessa *bridge* no sentido da *root bridge.*

As *root ports* e as *designate ports*, após uma fase de *learning port state* (estado de porta "aprendendo") na fase de configuração, estabilizam-se no *forwarding port state* (estado de porta "encaminhando").

Um *link backup* é um *link* de proteção a um ou mais *links* de serviço, sendo que todos esses *links* se situam entre um mesmo par de *bridges*.

Quando um *link backup* é ativado em substituição ao(s) *link*(*s*) de serviço, a *spanning tree* da rede não se altera.

As portas de *bridge* do *link backup* são denominadas portas *backup (backup ports*).

Um *link* alternativo (*alternate link*) é um *link* de proteção que se situa isoladamente entre um par qualquer de *bridges*, protegendo assim diferentes *links* da rede. A sua ativação altera a *spanning tree* da rede.

As portas do *link* alternativo são denominadas portas alternativas (*alternate ports*).

A todas as portas habilitadas (*enabled*) que não sejam *root ports* ou *designated ports* é atribuído o papel de *alternate ports*. Em uma ST já configurada e operando normalmente, as *root ports* e as *designated ports* encontram-se no estado *forwarding*, enquanto as *alternate ports* encontram-se no estado *blocking*.

A Figura 3-4 ilustra os conceitos de *alternate ports* e de *backup ports*.

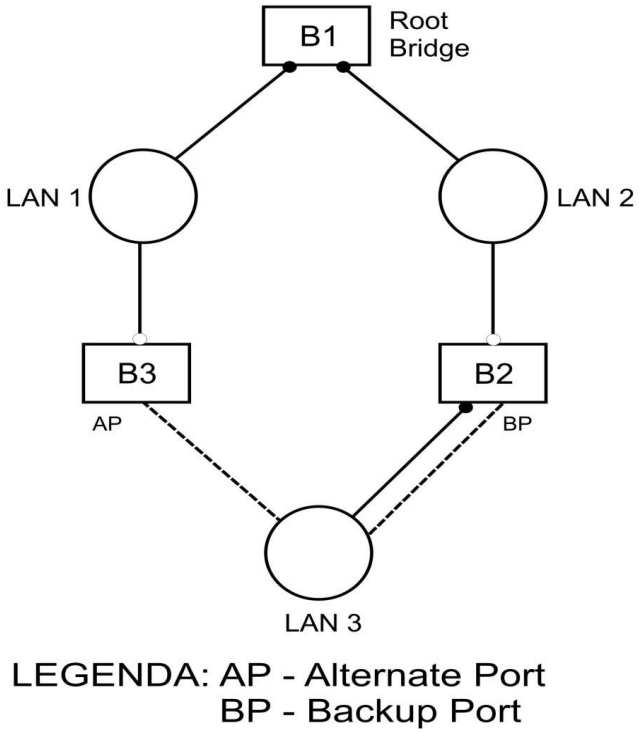


Figura 3-4– Alternate ports e backup ports.

Como se observa nessa figura, uma *backup port* é um tipo de porta de reserva da *designated bridge* voltada para um segmento de LAN na própria *bridge*. Uma *alternate port* é uma porta de reserva voltada para um segmento de LAN, mas que se encontra em outra *bridge* que não a *designated bridge*.

As *alternate ports* e as *backup ports* operam em *discarding port state* (estado de porta "descartando") após a sua estabilização. Ambas encontram-se disponíveis para uma eventual mudança do papel desempenhado por elas.

Um *port role* pode alterar-se se uma *bridge*, uma porta de *bridge* ou uma LAN falhar, for removida ou for inserida na *bridged* LAN. Uma *backup port* serve de proteção para o caso de falha na *designated port* ou no *link* local dessa porta, e entra em ação sem acarretar mudança na RST. As *alternate ports*, que servem de proteção para a rede como um todo, podem vir a participar de um novo ramo da RST, enquanto as *backup ports* não o podem.

Uma *alternate port* ou uma *backup port* pode passar a transmitir sem que os *timers* de protocolo expirem. Uma *alternate port* pode passar a operar como uma *root port* ou uma *designated port* sem que tenha recebido mensagens de outras *bridges*.

3.5.1.5. Spanning Tree Priority Vector

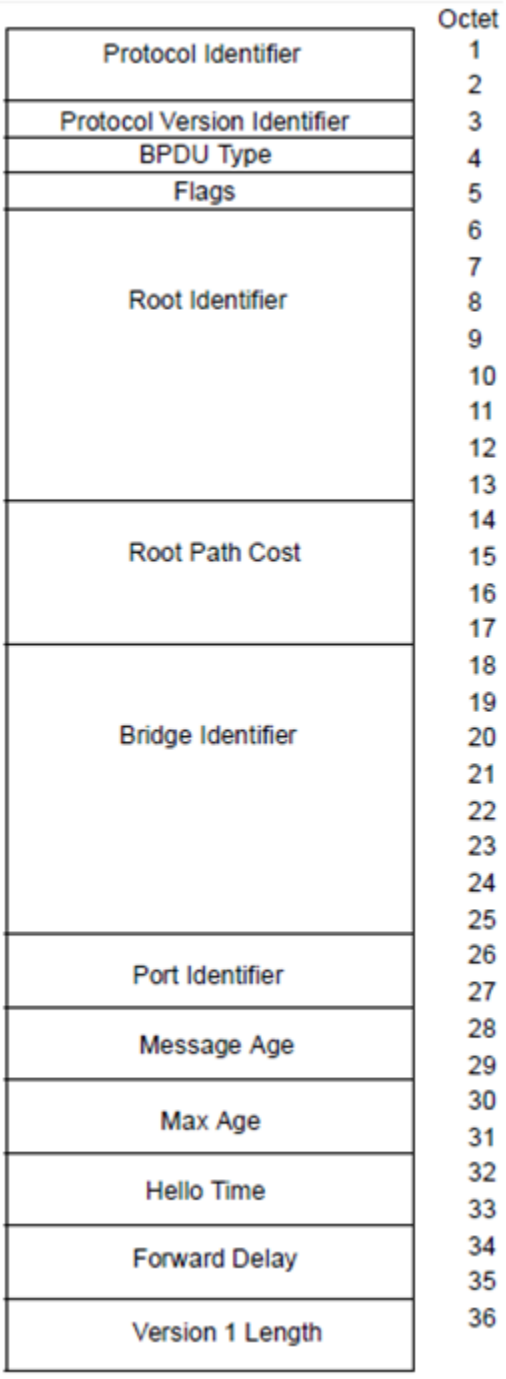
Spanning Tree Priority Vector é o conjunto de informações enviado em mensagens de configuração (*configuration messages*) entre RSTP *bridges*, para

selecionar a *root bridge* e o caminho mais curto (com menor custo) para chegar à *root bridge*.

3.5.2. **Formatação de RST BPDUs**

Este subitem apresenta a estrutura e a codificação das RST BPDUs conforme Seção 9.3.3 do padrão IEEE 802.1D-2004.

Ao contrário do STP, que possui dois tipos de BPDU, o RSTP apresenta uma



única BPDU, referida com RST BPDU, cujo formato encontra-se na Figura 3-5.

Figura 3-5– Formato de RST BPDUs.

O campo *protocol* ID utiliza o valor 0x0000, indicando a família de protocolos *spanning tree*.

O campo *Protocol Version Identifier* utiliza o valor 0x02, o que indica uma RST BPDU.

O campo *flags* possui os seguintes significados:

— TC *flag* (bit 1): indica *Topology Change Notification* BPDU;
— *Proposal flag* (bit 2);
— *Port Role flag* (bits 3 e 4);
— *Learning flag* (bit 5);
— *Forwarding flag* (bit 6);
— *Agreement flag* (bit 7);
— TCA *flag* (bit 8): indica *Topology Change Acknowledgement*.

O campo *Root* ID conduz a identificação de uma *bridge* definida como *root bridge* ou ainda na fase em que essa *bridge* supõe ser a *root bridge*.

O campo *Root Path Cost* indica o custo total relativo ao caminho para a *root bridge (root path cost)* da porta da *bridge* que originou a BPDU. O *root path cost* escolhido é o menor dos somatórios dos *port path costs* para atingir a *root bridge* a partir da porta da *bridge* originadora da BPDU.

Cada *port path cost* corresponde a um valor de custo atribuído ao *link* entre a *root port* da respectiva *bridge* e a *designated port* da próxima *bridge*. Esses valores de custo são tipicamente associados às velocidades de transmissão dos *links*.

O campo *Bridge* ID contém a identificação da *bridge* originadora da BPDU. Esse campo é codificado em 8 octetos.

O campo *Port ID* contém a identificação da porta pela qual a *bridge* originadora transmitiu a BPDU.

O campo *Message Age* representa o tempo decorrido entre o momento em que o quadro se encontra e o da transmissão inicial de uma RST BPDU pela *root bridge*. Ao ser emitida a RST BPDU, o valor contido neste campo é zero. Cada *bridge* ao longo do caminho incrementa o valor desse campo.

O campo *Maximum Age* (*Max Age*) representa o valor máximo admitido para o *message age*. Atingido o *maximum age*, a BPDU deve ser descartada.

O campo *Hello Time* representa o período de tempo entre emissões periódicas de RSTP BPDUs com a função de configuração por portas designadas.

O campo *Forward Delay* representa o período de tempo utilizado por STP *bridges* para a transição de *root ports* e de *designated ports* para o estado de porta encaminhando (*forwarding*).

O campo *Version 1 lenght* assume o valor 0x00, o que indica a inexistência da presença de informação de protocolo da versão 1.

3.5.3. **Procedimentos de Configuração**

Na fase de configuração, o RSTP realiza os seguintes procedimentos:

— Cada *bridge* supõe ser a *root bridge* e envia uma RST BPDU de configuração em cada um dos segmentos de LAN a que está conectada, notificando assim as demais *bridges* que estejam conectadas a esses segmentos. Essa notificação é enviada para o endereço MAC de grupo pré-estabelecido para as *bridged* LANs (01-80-C2-00-00-00), denominado *Bridge Group Address* (representando "*All Bridges on This* LAN*").*
— Essas notificações são processadas por todas as *bridges* da rede *bridged Ethernet*, que as retransmitem sucessivamente.
— As *bridges* receptoras comparam as RST BPDUs recebidas e concluem qual a *bridge* que consideram a *bridge* raiz. Esse processo se repete sucessivamente até que todas as *bridges* convirjam quanto ao conhecimento da identificação da *root bridge.*
— Para comparar dois *bridge* IDs, comparam-se os respectivos valores. Prevalece o valor mais baixo. Portanto, se os componentes de prioridade, correspondentes aos bits mais significativos dos *bridge* IDs, forem diferentes, prevalece o valor considerado como de maior prioridade (número mais baixo). Se forem iguais, comparam-se os bits restantes que compõem o *bridge* ID, prevalecendo o valor mais baixo.
— Cada uma das *bridges* (exceto a *root bridge*) identifica a sua porta com o menor *path cost* para a *root bridge* (*root path cost*). Essa porta é selecionada como sendo a *root port* da *bridge*. Em caso de igualdade de *root path costs* para diferentes portas da *bridge*, será selecionada aquela cuja *bridge* adjacente no sentido da *root bridge* tem o menor valor de identificação.
— Em cada LAN, as *bridges* a ela conectadas identificam qual delas possui o menor *root path cost* para alcançar a *root bridge* a partir dessa LAN. Essa *bridge* é a *designated bridge* para essa LAN. Em caso de igualdade de custos, será escolhida a *bridge* que tiver o menor valor de identificação.
— A porta de uma *designated bridge* para uma LAN é denominada *designated port* para essa LAN. Caso haja mais de uma porta nessa condição, será escolhida como *designated port* aquela com o menor valor de identificação.

3.5.4. **Exemplo de Configuração**

Será considerada a configuração física de uma *bridged Ethernet* LAN (antes da aplicação do RSTP), exposta na Figura 3-6.

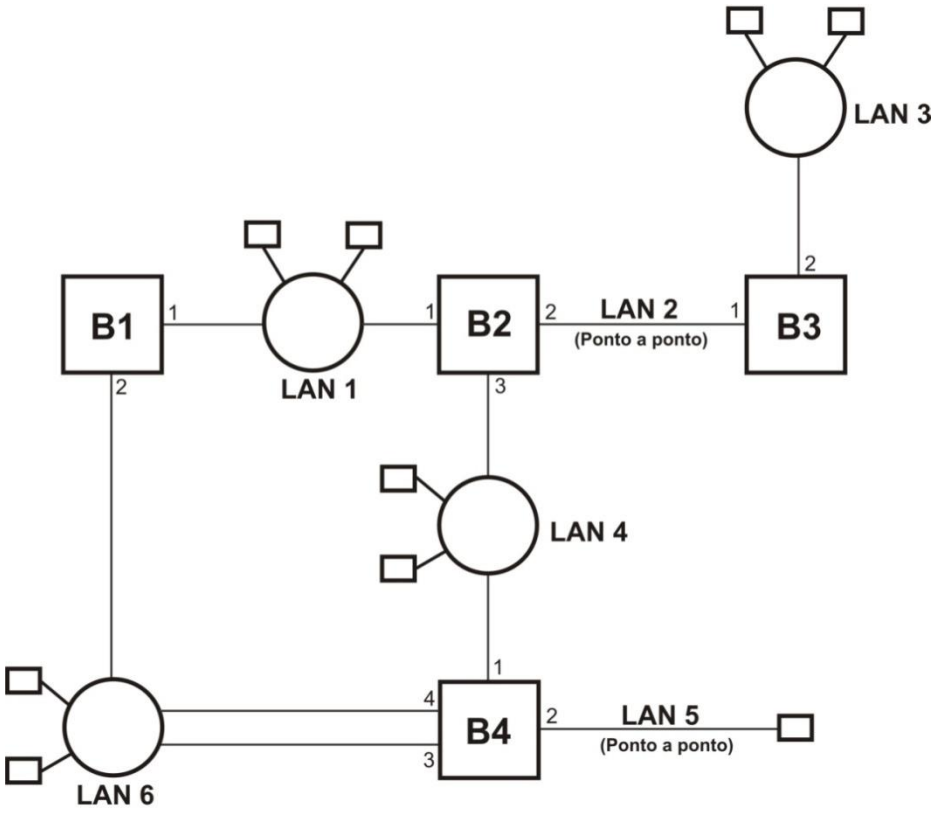


Figura 3-6– Configuração física de uma bridged Ethernet LAN.

Serão consideradas as seguintes hipóteses para a aplicação do RSTP nessa figura:

— A *bridge* B2, por ter o menor valor de identificação, ficou definida pelo algoritmo RSTP como a *root bridge*;
— O valor do *port path cost* é único em todos os segmentos de LAN.

Verifica-se, então, que a porta 1 da *bridge* B1 é a sua *root port*. Da mesma forma, conclui-se ser a porta 1 a *root port* da *bridge* B3. Para a *bridge* B4, verifica-se que a porta 1 é a *root port*.

A *bridge* B2 (*root bridge*) é naturalmente a *designated bridge* para as LAN 1, LAN 2 e LAN 4, sendo as portas 1, 2 e 3 da *bridge* B2 as *designated ports*, respectivamente, para as LAN 1, LAN 2 e LAN 4.

Para a LAN 3, a B3 é naturalmente a *designated bridge*, sendo a porta 2 da *bridge* B3 a sua *designated port*. Da mesma forma, a *bridge* B4 é a *designated bridge* para LAN 5, e a sua porta 2 a *designated port* para essa LAN.

Para a LAN 6, como há igualdade de *root path costs* para as *bridges* B1 e B4, será suposto que foi escolhida pelo algoritmo RSTP a *bridge* B4 como a *designated bridge*, por possuir o menor valor de identificação. Em consequência, a porta 2 da *bridge* B1 passa para o estado bloqueado.

Supõe-se a porta 4 da *bridge* B4 como a *designated port* para a LAN 6, por apresentar menor valor de identificação que a porta 3. Em consequência, a porta 3 da *bridge* B4 passa para o estado bloqueado.

Resulta, após a finalização da aplicação do algoritmo RSTP, a configuração lógica (ST) representada na Figura 3- 3.7.

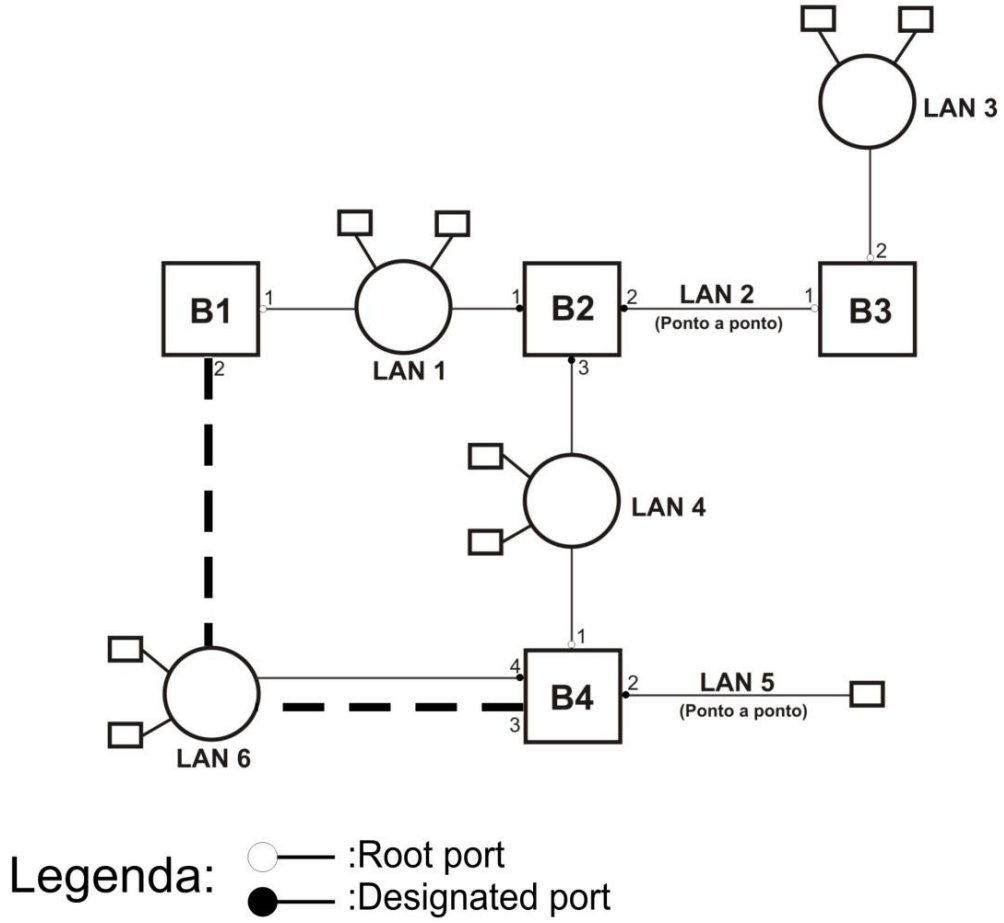


Figura 3-7– Configuração lógica resultante.

3.5.5. **Procedimentos de Reconfiguração**

No RSTP, caso a conectividade da rede após a estabilização de sua configuração venha a ser afetada por uma falha na rede ou por uma mudança de topologia, a rede obtém a sua configuração mais rapidamente que no STP, graças às novas facilidades introduzidas pelo algoritmo RSTP. A idéia básica que norteia essas novas facilidades é a descentralização de funções antes concentradas na *root bridge*.

No STP versão 0, quando uma BPDU TCN (*Topology Change Notification*) é gerada em um ponto da rede, essa BPDU é enviada para a *root bridge*. Cabe à *root bridge* a distribuição da mensagem contida nessa BPDU para a totalidade das *bridges* da rede. As *non-root bridges* repassam BPDUs quando as recebem em suas *root ports*.

No RSTP, diferentemente, cada *bridge* da *spanning tree* gera, em suas portas designadas, BPDUs a cada *hello time* (2 segundos, como *default*), independentemente do recebimento de BPDUs geradas por qualquer outra *bridge*.

Se uma *bridge* deixa de receber BPDUs de uma *bridge* vizinha por três *hello times*, ela apaga seus registros relativos ao *link* envolvido. As BPDUs atuam como um mecanismo *keep-alive*. Dessa forma, as falhas na rede são detectadas com maior rapidez.

No RSTP, portas que se encontram no estado *discarding* (*alternate* e *backup ports*) podem transicionar para o estado *forwarding* sem a necessidade de aguardar a expiração de temporizadores de protocolo.

Uma porta raiz (*root port*) pode transicionar para o estado *forwarding* sem intercambiar mensagens com outras bridges, enquanto uma porta designada conectada a uma LAN ponto a ponto pode transicionar quando ela é notificada da ocorrência de um acordo para o novo papel de porta (*port role*) pela outra *bridge* conectada a essa LAN.

Uma porta designada conectada a uma LAN compartilhada submete-se a um *delay* de transição após a notificação do acordo acima mencionado. Esse *delay* é longo o suficiente para que as demais *bridges* conectadas a essa LAN compartilhada recebam as mensagens transmitidas pela *bridges* de origem e atuem sobre essas mensagens.

No RSTP, uma *bridge* originadora de uma RST BPDU com a função TCN, distribui essa BPDU diretamente e de forma descentralizada, sem a intermediação da *root bridge*. Se uma *bridge* percebe uma falha no sentido da *root bridge*, ela imediatamente bloqueia a sua *root port* e estabelece uma negociação com uma *bridge* vizinha, através da respectiva *alternate port*, com o propósito de estabelecer uma nova *spanning tree.* Tal negociação tem como base os *flags* para isso destinados.

Fica assim definida uma nova *root port* para a *bridge* que percebeu a falha, por onde essa *bridge* passa a enviar e receber o tráfego de dados. Essa porta altera o seu papel de *alternate port* para *root port*.

Será agora examinado o que ocorre quando um novo *link* que vai participar da *spanning tree* é introduzido na rede. Para isso, será utilizado como exemplo a Figura 3-, que representa a configuração inicial de uma RST.

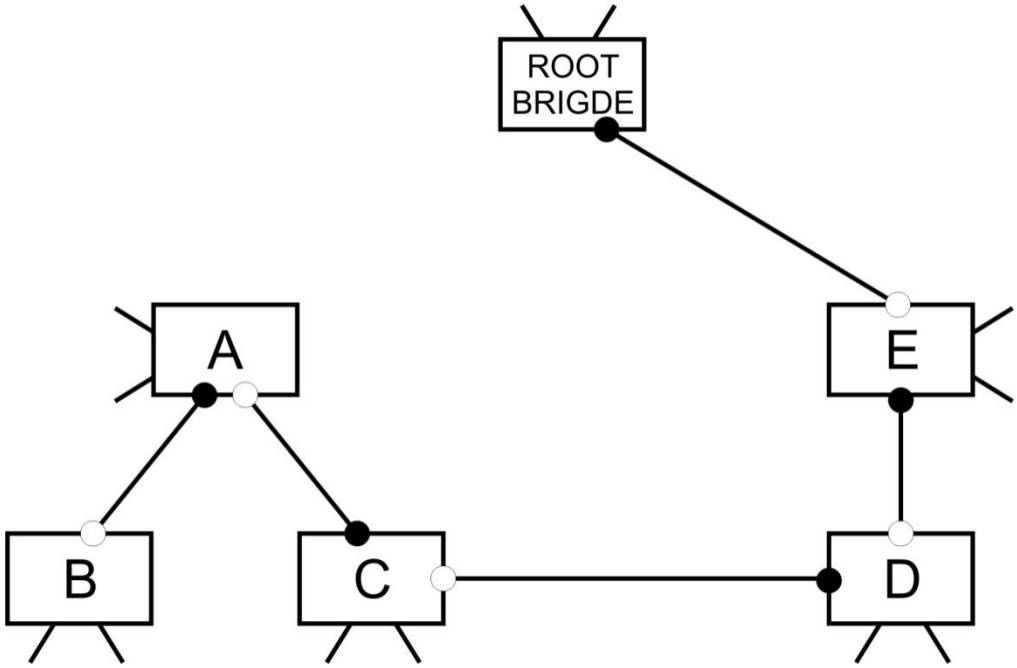


Figura 3-8– Configuração inicial de uma RST.

Nessa figura, os pequenos círculos claros representam portas raiz (RPs), enquanto os pequenos círculos pretos representam portas designadas (DPs).

Suponha-se que tenha sido introduzido um novo *link* entre a *bridge* A e a *root bridge* na figura.

As *bridges* conectadas pelo novo *link (bridge* A e a *root bridge*) entram inicialmente em estado de bloqueio e em um processo de negociação, iniciado pela *bridge* raiz, que envia uma RST BPDU para a *bridge* A com o *flag* TC setado e com os demais *flags* com os valores necessários à inclusão do novo *link.*

Ao receber essa RST BPDU da *bridge* raiz por esse novo *link*, a *brigde* A bloqueia todas as suas portas que não sejam *edge ports*, ou seja, bloqueia todas as portas que dão acesso a outras *bridges*. Essa negociação denomina-se *sync operation*. A *bridge* A coloca a sua porta conectada ao novo *link* no estado *forward*ing e no papel *root port*.

Uma vez bloqueadas essas *non-edge ports*, a *bridge* A autoriza a *bridge* na outra ponta do novo *link* (a *root bridge*, neste caso) colocar a sua porta conectada ao novo *link* no estado *forwarding* e no papel *designated port*. Nota-se que embora o novo *link* já entre em atividade, não existe a possibilidade de *loop* porque a *spanning tree* foi interrompida abaixo da *bridge* A.

Inicia-se, então, um processo de negociação com as portas das demais *bridges* adjacentes abaixo da *bridge* A (*bridges* B e C). Como a *bridge* B não será afetada pela negociação, a transição é imediata, por não haver necessidade de bloquear suas portas.

No caso da *bridge* C, a sua comunicação com a *bridge* D precisa ser bloqueada antes de qualquer transição.

A *bridge* C então define, com base nas informações de custo obtidas das *bridges* adjacentes (*bridge* A e *bridge* D), qual será o novo papel de suas portas. Por hipótese, estamos admitindo que a porta da *bridge* C para a *bridge* A será sua *root port*, e que o *link* entre as *bridges* C e D fique bloqueado, pois cada uma dessas *bridges* possui uma alternativa melhor para alcançar a *root bridge*.

Dessa forma, a configuração final da *spanning tree* foi alcançada rapidamente, sem necessidade de uso de nenhum *time-out*.

Note-se que essas negociações só são possíveis quando as *bridges* estão conectadas por *links* ponto a ponto.

A Figura 3- ilustra a RST resultante após a introdução do novo *link*.

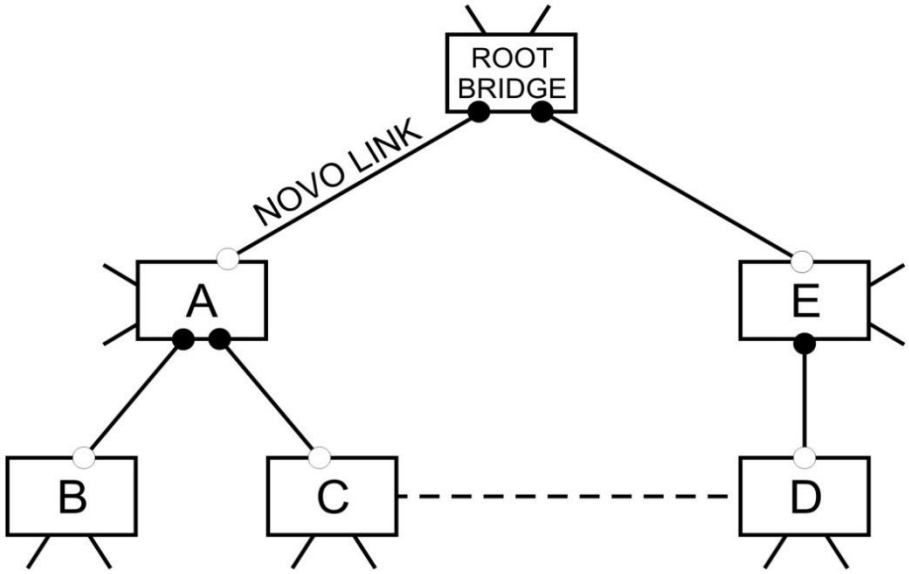


Figura 3-9– Configuração da RST resultante.

3.6. FAMÍLIA DE PROTOCOLOS MULTIPLE REGISTRATION (MRP)

A família de protocolos GARP (*Generic Attribute Registration Protocol*), definida pelo padrão IEEE 802.1D, foi obsoletada e substituída pela família de protocolos MRP, definida inicialmente pelo padrão IEEE 802.1ak, sendo que o padrão IEEE 802.1ak foi descontinuado e incorporado ao padrão 802.1Q (Seções 10 e 11).

Em consequência, os protocolos GVRP (GARP VLAN *Registration Protocol*) e GMRP (GARP *Multicast Registration Protocol*) foram obsoletados e substituídos, respectivamente, pelos protocolos MVPR (*Multiple* VLAN *Registration Protocol*) e MMRP (*Multiple* MAC *Registration Protocol*), também hoje definidos pelo padrão 802.1Q.

A finalidade da família MRP é possibilitar o registro dinâmico de diferentes atributos, tais como a participação de portas de *bridges* ou de estações terminais em VLANs. Um outro exemplo é o registro dinâmico de participação em endereços

MAC de grupo (endereços *multicast*) ou registro dinâmico de endereços MAC individuais.

Outros exemplos de uso do MRP são o MSP (*Multiple Stream Registration Protocol*), definido na Cláusula 35 do padrão IEEE 802.1Q-2014, e o MIRP (*Multiple I-SID Registration Protocol*), definido na Cláusula 39 desse padrão.

A Cláusula 35 do padrão IEEE 802.1Q-2014 define o *Stream Registration Protocol* (SRP), que utiliza três protocolos de sinalização, que são o MMRP, o MVRP e o MSRP (*Multiple Stream Registration Prococol*), para estabelecer reservas de recursos em uma *bridged* LAN para fluxos de dados.

O MSRP provê estações finais com a capacidade de reservar recursos de rede que garantirão a transmissão e a recepção de fluxos de quadros de dados através da rede, com a qualidade de serviço solicitada.

As estações finais que produzem fluxos de dados são referidas como *Talkers*, enquanto as estações finais que consomem esses fluxos de dados são denominadas *Listeners*.

Os *Talkers* declaram os atributos necessários à alocação de recursos para um fluxo de dados. Os *Listeners* declaram os atributos que requisitam a recepção desse fluxo de dados.

A operação do MSRP assemelha-se a operação do protocolo IP de reserva de recursos, que é o RSVP (*Resource Reservation Protocol*).

A arquitetura do MSRP e a formatação de suas PDUs são aquelas da família MRP, com a devida adequação.

O endereço de grupo adotado para o MSRP é o endereço *Individual LAN Scope group address*, *Nearest Bridge group address* (01-80-C2-00-00-2E).

O *EtherType* para o MSRP utiliza o valor 22-EA (hexadecimal).

A Cláusula 39 do padrão IEEE 802.1Q-2014 define o MIRP (*Multiple I-SID Registration Protocol*) que provê a habilidade de remover entradas de endereços MAC (*MAC Address Entries*) aprendidos que se encontravam registradas na FDB (*Forwarding Database*) de um *I-Component* de uma PBBN (*PBB Network*). Esses registros encontravam-se realizados por *I-SID*.

Um MIRP é uma extensão do do MRP para uso em *I-Components* e em *B-Components*, com a expansão dos respectivos valores de atributo de 12 bits para 24 bits.

O valor *EtherType* utilizado pelo MIRP é igual a 89-29. O endereço MAC-DA *multicast* utilizado pelo MIRP é o *Nearest Customer bridge group address* (01-80-C2-00-00-00).

3.6.1. **Arquitetura do MRP**

A arquitetura da família de protocolos MRP encontra-se representada na Figura 3-.

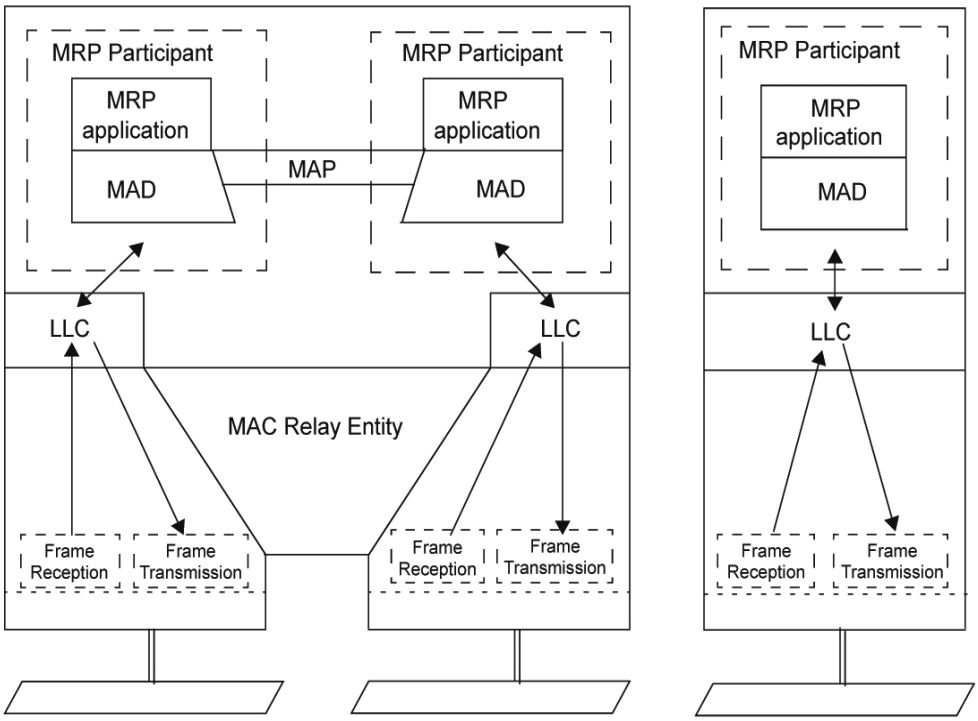


Figura 3-10– Arquitetura da família MRP.

Nessa figura encontram-se representados os componentes de *MRP Participants* em uma *bridge* com duas portas e em uma estação final.

Um *MRP Participant* de uma *bridge* consiste nos seguintes componentes:

— A aplicação MRP;
— O componente *MRP Attribute Declaration* (MAD);
— O component *MRP Attribute Propagation* (MAP).

No caso de uma estação final, o *MRP Participant* consiste apenas na aplicação MRP e no MAD.

O componente Aplicação MRP é o responsável pela semântica associada com os valores de atributo e os respectivos registros, incluindo o uso de novas declarações explicitamente sinalizadas.

O componente MAD executa o MRP, gerando mensagens MRP para transmissão e processando as mensagens MRP recebidas de outros *MRP Participants*.

A função *MRP Attribute Propagation* (MAP) possibilita a propagação de atributos registrados em portas de *bridges* constituintes da *bridged LAN* para outros *MRP Participants.* Cada aplicação MRP especifica a operação da respectiva função MAP.

Para melhor entendimento dessa arquitetura recomenda-se aos leitores consultar os sub-itens referentes às arquiteturas do MMRP e do MVRP, que se encontram descritas adiante neste capítulo.

3.6.2. **Operação do MRP**

O MRP permite que participantes de uma aplicação MRP (tais como o MMRP e o MVRP) possam registrar atributos junto a outros participantes dessa aplicação em uma *bridged* LAN. A definição dos tipos de atributos, de seus valores e da correspondente semântica, é específica para cada aplicação MRP.

Os participantes de uma aplicação MRP realizam e cancelam declarações de atributos, que causam a criação ou a remoção do registro (*register*/*deregister*) desses atributos junto a outros participantes dessa aplicação.

As declarações de atributos, realizadas pela transmissão de MRPDUs, utilizam os *links* da *bridged* LAN que participam de uma topologia ativa, ou seja, de uma *spanning tree*. O processo de declaração, que ocorre em cascata, atinge todas as *bridges* que compõem a rede ou que compõem uma VLAN.

Tais declarações alcançam também todas as estações conectadas às *bridges* por LANs compartilhadas ou por LANs ponto a ponto (conexão direta de uma estação à respectiva *bridge*).

Dessa forma, o processo de filtragem de quadros MAC *unicast*, *multicast* ou *broadcast*, proporcionado pela aplicação MRP às *bridges* da rede, estende-se às estações finais conectadas a essas *bridges* que estejam habilitadas a participar da aplicação MRP. Tais estações finais podem, inclusive, efetuar *source pruning*.

Source pruning é um comportamento que permite a estações finais o não envio de um quadro quando não há receptores registrados como participantes de uma aplicação MRP na porta pela qual esse quadro seria transmitido. Esse comportamento elimina a transmissão desnecessária de tráfego por estações finais.

A Figura 3-6 ilustra o resultado das declarações, por múltiplas estações finais, de algum atributo (participação em endereços de grupo ou em VLANs, por exemplo), onde se evidencia a propagação dessas declarações.

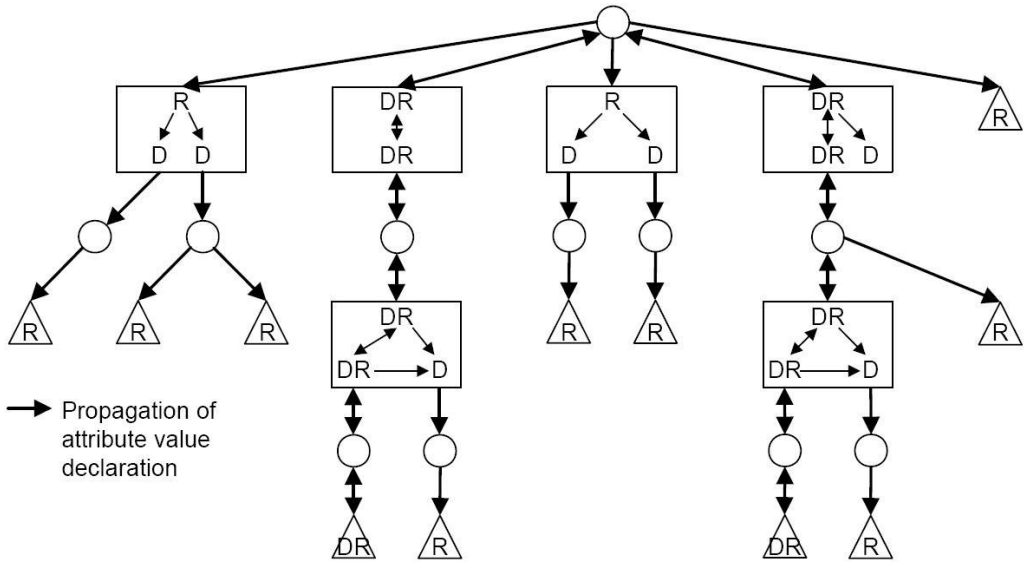


Figura 3-6– Propagação de declaração de atributos no MRP.

Nessa figura, R significa *registration*, D significa *declaration* e as setas indicam o sentido da propagação das declarações e também o sentido do tráfego de dados, caso esse tráfego de dados exista.

A figura representa declarações por apenas duas estações finais. É importante registrar que os efeitos dessas declarações ocorrem no tráfego no mesmo sentido das declarações (sentido contrário ao das solicitações de registro).

Uma estação final ou uma porta de *bridge* que registra uma participação MRP, está comunicando à porta da *bridge* ou à estação final vizinha, a sua habilitação para receber tráfego de dados na aplicação MRP considerada.

Uma estação final ou porta de *bridge* que declara uma participação MRP está notificando a porta de *bridge* ou a estação final vizinha, que registrou o pedido de registro, e que poderá então passar a enviar tráfego de dados no sentido da propagação da declaração na aplicação MRP, caso esse tráfego de dados exista.

Nessa figura, as estações finais que se limitam à solicitação de registro em uma aplicação MRP habilitam-se somente a receber tráfego de dados. Tais estações, por não receberem solicitação de registro por parte das correspondentes portas de bridge, não transmitem quadros de dados.

3.6.3. **Endereços e Valores EtherType no MRP**

Todas as aplicações MRP utilizam endereços MAC de grupo como endereços de destino para suas MRPDUs, selecionados de acordo com cada uma das aplicações. Tais endereços MAC encontram-se representados na Figura 3-7.

Atribuição	**Valor**
MMRP address	01-80-C2-00-00-20
MVRP address	01-80-C2-00-00-21
Reservado	01-80-C2-00-00-22 até 01-80-C2-00-00-2D
MSRP address	01-80-C2-00-00-2E
Reservado	01-80-C2-00-00-2F

Figura 3-7– Endereços para aplicações MRP.

A transmissão e a recepção de MRPDUs entre MRP *Participants* devem utilizar os procedimentos do protocolo LLC com uso do SNAP. Em decorrência, cada aplicação MRP utiliza um código *EtherType* próprio para a sua identificação. A Figura 3-8 especifica os valores *EtherType* atribuídos às aplicações MRP já definidas.

Atribuição	Valor
MMRP EtherType	88-F6
MVRP EtherType	88-F5
MSRP EtherType	22-EA
MIRP EtherType	89-29

Figura 3-8– Valores *EtherType* para o MRP.

3.6.4. Formato das MRPDUs

O formato das MRPDUs, válido para qualquer uma das aplicações do MRP, encontra-se representado na Figura 3-9.

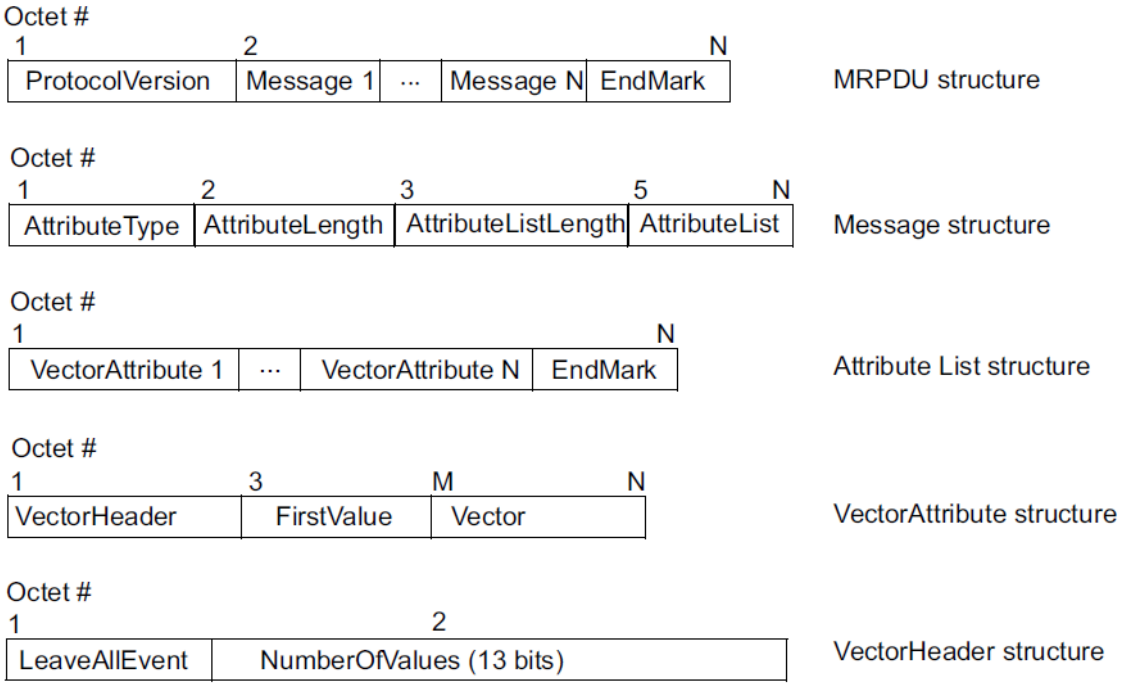


Figura 3-9– Formato das MRPDUs.

Como se observa nessa figura, o *payload* de uma MRPDU consiste em múltiplas mensagens. Tal *payload* é transportado entre um campo *Protocol Version* (cujo valor depende da aplicação MRP utilizada) e um campo final *End Mark*.

Cada mensagem transporta um tipo de atributo, especificado em um campo *AttributeType*, seguido pelos campos *AttributeLenght*, *AttributeListLenght* e *AttributeList*. O campo *AttributeList* é composto por múltiplos subcampos *VectorAttribute*, e por um campo final *EndMark*.

Cada subcampo *VectorAttribute*, por sua vez, compõe-se de um *VectorHeader*, de um campo *FirstValue* (definido pela aplicação MRP específica) e do campo *Vector*.

Finalmente, o campo *VectorHeader* é composto por um campo *LeaveAllEvent* e de um campo *NumberOfValues*.

Os números de octetos desses campos encontram-se especificados na figura.

3.7. MMRP (MULTIPLE MAC REGISTRATION PROTOCOL)

Conforme menção anterior, o MMRP representa uma aplicação do MRP que obsoletou e substituiu o GMRP a partir da emissão do padrão IEEE 802.1ak, tendo sido esse padrão obsoletado e incorporado ao padrão IEEE 802.1Q-2014.

O MMRP provê um mecanismo que possibilita a estações finais para isso habilitadas, registrar, e subsequentemente desregistrar, automaticamente a sua habilitação para receber quadros de dados com um determinado endereço de grupo, junto às portas de *bridge* às quais essas estações estão diretamente conectadas.

Possibilita também que portas de *bridges* possam registrar, e subsequentemente desregistrar, automaticamente a sua habilitação para receber quadros de dados com determinado endereço de grupo, junto a estações finais a elas (portas de *bridges* que solicitam registro) diretamente conectdas e para isso habilitadas, ou junto a portas de outras *bridges* a elas diretamente conectadas.

Em qualquer dessas hipóteses, a porta de *bridge* ou a estação final que tenha recebido uma solicitação de registro, deve responder com uma declaração de efetivação do registro solicitado, estando pronta, portanto, para enviar quadros de dados com o endereço de grupo registrado.

O MMRP permite, consequentemente, que as solicitações de registro sejam então disseminadas em cascata ao longo da rede, sendo tais solicitações registradas pelas portas de *bridge* e estações finais para isso habilitadas. Tais portas de *bridge* e estações finais declaram a realização dos registros solicitados, também em cascata, no sentido contrário ao das solicitações de registro.

3.7.1. Registro de Informações

As informações de registro e de cancelamento de registro disseminadas via MMRP são dos seguintes tipos:

- Informações de participação em grupos identificados por endereços MAC de grupo (endereços *multicast*);
- Informações de requerimento de serviço de grupo;
- Informações de endereços MAC individuais.

O registro das informações de participação em grupos identificados por endereços MAC de grupo garante que os quadros com esses endereços *multicast* sejam enviados exclusivamente nas direções onde efetivamente existam estações participantes dos grupos endereçados. Dessa forma, a utilização dos recursos das *bridged* LANs se torna mais eficiente.

3.7.2. **Arquitetura do MMRP**

A arquitetura do MMRP encontra-se ilustrada na Figura 3-10, que representa também o modo de operação desse protocolo.

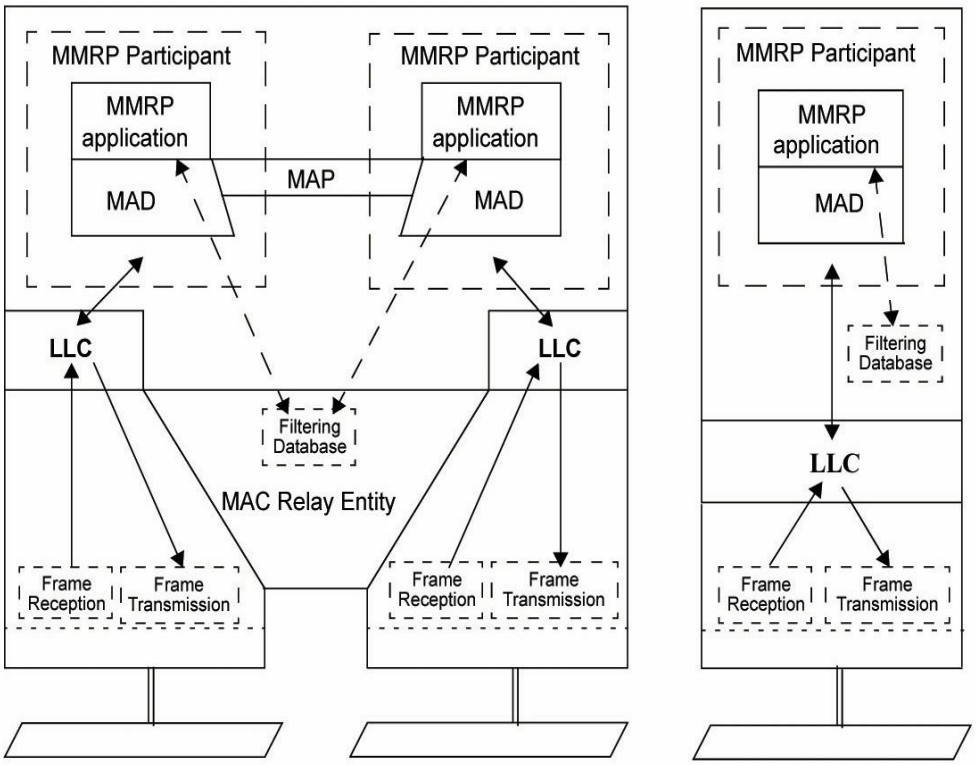


Figura 3-10– Arquitetura do MMRP.

3.7.3. **Exemplo de Aplicação do MMRP**

A Figura 3-11 representa um exemplo de aplicação do MMRP, com transmissão de um quadro MAC *multicast* correspondente a um endereço MAC de grupo.

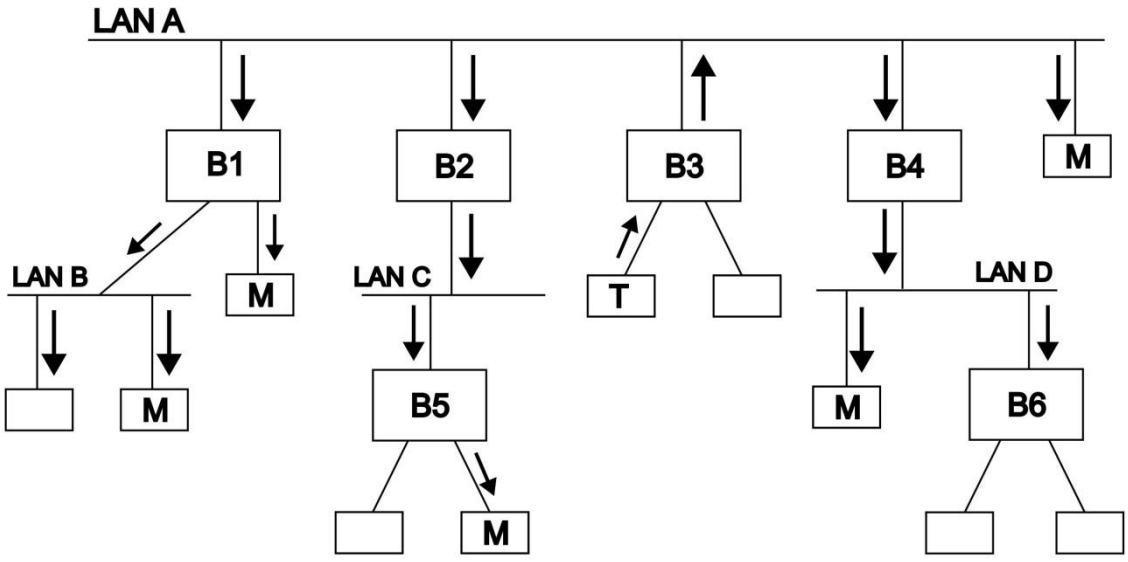


Figura 3-11– Exemplo de aplicação do MMRP.

As estações designadas como *M* são aquelas que registraram, pelo MMRP, a sua participação no grupo identificado pelo endereço MAC de grupo que representa o destino do quadro MAC *multicast* transmitido.

A estação final designada como *T* é a estação transmissora do quadro com endereço *multicast*. Essa estação final transmissora pode não ser membro do grupo. O que define a participação de uma estação final em um endereço de grupo é o fato dessa estação final ter solicitado o seu registro no endereço de grupo junto à porta de *bridge* correspondente.

As *bridges* da rede efetuam o encaminhamento do quadro de dados de acordo com as suas bases de dados de filtragem MMRP. Nesse exemplo, como se verifica na figura, as *bridges* B3, B5 e B6 não efetuam o encaminhamento do quadro de dados pelas portas por onde não ocorreram registros de participação no grupo.

3.7.4. **Utilização do IGMP**

Pressupondo-se que os quadros MAC, inclusive os quadros MAC *multicast*, transportem datagramas IP, é possível a utilização do IGMP (*Internet Group Management Protocol*), do TCP/IP, para desempenhar a função do MMRP. Tal fato já ocorria com o GMRP quando esse protocolo ainda era especificado pelo IEEE.

O IGMP realiza, no TCP/IP, a função desempenhada pelo MMRP em *LAN Bridging*. Torna-se então possível, por "intromissão" (*snooping*) nos pacotes IGMP, aproveitar o processo IGMP para desempenhar o papel do MMRP. As *bridges* podem, nesse caso, além de interceptar e sumarizar pacotes IGMP, gerar tais pacotes como se fossem roteadores.

Embora esse procedimento viole as regras básicas da modelagem OSI, a economia de recursos de rede, obtida pela não transmissão de MMRPDUs, representa uma justificativa válida e aceita para a sua adoção.

3.8. **MODOS DE FILTRAGEM POR BRIDGES IEEE 802.1D**

Os serviços de filtragens em IEEE 802.1D *bridging* baseiam-se em entradas na base de dados de filtragem (*filtering database*) de uma *bridge* IEEE 802.1D (isto é, de uma *VLAN-unaware bridge*) contendo informações específicas.

Existem duas modalidades de filtragem por parte de *bridges* IEEE 802.1D:

— Modo Básico de Filtragem;
— Modo Ampliado de Filtragem (*Extended Filtering Mode*).

O Modo Básico de Filtragem não leva em conta o processo MMRP. As *bridges* que operam nesse modo transmitem quadros *multicast* por inundação (dentro da *spanning tree*).

O modo básico de filtragem utiliza os seguintes tipos de entrada:

— Entradas de filtragem estáticas para quadros com endereços *unicast*;
— Entradas de filtragem dinâmicas que resultam do processo de aprendizagem de endereços MAC *unicast*.

No Modo Ampliado de Filtragem, que se superpõe ao Modo Básico de Filtragem, a filtragem utiliza tabelas de registros de endereços de grupo (que podem ser estáticas ou dinamicamente mantidas através da operação do MMRP), endereços esses associados às portas por onde ocorreram os registros.

O Modo Ampliado de Filtragem utiliza, complementarmente ao Modo Básico de Filtragem, os seguintes tipos de entrada:

— Entradas de filtragem estáticas para quadros com endereços *multicast*;
— Entradas de filtragem dinâmicas que resultam da aplicação do MMRP.

O MMRP pode operar também dentro do contexto de VLANs. Em VLAN *bridges* (denominadas VLAN-*aware bridges* pelo padrão IEEE 802.1Q-2014), só é possível a operação no Modo Ampliado de Filtragem.

As *bridges* IEEE 802.1D devem suportar o Modo Básico de Filtragem, e podem suportar o Modo Ampliado de Filtragem.

As *bridges* incapazes de operar no Modo Ampliado de Filtragem, do mesmo modo que aquelas operando no Modo Básico de Filtragem, não utilizam o MMRP.

3.9. DEFAULT GROUP FILTERING BEHAVIOR

Por *default*, existem três tipos de comportamento quanto à filtragem de quadros de dados com endereços de grupo que podem ser assumidos por uma porta de *bridge* que suporta o Modo Ampliado de Filtragem:

— Encaminhamento de Todos os Grupos (*Forward All Groups*);
— Encaminhamento de Grupos Não Registrados (*Forward Unregistered Groups*);
— Filtragem de Grupos Não Registrados (*Filter Unregistered Groups*).

No comportamento Encaminhamento de Todos os Grupos, as portas de *bridge* configuradas com esse tipo de comportamento encaminha todos os quadros de dados *multicast* sem considerar os respectivos endereços de grupo. Esse comportamento admite exceção quando existe uma entrada de filtragem estática que explicitamente especifique filtragem independente de qualquer informação de filtragem dinâmica.

No comportamento Encaminhamento de Grupos Não Registrados, todos os quadros de dados *multicast* cujos endereços de grupo não estejam registrados na porta da *bridge* são encaminhados, exceto quando há uma entrada de filtragem estática explícita que especifique filtragem independente de qualquer informação de filtragem dinâmica.

3.10. APRENDIZAGEM DE ENDEREÇOS MAC POR IEEE 802.1D BRIDGES

Uma vez construída a *spanning tree* de uma IEEE 802.1D *bridged LAN,* pode ter início o processo de aprendizagem de endereços MAC individuais (*endereços* MAC *unicast*) de origem. Esse processo, obedecida a *spanning tree*, possibilita a constituição de tabelas de roteamento MAC em todas as *bridges* envolvidas.

Para que a aprendizagem de endereços MAC de origem ocorra, é necessário que a porta de recepção esteja no *learning state* ou no *forwarding state.*

Existe a alternativa de montagem estática de tabelas de filtragem (encaminhamento) MAC, mas esse processo limita-se a redes de pequeno porte. A aprendizagem dinâmica de endereços MAC *unicast* é o método apropriado para redes de maior extensão. As tabelas de filtragem, estáticas os dinâmicas, são também referidas como tabelas CAM (*Content Addressable Memory*).

Quando da inicialização da rede *bridged Ethernet*, as *bridges* desconhecem totalmente a associação entre cada endereço MAC *unicast* de destino e a porta pela qual os respectivos quadros devem ser enviados.

Através do processo de aprendizagem dinâmica de endereços MAC, as *bridges* "aprendem" as associações entre endereços MAC *unicast* de origem e as portas de entrada, e registram-nas em tabelas *cache*, para fins de uso no processo de comutação, nas respectivas bases de dados de filtragem (*Filtering Databases*). As entradas referentes ao processo de aprendizagem de endereços MAC nas bases de dados de filtragens são referidas como entradas de filtragem dinâmicas (*Dynamic Filtering Entries*).

Como o tráfego é bidirecional com congruência entre transmissão e recepção, as associações entre portas e endereços MAC ocorrem através da observação de endereços MAC de origem nos quadros entrantes por essas portas. Uma vez preenchida uma entrada na base dados de filtragem, essa informação já pode ser usada para o encaminhamento de tráfego de dados.

Os endereços MAC que são inseridos nas tabelas da base dados de filtragem (tabelas de filtragem) de uma *bridge* podem ser de estações finais conectadas a essa *bridge*, ou de estações finais conectadas a qualquer sub-rede ligada nessa *bridge*, em *bridges* vizinhas ou em *bridges* remotas. O importante é que fique registrada a porta dessa *bridge* que será usada para atingir o endereço MAC de destino.

Observa-se que, de início ou após a eliminação dos respectivos registros na base de dados de filtragem, os endereços de origem são desconhecidos, e as tabelas de filtragem estão vazias. Até que ocorra a aprendizagem dos endereços MAC *unicast* de destino, os quadros são transmitidos em *broadcasting* pela rede, da mesma forma que ocorre com a transmissão de quadros com endereços *broadcast*. Esse processo é referido como *flooding* (inundação).

Como os endereços de origem utilizados na transmissão de quadros de dados com endereço de grupo ou com endereço *broadcast*, assim como os endereços de origem utilizado na transmissão de quadros de dados com endereço *unicast*, são

endereços *unicast*, todas essas transmissões propiciam a aprendizagem dos endereços MAC individuais de origem.

As entradas nas bases de dados de filtragem das *bridges* perduram por um certo período (o período *default* é de 5 minutos). As entradas não reatualizadas são eliminadas.

A reatualização das entradas nas bases de dados de filtragem ocorre com base no tráfego normal de dados, caso a estação cujo endereço esteja registrado envie tráfego antes da expiração do *time-out* da respectiva entrada em uma *bridge*.

3.11. TRANSMISSÃO DE QUADROS DE DADOS POR IEEE 802.1D BRIDGES

Será considerada, neste item, a transmissão de quadros de dados com endereços *broadcast*, quadros de dados com endereços *unicast* e quadros de dados com endereços *multicast* por IEEE 802.1D bridges, isto é, por bridges que operam em contextos onde não existem VLANs, e o protocolo *spanning tree* utilizado é o RSTP.

3.11.1. Quadros de Dados com Endereços Broadcast e com Endereço Unicast Não Registrados

Uma vez definida a *spanning tree* (a RST) em uma IEEE 802.1D *bridged* LAN, pode ter início, nessa rede, a transmissão de quadros de dados com endereços *broadcast* e quadros de dados com endereços *unicast* não registrados, por inundação. A inundação se realiza em conformidade com a *spanning tree* definida para a rede.

Essa transmissão por inundação possibilita o início do processo de aprendizagem de endereço MAC *unicast* de origem.

Como foi visto anteriormente neste livro, os quadros de dados com endereço *broadcast* utilizam o endereço MAC DA FF-FF-FF-FF-FF-FF (hexadecimal).

3.11.2. Quadros de Dados com Endereços de Destino Unicast Registrados

Como consequência da ocorrência da aprendizagem de um endereço MAC *unicast* de origem, esse endereço MAC é registrado na base de dados de filtragem da *bridge* e vinculado à porta pela qual o quadro entrou na *bridge.*

Existe a alternativa de montagem estática de tabelas de roteamento MAC, mas esse processo limita-se a redes de pequeno porte.

Os quadros passam a ser transmitidos para os endereços MAC registrados, com base nas entradas da base de dados de filtragem.

3.11.3. **Quadros de Dados com Endereços Multicast**

A transmissão de quadros de dados com endereço *multicast,* assim como a transmissão de quadros de dados com endereços *unicast* e *broadcast,* ocorre em conformidade com a *spanning tree* definida para a rede.

O comportamento de uma IEEE 802.1D *bridge* para o tráfego *multicast* depende do modo de filtragem utilizado pelas *bridges.*

3.11.3.1. Bridges Operando no Modo Básico de Filtragem

Nesse caso, as *bridges* são transparentes ao MMRP. Então, os quadros de dados com endereço *multicast* são transmitidos por inundação (como se fossem quadros *broadcast)*, pois as *bridges* tratam os quadros *multicast* como tendo endereços de destino desconhecidos.

3.11.3.2. Bridges Operando no Modo Ampliado de Filtragem

Nesse caso, uma estação MMRP-*aware* deve registrar, junto à porta de *bridge* a que se encontra diretamente conectada, a sua participação em um endereço de grupo para habilitar-se ao recebimento de quadros de dados com esse endereço de grupo.

A porta de *bridge* que recebe uma solicitação de registro em um endereço de grupo por parte de uma estação final, deve enviar uma declaração de realização de registo a essa estação final, estando então habilitada a transmitir quadros de dados com esse endereço de grupo.

Opcionalmente, essas estações finais podem também registrar, junto às portas de *bridges* às quais se encontram diretamente conectadas, a opção Encaminhamento de Todos os Grupos (*Forward All Groups*) ou a opção Encaminhamento de Grupos Não Registrados (*Forward Unregistered Groups*).

Como as estações finais MMRP-*aware* podem também receber solicitações de registro de participação em endereços de grupo oriundas das portas de *bridge* às quais elas se encontram diretamente conectadas, tais estações podem realizar o que se denomina *source pruning*.

No *source pruning*, uma estação final MMRP-*aware* envia quadros de dados *multicast* para a rede apenas se tiver recebido uma solicitação de registro de participação em um endereço de grupo enviada pela porta de *bridge* à qual ela se encontra diretamente conectada. É necessário também que essa estação final tenha acusado a realização do registro solicitado, enviando para isso uma declaração de registro para a porta de *bridge* que solicitou o registro.

3.12. VLAN BRIDGING (PADRÃO IEEE 802.1Q-2014)

A última versão do padrão IEEE 802.1D, de 2004, define a concepção de *LAN Bridging*, sem indroduzir, no entanto, formas de constituição de L2VPNs nas *bridges* LANs, ou seja, sem introduzir o conceito de VLANs.

Em 1998 foi emitida uma primeira versão do padrão IEEE 802.1Q-2014, que introduziu o conceito de VLANs. Essa versão foi atualizada em 2005, em 2008, em 2011 e finalmente em dezembro de 2014. A versão atual desse padrão é referida como IEEE 802.1Q-2014 (*Bridges e Bridged Networks*).

O padrão IEEE 802.1Q-2014 define a operação de *bridges* com a possibilidade de criação, operação e administração de VLANs dentro de *bridged* LANs.

Esse padrão suporta VLANs em diferentes tipos de *bridged* LANs (tais como aquelas apoiadas nos padrões IEEE 802.3 e 802.5).Será dada maior atenção, neste livro, a VLANs constituídas sobre redes *bridged Ethernet*, ou seja, às *Ethernet* VLANs.

Uma *Ethernet* VLAN é um tipo de L2VPN constituído sobre *bridged Ethernet* LANs, em que as estações que pertencem a um grupo logicamente identificado comunicam-se como em uma LAN única isolada do restante da rede.

A identificação lógica das VLANs ocorre por valores de VID, ou VLAN ID (VLAN *Identifer*), contidos no campo VLAN *Tag* transportado no interior dos quadros MAC *Ethernet* utilizados.

Com a aplicação do MSTP (*Multiple Spanning Tree Protocol*), que será abordada adiante neste capítulo, cada VLAN, ou cada grupo de VLANs, pode ser associado à sua própria *spanning tree* (MST). Uma MST representa a topologia ativa da correspondente VLAN ou do correspondente grupo de VLANs.

Processa-se, dentro de cada VLAN, o protocolo MVRP (*Multiple* VLAN *Registration Protocol*), cuja finalidade é registrar nas *bridges* a eventual presença dessa VLAN nas portas ativas da correspondente *spanning tree*, para evitar o envio desnecessários de dados.

As PDUs MVRP (MVRPDUs) transitam nos *links* pertinentes à topologia ativa da *multiple spanning tree* (MST) associada à VLAN a que pertencem. Fica assim definida a topologia VLAN correspondente a cada uma das diferentes VLANs, isto é, a topologia dessa VLAN definida de acordo com o protocolo MVRP.

A topologia VLAN de uma VLAN, definida pela aplicação do MSTP e do MVRP conforme parágrafos anteriores, delimita o domínio para a transmissão de quadros dentro dessa VLAN (quadros *unicast*, *multicast* ou *broadcast*) exclusivamente para onde essa transmissão é necessária.

A aprendizagem e a atualização de endereços MAC de origem, que vai direcionar o posterior envio de quadros *unicast* pelas *brigdes* de uma *bridged LAN*, ocorre dentro da topologia VLAN de cada uma das VLANs, ou seja, a topologia resultante da aplicação do protocolo MVRP. Para isso, os quadros de dados

transmitidos por inundação obedecem também topologia VLAN da respectiva VLAN.

O protocolo MMRP aplica-se também em VLAN *Bridging*, sendo que os quadros de dados com endereços *multicast* correspondentes a um grupo de endereços transitam na rede obedecendo a topologia VLAN da respectiva VLAN.

O funcionamento de *VLAN Bridging* é ilustrado na Figura 3-12.

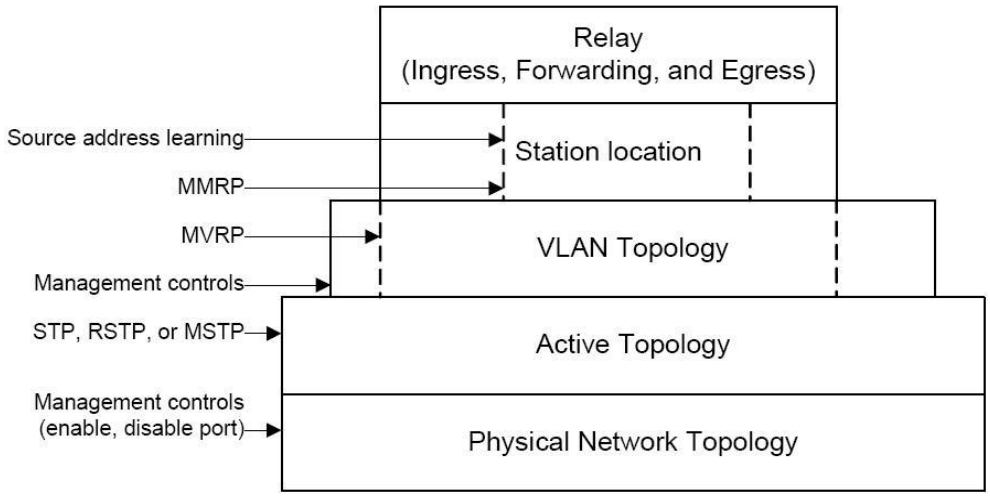


Figura 3-12– Funcionamento de VLAN Bridging.

3.12.1. **Formato de Quadros Contendo VLAN Tags**

Os quadros MAC contendo um VLAN *tag* apresentam o formato básico de um quadro MAC *Ethernet* IEEE 802.3, com a inserção de um VLAN *tag* entre o campo *Source* MAC *Address* e o campo *Length/Type* desses quadros.

A Figura 3-13 representa a formatação de um VLAN *tag* e a sua localização no cabeçalho do respectivo quadro MAC.

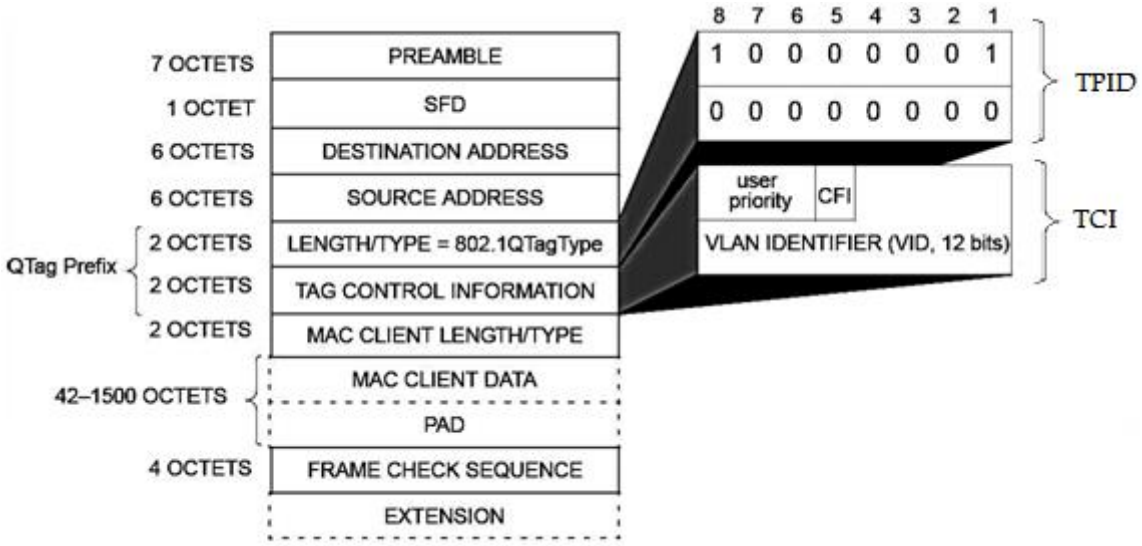


Figura 3-13– Formato e localização de VLAN tags.

Um VLAN *tag* é composto pelos seguintes elementos de informação:

— Um campo *Lenght/Type* contendo um *Tag Protocol Identifier* (campo TPID), com 2 octetos;
— Um campo contendo *Tag Control Informations* (campo TCI), com 2 octetos.

O campo TPID consiste em um valor utilizando a codificação *EtherType* para indicar o tipo de *tag* em utilização. No caso de existência de um único VLAN *tag*, característico de *bridged* LANs de usuários, o valor desse campo é igual a 0x8100. O VLAN *tag* identificado por esse valor, ou seja,o VLAN *tag* especificado no padrão IEEE 802.1Q-2014, é também referido como *Q-tag*.

Como será visto adiante neste livro, esse campo assume outros valores no caso de introdução de novas tecnologias de VLAN *Bridging* (antes definidas pelos padrões IEEE 802.lad, 802.1ah e 802.1Qay, e hoje incorporadas ao padrão IEEE 802.1Q-2014). Por exemplo, o PB (*Provider Bridges*) utiliza o valor 0x08A8, enquanto o PBB (*Provider Backbone Bridges*) utiliza o valor 0 x 88E7.

O campo TCI é dividido em três subcampos:
— Subcampo PCP (*Priority Code Point*), com 3 bits;
— Subcampo CFI (*Canonical Format Indicator*), com 1 bit;
— Subcampo VID (VLAN ID), com 12 bits.

O subcampo PCP, possibilita a indicação de até 8 níveis de prioridade, estabelecidos de acordo com o padrão IEEE 802.1p, o que será visto adiante neste capítulo.

O CFI indica se a informação contida na MSDU encontra-se no formato canônico. No caso de *Ethernet bridges*, o seu valor é sempre igual a zero.

O campo VID possibilita a atribuição de até 4094 valores de identificação de VLANs. Alguns desses valores têm significado especial, o que será visto adiante neste capítulo.

3.12.2. **Tipos de Ethernet Bridges**

As *Ethernet bridges* podem ser de três tipos:

— *VLAN-unaware bridges*;
— *Two-Port Media Access Control Relay* (TPMR) *bridges*;
— *VLAN-aware bridges*.

VLAN-*unaware bridges* são *bridges* que não processam o campo VLAN-*tag*, estando esse campo presente ou não nos quadros de dados que por elas transitam. Os quadros de dados são filtrados exclusivamente com base nos endereços MAC de destino contidos nos quadros, ainda que contenham VLAN *tags*.

TPMR *bridges* são VLAN-*unaware bridges* que possuem exatamente duas portas externamente acessáveis e que suportam um subconjunto da funcionalidade de uma MAC *bridge*, como especificado na Seção 5.14 do padrão IEEE 802.1Q-2014.

VLAN-*aware bridges* são *bridges* que reconhecem a presença de VLAN *tags* nos quadros de dados que por elas transitam, podendo inserir ou retirar *tag headers*. Os quadros de dados podem conter ou não VLAN *tags*. Quadros contendo VLAN *tags* são filtrados nesse tipo de *bridge,* com base nos endereços MAC de destino e nos valores de VID neles contidos.

As VLAN-*aware bridges* podem ser subdivididas nos seguintes tipos de *bridge*:

— *Shared VLAN Learning* (SVL) *bridges*;
— *Independente VLAN Learning* (IVL) *Bridges;*
— *SLV/IVL bridges*.

Tais tipos de VLAN-*aware bridges* serão apresentados adiante neste capítulo.
Uma implementação de VLAN-*aware bridge* pode:

— Suportar o MSTP;
— Suportar a classificação *Port-and-Protocol-based* de VLANs;
— Suportar *Extended Filtering Services* e o MMRP;
— Suportar o MVRP;
— Permitir que os bases de dados de filtragem possam conter registros de VLANs de modo estático ou dinâmico, até um máximo de 4094 VLANs.

A organização interna da subcamada MAC de uma VLAN-*aware bridge* encontra-se representada na Figura 3-.

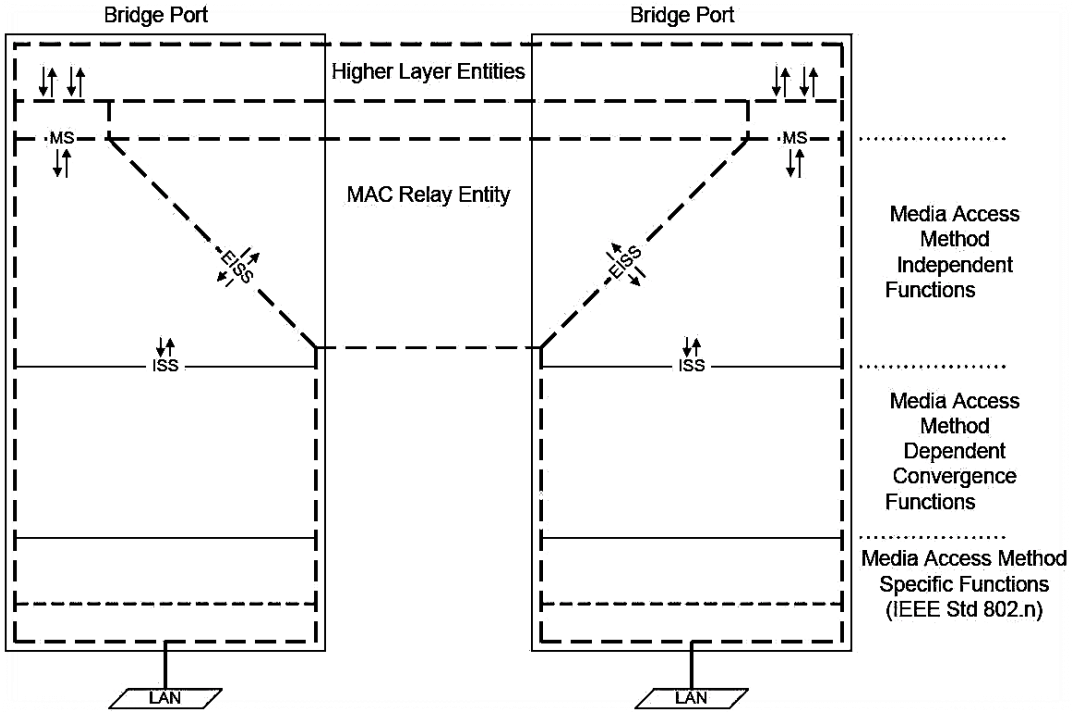


Figura 3-19– Organização interna da subcamada MAC de uma VLAN-*aware bridge*.

Observa-se nessa figura a utilização do *Enhanced Internal Sublayer Service* (EISS) em sobreposição ao *Internal Sublayer Service* (ISS). O EISS é derivado do ISS, com a ampliação de sua especificação com elementos necessários à operação de funções de *tagging* e de *untagging*.

As primitivas utilizadas pelo EISS são as seguintes:

— EM_UNITDATA_*indication*;
— EM_UNITDATA_*request*.

Os parâmetros dessas primitivas de serviço são os mesmos que os das primitivas de serviço do ISS, exceto pelo inclusão do parâmetro *vlan_identifier* em ambas primitivas de serviço.

Para o funcionamento do EISS, cada porta de uma *VLAN-aware bridge* deve suportar os seguintes parâmetros:

— Um parâmetro *Acceptable Frame Types*, com no mínimo um dos seguintes valores: *Admit Only VLAN-Tagged Frames*, *Admit Only Untagged and Priority-Tagged Frames* ou *Admit All Frames*;
— Um parâmetro PVID para a classificação *Port-based* de VLANs;
— Opcionalmente um parâmetro VID *Set* para a classificação *Port-and-Protocol-based* de VLANs.

3.12.3. **Tipos de Quadros**

No que diz respeito aos respectivos formatos, os quadros que transitam em VLAN-*aware bridges* podem ser de três tipos:

— Quadros *untagged*;
— Quadros priority-tagged;
— Quadros VLAN-*tagged*.

Quadros *untagged* não conduzem VLAN-*tags*, mas podem ser implicitamente associados a uma determinada VLAN.

Quadros *priority-tagged* conduzem VLAN-*tags* com informações de prioridade, mas com valor de VID nulo. Como os quadros *untagged*, os quadros *priority-tagged* podem ser implicitamente associados a uma VLAN.

Quadros VLAN-*tagged* conduzem VLAN-*tags* com valores de VID que efetivamente identificam as respectivas VLANs, contendo também informações de prioridade.

3.12.4. **Tipos de Estações Finais e de Portas de Bridges**

Estações finais (*end stations*), também referidas como *guest virtual machines* por fabricantes, podem ser de dois tipos:

— Estações finais VLAN-*unaware*: são estações finais que transmitem e recebem apenas quadros *untagged* e utilizam consequentemente os VIDs atribuídos pelas portas de *bridge* que classificam esses quadros;
— Estações finais VLAN-*aware: s*ão estações finais que suportam VLAN *tagging.*

Embora não processem normalmente VLAN *tags*, estações finais VLAN-*unware* podem pertencer a uma ou mais VLANs, ocorrendo a classificação *Port-based* ou *Port-and-Protocol-based* dessa VLAN ou dessas VLANs.

Quadros de dados são transmitidos por estações finais VLAN-*unware* normalmente *untagged* ou *priority-tagged*, embora tais quadros de dados possam receber VLANs *tags* com VIDs especiais (PVIDs ou *native* VIDs), quando de suas entradas em suas respectivas portas de *bridge.* Quadros de dados *s*ão recebidos *untagged* por esse tipo de estação.

Estações finais VLAN-*aware* podem ser estações finais MMRP-*aware* e/ou estações finais MVRP-*aware*.

Estações finais MMRP-*aware* são aquelas habilitadas a informar a sua participação em endereços MAC de grupos e de registrar a participação das respectivas portas de VLAN-*aware bridges* nesses endereços.

Estações finais MVRP-*aware* são aquelas habilitadas a informar a sua participação em VLANs e de registrar a participação das respectivas portas de VLAN-*aware bridges* nessas VLANs.

Os tipos de porta definidos são os seguintes:

— Portas de acesso;
— Portas tronco.

Portas de acesso são aquelas utilizadas para a conexão de estações finais VLAN-*unware*. De um modo geral, os quadros que ingressam por esse tipo de porta são os quadros *untagged* e os quadros *priority-tagged*.

Portas tronco são aquelas utilizadas para a conexão de estações finais VLAN-*aware.* De um modo geral, os quadros que transitam por esse tipo de porta são os quadros VLAN-*tagged*.

Observa-se que as portas tronco, além de se destinarem à conexão de estações finais VLAN-*aware* como acima especificado, são também utilizadas para a conexão entre *bridges* que constituem uma *bridged* LAN.

3.12.5. **VIDs Especiais**

As *bridged* LANs utilizam VIDs especiais, sendo parte dos valores desses VIDs atribuídos de modo reservado, como mostra a Figura 3- (atribuições conforme o padrão 802.1Q).

VID value (hexadecimal)	Meaning/Use
0	The null VID. Indicates that the tag header contains only priority information; no VID is present in the frame. This VID value shall not be configured as a PVID or a member of a VID Set, or configured in any Filtering Database entry, or used in any Management operation.
1	The default PVID value used for classifying frames on ingress through a Bridge Port. The PVID value of a Port can be changed by management.
2	The default SR_PVID value used for SRP (35.2.1.4(i)) Stream related traffic. The SR_PVID value of a Port can be changed by management.
FFF	Reserved for implementation use. This VID value shall not be configured as a PVID or a member of a VID Set, or transmitted in a tag header. This VID value may be used to indicate a wildcard match for the VID in management operations or Filtering Database entries.

Figura 3-20– Valores reservados de VID.

3.12.6. **Classificação de VLANs**

Um PVID (*Port* VLAN ID) é um valor de VID atribuído a uma porta de uma VLAN-*aware bridge* (porta tronco ou de acesso), de forma tal que quadros *untagged*, quadros *priority-tagged* ou quadros VLAN-*tagged* com o valor de VID igual ao valor do PVID, transmitidos por estações finais associadas a essa porta e por ela recebidos, sejam por ela transmitidos para o interior da rede como quadros VLAN-*tagged* com esse VID (ou seja, com o PVID).

Na *bridge* de destino os quadros assim recebidos são transmitidos *untagged* ou *tagged* para as estações finais.

Caso a uma porta de uma VLAN-*aware bridge* não tenha sido atribuído explicitamente um PVID, será implicitamente atribuído a essa porta o valor do *default* PVID da respectiva *bridge*.

Conforme a figura anterior, o padrão IEEE 802.1Q-2014 estabelece o valor 1 como *default* PVID. Fabricantes, no entanto, vêm utilizando outros valores de VID como PVID (valor 97, como exemplo).

Fabricantes utilizam a denominação *native VID* para o PVID. A VLAN indentificada na rede pelo *native VID* é referida como *native VLAN*. Observa-se que a denominação *native VID* não consta do padrão IEEE 802.1Q-2014, sendo introduzida por fabricantes.

Essa forma de classificação de VLANs com base em PVIDs (ou *native* VIDs) configurados em portas de *bridges*, é referida como classificação *Port-based*.

No caso de portas tronco onde são configurados valores de VID que constam de quadros de dados recebidos nessas portas e que serão utilizados quando da transmissão de quadros de dados para o interior da rede, o tipo de classificação utilizado é referido como classificação VLAN-*based*, classificação *Tag-based* ou classificação por VLAN-*tagging*.

O padrão IEEE 802.1Q-2014 define uma outra forma de classificação de VLANs, referida como classificação *Port-and-Protocol-based*.

Na classificação baseada em porta e protocolo (*Port-and-Protocol-based)*, de caráter opcional, a cada quadro *untagged* ou *priority-tagged* recebido em uma porta de acesso de uma VLAN-*aware bridge* é atribuído um valor de VID, que é função da porta de entrada (ou seja, é do tipo *port-based*) e também do grupo de protocolos de camada 2 utilizado no envelopamento do *payload* do quadro recebido (*Tagged, Ethernet, LLC_Other,*RFC_1042*, SNAP_Other)*.

Os valores de VID para isso utilizados são parte do VID *Set* implementado em uma *VLAN-aware bridge* para cada uma de suas portas.

3.12.7. **Endereços MAC Reservados**

A Seção 8.6.3 do padrão IEEE 802.1Q-2014 relaciona endereços reservados para quadros de controle para os diferentes componentes de *bridges.* Cada um dos endereços relacionados deve estar permanentemente configurado nas bases de dados de filtragem dos respectivos componentes.

Os endereços reservados não identificam os protocolos de controle, cuja identificação ocorre pelo campo *EtherType* subsequente aos campos de endereços MAC. Esses endereços reservados determinam o alcance da conectividade das PDUs dos protocolos que os utilizam.

O padrão IEEE 802.1Q-2014 reservou conjuntos de endereços MAC de grupo para componentes de VLAN-*aware bridges*, nas seguintes opções:

— Componente C-VLAN (*Customer VLAN*);
— Componentes S-VLAN (*Service VLAN*);
— Componentes TPMR (*Two-Port Media Access Control Relay*).

A Figura 3-14 apresenta os valores de endereço de grupo reservados para componentes C-VLAN.

Assignment	Value
Bridge Group Address, Nearest Customer Bridge group address	01-80-C2-00-00-00
IEEE MAC-specific Control Protocols group address	01-80-C2-00-00-01
IEEE 802.3 Slow_Protocols_Multicast address	01-80-C2-00-00-02
Nearest non-TPMR Bridge group address	01-80-C2-00-00-03
IEEE MAC-specific Control Protocols group address	01-80-C2-00-00-04
Reserved for future standardization	01-80-C2-00-00-05
Reserved for future standardization	01-80-C2-00-00-06
Metro Ethernet Forum ELMI protocol group address	01-80-C2-00-00-07
Provider Bridge Group Address	01-80-C2-00-00-08
Reserved for future standardization	01-80-C2-00-00-09
Reserved for future standardization	01-80-C2-00-00-0A
Reserved for future standardization	01-80-C2-00-00-0B
Reserved for future standardization	01-80-C2-00-00-0C
Provider Bridge MVRP Address	01-80-C2-00-00-0D
Individual LAN Scope group address , Nearest Bridge group address	01-80-C2-00-00-0E
Reserved for future standardization	01-80-C2-00-00-0F

Figura 3-21– Endereços MAC reservados para componentes C-VLAN.

A Figura 3-22 contém a relação de endereços MAC reservados para o contexto dos componentes S-VLAN das *bridges* de provedor, o que será visto como maiores detalhes no capítulo deste livro referente a redes Ethernet de provedores (*Provider Bridges*).

Assignment	Value
IEEE MAC-specific Control Protocols group address	01-80-C2-00-00-01
IEEE 802.3 Slow_Protocols_Multicast address	01-80-C2-00-00-02
Nearest non-TPMR Bridge group address	01-80-C2-00-00-03
IEEE MAC-specific Control Protocols group address	01-80-C2-00-00-04
Reserved for future standardization	01-80-C2-00-00-05
Reserved for future standardization	01-80-C2-00-00-06
Metro Ethernet Forum ELMI protocol group address	01-80-C2-00-00-07
Provider Bridge Group Address	01-80-C2-00-00-08
Reserved for future standardization	01-80-C2-00-00-09
Reserved for future standardization	01-80-C2-00-00-0A
Individual LAN Scope group address , Nearest Bridge group address	01-80-C2-00-00-0E

Figura 3-22 - Endereços MAC reservados para componentes S-VLAN.

A Figura 3-23 relaciona os endereços reservados para o componente TPMR.

Assignment	Value
IEEE MAC-specific Control Protocols group address	01-80-C2-00-00-01
IEEE 802.3 Slow_Protocols_Multicast address	01-80-C2-00-00-02
IEEE MAC-specific Control Protocols group address	01-80-C2-00-00-04
Individual LAN Scope group address , Nearest Bridge group address	01-80-C2-00-00-0E

Figura 3-23 - Endereços MAC reservados para componentes TPMR.

3.12.8. **Valores EtherType Atribuídos pelo IEEE**

A Figura 3-24 apresenta alguns exemplos de valores *EtherType* atribuídos pelo IEEE.

CÓDIGOS	PADRÃO IEEE	DESTINAÇÕES
0x8100	802 .1Q	Customer VLAN Tag Type (C-Tag)
0x8808	802 . 3	Ethernet Passive Optical Network (EPON)
0x888E	802 . 1X	Port-based network access control
0x88A8	802 . 1Q	Service VLAN tag identifier (S-Tag)
0x88B5	802	Local Experimental Ethertype
0x88B6	802	Local Experimental Ethertype
0x88B7	802	OUI Extended Ethertype
0x88C7	802 . 11i	Pré-Authentication
0x88CC	802 . 1AB	Link Layer Discovery Protocol (LLDP)
0x88E5	802 . 1AE	Media Access Control Security
0x88F5	802 . 1ak	Multiple VLAN Registration Protocol (MVRP)
0x88F6	802 . 1Q	Multiple Multicast Registration Protocol (MMRP)
0x890D	802 . 11r	Fast Roaming Remote Request

Figura 3-24 - Exemplos de códigos *EtherType* especificados pelo IEEE 802.

PDUs com endereços de grupo na primeira dessas tabelas (relativa a componentes C-VLAN) são bloqueadas por componentes C-VLAN (inclusive C-VLAN *bridges)* e por *bridges* IEEE 802.1D.

Os endereços de grupo da tabela seguinte, referente a componentes S-VLAN, tem o mesmo tratamento relativamente a esses componentes, ou seja, PDUs com esses endereços de grupo são bloqueadas por componentes S-VLAN.

O mesmo ocorre com os endereços de grupo da terceira tabela, com realção a componentes TPMR, sendo que as PDUs com os endereços de grupo dessa tabela são bloqueadas por esse tipo de componente.

Em consequência, um endereço de grupo que é bloqueado por qualquer tipo de componente ou de *bridge* deve constar dessas três tabelas. Um exemplo

importante desse tipo de endereço de grupo é o *Individual LAN Scope group address*, também referido como *Nearest Bridge group address* (01-80-C2-00-00-0E).

Outros endereços de grupo têm as suas PDUs bloqueadas por todos os componentes e *bridges* à exceção de componentes TPMR. Assim, tais endereços de grupo constam da primeira e da segunda tabelas, mas não constam da terceira tabela. Um importante exemplo é o endereço de grupo *Nearest non-TPMR Bridge group address* (01-80-C2-00-00-03).

Finalmente, existem endereços de grupo cujas PDUs são bloqueadas exclusivamente por componentes C-VLAN (inclusive *bridges*) e por *bridges* IEEE 802.1D, ou seja, por equipamentos de usuário. Tais endereços de grupo constam da primeira tabela, mas não constam da segunda e da terceira tabelas. Um importante exemplo é o endereço de grupo *Nearest Customer Bridge group address* (01-80-C2-00-00-00).

3.12.9. **Prioridade e Precedência para Descarte**

Os parâmetros prioridade e elegibilidade para descarte são codificados no campo *Priority Code Point* (PCP) do VLAN *tag*.

A Figura 3-25 ilustra como deve ser interpretado e decodificado o campo PCP segundo a alternativa de codificação escolhida, com as classes de tráfego associadas.

PCP		7	6	5	4	3	2	0	1
priority drop eligible	8P0D	7	6	5	4	3	2	0	1
	7P1D	7	6	4	4DE	3	2	0	1
	6P2D	7	6	4	4DE	2	2DE	0	1
	5P3D	7	6	4	4DE	2	2DE	0DE	0
number of traffic classes	1	BE							
	2	VO				BE			
	3	NC		VO		BE			
	4	NC		VO		CA		BE	
	5	NC	IC	VO		CA		BE	
	6	NC	IC	VO		CA		BE	BK
	7	NC	IC	VO		CA	EE	BE	BK
	8	NC	IC	VO	VI	CA	EE	BE	BK

Figura 3-25 – Decodificação do campo PCP.

A prioridade aparece como um número na tabela, e as letras "DE" após esse número indicam que o parâmetro *drop_eligible* possui o valor *True*.

Foram definidas quatro alternativas de codificação, designadas por 8P0D (*default*), 7P1D, 6P2D e 5P3D, onde P significa prioridade e D significa eligibilidade para descarte.

Por exemplo, na alternativa 8P0D (*default*) são suportadas 8 classes de prioridade, designadas como, em ordem de prioridade menor a maior, de 0 a 8, correspondentes, respectivamente, às classes de tráfego BE (*best effort*), BK (*background*), EE (*excellent effort*), CA (*critical applications*), VI (*video*), VO (*voice*), IC (*internetwork control*) e NC (*network control*). Entretanto, não há indicação de elegibilidade de descarte.

Já na alternativa 5P3D, como outro exemplo, as classes de prioridade ficam reduzidas a 5, sendo que os quadros pertencentes a três delas podem ser identificados como elegíveis para descarte. As cinco classes de prioridade remanescentes são as classes 0 (BE), 2 (CA), 4 (VO), 6 (IC) e 7 (NC).

Os quadros das classes 0, 2 e 4 na alternativa 5P3D podem ser indicados como sendo, adicionalmente, elegíveis ou não para descarte. Assim, por exemplo, o valor de PCP igual a 5 representa, na realidade, a classe 4 (VO) não elegível para descarte. O valor de PCP igual a 4 representa também a classe 4 (VO), porém com elegibilidade para descarte.

3.13. MULTIPLE SPANNING TREE PROTOCOL (MSTP)

O MSTP foi definido inicialmente pelo padrão IEEE 802.1s, que foi obsoletado e incorporado posteriormente ao padrão IEEE 802.1Q-2014. Esse protocolo é uma nova versão da família de protocolos STP, com características operacionais baseadas no RSTP, mas que possibilita a constituição de múltiplas *spanning trees* independentes em uma *bridged* LAN.

As *spannings trees* independentes destinam-se ao atendimento de uma VLAN ou de subconjuntos de VLANs.

O MSTP possibilita, com base na constituição de múltiplas STs (MSTP *instances*), balanceamento da carga de tráfego nos *links* da rede. Além disso, como uma falha na rede não atinge normalmente todas as MSTs constituídas, o MSTP permite que parte das VLANs não seja interropida por essa falha.

A associação entre VLANs e *spanning trees* realiza-se administrativamente, selecionando-se a ST mais apropriada para cada VLAN ou para determinados subconjuntos de VLANs.

Ocorre no MSTP a hierarquização das *spanning trees* por meio da definição de *Multiple Spanning Tree Regions* (Regiões MST). Uma Região MST é administrativamente composta por um número de MST *bridges* e de LANs (ponto a ponto ou multiponto) conectadas a essas MST *bridges*.

Uma Região MST possui funcionalidades para a comunicação com as demais Regiões MST e com as *Single Spanning Trees* (SSTs) da *bridged* LAN. As SST *bridges* são também referidas como *non*-MST *bridges*.

Cada Região MST recebe um identificador próprio, referido como MST *Configuration Identifier.* O número de Regiões MST em redes *bridged Ethernet* é teoricamente ilimitado.

Os *links* entre regiões e as portas de *bridge* onde esses links são conectados, são referidos como *links* e portas de fronteira (*boundary links* e *boundary ports*), e as *bridges* que possuem essas portas são referidas como *bridges* de fronteira (*boundary bridges*).

O MSTP incorpora a existência de uma Região MST global, cuja *spanning tree* é de uso comum a todas as MST *bridges* e a todas as SST *bridges* existentes na *bridged* LAN. Essa *spanning tree* é denominada CST (*Common Spanning Tree*). Observa-se que a CST é do tipo RSTP.

Adicionalmente, o MSTP constitui, em cada Região MST, um conjunto de MSTIs (*Multiple Spanning Tree Instances*). Uma MSTI é identificada por um MSTI *Identifier* (MSTID). Existe a possibilidade de configuração de até 16 MSTIDs em uma Região MST.

Um determinado MSTID (normalmente o MSTID igual a zero) é utilizado para a constituição, em cada Região MST, de uma *spanning tree* especial, referida como IST (*Internal Spanning Tree*), que interconecta todas as MST *bridges* da Região MST.

A função de uma IST é a de conduzir, nos campos que lhe cabem nas BPDUs, informações comuns a todas as MSTIs da respectiva Região MST, particularmente o identificador da Região MST (ou seja, o valor do campo *MST Configuration Identifier*).

As ISTs das Regiões MST da rede integram-se à CST, como sub-árvores, formando uma *spanning tree* englobando a totalidade das MST *bridges* e *non*-MST *bridges* da rede *bridged Ethernet*. Essa ST global é referida como CIST (*Common and Internal Spanning Tree).* A CIST (ou seja, a CST mais as ISTs) suporta indiscriminadamente todas as VLANs da *bridged* LAN.

Os demais MSTIDs (normalmente com valores de 1 a 15) são utilizados para a constituição de MSTIs, independentemente para cada Região MST.

A cada MSTI de uma Região MSTP são atribuídos sub-conjuntos de VLANs ou mesmo VLANs individuais, identificados pelos respectivos VIDs. Em qualquer hipótese, uma VLAN só pode ser atribuída a uma única MSTI em cada Região MST.

As informações necessárias para a constituição das diferentes MSTIs de uma Região MST são conduzidas nas mesmas MSTPBPDUs. Compete à IST de uma Região MST transportar essas MSTPBPDUs contendo as informações das MSTIs dessa Região MST sob a forma de *piggybacking* nas BPDUs da IST. Resulta uma elevada eficiência no processo de constituição das MSTIs em cada Região MST.

Registra-se a possibilidade de configurações MSTP abrangendo uma única Região MST (*single-region* MSTP), além das configurações de MSTP envolvendo múltiplas regiões MST (*multiple-region MSTP*).

3.13.1. **Constituição da CIST**

A CIST é constituída pelo MSTP utilizando para isso uma versão *Force Protocol* do RSTP, baseada em um parâmetro próprio, controlado pela supervisão da rede. Tal parâmetro, denominado *Force Protocol Version parameter*, possibilita que o MSTP seja instruído para emular aspectos adicionais do RSTP. Nessa condição, as portas de uma MST *bridge* passam a operar como portas de uma RST *bridge*.

No MSTP, o campo *Protocol Version Identifier* das BPDUs de configuração assume o valor 3. Para a versão *Force Protocol* emular o STP, esse campo assume o valor zero, enquanto que na emulação do RSTP esse valor passa a ser igual a 2.

A constituição da CIST engloba, em um mesmo processo, a constituição da CST e das ISTs das Regiões MST. Todo esse processo tem como suporte uma versão *Force Protocol* do MSTP.

A CST é composta por uma CIST *Root Bridge*, e, em cada Região MST, pelas CIST *Regional Root Bridges* (também denominadas IST *Master Bridges* por fabricantes). As SST (*non*-MST) *bridges* são como Regiões MST em si mesmas, participando do processo de constituição da CST.

Em uma *bridged LAN* com múltiplas Regiões MST, em cada Região MST a CIST *Regional Root Bridge* é também a *root bridge* da respectiva IST. Caso a *bridged LAN* contenha uma única Região MST, a IST é constituída independentemente, de acordo com os processos normais do RSTP.

Na Região MST onde se localiza a CIST *Root Bridge*, essa *bridge* é também a CIST *Regional Root Bridge* dessa Região MST (e também a *root bridge* da IST dessa Região MST).

Quando o MSTP inicializa, todas as *bridges* da *bridged* LAN propagam BPDUs atribuindo-se a função de CIST *Root Bridge*, sendo que as MSTP *bridges* atribuem-se também a função de CIST *Regional Root Bridge* das respectivas Regiões MST.

A propagação de BPDUs em portas internas contém todas as informações das BPDUs. Nas portas de fronteira, contudo, as *bridges* de fronteira propagam exclusivamente o *CIST Root Bridge ID* e o *CIST External Root Path Cost*, omitindo, portanto, a topologia interna das respectivas regiões MST.

Fica assim definida a CIST *Root Bridge*. A *root bridge* é escolhida arbitrariamente, com base nos valores de identificação das *bridges*, nos moldes utilizados para os protocolos STP.Como visto para o RSTP, é escolhida a *bridge* com o menor valor de identificação, não importando a Região MST em que se encontre essa *bridge*. A CIST *Root Bridge* pode ser inclusive uma *non-MST Bridge.*

Para cada Região MST, à exceção daquela onde se localiza a CIST *Root Bridge*, é escolhida uma *bridge* de fronteira para ser a CIST *Regional Root Bridge* dessa Região MST. A *bridge* escolhida é aquela que apresenta o menor *CIST External Root Path Cost*.

CIST External Root Path Cost é o custo para que uma *bridge* de fronteira atinja a *CIST Root Bridge* através de *links* conectando portas de fronteira, ou seja, de *links*

entre regiões MST. Caso a CIST *Root Bridge* não seja uma *bridge* de fronteira, os custos internos para alcançá-la não são computados no cálculo no *CIST External Root Path Cost.*

A porta de fronteira de uma CIST *Regional Root Bridge* que apresenta o menor *CIST External Root Path Cost* para atingir a CIST *Root Bridge,* constitui-se na CIST *Root Port* dessa CIST *Regional Root Bridge,* sendo as demais portas de fronteira dessa CIST *Regional Root Bridge* bloqueadas. A essa porta, por onde escoa todo o tráfego de outras Regiões MST ou para outras Regiões MST, atribui-se um novo papel (*port role*), denominado *Master Port* (os demais *port roles* do MSTP são aqueles do RSTP).

As *bridges* de fronteira que não se tornam uma CIST *Regional Root Bridge* bloqueiam todas as suas portas de fronteira.

Uma SST *Bridge* existente na *bridged* LAN, que é considerada como uma Região MST em si mesma, desempenha ela própria a função de CIST *Regional Root Bridge* dessa "Região".

A CST, constituída conforme os parágrafo anteriores, é única na *bridged* LAN, sendo que por ela transita todo o tráfego MST e *non*-MST entre as Regiões MST.

Para a constituição da IST em uma *bridged LAN* com apenas uma Região MST, aplica-se a versão *Force Protocol* do MSTP como em uma RSTP. A CIST é formada apenas por uma IST. A IST *Root Bridge* é definida entre todas as *bridges* da rede (ou seja, todas as *bridges* da região), sendo aquela com o menor valor de identificação.

Caso a *Bridged* LAN seja dividida em múltiplas regiões MST, a IST *Root Bridge* de cada IST se confunde com a *CIST Regional Root Bridge* na mesma Região MST. Cada uma dessas ISTs é constituída pela versão *Force Protocol* do RSTP, de forma integrada com a constituição da CST, a partir da respectiva *IST Root Bridge* (que, como dito anteriormente, se confunde com a *CIST Regional Root Bridge* da respectiva Região MST).

Na configuração de uma IST, os custos para atingir a IST *Root Bridge* são referidos como *CIST Internal Root Path Costs*.

Uma vez constituídas as CST e as ISTs das Regiões MST, fica constituída a CIST da *bridged LAN.*

3.13.2. **Constituição de uma MSTI**

As MSTIs podem ser constituídas de forma integrada com as CIST, ou podem ser constituídas independentemente.

Uma MSTI em uma Região CST é identificada pelo correspondente MSTID (campo *MSTI Regional Root Identifier*, contido nas *MSTI Configuration Message*s) e pela identificação da Região MST (campo *MST Configuration Identifier*). É necessário que todas as *bridges* da Região MST concordem com a correspondência entre valores de VID e de MSTID das respectivas MSTIs.

Com a aplicação do protocolo MSTP restrito a uma Região MST e a um valor de MSTID, resulta a montagem de uma MSTI, válida para uma VLAN ou para um subconjunto de VLANs a ela associados.

Os identificadores de *bridges* e de portas, incluindo os valores de prioridade, são alterados de forma aleatória para cada MSTI de um Região MST, de forma a possibilitar a montagem de MSTIs utilizando *bridges* e *links* dessa Região MST de modo independente. Com esse critério obtém-se melhor utilização dos recursos da rede, que passam a ser distribuídos entre as diferentes MSTIs, além de possibilitar que parte das MSTIs possa não ser afetada por uma falha na rede.

Dentro da limitação imposta pela identificação da Região MST e de um valor de MSTID, o algoritmo MSTP define uma *bridge* como sendo a MSTI *Regional Root Bridge* de uma MSTI na Região MST. Essa *bridge* é aquela de menor valor de identificação dentro da MSTI em questão (diferentes MSTIs possuem diferentes MSTI *Regional Root Bridges*), indicada de forma independente da escolha da CIST *Regional Root Bridge* da mesma Região MST.

Mantida a mesma limitação, realiza-se a definição das MSTI *Root Ports* de cada MSTI nas diferentes *bridges* que constituem a Região MST. Tais portas são aquelas que apresentam o menores MSTI *Internal Root Path Costs* para alcançar a MSTI *Regional Root Bridge* da MSTI na Região MST.

São também definidas, ainda dentro dessa limitação, as *Designated Bridges* e as *Designated Ports* da MSTI para cada LAN conectada às *bridges* da respectiva Região MST.

Dessa forma, fica constituída uma MSTI em uma Região MST, contida totalmente nessa Região MST, de forma totalmente independente das demais MSTIs também constituídas nessa Região MST.

3.13.3. **Exemplo de Configuração**

A Figura 3-26 representa um exemplo de configuração de uma *bridged* LAN operando com MSTP, que será utilizada como referência para melhor entendimento dos conceitos acima apresentados.

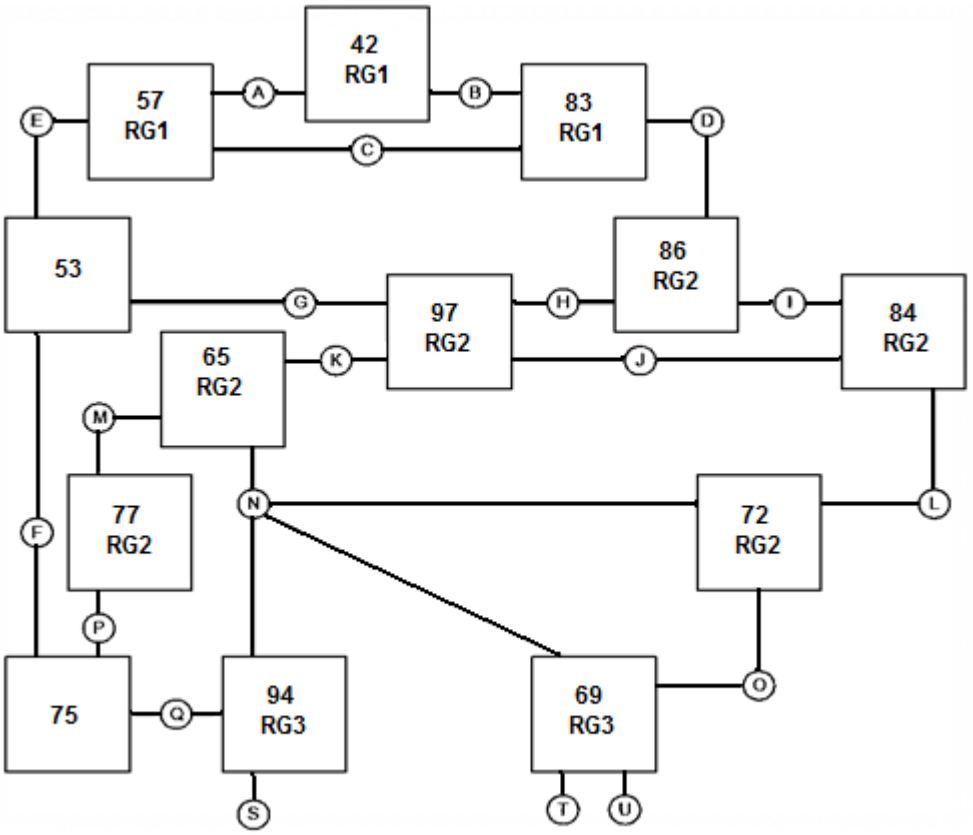


Figura 3-26– Exemplo de configuração de uma bridged LAN.

Nessa figura, os quadrados representam *bridges*, enquanto os círculos representam LANs.

A CIST representada é composta por uma CST e pelas ISTs das diferentes regiões existentes.

3.13.3.1. Configuração da CIST (CST e ISTs)

Considerando-se a aplicação do MSTP na *bridged* LAN dessa figura resulta a CST apresentada na Figura 3-27.

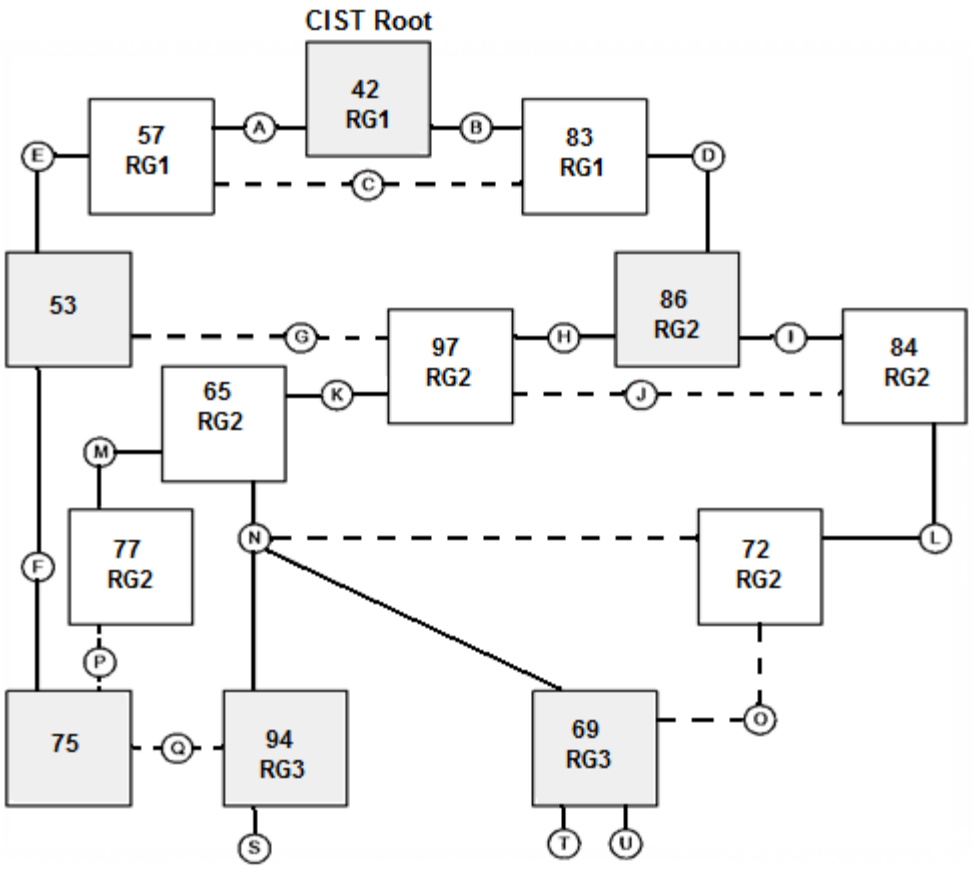


Figura 3-27– CST resultante da figura anterior.

Foram supostas para a constituição dessa CST os seguintes pontos:

— A *bridge* 42 foi definida como a *CIST Root Bridge* por apresentar o menor valor de identificador de *bridge* (essa *bridge* é também a *CIST Regional Bridge* da RG1, ou seja, da Região MST 1);
— As *bridges* 57 e 83 estão na RG1, onde também se encontra a CIST *Root Bridge*;
— As LANs A, B, C e D encontram-se na RG1 porque as respectivas *CIST Designated Bridges* estão nessa Região MST, e elas não se conectam a nenhuma *non*-MST *Bridge*;
— As LANs E, F, P e Q não pertencem a qualquer Região MST, por estarem conectadas a *non*-MST *Bridges*;
— As *bridges* 65, 72, 77, 84, 86 e 97 encontram-se na RG2, por terem o mesmo *MST Configuration Identifier* e por serem todas interconectadas por LANs para as quais uma dessas *bridges* é a *CIST Designated Bridge*;
— A *bridge* 86 foi definida como a *CIST Regional Root Brigde* para a RG2, por apresentar o menor *CIST External Root Path Cost* para alcançar a *CIST Root Bridge*;
— As LANs H, I, J, K, L, M, N e O encontram-se na RG2;

— A *bridge* 69 (e as LANs T e U) e a *bridge* 94 (e a LAN S) encontram-se em diferentes Regiões MST, mesmo tendo o mesmo MST *Configuration Identifier*, porque a LAN que as conecta (LAN N) encontra-se em uma diferente Região MST. A LAN N encontra-se na RG2 porque a sua *CIST Designated Bridge*, que é a *bridge* 65, encontra-se na RG2.

Assim sendo, a CST é composta pela *bridge* 42 (*CIST Root Bridge*), pelas MST *Bridges* 86 (*CIST Regional Root Bridge* da RG2), 69 e 94, e pelas *non*-MST *Bridges* 53 e 75. Essas *bridges* encontram-se escurecidas na figura anterior.

As ISTs da RG1 e da RG2 foram configuradas juntamente com a CST. Assim, por hipóteses, foram configuradas:

— A IST da RG1, cuja *IST Root Bridge* é a *bridge* 42 (que é também a *CIST Root Bridge* da CST), sendo a LAN C bloqueada;
— A IST da RG2, cuja *IST Root Bridge* é a bridge 86 (que é também a *CIST Regional Root Bridge* da RG2), sendo bloqueadas as LANs G, J, N, O, P e Q.

Cada IST é composta por todas as *bridges* da respectiva Região MST (onde uma dessas *bridges* é a *IST Root Bridge*) e por todas as LANs não bloqueadas nessa IST.

3.13.3.2. Configuração de uma MSTI na RG2

As MSTIs nas diferentes regiões MST podem ser configuradas juntamente com a CIST ou configuradas posteriormente à configuração da CIST. Em qualquer hipótese, a configuração de uma MSTI ocorre com base no campo *MSTI Configuration Messages* das BPDUs, considerando-se também o campo *MST Configuration ID*.

Será considerada, como exemplo, a configuração de uma MSTI na RG2, que atende a uma VLAN ou a um subconjunto de VLANs, representada na Figura 3-28.

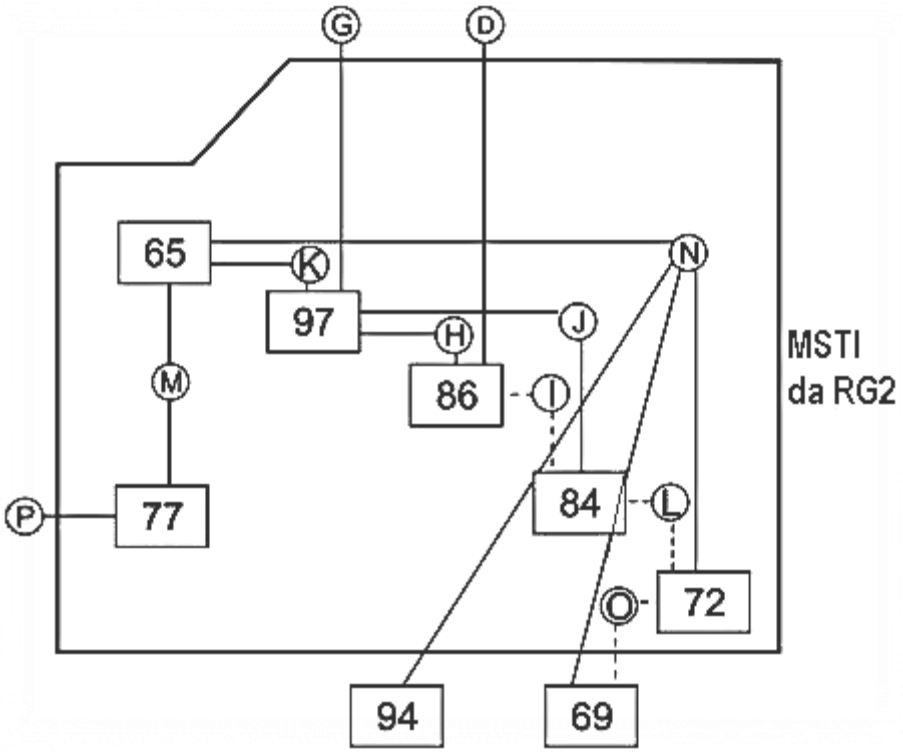


Figura 3-28– Configuração de uma MSTI da RG2.

Nessa figura, são considerados os seguintes pontos:

— A *bridge* 65 foi definida como a MSTI *Regional Root Bridge* da RG2, por possuir o menor valor de *Bridge Identifier* na Região MST;
— A conectividade entre a totalidade da RG2 (ou de qualquer outra Região MST) com as *bridges* e com as LANs da *bridged* LAN exteriores a essa Região MST é similar àquela que resultaria da substituição da Região MST inteira por uma única SST *Bridge* (*non-MST bridge*), posto que toda a RG2 (assim como as demais Regiões MST) possui uma única CIST *Root Port* (que é a sua *MasterPort*).

3.13.4. **Formato de MSTP BPDUs**

As MSTP BPDUs são utilizadas para transmitir três tipos de mensagens *spanning tree:*

— Mensagens de Configuração;
— Mensagens de Notificação de Mudanças de Topologia (TCNs);
— Identificadores de Configuração MST.

Conforme menção anterior neste capítulo, o STP básico utiliza formatos próprios para as Mensagens de Configuração e as Mensagens TCN. O MSTP, assim como o RSTP, ao contrário, utiliza um único formato de BPDU para essas mensagens. Tais BPDUs de formato único são referidas como MST BPDUs e RST BPDUs no MSTP e no RSTP, respectivamente.

Uma MST BPDU codifica também os Identificadores de Configuração MST, como será visto a seguir.

O formato de uma MST BPDU encontra-se na Figura 3-15.

	Octet
Protocol Identifier	1–2
Protocol Version Identifier	3
BPDU Type	4
CIST Flags	5
CIST Root Identifier	6–13
CIST External Path Cost	14–17
CIST Regional Root Identifier	18–25
CIST Port Identifier	26–27
Message Age	28–29
Max Age	30–31
Hello Time	32–33
Forward Delay	34–35
Version 1 Length = 0	36
Version 3 Length	37–38
MST Configuration Identifier	39–89
CIST Internal Root Path Cost	90–93
CIST Bridge Identifier	94–101
CIST Remaining Hops	102
MSTI Configuration Messages (may be absent)	103–*39* + *Version 3 Length*

Figura 3-15– Formato de uma MST BPDU.

Como se observa, esse formato é compatível com aquele especificado para as RST BPDUs, com adição de campos objetivando o transporte de informações relativas à constituição da CST e de cada MSTI das Regiões MST.

As MST BPDUs são transportadas como LLC *Type* 1 SDUs, contidas em quadros IEEE 802 MAC.

3.13.4.1. Campos Protocol e Protocol Version Identifier

O campo *Protocol Identifier* apresenta o valor 0X0000, significando tratar-se da família de protocolos STP.

O campo *Protocol Version Identifier*, que objetiva identificar o membro dessa família de protocolos, assume os seguintes valores:

— Para o STP: valor 0x00;

— Para o RSTP: valor 0x02;
— Para o MSTP: valor 0x03.

3.13.4.2. Campo BPDU Type

Esse campo assume os seguintes valores:

— Para STP *Configuration* BPDUs: valor 0x00;
— Para STP TCN BPDUs: valor 0x80;
— Para RST BPDUs: valor 0x02;
— Para MST BPDUS: valor 0x02.

3.13.4.3. Campo CIST Flags

Os sub-campos desse campo têm os seguintes significados:

— Bit 1 (MSD bit): *CIST Topology Change Flag*;
— Bit 2: *CIST Proposal Flag*;
— Bits 3 e 4: *CIST Port Role*;
— Bit 5: *CIST Learning Flag*;
— Bit 6: *CIST Forwarding Flag*;
— Bit 7: *CIST Agreement Flag*;
— Bit 8: *Topology Change Acknowledge Flag*.

3.13.4.4. Campos Seguintes até o Campo Message Age

Os campos *CIST Root Identifier*, *CIST External Root Path Cost*, *CIST Regional Root ID* e *CIST Port ID* objetivam a definição da CST (*CIST Root Bridge* e *CIST Regional Root Bridges*).

O campo *CIST External Path Cost* representa o custo para atingir a *CIST Root Bridge* através dos *links* conectando as portas de fronteira, isto é, dos *links* entre Regiões MST.

3.13.4.5. Campos Message Age até Forward Delay

Esses campos têm o mesmo significado que no RSTP.

3.13.4.6. Campos Version 1 Length e Version 3 Length

O campo *Version 1 Length* apresenta o valor 0x00 por não existir essa versão no MSTP.

O campo *Version 3 Length* representa o número de octetos correspondentes aos parâmetros que se seguem no quadro.

3.13.4.7. Campos MST Configuration Identifier até Campo CIST Remaining Hops

Esses campos representam as informações contidas nas BPDUs referentes a cada uma das ISTs.

- **Campo MST Configuration Identifier**

Esse campo identifica a Região MST, e apresenta os seguintes significados para os elementos que o compõem:

— Octeto 39: *Configuration Identifier Format Selector*, que assume o valor 0x00;
— Octetos 40-71: *Configuration Name*;
— Octetos 72-73: *Revision Level*;
— Octetos 74-89: *Configuration Digest*.

- **Campos CIST Internal Root Path Cost e CIST Bridge ID**

Esses campos aplicam-se na constituição da IST de cada Região MST. Em *bridged* LANs com múltiplas Regiões MST, a *CIST Regional Root Bridge* de uma Região MST é também a IST Root Bridge da IST dessa Região MST.

O campo *CIST Internal Root Path Cost* informa o *path cost* acumulado para alcançar a *IST Root Bridge* da Região MST a partir da *bridge* transmissora da BPDU, através de cada porta dessa *bridge* transmissora.

O campo *CIST Bridge Identifier* conduz a identificação completa (endereço MAC e prioridade)da *bridge* transmissora.

- **Campo CIST Remaining Hops**

A *IST Root Bridge* de uma Região MST envia BPDUs com esse campo contendo o valor de *hop count* igual ao valor *MaxHops* configurado (*default* igual 20). Cada *bridge downstream* decrementa esse valor de *hop count* quando da recepção de uma BPDU que o contém, e a retransmite.

Assim que o *hop count* atinge o valor zero, as *bridges* receptoras passam a ignorar as informações contidas na BPDU recebida, e a *bridge* começa a se declarar como sendo a nova *IST Root Bridge*.

O parâmetro *MaxHops* desempenha, na IST e nas MSTIs de uma Região MST, o mesmo papel que aquele desempenhado pelos parâmetros *MaxAge* e *Message Age* no RSTP.

3.13.4.8. Campo MSTI Configuration Messages

O campo *MSTI Configuration Messages* traz uma sequência de até 64 mensagens *MSTI Configuration*, cada uma referente a uma MSTI. Cada uma dessas mensagens possui o formato representado na Figura 3-.

MSTI Flags	1
MSTI Regional Root Identifier	2–9
MSTI Internal Root Path Cost	10–13
MSTI Bridge Priority	14
MSTI Port Priority	15
MSTI Remaining Hops	16

Figura 3-30– Formato de uma MSTI Configuration Message.

O campo *MSTI Flags* utiliza, para cada MSTI sinalizada na BPDU, a mesma semântica que o campo *CIST Flags* apresentado no subitem 3.13.4.3 anterior.

O campo *MSTI Regional Root Identifier*, além de identificar a MSTI *Regional Root Bridge* de cada MSTI sinalizada na BPDU, conduz o MSTID dessa MSTI.

O campo *MSTI Internal Root Path Cost* foi já definido anteriormente neste capítulo com relação a ISTs, dizendo respeito agora a custos para atingir a *MSTI Regional Root Bridge*.

Os campos *MSTI Bridge Priority* e *MSTI Port Priority*, relativos à *bridge* transmissora da BPDU, dizem respeito a cada MSTI sinalizada na BPDU.

O campo *MSTI Remaining Hops* desempenha, para cada MSTI, o mesmo papel desepenhado pelo campo *CIST Remaining Hops* na constituição de uma IST.

Cada conjunto de valores contidos no campo *MSTI Configuration Messages* é referido, por fabricantes, como sendo um *MRecord*. Em todas as Regiões MST, cada *bridge* envia *MRecords* referentes a cada uma das MSTIs até a total convergência dessa MSTI.

Uma exceção à regra são as ISTs, que não utilizam *MRecords* por utilizarem BPDUs próprias.

As *bridges* são configuradas com alguns parâmetros, tais como a identificação da Região MST (*MST Configuration Identifier*), identificação de cada uma das MSTIs e um mapa associando valores de VLAN ID a identificadores de MSTI. Para a IST, são associados todos os valores de VLAN ID a serem utilizados na respectiva Região MST.

O *MST Configuration Identifier*, por ser comum a todas as MSTIs, é transportado nas BPDUs da IST.

O mapa de associação entre VLAN IDs e identificadores de MSTI é de natureza estática (isto é, não é transportado por BPDUs), sendo utilizado, no plano

de dados, para identificar a MSTI a ser utilizada a partir do VLAN ID contido no quadro de dados.

3.14. MULTIPLE VLAN REGISTRATION PROTOCOL (MVRP)

Conforme o padrão IEEE 802.1Q-2014 (que obsoletou e incorporou o padrão IEEE 802.1ak), existem dois mecanismos que possibilitam a configuração de informações relativas à participação de componentes de uma *bridged* LAN em uma determinadada VLAN:

— Configuração estática dessas informações via ações da gerência da rede;
— Configuração e distribuição dinâmicas dessas informações com base em um protocolo específico.

As informações relativas à participação de componentes da rede (estações finais e *bridges*) em uma VLAN são configuradas (registradas), estática ou dinamicamente, em portas desses componentes. O conjunto de portas com registro de uma VLAN constitui o *Member Set* dessa VLAN.

Os registros de participação de portas em uma VLAN representam um mecanismo efetivo de filtragem de quadros *unicast*, *multicast* ou *broadcast* dessa VLAN por essas portas. Tais quadros só podem ser transmitidos por uma porta que seja membro do *Member Set* dessa VLAN.

Na configuração estática tais registros efetivam-se manualmente, constituindo assim as *Static* VLAN *Registration Entries* relativas às VLANs.

O protocolo utilizado para a configuração dinâmica dos *Member Sets* de VLANs é o MVRP. A operação do MVRP resulta na constituição das *Dynamic* VLAN *Registration Entries*, que indicam dinamicamente quais as VLANs registradas em cada porta.

Os mecanismos de configuração estática e dinâmica de VLANs podem ser utilizados de forma combinada em uma *bridged* LAN, de forma tal que os *Member Sets* de determinadas VLANs sejam constituídos de modo diferente que os das demais VLANs.

Conforme menção anterior, o MVRP é uma aplicação do MRP que obsoletou e substituiu o GVRP a partir da emissão do padrão IEEE 802.1ak.

O MVRP provê um mecanismo para registro e manunteção dinâmicos do conteúdo das *Dinamic* VLAN *Registration Entries* para cada VLAN, e para a propagação dessas informações para os demais componentes VLAN-*aware* da rede. Tais informações possibilitam que componentes VLAN-*aware* registrem e mantenham o conhecimento das VLANs que tenham membros ativos, e por que portas esses membros ativos podem ser alcançados.

Recorda-se que o tráfego de dados de uma VLAN flui de acordo com os registros MVRP que ocorreram no sentido oposto a esse tráfego.

3.14.1. **Visão Geral do MVRP**

Para o MVRP, os valores de atributo transportados pelas MVRPDUs são os valores de VID (identificadores de VLANs), que, como visto anteriomente neste capítulo, utilizam um campo de 12 bits.

3.14.1.1. Operação do MVRP

A operação básica do MVRP, por consistir de uma aplicação do MRP, tem como base a Figura 3.11 deste capítulo.

Essa figura é complementada pela Figura 3-, onde podem ser visualizados a arquitetura e o modo de operação básica do MVRP, para o caso de uma *bridge* com duas portas e uma estação final.

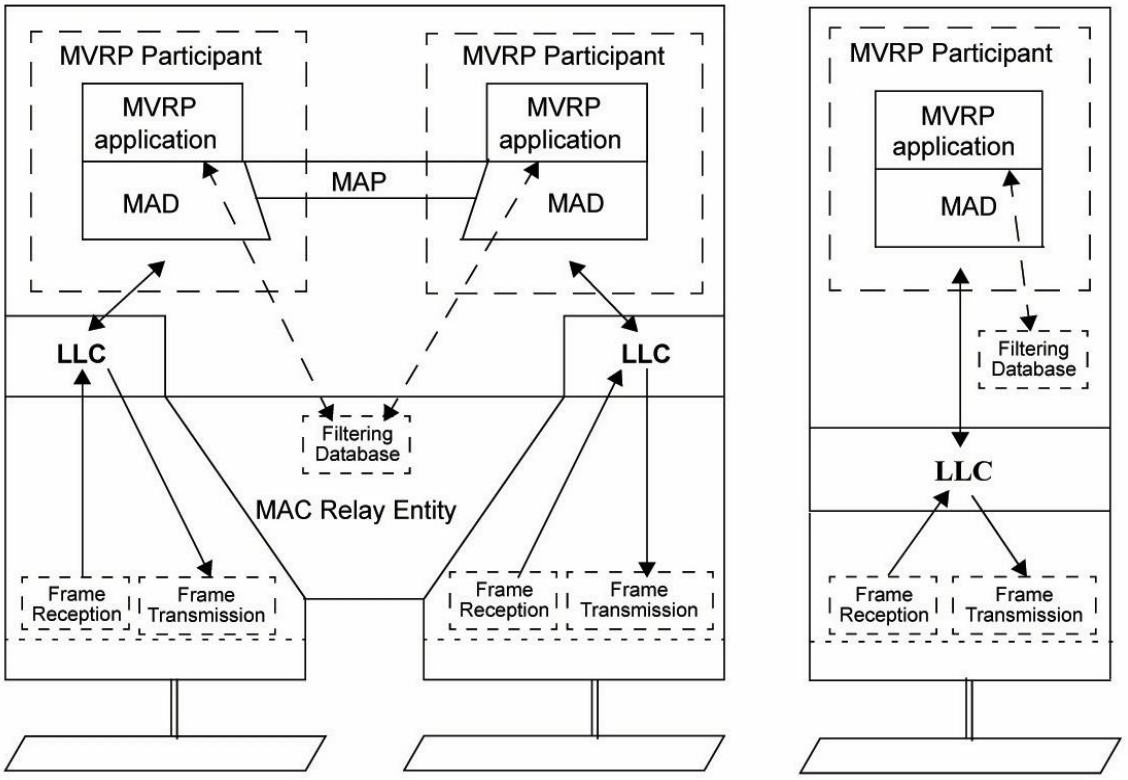


Figura 3-31– Arquitetura do MVRP.

Como se verifica nessa última figura, a operação dos MVRP *Participants* envolvem os seguintes componentes:

— A aplicação MVRP;
— A *MRP Attribute Propagation* (MAP);
— A *MRP Attribute Declaration* (MAD).

3.14.1.2. Comportamento de Estações Finais e de Bridges

Quando uma estação final VLAN-*aware* deseja se tornar membro do *Member Set* de uma VLAN (por meio do serviço VLAN *Registration*, que será visto adiante), isto é, habilitar-se a receber quadros nessa VLAN, ela registra esse fato na porta de *bridge* conectada à LAN à qual essa estação se encontra também conectada.

O componente VLAN-*aware* da *bridge* que recebe essa solicitação de registro por uma de suas portas, encaminha essa solicitação para o MVRP *Participant* dessa porta. Esse MVRP *Participant* armazena essa declaração em sua MAD, e a envia para constituir uma *Dynamic* VLAN *Registration Entry* na base de dados de filtragem (*Filtering Database*) da entidade *MAC Relay* da *bridge*.

Por comando da MAD, a porta da *bridge* envia uma declaração de ocorrência de registro na VLAN para a estação final que solicitou o registro, estando essa estação final assim habilitada para receber tráfego de dados na VLAN registrada.

A partir do recebimento da declaração por parte da porta de *bridge*, a estação final passa a se considerar membro do *Member Set* da VLAN.

O MVRP *Participant* receptor da solicitação de registro transfere essa solicitação, via MAP, para as outras portas da *bridge* que sejam portas ativas da MST correspondente à VLAN declarada. O uso do MAP de forma específica para cada VLAN significa a utilização de diferentes contextos MAP.

A solicitação de registro é então transmitida por essas portas, alcançando assim um conjunto de portas situadas nas *bridges* receptoras dessa solicitação.

Cada porta assim alcançada realiza o registro solicitado e envia uma declaração de realização desse registro para a porta que o solicitou. Isso habilita a porta solicitante a transmitir quadros de dados nessa VLAN, ou seja, a se tornar um membro do *Member Set* dessa VLAN.

Os procedimentos acima se repetem em todas as VLAN-*aware bridges* da *bridged LAN*, em cascata, podendo inclusive atingir estações finais.

Uma estação final VLAN-*aware* receptora, ao receber uma solicitação de registro em uma VLAN e registrá-la em sua base de dados de filtragem, habilita-se a realizar o que se denomina *source prunning*. Isso significa que essa estação final só transmite um quadro de dados em uma VLAN por uma de suas portas, caso essa porta tenha recebido uma solicitação de registro relativa a essa VLAN, e tenha declarado a realização desse registro à porta de *bridge* que a solicitou.

3.14.1.3. Formato e Codificação de MVRPDUs

As MVRPDUs são transmitidas em quadros *Ethernet* no modo *untagged*. Esses quadros utilizam como endereço MAC de destino o endereço de grupo 01-80-C2-00-00-21.

A transmissão de quadros MAC contendo as MVRPDUs utiliza o LLC (o LLC-1, a rigor).Para a indicação do MVRP, o LLC utiliza, via SNAP, a codificação *EtherType*. Foi adotado o código *EtherType* 0x88F5 para o MVRP.

Para as MVRPDUs, formatada conforme a Figura 3.14 deste capítulo, aplicam-se as seguintes particularidades:

— O campo *ProtocolVersion*, relativo à versão do MVRP definida pelo padrão IEEE 802.1ak-2007 e mantido pelo padrão 802.1Q, tem o valor 0x00;
— O MVRP define um único *Attribute Type*, que é o VID *Vector Attribute Type*, cujo valor de identificação é igual a 1;
— Os atributos identificados pelo VID *Vector Attribute Type* são instâncias de *VectorAttributes*, utilizadas para identificar uma sequência de valores de VID;
— O valor *FirstValue* de cada uma dessas instâncias, codificado em dois octetos, representa um valor de VID declarado, que pode se encontrar na faixa de valores de 1 a 4094.

Observa-se que uma estação final *non*-VLAN-*aware* que se torna membro de uma VLAN, não necessita declarar esse fato via MVRP. Na realidade, essa declaração é mesmo impossível.

Nesse caso, em que ocorre classificação *port-based* de VLANs na porta de acesso da VLAN-*aware bridge* na qual a estação final *non*-VLAN-*aware* está conectada, os quadros enviados por essa estação final são retransmitidos pela *bridge* receptora na VLAN correspondente ao PVID configurado.

Assim, o MVRP ocorre no interior da rede, com o VID igual ao PVID, a partir de cada *bridge* onde tenha sido configurado esse PVID.

3.14.2. **Serviço VLAN-Registration**

A solicitação para o registro de participação de um componente VLAN-*aware* (tipicamente uma estação final) no *Member Set* de uma VLAN efetiva-se por meio do serviço VLAN *Registration*. Esse serviço possibilita ao usuário do MAC *Service* indicar ao provedor do MAC *Service* o desejo de receber quadros pertinentes a uma dada VLAN (identificada pelo respectivo VID).

As primitivas de serviço para a solicitação de registro e de deregistro de participação no *Member Set* de uma VLAN são respectivamente as seguintes:

— ES_REGISTER_VLAN_MEMBER (VID);
— ES_DEREGISTER_VLAN_MEMBER (VID).

3.15. APRENDIZAGEM DE ENDEREÇOS MAC POR VLAN-AWARE BRIDGES

Ressalvadas a existência de VLANs e a utilização do MSTP e do MVRP, o processo de aprendizagem de endereços MAC por VLAN-*aware bridges* é o mesmo que o utilizado por *bridges* IEEE 802.1D (VLAN-*unaware bridges*).

Quando evocado, o processo de aprendizagem em VLAN-*aware bridges* deve criar ou atualizar uma entrada de filtragem dinâmica (*Dynamic Filtering Entry*), que

especifica a porta de recepção associada ao endereço MAC *unicast* de origem e ao VID dos quadros de dados recebidos nessa porta.

O anexo F do padrão IEEE 802.1Q-2014 especifica as seguintes abordagens para aprendizagem de endereços MAC *unicast* em função das VLANs identificadas nos quadros de dados:

— Exclusivamente *Shared VLAN Learning* (SVL);
— Exclusivamente *Independent VLAN Learning* (IVL);
— SVL/IVL (SVL e IVL).

Shared VLAN *Learning* (SVL) é um mecanismo para configuração e para operação do processo de aprendizagem de endereços MAC, de modo tal que, se um dado endereço MAC individual for aprendido em uma VLAN associada a uma MST, essa informação aprendida será utilizada para encaminhamento de tráfego para aquele endereço em todas as VLANs que compõem o conjunto constituído por todas as VLANs associadas a essa MST.

Uma VLAN-*aware bridge* que suporta exclusivamente SVL é referida como uma SVL *bridge*.

Independent VLAN *Learning* (IVL) é um mecanismo para configuração e para operação do processo de aprendizagem de endereços MAC, de modo tal que, se um dado endereço MAC individual for aprendido em uma VLAN associada a uma MST, essa informação aprendida será utilizada para encaminhamento de tráfego para aquele endereço exclusivamente nessa VLAN.

Uma VLAN-*aware bridge* que suporta exclusivamente IVL é referida como uma IVL *bridge*.

Uma VLAN-*aware bridge* que suporta SVL e IVL é referida como uma SVL/IVL *bridge.*

Em SVL e IVL, os *Filtering Databases* são constituídos por FIDs (*Filtering Identifiers*). Os FIDs são diferentes instâncias de uma base de dados de filtragem, cuja funcionalidade depende do tipo de aprendizagem de endereço utilizado.

No SVL, um único FID atende, em uma *bridge*, o conjunto de VLANs associado a uma MST, enquanto no IVL cada FID corresponde, na *bridge*, a uma VLAN independente.

No modelo SVL/IVL, que é uma forma mista de operação entre o SVL e o IVL, uma das FIDs de uma SVL/IVL *bridge* pode corresponder a um sub-conjunto de VLANs associadas a uma MST, enquanto as demais FIDs correspondem a VLANs independentes.

A escolha do tipo de processo de aprendizagem de endereços MAC depende de alguns fatores, uma vez que existem prós e contras para cada um deles.

Uma condição para o uso de SVL ou de SVL/IVL é a de que os endereços MAC no conjunto de todas as VLANs sejam globalmente únicos nesse conjunto.

O padrão IEEE 802.1Q-2014 define normas para alocação de VIDs em FIDs, sendo que tal alocação pode ocorrer de modo fixo ou por modos dinâmicos.

3.16. TRANSMISSÃO DE QUADROS DE DADOS EM VLAN-AWARE BRIDGES

Quando da transmissão de um quadro de dados com endereço *unicast*, *multicast ou broadcast* em uma VLAN, a indicação dessa VLAN (ou seja, o respectivo VLAN ID) contida nesse quadro de dados é associada, na MST *Bridge* de origem, a uma MSTI.

O curso de um quadro de dados com endereço *unicast* depende de uma das seguintes condições:

— O endereço MAC *unicast* de destino encontra-se na mesma Região MST que o endereço *unicast* de origem, e a sua aprendizagem encontra-se registrada nas bases de dados de filtragem correspondentes à MSTI nas MST *bridges* dessa Região MST;
— O endereço MAC *unicast* de destino encontra-se em outra Região MST e a sua aprendizagem está registrada na base de dados de filtragem correspondente à CST na CIST *Regional Root Bridge* da Região MST de origem;
— Nenhuma das condições acima é atendida.

Na primeira condição o quadro de dados trafega, de modo direcionado, exclusivamente pela MSTI que corresponde à VLAN indicada nesse quadro, no interior da Região MST.

Na segunda condição, o quadro de dados trafega pela MSTI de origem até alcançar a *CIST Regional Root Bridge* da respectiva Região MST. Essa *bridge* interrompe a transmissão do quadro de dados na MSTI, e o encaminha, via CST, à *CIST Root Bridge* através da sua *Master Port*.

Cabe à CIST *Root Bridge* reencaminhar esse quadro para a CIST *Regional Root Bridge* da Região MST de destino. Tal envio ocorre em obediência à CST, o que significa que o quadro deverá cursar pelas eventuais CIST *Regional Root Bridges* intermediárias indicadas pela CST.

Ao receber o quadro de dados, a *CIST Regional Root Bridge* de destino verifica, com base no VLAN ID contido no quadro de dados, qual MSTI da respectiva Região MST deverá ser utilizada. O quadro de dados é inserido na MSTI indicada, sendo então conduzido ao seu destino.

Na terceira condição, o quadro de dados com endereço *unicast* é tratado como se possuísse endereço *broadcast*. Ocorre a inundação da MSTI correspondente, e também de toda a MST, via CST.

A CST encaminha o quadro para todas as regiões MST por inundação, sendo esse quadro de dados distribuído, também por inundação, nas MSTIs das Regiões MST associadas ao VLAN ID contido no quadro de dados.

A transmissão de quadros de dados com endereços *unicast* ou endereços *broadcast* nas respectivas MSTIs ocorre de acordo com as tabelas de registro criadas pelo MVRP aplicáveis à VLAN indicada no quadro de dados.

Observa-se que, embora uma MSTI possa estar associada a múltiplas VLANs, os registros criados pelo MVRP referem-se a cada uma das VLANs independentemente. Uma porta de *bridge* em uma MSTI pode estar bloqueada por falta de registro pelo MVRP para uma determinada VLAN, encontrando-se, contudo desbloqueada para uma outra VLAN também associada a essa MSTI (uma VLAN para a qual exista registro MVRP).

Os quadros de dados com endereço *multicast* são distribuídos na MST (CIST e MSTIs) de acordo com as tabelas de registro de endereços de grupo criadas pelo MMRP. As MSTIs utilizadas são também definidas com base no VLAN ID contido em cada quadro de dados.

LAN BRIDGING COM IS-IS

4.1. PREÂMBULO

Conforme palavras de Radia Perlman, a própria inventora do mecanismo *spanning tree*, a utilização de *spanning trees* representa uma baixa eficiência no uso dos recursos das redes Ethernet que as utilizam.

Dentro dessa perspectiva, o IEEE e o IETF passaram a desenvolver novas tecnologias para *LAN bridging*, agora utilizando extensões do protocolo IS-IS.

Assim, surgiram o SPB (*Shortest Path Bridging*), especificado pelo IEEE mas com a participação do IETF em seu plano de controle, e o TRILL (*Transparent Interconection of Lots of Links*), definido inteiramente pelo IETF.

O presente capítulo apresenta o SPB e o TRILL.

4.2. UTILIZAÇÃO DO IS-IS PARA ROTEAMENTO DE CAMADA 2

O protocolo IS-IS foi ampliado no sentido de possibilitar a sua utilização para roteamento de camada 2, como será visto adiante neste capítulo.

4.2.1. Protocolo IS-IS

O protocolo IS-IS (*Intermediate System to Intermediate System*), definido no padrão ISO/IEC 10589 e na RFC 1195 (*Use of OSI IS-IS for Routing in TCP/IP and Dual Environments*), é um protocolo *link-state* que equivale ao protocolo OSPF, sendo que ambos objetivam especificar os processos de roteamento no interior de domínios administrativos (roteamento do tipo IGP) de inter-redes.

Como no OSPF, o IS-IS é constituído por domínios de roteamento IS-IS (*IS-IS Routing Domains*) e, no interior de um domínio, por diferentes áreas.

O IS-IS utiliza os seguintes tipos de PDU:

- LSPs (*Link State Packets* ou *Link State Protocol Data Units*);
- IIH (*IS-IS Hello*) PDUs;
- CSNPs (*Complete Sequence Number PDUs*);
- PSNPs (*Partial Sequence Number PDUs*).

Cada roteador IS (*intermediate system*) distribui LSPs com informações de roteamento. Os LSPs (não confundir com os LSPs do MPLS) são equivalentes às LSAs (*link state advertisements*) do OSPF.

Os LSPs são compostos por um cabeçalho fixo e por um conjunto de TLVs, onde são transmitidos os parâmetros necessários ao roteamento, inclusive as rotas consideradas. Os TLVs podem ser compostos por sub-TLVs, onde mais especificamente são codificados os parâmetros transportados.

As IIH PDUs são utilizadas no processo inicial de autodescobrimento dos sistemas intermediários e na manuntenção de adjacências entre eles.

As CSNPs são utilizadas para garantir que ISs vizinhos tenham o mesmo conhecimento de quais foram os últimos LSPs trocados entre eles e que estejam sincronizados.

As PSNPs são utilizadas para requisitar um LSP (ou LSPs) e acusar o recebimento de um LSP (ou LSPs).

4.2.2. **Extensões do IS-IS para Roteamento de Camada 2**

A RFC 6165 (*Extensions to IS-IS for Layer-2 Systems)* objetiva definir extensões do protocolo IS-IS para suportar roteamento *link state* em diferentes tecnologias de camada 2. São descritos dois novos TLVs de caráter genérico, sendo que adições específicas julgadas necessárias devem ser objeto de extensões específicas para cada tecnologia de camada 2 atendida.

O paradigma atual para o funcionamento de redes *bridged Ethernet*, baseado nos padrões IEEE 802.1D-2004 e IEEE 802.1Q-2014, apresentados anteriormente, fundamenta-se na constituição de *spanning trees*.

A despeito do advento das MSTs (*multiple spanning trees*), que possibilitam a distribuição de tráfego nas *bridged LANs* com base nos VLAN IDs contidos nos quadros de dados, a solução baseada em *spanning trees* permanece inadequada no que concerne o uso dos recursos da rede.

As *spanning trees* são montadas de forma arbitrária. O critério para a escolha das *bridges* raiz (*root bridges*) das *spanning trees,* aleatoriamente situadas, conduz à constituição de árvores que levam à utilização desnecessária de recursos da rede, além de causar *delays* mais elevados no curso dos quadros de dados.

Além do mais, a metodologia utilizada nas *spanning trees* não proporciona a possibilidade de utilização de *multi-topologies*, o que representa uma razão adicional para a baixa eficiência de utilização de recursos de rede.

Por essas razões, o IEEE e o IETF vêm desenvolvendo novas alternativas de operação de *bridged LANs* baseadas em protocolos de roteamento IP, prevalecendo o IS-IS como base para isso. As árvores proporcionadas por esses novos protocolos são constituídas por *links* de menor custo diretamente entre a origem e o destino dos quadros de dados.

Acresce-se o fato de que a possibilidade de uso de roteamento ECMP (*Equal-Cost Multi-Path*) nesses protocolos proporciona melhor distribuição de tráfego na rede.

Assim, foram definidas duas novas tecnologias para *LAN Bridging* baseadas no IS-IS:

— SPB (*Shortest Path Bridging*), inicialmente especificado no padrão IEEE 802.1aq, tendo sido esse padrão obsoletado e o seu conteúdo incorporado às Cláusulas 27 e 28 do padrão IEEE 802.1Q-2014;
— TRILL (*Transparent Interconnection of Lots of Links*), pelo IETF .

Para o plano de controle do SPB o IETF emitiu a RFC 6329 (*IS-IS Extensions Supporting IEEE 802.1aq Shortest Path Bridging*), enquanto que para o TRILL foram definidas diferentes RFCs específicas.

Como suporte à definição dessas tecnologias, são utilizadas as seguintes RFCs:

— RFC 5120, intitulada M-ISIS: *Multi-Topology* (MT) *Routing in Intermediate System to Intermediate Systems* (IS-ISs);

— RFC 5305 (*IS-IS Extensions for Traffic Engineering*);
— RFC 6165 (*Extensions to IS-IS for Layer-2 Systems*).

O propósito da RCF 5120 é o de descrever a constituição, dentro de um domínio IS-IS, de um conjunto de topologias independentes, referidas como *multi-topologies*.

A RFC 5305, vista anteriormente neste livro, foi emitida com o propósito de prover mecanismos de roteamento com engenharia de tráfego, particularmente para o MPLS-TE.

A RFC 6165 define os seguintes dois tipos específicos de TLV:

— *Multi-Topology aware Port Capability* (MT-PORT-CAP) TVL;
— MAC – *Reachability* (MAC-RI) TVL.

O MT-PORT-CAP TLV utiliza o código 143 para o tipo de TLV, e objetiva transportar, além do MT-ID da topologia sendo anunciada, um conjunto de sub-TLVs válidos para a tecnologia de camada 2 específica. A formatação desses sub-TLVs é definida na RFC 5305, e suas definições são específicas para cada tecnologia que os utiliza.

O transporte de MT-PORT-CAP PDUs ocorre no interior de IIH PDUs (*IS-IS Hello* PDUs), ou seja, em mensagens *Hello*. Uma IIH PDU pode conter mútiplos MT-PORT-CAP TLVs.

O MAC-RI TVL utiliza o código 147 para o tipo de TLV, e tem como função primordial transportar os endereços MAC alcançáveis a partir do IS que anuncia o TLV. Transporta também o VLAN ID da VLAN roteada.

Esse tipo de TLV é transportado em um LSP (*Link State Packet*) convencional, e pode estar contido múltiplas vezes em uma ou em múltiplas LSPs . O seu funcionamento detalhado depende da tecnologia de camada 2 específica.

4.3. SHORTEST PATH BRIDGING

O SPB (*Shortest Path Bridging*), definido inicialmente no padrão IEEE 802.1aq (*Shortest Path Bridging*), teve a sua especificação incorporada à nova versão do padrão IEEE 802.1Q emitida em dezembro de 2014, em suas Cláusulas 27 e 28.

As adaptações do padrão IEEE 802.1aq incorporadas ao padrão IEEE 802.1Q-2014 foram inicialmente definidas no *amendment* IEEE 802.1Qbp-2014, denominado *Equal Cost Multipath*.

Foram especificados dois modos de operação do SPB:

— SPBV (*Shortest Path Bridging* – VID);
— SPBM (*Shortest Path Bridging* –MAC).

O SPBV e SPBM eram referidos no início do processo de definição do padrão IEEE 802.1aq como SPB e SPBB (*Shortest Path Backbone Bridging*), respectivamente. Na nova terminologia, a sigla SPB passou a referir-se a ambas tecnologias SPB definidas.

Na fase inicial desse processo de definição, ocorreu uma proposta, por fabricantes, da tecnologia PLSB (*Provider Link State Bridging*), que não foi adotada em sua totalidade, mas que teve os seus conceitos chaves incorporados na definição do SPB.

Os objetivos práticos para projeto dos SPBV e SPBM são, respectivamente, redes com 100 nós e redes com 1000 nós.

Os modos de operação do SPB adotam planos de dados próprios, mas utilizam, no plano de controle, um único conjunto de extensões do protocolo IS-IS, extensões essas referidas como protocolo ISIS-SPB. As bases de dados de filtragem são preenchidas como consequência da aplicação do protocolo ISIS-SPB. O ISIS-SPB foi definido na RFC 6329 (*IS-IS Extension Supporting IEEE 802.1aq Shortest Path Bridging*).

O plano de dados do SPBV tem como base o encapsulamento utilizado no plano de dados das redes PB, ou mesmo das redes PBB. As bases de dados de filtragens são calculadas e instaladas para VIDs unidirecionais referidos como SPVIDs.

Como no SPBV, diferentemente do que ocorre no SPBM, as rotas utilizadas pelo ISIS-SPB não constam dos quadros de dados enviados através da rede SPBV, a aprendizagem de endereços MAC ocorre de forma idêntica àquela utilizada nas *bridged LANs* convencionais, tanto nas nas BEBs quanto nas BCBs. Os registros nas bases de dados de filtragem dessas *bridges* associam endereços C-MAC de origem dos quadros de dados às portas de *bridge* por onde esses quadros foram recebidos.

O SPBM, por sua vez, tem o seu plano de dados baseado no plano de dados das redes PBB, sendo o seu uso indicado para aplicações onde o completo isolamento entre os endereços de usuários e os endereços da rede *backbone* é desejável.

A transmissão de quadros de dados no SPBM requer a associação entre o endereço MAC da estação final de destino e o endereço B-MAC da BEB de destino, estando essa associação registrada na BEB de origem. Como o endereço B-MAC é a rota de destino para o ISIS-SPB, torna-se desnecessária a aprendizagem de endereços pelas BCBs.

Os registros dessas associações ocorrem no plano de dados, com base no tráfego de dados no sentido inverso, da mesma forma do que ocorre no PBB.

Quando da inexistência dessa associação, tanto para o SPBV quanto para o SPBM, o quadro deve ser transmitido por inundação, respeitadas as respectivas SPTs e suas eventuais filtragens.

O SPBV e o SPBM podem coexistir simultaneamente em uma mesma rede SPB, compartilhando uma única SPT ou um conjunto de até 16 ECTs.

4.3.1. **Roteamento no SPB**

O protocolo ISIS-SPB, especificado pela RFC 6329, destina-se a prover roteamento no SPB definido pelo padrão IEEE 802.1Q-2014.

Uma SPB *bridge* deve especificar o valor de NLPID (*Network Layer Protocol Identifier*) para o SPB em suas mensagens *Hello* quando da constituição de uma instância MT. Tal valor foi definido como sendo igual a 0xC1. *Links* que não suportam esse valor de NLPID devem ser excluídos na constituição de instâncias MT.

A RFC 6329 define nove novos sub-TLVs adicionais ao IS-IS, baseados em novos TLVs definidos na RFC 5120, na RFC 5305, na RFC 6165 e na própria RFC 6329.

Tais TLVs são os seguintes:

— *Multi-Topology-aware Port Capability (MT-Port-Cap) TLV;*
— *Multi-Topology Capability (MT-Capability) TLV;*
— *Multi-Topology Intermediate System Neighbor (MT-ISN) TLV;*
— *Extended IS Reachability TLV*.

O *MT-Port-Cap* TLV, cujo códido de identificação é 143, foi definido na RFC 6165. Esse TLV é transportado por IIH (*IS-IS Hello*) PDUs.

O *MT-Capability* TLV, cujo código de identificação é 144, foi definido na RFC 6329. Esse TLV é transportado por LSPs (*Link State Protocol Data Units*).

O MT-ISN TLV, cujo código de identificação é 222, foi definido na RFC 5120. Esse TLV é transportado por LSPs.

O *Extended IS Reachability* TLV, cujo código de identificação é 22, foi definido na RFC 5305, sendo também transportado por LSPs.

Os sub-TLVs definidos na RFC 6329 foram classificados da seguinte forma:

— Extensões do protocolo *Hello* (TLV 143) : SPB MCID (SPB-MCID, com código 4) *sub-TLV*, *SPB Digest* (*SPB-Digest,* com código 5) *sub-TLV* e *SPB Base VLAN-Identifiers* (SPB-B-VID, com código 6) *sub-TLV*;
— Extensões de informações nodais (TLV 144): *SPB Instance* (*SPB Inst,* com código 1) *sub-TLV* e *SPB Instance Opaque ECT-ALGORITHM* (SPB-I-OALG, com código 2) *sub-TLV*;
— Extensões de informações de adjacências (TLV 222 ou TLV 22): SPB *Link Metric* (SPB-*Metric*, com código 29) *sub-TLV* e SPB *Adjacency Opaque ECT-ALGORITHM* (SPB-A-OALG, com código de 30) *sub-TLV*;
— Extensões de informações de serviços (TLV 144): S*PBM Service Identifier and Unicast Address* (SPBM-SI, com código 3) *sub-TLV* e *SPBV MAC Address* (SPBV-ADDR, com código 4) *sub-TLV*.

Observa-se a existência de dois sub-TLVs (SPB-MCID *sub-TLV*) e SPBV-ADDR *sub-TLV*) com o mesmo código (código 4).

Essas informações encontram-se sumarizadas na Figura 4-1.

PDU	TLV/SUB-TLV	Tipo de TLV	Tipo de Sub-TLV
	MT-Port-Cap	143	-

IIH	SPB-MCID	-	4
	SPB-Digest	-	5
	SPB-B-VID	-	6
LSP	MT-Capability	144	-
	SPB-Inst	-	1
	SPB-I-OALG	-	2
	SPBM-SI	-	3
	SPBV-ADDR	-	4
	MT-ISN ou Extended IS Reachability	222 ou 22	-
	SPB-Metric	-	29
	SPB-A-OALG	-	30

Figura 4-1 – TLVs e Sub-TLVs adicionais para o ISIS-SPB.

4.3.2. **Constituição de SPTs (Shortest Path Trees)**

Pela aplicação do ISIS-SPB, constitui-se, no SPB, uma árvore ponto a multiponto a partir de cada uma das SPB *bridges* da rede, com suporte nos NNI *links* que apresentam menores custos para alcançar as demais SPB *bridges* da rede. Cada uma dessas árvores, centradas em SPB *bridges,* é referida como uma SPT (*Shortest Path Tree*). Uma SPT é também denominada MDT (*Multicast Distribution Tree*).

Na versão atual do SPB apresentada no padrão IEEE 802.1Q-2014, apenas NNI *links* ponto a ponto são considerados, embora NNI *links* multiponto (*Shared LANs*) possam ser suportados, o que poderá ser contemplado em futuras versões do SPB.

O ISIS-SPB é aplicável em redes SPB dentro de duas condições, que são extensões diretas das propriedades fundamentais da operação de *bridged Ethernet* LANs:

— Congruência entre caminhos reversos;
— Congruência entre transmissão *unicast* e transmissão *multicast*.

O conjunto de entidades SPB nas *bridges* habilitadas para o SPB (ou seja, nas SPB *bridges*) da rede que suporta o SPB (uma PBBN, por exemplo), é referido como uma Região SPB.

A Figura 4-2 apresenta um exemplo de constituição de uma SPT (ou MDT) em uma rede (Região) SPB.

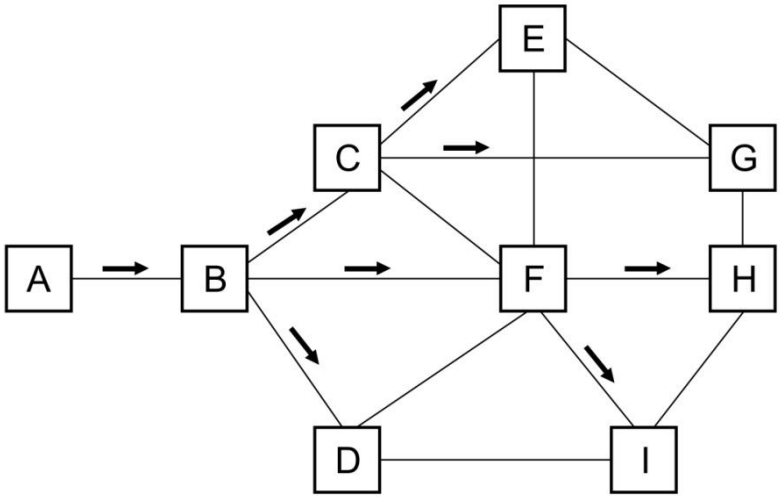


Figura 4-2 - Exemplo de constituição de SPT.

Observa-se que, de acordo com a segunda condição acima, uma SPT (como a exemplificada nessa figura) é obedecida tanto para tráfego *unicast* quanto para tráfego *multicast/broadcast*.

A Figura 4-3 representa a aplicação global do ISIS-SPB na rede SPB correspondente à figura anterior, onde se verifica a utilização *full-mesh* e bidirecional dessa rede.

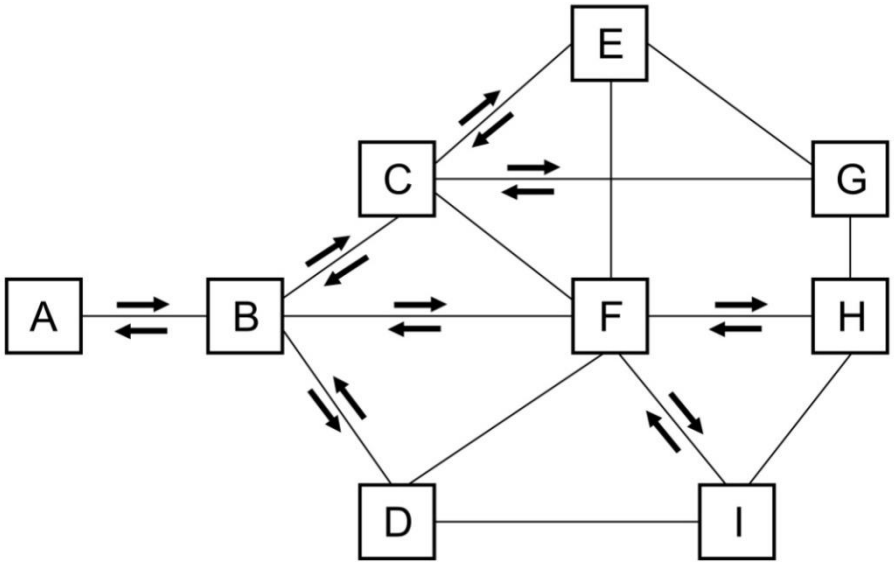


Figura 4-3 – Utilização *full-mesh* e bidirecional da rede da figura anterior.

Observa-se nessa figura a aplicação da condição relativa à congruência de caminhos reversos.

Para a utilização mais eficiente de redes SPB pela melhor distribuição de tráfego em seus NNI *links*, podem ser constituídas múltiplas ECTs (*Equal Cost Trees*) a partir das SPB *bridges* da rede, o que será aprofundado a seguir.

4.3.3. **Uso de ECTs**

O padrão IEEE 802.1Q-2014 possibilita a constituição de múltiplas ECTs a partir das SPB *bridges* de uma rede SPB, do que resulta a disponibilização de múltiplos caminhos com custos menores e iguais entre os pares de nós da rede.

Para possibilitar a constituição de múltiplas ECTs conforme o parágrafo anterior, todas as informações de controle relativas ao SPB devem estar, implícita ou explicitamente, associadas com um *Multi-Topology Identifier* (MT ID). As informações SPB relativas a um dado MT ID constituem uma única instância SPB lógica (*logical SPB instance*).

Tráfego no plano de dados relativo a um determinado MT ID é intrinsicamente isolado pelo VID atribuído à instância SPB em questão.

O MT ID #0 é o *default* MT ID.

Mediante a aplicação de algoritmos especificamente definidos, o nó de origem pode distribuir o tráfego entrante na rede entre ECTs, de modo a otimizar a utilização dos seus NNI *links*, utilização essa já eficiente em redes *link state* comparativamente aos processos tradicionais de *bridged* LANs baseados em *spanning trees*.

No presente estado de definição do padrão IEEE 802.1Q (versão de 2014), é possível a constituição de até 16 caminhos com custos menores e iguais entre dois nós da rede, o que possibilita a distribuição de tráfego com base em VLANs.

A constituição de caminhos alternativos com custos menores e iguais efetiva-se mediante a utilização de formas de manipulação na definição (*tie breaking*) dos múltiplos caminhos mais curtos com custos iguais.

Para isso, cada SPB *bridge* é univocamente identificada por um *bridge* ID. Um *bridge* ID, com 64 bits (8 octetos), é composto por um valor de *bridge priority* com dois octetos, e pelo SYSID (*System ID*) com 6 octetos.

SYSIDs são endereços MAC bem conhecidos (*well know*) de equipamentos SPB.

Um caminho escolhido é definido pela sucessão de *bridge* IDs que o compõem, sucessão essa referida como o PATHID desse caminho. Na constituição de um PATHID são excluídos os SYSIDs dos pontos de terminação do respectivo caminho.

A definição de ECTs a partir de uma SPB *bridge* raiz é feita sobre PATHIDs de menor custo para alcançar as demais SPB *bridges* da rede, podendo esses caminhos serem escolhidos em função dos valores de *bridge* IDs das SPB *bridges*, desde que com PATHIDs que tenham o mesmo menor custo para alcançar as rotas definidas.

Mediante transformações dos valores de parâmetros de *tie breaking* na rede, isto é, utilizando-se diferentes ECT-ALGORITHMS, podem surgir outros PATHIDs com o mesmo menor custo para alcançar as demais SPB *bridges* da rede, do que resulta SPTs alternativas equivalentes, que são as ECTs. A variação aleatória dos valores de *bridge priority* representa um parâmetro de *tie-breaking*.

Algoritmos ECT padronizados foram propostos inicialmente com valores de identificação na faixa de 00-80-C2-01 a 00-80-C2-10, o que representa 16 valores.

Esses valores de identificação consistem em um valor único de OUI (*Organizationally Unique Identifier*), igual a 00-80-C2, concatenado com os índices que variam de 0x01 a 0x10.

O algoritmo *default* é o ECT-ALGORITHM igual a 00-80-C2-01.

Para suportar novos mecanismos de *tie-breaking* no futuro, foram definidos dois novos ECT TLVs opacos, que podem ser utilizados para a especificação de parâmetros para ECT-ALGORITHMS fora do espaço definido quando da emissão da RFC 6329.

A Figura 4-4 ilustra a constituição de múltiplos caminhos com custos menores e iguais entre pares de SPB *bridges*.

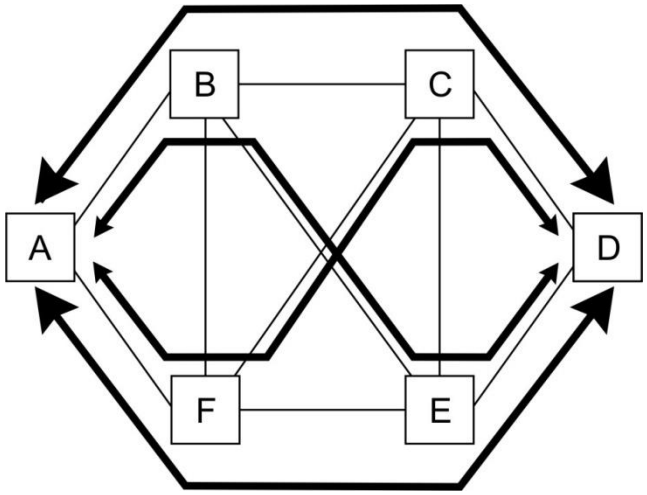


Figura 4-4 - Caminhos com custos menores e iguais.

Observa-se, nessa figura, a existência de quatro ECTs bidirecionais de menor custo entre a SPT *bridge* A e a SPT *bridge* D:

— [A/B/C/D];
— [A/F/E/D];
— [A/B/E/D];
— [A/F/C/D].

Essa figura representa um *SPT Set*. Um *SPT Set* é então um conjunto de SPTs determinadas por um *ECT Algorithm* específico e capacitadas para suportar uma ou mais VLANs dentro de uma Região SPT. Esse conjunto é formado pela união de todas as SPTs com a totalidade de *bridges* que utilizam o *ECT Algorithm*.

Uma instância de topologia IS-IS suporta múltiplos *SPT Sets*, tanto no SPBV quanto no SPBM. A cada *ECT Algorithm*, e consequentemente a cada *SPT Set*, corresponde um valor de *Base VID*.

No SPBM particularmente, o *Base VID* corresponde ao B-VID da Região SPT, no qual são mapeados os I-SIDs, que por sua vez mapeiam os C-VIDs das redes Ethernet usuárias do SPBM.

Como as ECTs são bidirecionais, não importa a ordem de indicação das SPTs *bridges* acima apresentadas. Assim, por exemplo, [A/B/C/D] representa também a transmissão de tráfego a partir da SPT *bridge* D.

A forma de utilização de ECTs na rede é função do tipo de SPB utilizado (SPBV ou SPBM), como será visto adiante neste item. Em qualquer hipótese, é necessária a definição de um algoritmo de distribuição de tráfego na SPB *bridge* raiz das ECTs.

4.3.4. **Mitigação de Loops**

O SPB utiliza um mecanismo de mitigação de *loops,* que resulta da condição de congruência de caminhos reversos adotada, mecanismo esse referido como *ingress checking*.

Por esse mecanismo, um quadro de dados com um endereço MAC-SA deve ingressar em uma SPB *bridge* pela mesma interface pela qual sairia um quadro de dados destinado a esse endereço MAC.

Dessa forma, cada SPB *bridge* verifica o endereço MAC-SA de cada quadro de dados recebido, e verifica se a entrada ocorreu pela interface esperada. Caso isso não ocorra o quadro deve ser descartado.

A prevenção de *loops* ocorre em conformidade com os procedimentos correspondentes ao protocolo IS-IS.

4.3.5. **Shortest Path Bridging – VID (SPBV)**

Conforme menção anterior, o SPBV apresenta o plano de dados baseado no padrão 802.1Q-2014, sendo indicado para configurações de rede SPB de menor porte (até 100 SPB bridges) e para aplicações onde a separação entre endereços da rede SPB e endereços de usuários não é considerada imprescindível.

O SPBV constitui árvores de menor custo (*shortest path trees*) mas também interage com *bridges* legadas utilizando RSTP e MSTP. Para isso, o SPBV contrói uma *spanning tree*, incluindo a CIST (*Common and Internal Spanning Tree*), que é a árvore *default* utilizada para a interação do SPBV com outras *bridges* legadas. Isso também serve como um mecanismo de *backup* com *spanning tree*, se houver algum problema de configuração no SPBV.

Para melhor entendimento da operacionalização do SPBV, particularmente no que concerne o plano de controle, serão abordados os seguintes aspectos neste subitem:

— Identificação de SPTs;
— Transmissão de quadros de dados.

4.3.5.1. Identificação de SPTs (ECTs)

O SPBV utiliza um *shortest path* VID (referido como SPVID) para identificar cada uma das ECTs dentro de uma Região SPB.

A partir de cada SPB *bridge*, será atribuído um valor de SPVID com o objetivo de identificar cada uma das ECTs constituídas a partir dessa SPB *bridge*. A um valor de SPVID podem ser associadas múltiplas VLANs.

A atribuição dos valores de SPVID e a manutenção das tabelas para registros e transformações de valores de VID (*base* VID para SPVID), são realizadas automaticamente pelo ISIS-SPB.

Quando um quadro de dados entra em uma SPB *bridge*, é definido o SPVID com base no valor de VID contido nesse quadro de dados, do que resulta a definição da ECT a ser utilizada.

O quadro de dados, agora contendo o valor de SPVID obtido, será transmitido pela ECT definida (identificada pelo valor de SPVID ao longo da rede SPBV). A forma de transmissão é função do tipo de quadro de dados (*unicast*, *broadcast* ou *multicast*) e de ocorrência (ou não) de registro do endereço MAC de destino (por aprendizagem de endereços MAC para quadros *unicast*, por exemplo).

Quando o quadro de dados atinge o seu limite na SPT utilizada, a SPB *bridge* de destino retira o valor de SPVID e o quadro de dados é transmitido para a estação final de destino.

4.3.5.2. Transmissão de Quadros de Dados

O SPBV utiliza os processos de transmissão de quadros *unicast* e de quadros *broadcast* das redes IEEE 802.1Q-2014, ressalvada a inexistência de *spanning trees*.

A transmissão de um quadro de dados, de forma direcionada ou por inundação, realiza-se dentro da SPT indicada pelo respectivo valor de SPVID. Dessa forma, é assegurado o uso eficiente da rede e a prevenção de ocorrência de *loops* na fase de sua operação estável.

Quadros *unicast* cujos endereços MAC-DA não tenham sido aprendidos, assim como quadros *broadcast*, são transmitidos por inundação da ECT correspondente, ressalvado o uso de mecanismos de filtragem.

Após a aprendizagem do respectivo endereço MAC-DA, um quadro *unicast* será transmitido de forma direcionada através da SPT correspondente.

Como a aprendizadem de quadros de dados MAC *unicast* em portas das SPB *bridges* ocorre com base nos valores de SPVID contido nos quadros de dados, e não em valores de cada VLAN ID em particular, essa aprendizagem ocorre no modo *shared learning*.

A transmissão direcionada de um quadro de dados para um endereço MAC *unicast* aprendido pela BEB de origem depende apenas da existência do registro da associação da porta de saída com o endereço MAC de destino contido nesse quadro, e da definição do valor de SPVID a ser utilizado. Ao longo da rede SPBV

as BCBs transmitem os quadros de dados dentro da árvore de distribuição identificada pelo SPVID, de forma direcionada com base nos registros de associação entre o endereço C-MAC de destino e as respectivas portas de saída.

A transmissão de quadros *multicast* no SPBV ocorre com base no SPBV (S,G), codificado por um SPVID e por um endereço MAC de grupo padrão. Os registros de participação em endereços MAC de grupo realizam-se com suporte no protocolo ISIS-SPB ou no protocolo MMRP.

4.3.6. **Shortest Path Bridging – MAC (SPBM)**

No SPBM, o ISIS-SPB distribui exclusivamente os endereços B-MAC como rotas.

O SPBM utiliza o encapsulamento *MAC-in-MAC* do PBB. Dessa forma, ocorre o total isolamento entre os endereços MAC de usuários (C-MAC *addresses*) e os endereços da rede SPBM (B-MAC *addresses*).

Como no PBB, a associação entre endereços C-MAC de destino e endereços B-MAC de destino ocorre apenas nas *bridges* de borda da rede *backbone*. Essa associação resulta do fato de que os quadros de dados na rede SPBM contêm os endereços C-MAC de origem e os correspondentes endereços B-MAC, sendo que os endereços C-MAC só têm significado para as *bridges* de borda.

Como as STPs se encontram constituídas previamente ao tráfego de dados pelo ISIS-SPB, tendo como rotas de destino os endereços B-MAC, torna-se desnecessária a aprendizagem de endereços MAC pelas *bridges* internas da Região SPBM para transmissão direcionada de quadros de dados com endereços *unicast* aprendidos.

Os quadros de dados SPBM são transmitidos por essas *bridges* internas com base nas associações {B-DA,B-VID} nelas registradas para uma interface de saída, registros esses decorrente da computação do ISIS-SPB.

4.3.6.1. Identificação de SPTs

A uma instância SPBM, identificada por um IS-IS MT-ID (*IS-IS Multi Topology ID*), podem ser atribuídos um ou mais B-VIDs. O uso de múltiplos B-VIDs para uma instância SPBM aplica-se quando se deseja mais de uma STP com igualdade de *shortest path costs*. Um valor de B-VID é aplicável ao longo de toda a SPT a que corresponde.

Sobre o assunto, o IETF emitiu a RFC 5120, intitulada *Multi Topology* (MT) *Routing in Intermediate to Intermediate Systems* (IS-IS).

4.3.6.2. Filtragem de Tráfego

O SPBM utiliza o B-VID e o I-SID nos quadros de dados transmitidos.

Como não se utiliza o MVRP no SPBM, a filtragem para a transmissão eficiente de quadros efetiva-se por um mecanismo específico definido pelo padrão

IEEE 802.1Q-2014. Por esse mecanismo, os valores de I-SID são incluídos em um novo sub-TLV *(Service Identifier and Unicast Address sub*-TLV) em mensagens IS-IS utilizadas na constituição de uma SPT (ou seja, mensagens ISIS-SPB), a partir de cada BEB da rede.

Cada BEB informa, por esse sub-TLV, os valores de I-SID nela configurados. Dessa forma, as *bridges* BEBs e BCBs que compõem uma SPT registram os I-SIDs informados pelas mensagens IS-IS, associando-os às portas de recepção dessas informações.

Esse mecanismo de filtragem aplica-se à transmissão de quadros de dados *unicast*, *multicast* ou *broadcast*, em qualquer condição que se encontre os respectivos endereços MAC de destino. Quando ocorre, por exemplo, a inundação de uma SPT, o quadro de serviço é transmitido apenas pelas portas que registraram (nas FIBs das *bridges* dessa SPT) o valor de I-SID contido nesse quadro.

A Figura 4-5 ilustra o funcionamento desse mecanismo de filtragem, onde se visualiza a sub-árvore da STP da Figura 4-3 anterior com o resultado da filtragem relativa ao I-SID com valor igual a K.

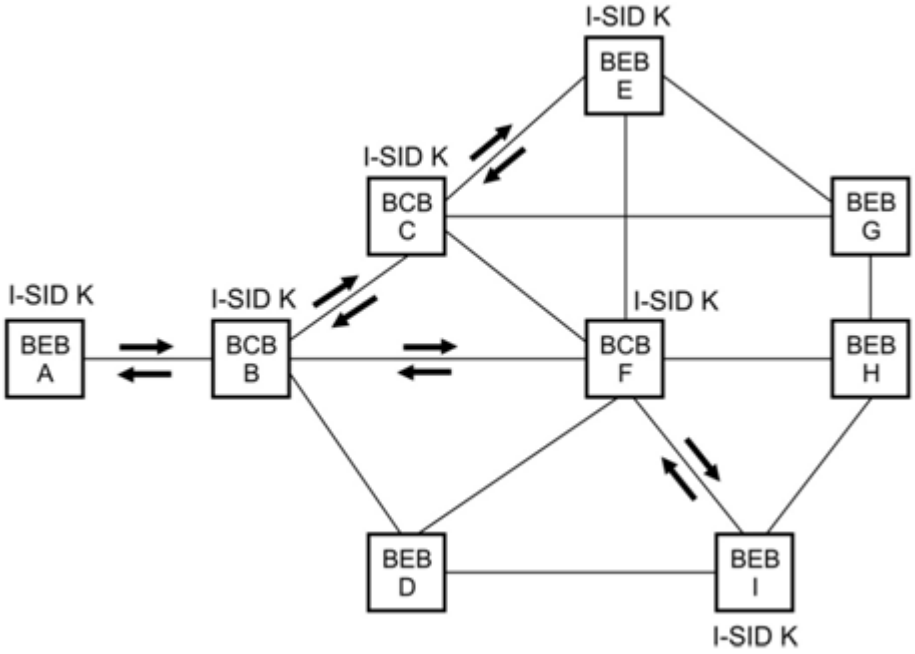


Figura 4-5 - Sub-árvore SP para o I-SID K.

4.3.6.3. Transmissão de Quadros Unicast e Broadcast

Um quadro de dados *broadcast*, ou um quadro de dados *unicast* com endereço C-MAC de destino não associado ao respectivo endereço B-MAC de destino, é transmitido pela BEB receptora por inundação da rede.

Essa inundação ocorre dentro da SPT utilizada (identificada pelo valor de B-VID inserido no quadro), e passa pela filtragem baseada no valor de I-SID como visto anteriormente. O endereço B-DA é o endereço *broadcast* ou um endereço

multicast que atinja a totalidade das BEBs de destino que participam da árvore de distribuição identificada pelo B-VID.

Para um quadro *unicast* com endereço C-MAC de destino associado ao respectivo endereço B-MAC de destino, associação essa registrada na BEB de origem, esse quadro será transmitido na rede SPBM contendo o endereço B-MAC de destino (envelopando o endereço C-MAC de destino), o B-VID identificador da SPT e o respectivo valor de I-SID. Dessa forma, o quadro alcançará a BEB de destino, direcionadamente, pela devida SPT, sem que tenha havido aprendizagem (de endereços C-MAC ou de endereços B-MAC) pelas BCBs atravessadas.

A transmissão conforme o parágrafo anterior por cada porta de *bridge*, assim como no caso de inundação, é também condicionada à existência do registro do I-SID nessa porta.

4.3.6.4. Transmissão de Quadros Multicast

O SPBM utiliza endereços MAC *multicast* (endereços MAC de grupo) de destino (B-DAs) especiais, constituídos pela aplicação de um algoritmo específico para esse fim. Um endereço desse tipo, construído para um nó particular (uma BEB de origem), consiste em um identificador nodal (da BEB de origem), denominado SP*Source*ID (com 20 bits), combinado com o I-SID (24 bits) que define as BEBs de destino associadas ao endereço *multicast*.

A transmissão de quadros de dados com endereços *multicast* ocorre dentro da topologia de uma árvore de distribuição identificada por um B-VID.

Na Figura 4-6 encontra-se representado o formato correspondente à construção de um endereço *multicast* MAC DA utilizado em redes SPBM.

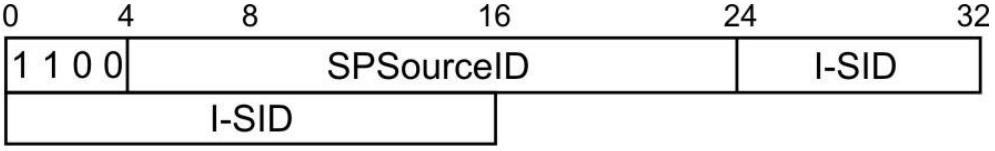


Figura 4-6 - Formato de um endereço multicast MAC DA no SPBM .

O primeiro bit à esquerda representa o *Multicast bit* (M *bit*), sempre setado para 1 em endereços *multicast* de destino.

O segundo bit à esquerda representa o *Local bit* (L *bit*), sempre setado para 1 em endereços *multicast* SPBM de destino.

Os dois bits subsequentes representam os *Type bits* (TYP *bits*). 00 é o único tipo suportado no presente momento.

O SPSrc (*SPSourceID*) é um campo com 20 bits que identifica, de forma única, um nó SPBM para todos os B-VIDs alocados para a operação SPBM. Esse campo corresponde ao *SPSourceID* distribuído no sub-TLV *SPB Instance* (*SPB-Inst*) do ISIS-SPB.

O I-SID, como dito anteriomente, é o *I-Component Service ID* distribuído no *TLV SPBM Service Identifier* (código 144).

A Figura 4-7 representa a transmissão de um quadro de dados *multicast* em uma rede SPBM, cujo endereço é constituído pelo SP*Source*ID igual a A e pelo valor de I-SID igual a K, endereço esse que pode ser representado sob a forma DA=<A,K>, que, de forma única, identifica a sub-árvore de distribuição de tráfego (S,G), dentro de uma determinada árvore de distribuição.

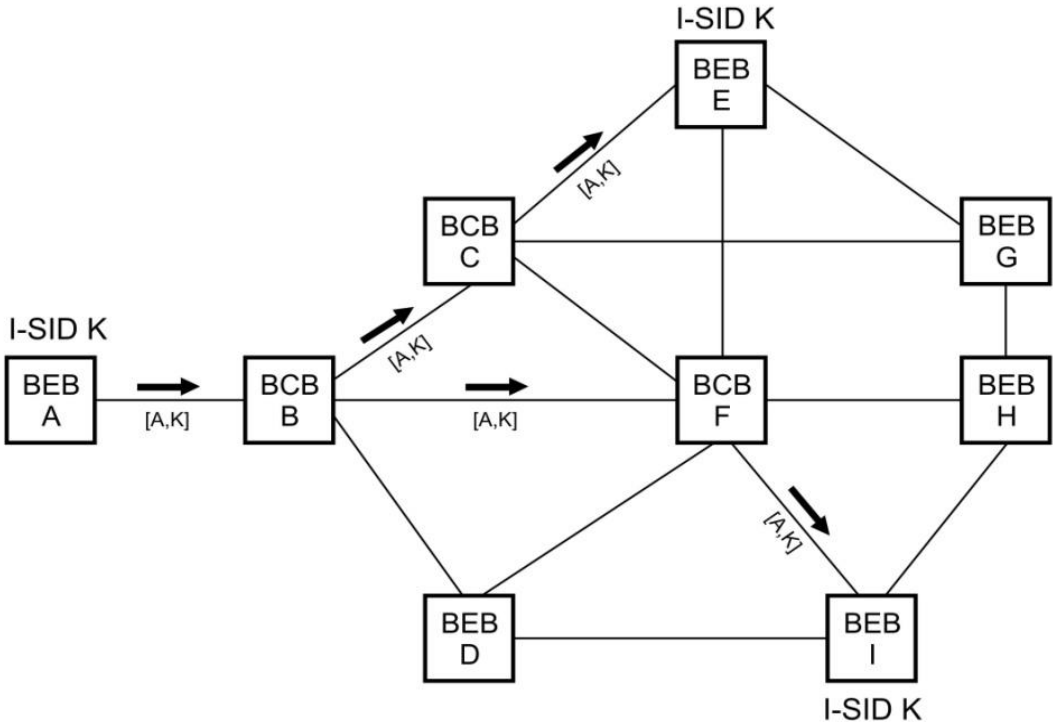


Figura 4-7 - Transmissão de quadro multicast em uma rede SPBM.

Alternativamente, o SPBM possibilita o uso de uma forma de transmissão de quadros *multicast* pela replicação dos quadros pelas BEBs de entrada, e pela transmissão *unicast* subsequente desses quadros. Essa forma é indicada para o caso onde existe pouco tráfego *multicast* para um I-SID.

4.3.7. **Distribuição de Tráfego por SPBM ECMP**

O SPB admite duas formas de distribuição da carga de tráfego por *links* com iguais e menores custos (*equal costs shortest path*) em uma Região SPT:

— Pela constituição de SPTs com iguais e menores custos (ou seja, pela constituição de ECTs);
— Por meio do uso de algoritmos ECMP ECT, onde ECMP significa *Equal Cost Multi-Path*.

A constituição de ECTs, vista nos subitens anteriores deste item, é comum ao SPBV e ao SPBM.

A distribuição de tráfego por ECMP ECTs, todavia, é utilizada exclusivamente pelo SPBM, conforme a Cláusula 44 do padrão IEEE 802.1Q-2014.

A operação SPBM ECMP provê um mecanismo para distribuição da carga de tráfego através de um conjunto de *links* com iguais e menores custos em uma Região SPT, utilizando um único VID. Essa operação se baseia em algoritmos para distribuição de tráfego por portas com iguais e menores custos.

O padrão IEEE 802.1Q-2014 contempla critérios para a distribuição de tráfego de dados por ECMP com endereços individuais e com endereços *multicast*.

- **Endereços Individuais (Unicast)**

Existem duas alternativas de critérios para o tratamento de endereços individuais por ECMP, dependendo do suporte ou não da função *Flow Filtering*.

Quando o ECMP é utilizado sem filtragem de fluxo (nesse caso, o campo ECT-ALGORITHM do Sub-TLV *SPB Base VLAN-Identifiers* possui o valor 00-80-C2-11), o *ECMP ECT Algorithm* seleciona, em cada *bridge* da Região SPT, uma porta de saída dentre o conjunto de portas de saída com custos iguais e menores. É criada então uma *Dynamic Filtering Entry* na FDB de cada *bridge*.

Quando o ECMP é prestado com filtragem de fluxo (nesse caso, o campo ECT-ALGORITHM do Sub-TLV *SPB Base VLAN-Identifiers* possui o valor 00-80-C2-12), o protocolo ISIS SPB é capaz de distribuir o tráfego por fluxos de dados, para cada porta de saída, o que garante o correto sequenciamento dos quadros de dados em cada fluxo de dados.

- **Endereços Multicast**

Os quadros de dados com endereços *multicast,* diferentemente dos quadros de dados com endereços individuais, são distribuídos pela atribuição de um modo *multicast* (*multicast mode*) e de um *tie-breaker* para cada endereço de grupo. Dessa forma, podem ser utilizadas diferentes SPTs para cada endereço de grupo ou para cada conjunto de endereços de grupo. Como visto anteriormente neste item, um endereço de grupo é identificado por um *SPSourceID* e por um valor de I-SID. Assim, cada escolha de um modo *multicast* e de um *tie-breaker* é associada ao valor de I-SID que compõe o endereço *multicast*.

A distribuição de tráfego *multicast* pelo SPBM ECMP pode se realizar também pelo uso de *shared trees* (*spanning trees*).

Tanto no caso de SPTs quanto no de *spanning trees*, podem ser utilizados até 16 diferentes *tie-breakers*, o que significa a possibilidade de até 16 caminhos alternativos para quadros de dados com endereços *multicast*.

4.4. TRANSPARENT INTERCONNECTION OF LOTS OF LINKS (TRILL)

O TRILL (*Transparent Interconnection of Lots of Links)* representa uma iniciativa do IETF na mesma linha do SPB definido pelo padrão IEEE 802.1aq, onde o protocolo IS-IS é utilizado no roteamento de *bridged Ethernet* LANs.

Com o propósito de definir os aspectos concernentes ao TRILL, o IETF emitiu um conjunto de RFCs, dentre as quais podem ser destacadas as seguintes:

- RFC 5556 - *Transparent Interconnection of Lots of Links* (TRILL): *Problem and Applicability Statement;*
- RFC 6325 – *Routing Bridges* (*RBridges*): *Base Protocol Specification*;
- RFC 6439 – *RBridges*: *Appointed Forwarders*;
- RFC 7174 – *TRILL OAM Framework;*
- RFC 7176 – *TRILL Use of IS-IS;*
- RFC 7180 – *TRILL: Clarifications, Corrections, and Updates*;
- RFC 7357 – *TRILL: End Station Address Distribution Information (ESADI) Protocol*.

Esta introdução ao TRILL tem como base a RFC 6325 e o artigo *Introduction to TRILL*, de autoria de *Radia Perlmam* e *Donald Eastlake*, publicado em *The Internet Protocol Journal*, *Volume 14, no. 3*.

Radia Perlmam é a inventora do algoritmo *Spanning Tree,* do protocolo IS-IS e do conceito original para o TRILL. Além disso, é co-autora de diversas RFCs relativas ao TRILL, incluindo a RFC 6325.

Donald Eastlake é também co-autor de diversas RFCs, desempenhando um importante papel nos trabalhos do IETF.

O TRILL é um protocolo para roteamento de quadros de forma transparente e baseada em custos menores (*least-costs*) em redes *multi-hop* com topologias e tecnologias de *link* arbitrárias, utilizando roteamento *link-state* e um *hop-count*.

O TRILL suporta a existência de VLANs, a habilidade para autoconfiguração (embora permita opcionalmente a configuração manual) e o tráfego *multicast/broadcast* em adição ao tráfego *unicast*.

A concepção TRILL define um novo tipo de *bridge*, em adição às IEEE802.1D *bridges* e as IEEE 802.1Q *bridges*, referido como *RBridges* (*Routing Bridges*).

As *RBridges* realizam roteamento IS-IS (ou seja, são ISs) e realizam o roteamento de quadros MAC de dados envelopados em um TRILL *header* contendo um *hop count*. As *RBridges* são também referidas como *TRILL switches*.

Podem ser utilizadas tecnologias e topologias arbitrárias entre *RBridges*, como, por exemplo, Ethernet, PPP e PWs.

Um *RBridge Campus* é uma rede formada por *RBridges*, por *links* e por quaisquer *bridges* de outros tipos envolvidas, rede esse delimitada por estações finais ou por roteadores de camada 3.

A RFC 6325 define a operação de um *RBridge Campus* em uma única área IS-IS, no pressuposto de que um único operador gerencia esse *RBridge Campus*.

Os quadros de dados Ethernet nativos recebidos são encapsulados na *RBridge* de ingresso pelo *TRILL header*, roteados na rede utilizando-se o IS-IS e desencapsulados na *RBridge* de egresso.

Foram definidos duas formas de encapsulamento de quadros nativos:

— Formato VL (*VLAN-Labeled*);
— Formato FGL (*Fine-Grained Label*).

Toda a concepção inicial do TRILL, e consequentemente todas as RFCs iniciais, como principalmente a RFC 6325, destinam-se ao TRILL no formato VL (ou seja, ao VL TRILL).

A formatação FGL, que objetiva ampliar a escalabidade do TRILL, foi introduzida a partir da RFC 7172 (TRILL: *Fine-Grained Label*), de maio de 2014.

Da mesma forma que no SPB, o TRILL utililiza a constituição de *multi-topology links* com base na RFC 5120. Dessa forma, o tráfego de dados pode ser distribuído por diferentes árvores selecionadas pela *RBridge* de ingresso, com base particularmente nos valores de VID contidos nos quadros de dados recebidos do usuário.

Tais árvores de distribuição são bidirecionais e alcançam a totalidade de um *RBridge Campus*. Diferentemente de IEEE 802.1 *Bridging*, as redes TRILL não impõem congruência entre caminhos *unicast* e *multicast* entre um dado par de *RBridges*. Além disso, o TRILL não impõe simetria nos caminhos para os dois sentidos de transmissão entre um dado par de *RBridges*.

Existem mecanismos para filtragem do tráfego de dados nessas árvores, baseados em VIDs e em endereços de grupo.

A RFC 7357, denominada TRILL: *End Station Address Distribution Information* (ESADI) *Protocol*, define o protocolo ESADI, cujo objetivo é proporcionar a aprendizagem de endereços MAC por meio de protocolo, de forma independente da ocorrência eventual de tráfego de dados. O SPB não dispõe de um mecanismo com essa funcionalidade.

4.4.1. **Tipos de Quadros do TRILL**

A RFC 6325 considera os seguintes tipos de quadro do TRILL:

— Quadros de controle do TRILL (ou quadros do IS-IS TRILL): são quadros utilizados para controle entre *RBridges;*
— Quadros de dados do TRILL: são quadros utilizados para o encapsulamento de quadros de dados do usuário;
— Outros quadros do TRILL (*TRILL Other Frames*): são quadros com endereços *multicast* reservados para o TRILL que não sejam os endereços *All-RBridges* ou *All-IS-IS-RBridges*, que devem ser descartados;
— Quadros de controle de camada 2: são quadros de controle que não são utilizados por *RBridges*, por dizerem respeito à rede de usuário;
— Quadros nativos: são quadros que não se enquadram nos tipos anteriores;
— Quadros do protocolo ESADI (*End Station Address Distribution Information*): são os quadros utilizados no protocolo ESADI, que serão vistos adiante neste item.

Para quadros de controle do TRILL (ou quadros do IS-IS TRILL) foi atribuído o valor *L2-IS-IS Ethertype* igual a 0x22F4. Para o protocolo *Hello* do TRILL L2-IS-IS foi atribuído o valor de NLPID (*Network Layer Protocol Identifier*) igual a 0xC0.

Para quadros de dados do TRILL foi atribuído o valor *TRILL Ethertype* igual a 0x22F3.

Os endereços de grupo atribuídos ao TRILL encontram-se na faixa de 01-80-C2-00-00-40 a 01-80-C2-00-00-4F.

O IETF apresenta, em diferentes RFCs, a formatação dos seguintes quadros do TRILL:

— Quadros de dados do VL TRILL;
— Quadros de dados do FGL TRILL;
— Quadros do protocolo ESADI;
— Quadros do *RBridge Channel*;
— Quadros de OAM do TRILL;
— Quadros de controle do TRILL.

Os formatos de cada um desses quadros serão apresentados ao longo dos subitens subseqüentes deste item.

4.4.2. **Encapsulamento de Quadros de Dados do VL TRILL**

Os quadros de dados do TRILL são estruturados conforme a Figura 4-8.

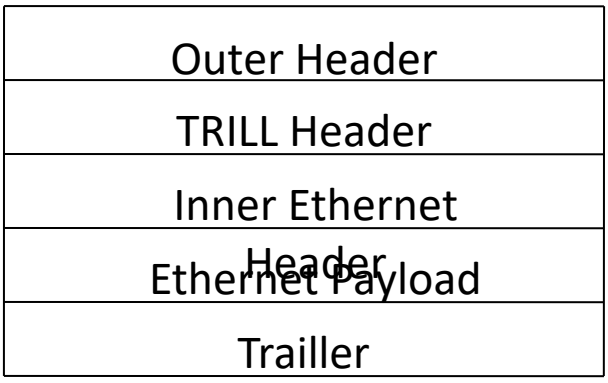


Figura 4-8 - Estrutura de quadros de dados do TRILL.

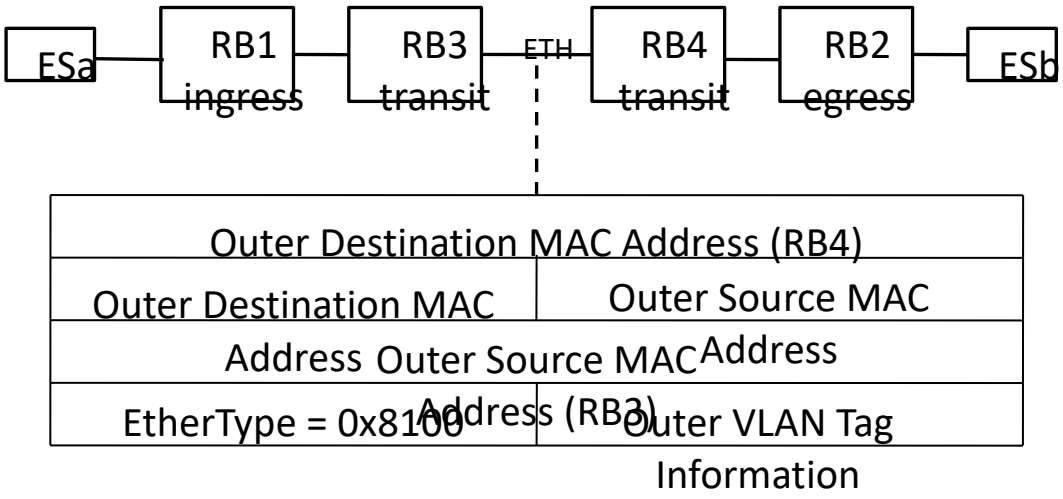
4.4.2.1. Outer Header (Cabeçalho Exterior)

O cabeçalho exterior que envelopa o quadro TRILL corresponde à rede que interliga *RBridges*, sendo formatado conforme a tecnologia dessa rede. Os casos considerados na RFC 6325 são o encapsulamento do quadro TRILL por Ethernet e o encapsulamento pelo PPP.

O envelopamento por Ethernet, que é a forma especificada com maiores detalhes na RFC 6325, ocorre tipicamente quando se verifica a substituição parcial de *bridges* por *RBridges*, restando assim nuvens Ethernet entre *RBridges*.

O endereço MAC da porta de uma *RBridge* que se conecta a rede Ethernet externa (que pode ser inclusive um *link* Ethernet ponto a ponto), é referido como o endereço MAC do SNPA (*Sub-Network Point of Attachment*) dessa porta da *Rbridge*.

A Figura 4-9 ilustra a formatação do cabeçalho exterior para o caso de envelopamento do quadro TRILL por Ethernet. Nesse caso, o cabeçalho exterior é referido como *Outer Ethernet Header*.



F

F

Figura 4-9 – Formato do Outer Ethernet Header.

Observa-se que, nessa figura, existe uma nuvem Ethernet entre as RB3 e RB4.

O *Outer Ethernet Header* é destinado à condução de quadros do TRILL exclusivamente no interior da nuvem Ethernet, tendo como endereço de destino o endereço MAC de uma das *RBridges* RB3 ou RB4 (RB4 no exemplo da figura), e como endereço de origem o endereço MAC da outra *RBridge* (RB3 no exemplo da figura).

O *Outer VLAN tag* tem significado exclusivamente no interior da nuvem Ethernet.

Serão apresentados, adiante, neste capítulo, outros exemplos já especificados pelo IETF de uso de tecnologias para constituir o *outer header*.

4.4.2.2. TRILL Header

A Figura 4-10 apresenta o formato do TRILL *header*.

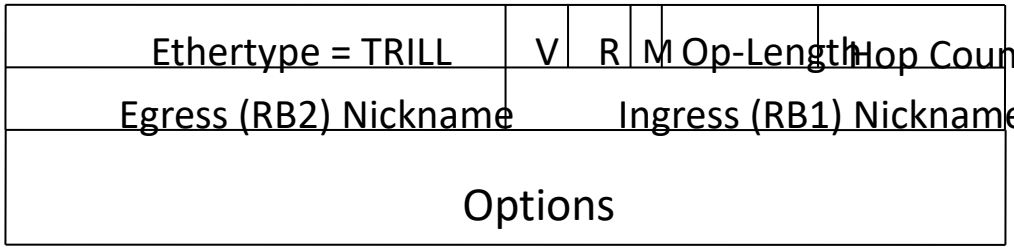


Figura 4-10 – Formato do TRILL header.

- **Campos Iniciais**

O valor *Ethertype*, que indica um quadro de dados do TRILL, é igual a 0x22F3 como visto anteriormente.

O campo *Version* (V), com dois bits, indica a versão do TRILL, sendo igual a zero no presente momento. O campo *Reserved* (R), também com dois bits, encontra-se reservado, sendo seu valor atual igual a 00.

O *multi-destination bit* (M), indica a forma de transmissão do seguinte modo:

— M=0 (*FALSE*): significa que o quadro de dados possui endereço *unicast* conhecido e que o *nickname* de egresso possui um valor que especifica o ramo da árvore de distribuição que conduz ao egresso do quadro da rede TRILL. Esse valor de *nickname* é aprendido, ou configurado, juntamente com o endereço MAC de destino para a VLAN indicada no quadro;

— M=1 (*TRUE*): significa a transmissão multidestino de um quadro de dados por um árvore de distribuição. A árvore de distribuição é identificada pelo valor de *nickname* de ingresso definido pela *RBridge* de ingresso para a VLAN correspondente ao quadro de dados. O valor do *nickname* de egresso não tem significado, sendo normalmente igualado ao *nickname* de ingresso.

- **Campos OP-Length e Campo Hop Count**

O campo *OP-Length*, com cinco bits, indica o comprimento das opções (extensões) do *TRILL Header*, medido em número de palavras de quatro octetos (32 bits). Dessa forma, o número máximo de palavras é igual a 31, que representa um máximo de 124 octetos.

Se o valor desse campo for zero, não existem extensões presentes no *TRILL Header*. Se esse valor for positivo, a primeira palavra, referida como *Extended Header Flags*, encontra-se presente, como será visto na apresentação do campo *Options*.

O campo *Hop Count* representa um valor inteiro com 6 bits.

Uma *RBridge* que receba um quadro com um valor de *hop count* positivo, deve decrementá-lo de uma unidade e reenviá-lo. Quando o *hop count* chegar ao valor zero, o quadro será descartado.

Quadros multidestino estão sujeitos ao rigoroso mecanismo *Reverse Path Forwarding* (RPF), em virtude da possibilidade de que um *loop* envolvendo uma ou mais bifurcações de árvores de distribuição possa resultar em rápida geração de múltiplas cópias do quadro mesmo com o funcionamento do mecanismo *hop count.*

A RFC 6325 define adicionalmente mais três alternativas de mecanismos de controle para quadros multidestino, que podem causar o descarte desses quadros.

- **Campos Nickname**

Os campos *nickname* possuem valores com 16 bits associados a *RBridges* e a árvores de distribuição no TRILL. Esse tamanho de *nickname* possibilita que um *TRILL Campus* possa suportar até 64.000 *RBridges*.

O valor 0x0000 é reservado para indicar a não especificação de *nickname*. Os valores na faixa de 0xFFC0 a 0xFFFE estão reservados para uso futuro, enquanto o valor 0xFFFF encontra-se permanentemente reservado.

Cada *RBridge* escolhe aleatoriamente um ou mais valores de *nickname*, evitando valores já adquiridos por outras *RBridges*. A coincidência de valores é descoberta pelo exame das bases de dados de LSP.

Em caso de coincidência de valores, existe um critério *tie-breaker* para definir a atribuição do valor coincidente entre as *RBridges* envolvidas. Esse critério baseia-se nos valores de prioridade e nos SYSIDs configurados, ou seja, nos *Bridge IDs* dessas *RBridges*.

É possível a configuração de valores de *nickname* para *RBridges*. Nesse caso, esses valores têm prioridade sobre os valores escolhidos aleatoriamente.

Um valor de *nickname* de ingresso é inserido no TRILL *Header* de um quadro do TRILL na *RBridge* de ingresso.

Existem duas hipóteses para o valor de *nickname* de egresso.

Para um quadro de dados do TRILL com endereço *unicast* de destino conhecido, o bit M é setado para zero (M=0), e o *nickname* de egresso é o *nickname* da *RBridge* de egresso aprendido para a VLAN indicada no quadro de dados. Esse *nickname* é aprendido de forma associada com o endereço MAC de um quadro de dados nessa VLAN.

Para quadros de dados do TRILL multidestino (com endereço *unicast* não aprendido, com endereço *multicast* ou com endereço *broadcast*), o bit M é setado para 1 (M=1), e o valor do *nickname* de ingresso é utilizado para identificar a árvore de distribuição utilizada. A *RBridge* de origem insere no quadro um valor de

nickname de egresso que não tem significado para a rede TRILL, e que é normalmente igualado ao valor do *nickname* de ingresso.

A atribuição de múltiplos valores de *nickname* para uma *RBridge* objetiva possibilitar a identificação de múltiplas árvores de distribuição com *links* de menores custos centradas nessa *RBridge.*

A multiplicidade de árvores almeja a melhor distribuição da carga de tráfego no *TRILL Campus*. As diferentes VLANs da rede Ethernet envelopada pelo TRILL têm os respectivos tráfegos atribuídos a diferentes árvores de distribuição, o que proporciona o balanceamento da carga de tráfego.

- **Campo Options**

O campo *Options* é constituído por até 31 palavras de quatro octetos, totalizando assim o máximo de 124 octetos, cujo objetivo é possibilitar adições de extensões ao *TRILL Header*. Como visto anteriormente, o número de palavras utilizadas com extensões é expresso pelo valor do campo *OP-Length*.

Como a RFC 6325 define o campo *Options* em linhas gerais, foi emitida a RFC 7179 (*TRILL: Header Extension*), que especifica esse campo com maiores detalhes.

Embora a RFC 7179 defina a estrutura da primeira palavra de quatro octectos do campo *Options* (palavra essa referida como *Extended Header Flags*), essa RFC não apresenta exemplos de valores para os *flags* definidos.

Além da limitação do parágrafo anterior, a RFC 7179 não apresenta outros tipos de extensão para o *TRILL Header*. Para suprir essa lacuna, o IETF encontra-se em fase de definição de novas RFCs com esse propósito, como atestam a existência do *draft RBridge: Further TRILL Header Extensions* e do *draft RBridges: More Proposed TRILL Header Options*.

A Figura 4-11 apresenta o formato da palavra *Extended Header Flags*.

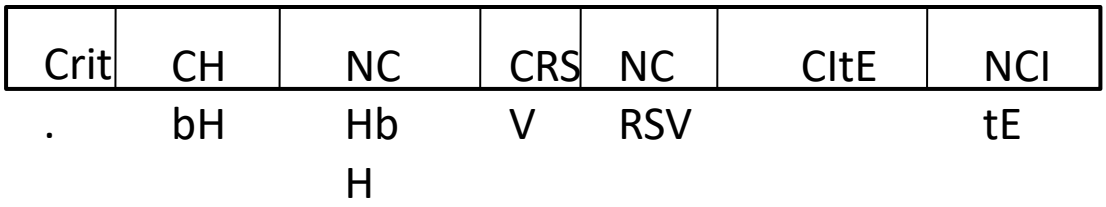


Figura 4-11 – Formato da palavra Extended Header Flags.

Os três primeiros bits (*Crit.*) representam o *Critical Summary*, cujo significado será visto adiante.

Os demais campos têm a seguinte semântica:

— CHbH: *Critical Hop-by-Hop extended flag bits* (bits de 3-7);
— NCHbH: *Non-critical Hop-by-Hop extended flag bits* (bits de 8-13);
— CRSV: *Critical Reserved extended flag bits* (bits de 14-16);
— NCRSV: *Non-critical Reserved extended flag bits* (bits de 17-20);

— CItE: *Critical Ingress-to-Egress extended flag bits* (bits 21-26);
— NCItE: *Non-critical Ingress-to-Egress extended flag bits* (bits de 27-31).

Os três bits do *Critical Summary* indicam a presença de extensões críticas da seguinte forma:

— Bit CHbHS: *Critical Hop-by-Hop extension(s):* indica, com o valor igual a 1, a presença de uma ou mais extensões críticas *hop-by-hop* no campo *Options* (indica até mesmo a presença do CHbH *flag*);
— Bit CItES: *Critical Ingress-to-Egress extension(s):* indica, com o valor igual a 1, a presença de uma ou mais extensões críticas *Ingress-to-Egress* no campo *Options* (indica até mesmo a presença do CItE *flag*);
— Bit CRSVS: *Critical Reserved extension(s)*: indica, com o valor igual a 1, a presença de uma ou mais extensões críticas reservadas para a aplicação de uma classe de *RBridges* no futuro.

4.4.2.3. Inner Ethernet Header

O *Inner Etherner Header* é o cabeçalho do quadro *MAC Ethernet* nativo envelopado em um quadro de dados do TRILL.

4.4.3. Fine-Grained Labeling (FGL) TRILL

Como visto no item anterior, no VL TRILL um quadro de dados MAC nativo, com seu *VLAN tag*, é envelopado em um quadro de dados do *TRILL* sem o acréscimo de qualquer tipo de *label* (*VLAN tag*). Dessa forma, os únicos *labels* utilizados são os *VLAN tags* do *Inner Ethernet Header* (que representa o cabeçalho do quadro MAC nativo). Como se sabe, esses *VLAN tags* suportam até 4096 (2^{12}) VLAN IDs, ou 4K VLAN IDs de forma arredondada.

Essa escala de VLANs IDs, que possibilita um número máximo de 4K VLANs atendidas pela rede VL TRILL, é julgada insuficiente para determinadas aplicações, particularmente aquelas destinadas ao uso em grandes centros de dados, que requerem o atendimento a um número elevado de redes e de VLANs.

Assim, a RFC 7172 (TRILL: *Fine-Grained Labeling*) desenvolveu o conceito de FGL, que possibilita a utilização de um *data label* de 24 bits, sendo esse *data label* referido como *fine-grained label* (FGL).

Um *fine-grained label* (FGL) resulta da inclusão de um Segundo VLAN *tag* do quadro de dados do TRILL, inclusão essa ocorrida no ingresso de um quadro de dados MAC nativo na rede TRILL.

Esse segundo VLAN *tag* é referido como *Inner Ethernet High Part*. A *Inner Ethernet Low Part* representa o VLAN *tag* refletido do quadro de dados MAC nativo que ingressa na rede TRILL, e que é restaurado no egresso.

Em cada VLAN ID da *Inner Ethernet High Part* são mapeados grupos de até 4K VLAN IDs da *Inner Ethernet Low Part*, do que resultam os FGLs. Dessa forma, torna-se possível o suporte de até 2^{24} FGLs (cerca de 16 milhões de FGLs).

O FGLs são representados na forma (X,Y), onde X é o VLAN ID superior (*High*) e Y o VLAN ID inferior.

4.4.3.1. Formato do Quadro de Dados do FGL TRILL

A Figura 4-12 representa o formato de um quadro de dados do FGL TRILL.

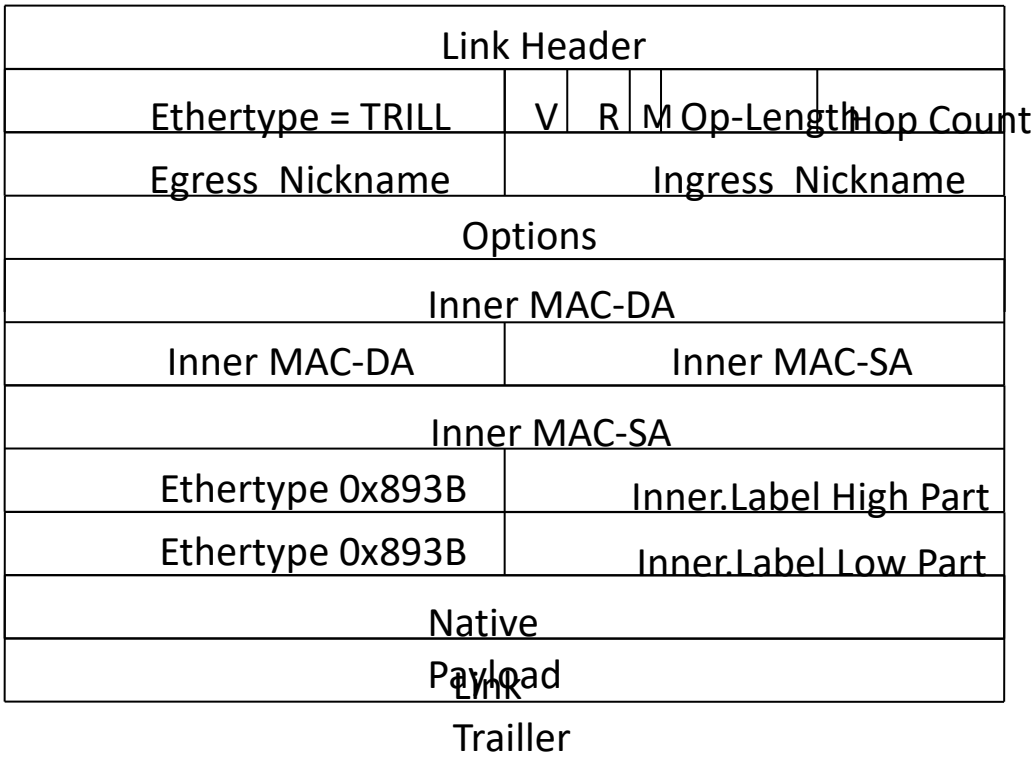


Figura 4-12 – Quadro de dados do FGL TRILL.

A *Inner Ethernet Low Part*, constituída pelo valor *Ethertype* 0x893B e pela *Inner.Label Low Part*, representa o *VLAN tag* do quadro Ethernet nativo, ressalvada a alteração do valor *Ethertype*, agora igual a 0x893B.

A *Inner Ethernet High Part*, constituída pelo valor *Ethertype* 0x893B e pela *Inner.Label High Part*, por sua vez, representa o *VLAN tag* adicional relativo ao FGL TRILL, em cujas instâncias individuais são mapeados os VLAN IDs da *Inner.Label Low Part*. A *Inner Ethernet High Part* representa para o FGL TRILL, em termos de escala, o mesmo que o *S-tag* representa para o PB.

A *Inner.Label High Part e a Inner.Label Low Part*, com 16 bits, são estruturadas na forma utilizada pelo campo TCI (*Tag Control Information*) dos *S-tags* (*Priority* com 3 bits, DEI com 1 bit e VLAN ID com 12 bits).

O valor do campo de prioridade da *Inner.Label High Part* indica a prioridade utilizada para o transporte do respectivo quadro de dados do FGL TRILL através do *TRILL Campus*, do ingresso até o egresso. O seu valor pode ser copiado da *Inner.Label Low Part*. A prioridade da *Inner.Label Low Part*, por sua vez, é o valor de prioridade copiado do quadro ethernet nativo no ingresso e restaurado no egresso.

4.4.3.2. Aspectos Operacionais do FGL TRILL

Não é possível que uma *VL RBridge* (ou seja, um *VL TRILL switch*) em sua configuração normal possa trafegar quadros de dados do FGL com segurança, mesmo que essa *VL RBridge* esteja apenas operando com capacidade de trânsito. Para que uma *VL RBridge* possa transitar tráfego FGL, é necessário que essa *VL RBridge* seja configurada como um *FGL-safe switch* e que notifique essa nova configuração em seus LSPs.

Se um quadro Ethernet nativo com um endereço MAC de destino *broadcast*, *multicast* ou unicast desconhecido ingressa em uma rede FGL TRILL, ele pode ser conduzido de duas formas alternativas. A escolha da alternativa pode variar para cada quadro de dados, a critério da *FGL RBridge* de ingresso.

Na primeira alternativa, o quadro é transmitido como um quadro de dados FGL multidestino (*TRILL Header M bit = 1*) em uma árvore de distribuição centrada em uma *RBridge* raiz (provavelmente a FGL *RBridge* de ingresso), com um *nickname* atribuído pela *FGL RBridge* de ingresso ou pelo pseudonó (*pseudonode*) de ingresso correspondente a um *link* FGL de acesso compartilhado.

Um quadro de dados FGL TRILL não pode ingressar com um *nickname* atribuído por uma *VL RBridge* ou por um *pseudonode* correspondente a um *link* compartilhado VL. O conceito de *pseudonode* será apresentado mais adiante neste item.

Na segunda alternativa, o quadro multidestino é enviado de forma serialmente direcionada para cada uma das *FGL RBridges* de egresso relevantes, utilizando cada um dos respectivos *nicknames* atribuídos pela *RBridge* de ingresso (com o *TRILL Header M bit = 0* em cada uma das transmissões).

De início, as FGL *RBridges* relevantes são aquelas que anunciaram interesse no *label* (x,y) do quadro de dados a transmitir, ou seja, interesse no *VLAN ID High Part* e no *VLAN ID Low Part* desse quadro. Para quadros de dados com endereço de destino *broadcast* ou *unicast* desconhecido, essa filtragem é suficiente.

Para quadros de dados com endereço de destino *multicast*, a definição das FGL *RBridges* relevantes deveria passar por uma filtragem adicional baseada em anúncios de endereços de grupo realizados por meio de LSPs.

4.4.4. Outras Alternativas de Cabeçalho Externo

Até o momento, o IETF especificou a constituição de cabeçalhos externos (*outer headers*) com as seguintes tecnologias:

— Ethernet, conforme a RFC 6325;
— PPP, conforme a RFC 6361 (*PPP TRILL Protocol Control Protocol*);
— PWs (*Pseusdowires*), conforme a RFC 7173 (*TRILL Transport Using PWs*).

O encapsulamento de quadros TRILL por Ethernet descrito na RFC 6325 foi já abordado anteriormente neste item.

4.4.4.1. Conexão de RBridges via PPP (RFC 6361)

A RFC 6361 descreve o uso do PPP (*Point-to-Point Protocol*) para o transporte de quadros TRILL entre *RBridges*, ou seja, via *links* PPP *inter-RBridges.*

A Figura 4-13 apresenta a estrutura de um quadro do TRILL encapsulado pelo PPP.

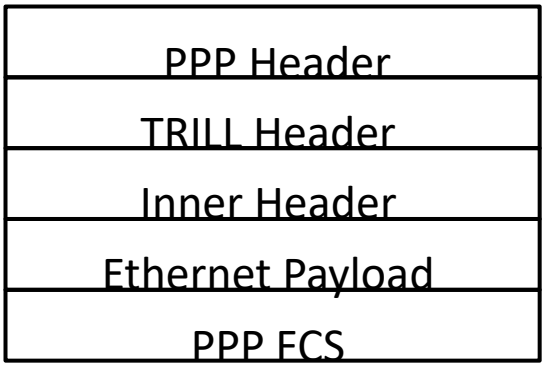


Figura 4-13 – Estrutura de um quadro do TRILL encapsulado pelo PPP.

O PPP define o LCP (*Link Control Protocol*) e um método para negociar o uso de tráfego multiprotocolo sobre *links* ponto a ponto.

O PPP utiliza também o protocolo BCP (*Bridging Control Protocol*), que permite o uso de *links* PPP para a interconexão de Ethernet *bridges*. Esse mecanismo é inadequado e ineficiente para o TRILL, razão pela qual foi emitida a RFC 6361.

Para o TRILL, foram definidos três novos números de protocolo PPP:

— 0x805D para o TNCP (*TRILL Network Control Protocol*);
— 0x005D para o TNP (*TRILL Network Protocol*);
— 0x405D para o TLSP (*TRILL Link State Protocol*).

O protocolo TNCP, que representa a adequação do LCP e do BCP para o TRILL, é responsável por negociar o uso do TNP e do TLSP em um *link* PPP.

Os quadros de dados do TRILL, em seu formato convencional, são envelopados pelo PPP com o código 0x005D.

O TNP destina-se ao transporte de quadros de dados do TRILL sobre *links* PPP. O TLSP, por sua vez, representa o protocolo *IS-IS for TRILL* com as respectivas PDUs envelopadas pelo PPP com o código 0x405D.

4.4.4.2. Conexão de RBridges via PWs

A RFC 7173 descreve a utilização de PWs para o transporte de quadros de dados do TRILL entre *RBridges*. Tais *pseudowires* são constituídos conforme a arquitetura PWE3 (*Pseudowire Emulation Edge-to-Edge*). Essa RFC aborda, na realidade, o uso de MPLS PWs com esse propósito.

Para a interconexão de *RBridge*s por *pseudowires* são utilizados PPP *pseudowires* (definidos no RFC 4618). Tais *pseudowires* podem ser sinalizados (RFC 4447) ou podem ser configurados manualmente.

Um *pseudowire* pode ser transportado sobre diversos tipos de túneis constituídos por *packet switching*, como, por exemplo, um túnel MPLS ou um túnel MPLS-TP.

A porta de PW da *RBridge* que envia o quadro de dados deveria mapear o valor de prioridade do quadro Ethernet interior (*inner Ethernet*) no campo *Traffic Class* do PW (RFC 5462).

A Figura 4-14 representa o envelopamento de um quadro de dados do TRILL por um PPP PW constituído sobre uma rede MPLS.

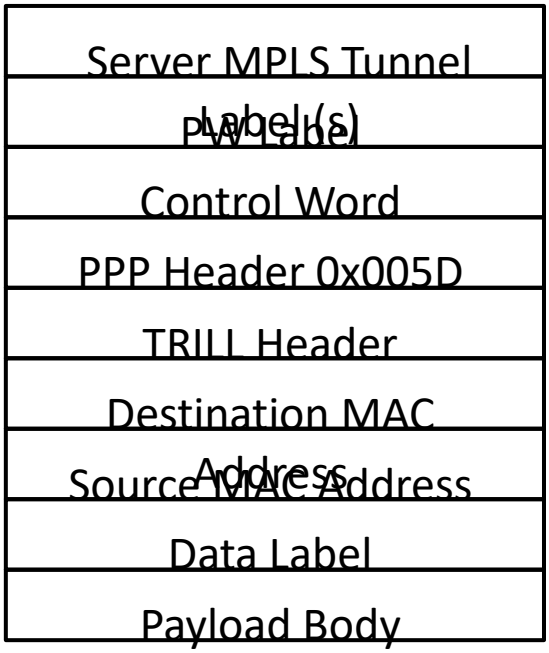


Figura 4-14 – Encapsulamento do TRILL por um MPLS PPP PW.

A Figura 4-15, por sua vez, apresenta um quadro do TRILL IS-IS envelopado por um PPP PW constituído sobre uma rede MPLS.

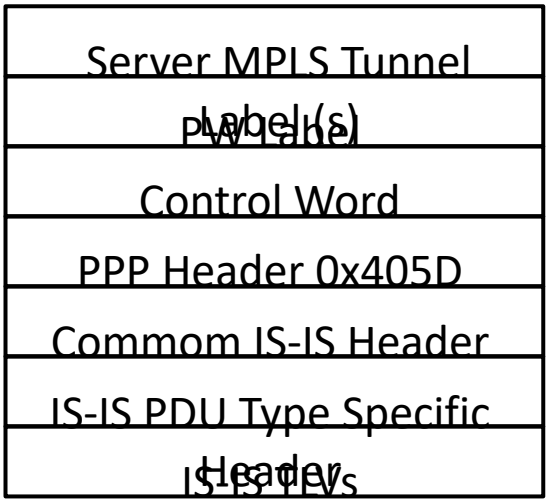


Figura 4-15 – Encapsulamento de quadros do TRILL IS-IS por um MPLS PPP PW.

Como se observa nessas figuras, são utilizados os números de protocolo PPP definidos na RFC 6361, que são o valor 0x005D para quadros de dados do TRILL e o valor 0x405D para quadros do TRILL IS-IS.

4.4.5. **Aprendizagem da Localização de Nós Finais**

Conforme a RFC 6325, existem cinco alternativas para a aprendizagem da localização de nós finais, que, a rigor, se resumem em dois mecanismos.

O mecanismo *default* consiste na aprendizagem de endereços *unicast* baseada no tráfego de origem nos moldes do que ocorre em *LAN bridging* tradicional. Uma *RBridge* de egresso registra ou atualiza a associação entre o valor do *nickname* de ingresso e o endereço MAC de origem de um quadro de dados do TRILL por ela recebido em uma determinada VLAN.

Essa associação é utilizada na transmissão de quadros de dados, no sentido inverso, para o endereço MAC *unicast* conhecido (aprendido).

Caso não tenha ocorrido a aprendizagem prévia do endereço MAC *unicast* de destino, assim como na hipótese de um quadros de dados com endereço MAC multidestino, a *RBridge* de ingresso envia esse quadro MAC por uma árvore de distribuição.

Um mecanismo opcional alternativo, denominado ESADI, foi definido na RFC 7357 denominada *TRILL: End Station Address Distribution Information* (ESADI) *Protocol*.

No ESADI, cada *RBridge* que tenha nós finais a ela conectados anuncia alguns (ou todos) desses nós finais com os respectivos endereços MAC. O resultado é que as demais *RBridge*s passam a possuir tabelas de associação entre *nicknames* e endereços MAC, como no mecanismo *default*, sem dependerem contudo da

aleatoriedade do tráfego de dados. O ESADI evita inundações em árvores de distribuição.

Em qualquer dos mecanismos, é possível a constituição de diretórios que associam também endereços MAC a endereços IP dos nós finais, de forma associada à aprendizagem de endereços MAC. Dessa forma, tornam-se desnecessários os dispendiosos processos do *IPv4 Address Resolution Protocol* (ARP) e do *IPv6 Neighbor Discovery* (ND).

4.4.6. **Protocolo ESADI**

O protocolo ESADI foi especificado pela RFC 7357 (*TRILL: End Station Address Distribution Information*).

A RFC 6325 define cinco formas pelas quais uma *RBridge* pode aprender os endereços de estações finais, conforme menção anterior neste item. O protocolo ESADI é uma dessas formas.

O ESADI é um protocolo opcional para a aprendizagem de endereços MAC, pelo qual as *RBridges* podem comunicar informações tais como endereços MAC de estações finais associados com o *nickname* da *RBridge* na qual as estações finais com esses endereços estão conectadas, dentro do escopo de um *data label* (VLAN ID ou FGL) ou de um grupo de *data labels*.

A *RBridge* à qual se conectam estações finais pode anunciar, pelo ESADI, a totalidade ou parte dos endereços MAC dessas estações finais, nos termos do parágrafo anterior.

Utilizando-se o protocolo ESADI como mecanismo de aprendizagem de endereços MAC, e não pela aprendizagem com base no tráfego de dados (ou seja, no plano de dados), inundações da rede podem ser evitadas.

Além disso, o ESADI pode ser utilizado para anunciar estações finais de forma autenticada, utilizando, por exemplo, o EAP (*Extensible Authentication Protocol*), que se baseia em criptografia. Obtém-se, dessa forma, maior segurança para as informações registradas nas tabelas de associação do ESADI.

A Figura 4-16 apresenta a estrutura de quadro ESADI.

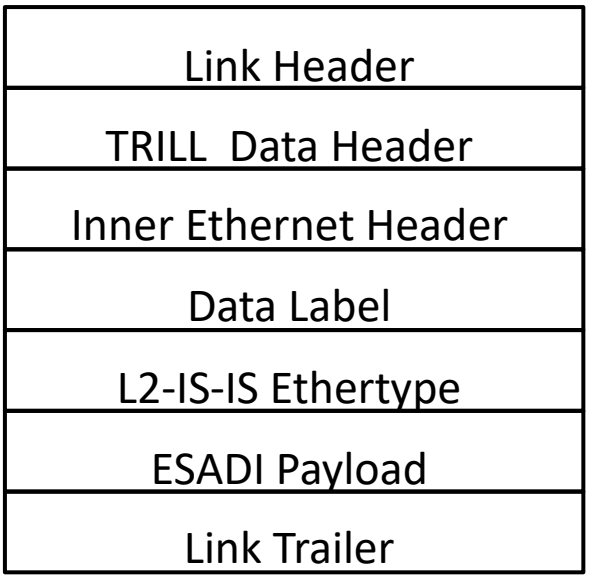


Figura 4-16 – Estrutura de um quadro ESADI.

Nessa figura, caso seja utilizada *Ethernet* como protocolo de *link*, o *Next Hop Destination Address* do *Link Header* (ou seja, do *Outer Header*) é o endereço *multicast All-RBridges*. Caso o quadro MAC do *link* seja enviado como *unicast*, esse endereço passa a ser o endereço *unicast* da *Next-Hop RBridge.*

O MAC-DA do *Inner Ethernet Header* é o endereço *All-Egress-RBridges.*

O ESADI *payload* é formatado em conformidade com o *IS-IS for TRILL*, razão pela qual é precedido pelo *Ethertype* correspondente a esse protocolo (*Ethertype* 0x22F4).

Assim, o ESADI *payload* é constituído pelos seguintes campos:

— *IS-IS Common Header;*
— *IS-IS PDU Specific Fields;*
— *IS-IS TLVs.*

O ESADI *payload* é estruturado como *IS-IS Extended Level 1 Circuit Scope* (E-L1CS), nas opções LSP, CSNP (*Complete Sequence Number PDU*) e PSNP (*Partial Sequence Number PDU*), conforme a RFC 7356 (*IS-IS Flooding Scope Link State PDUs*), com exceções especificadas na RFC 7357.

As ESADI PDUs são sempre encapsuladas como se fossem quadros de dados do TRILL, sendo tunelados transparentemente através de *RBridges* de trânsito.

A informação primária em ESADI PDUs consiste em *MAC-Rechability* (MAC-RI) TLVs, especificados na RFC 6165 (*Extensions to IS-IS for Layer-2 Systems*).

Os MAC-RI-TLVs contêm um ou mais endereços MAC *unicast* de estações finais conectadas a uma porta de *RBridge* e associadas a uma VLAN, endereços esses para os quais essa *RBridge* é o ingresso (possivelmente como um *Appointed Forwarder*).

A cada MAC-RI transmitido corresponde um valor de Nível de Confiança (*Confidence*), com um octeto, na faixa de zero a 254.

No mecanismo de Nível de Confiança, compete ao gerente de um *RBridge Campus* definir que certas fontes de aprendizagem de endereços possam prevalecer sobre outras. Informações de aprendizagem autenticadas poderiam ser, de modo razoável, anunciadas com um elevado nível de confiança.

4.4.7. **Constituição de Múltiplas Árvores de Distribuição no TRILL**

As redes TRILL não exigem a congruência entre caminhos *unicast* e caminhos *multicast* entre um dado par de *RBridges*.

Uma *RBridge* pela qual ingressa/egressa tráfego nativo calcula caminhos *unicast* direcionados com menor custo nela centrados. Basicamente, essa *RBridge* calcula uma árvore constituída por esses caminhos, sendo cada um dos caminhos constituído de forma independente.

Para tráfego multidestino (quadros de dados com endereços MAC *unicast* desconhecidos com endereços MAC *multicast* ou com endereços MAC *broadcast*), são constituídas árvores de distribuição de tráfego centradas em uma ou mais *RBridges* raiz, que têm como folhas (*leafs*) todas as demais *RBridges* pelas quais egressam quadro de dados.

Cada uma dessas *RBridges* raiz adquire diferentes valores de *nickname* e a cada um desses valores de *nickname* é associado um subconjuntos de *data labels*. Pelo protocolo *IS-IS for TRILL* são constituídas árvores de distribuição associadas a cada um desses valores de *nickname*.

Tais árvores de distribuição são bidirecionais e constituídas por ramos com menor custo.

Todas as *RBridges* em um *TRILL Campus* tomam conhecimento das *RBridges* raiz, ou da *RBridge* raiz, de todas as árvores de distribuição contituídas. Ademais, todas essas *RBridges* computam extamente as mesmas árvores que a(s) *RBridge*(*s*) raiz.

Quando uma *RBridge* de ingresso envia um quadro de dados do TRILL multidestino (M=1), esse quadro é enviado por todas as portas das *RBridges*, antes ou a partir da *RBridge* raiz, portas essas que estejam na árvore de distribuição associada ao *data label* do quadro. O envio do quadro no sentido da *RBridge* raiz é possibilitado pela bidirecionalidade da árvore de distribuição.

O *nickname* de egresso desse quadro de dados e o *nickname* da *RBridge* raiz, que é de conhecimento de todas as *RBridges* do *TRILL Campus.*

A Figura 4-17 representa a constituição de duas árvores de distribuição em uma rede TRILL, centradas em uma *RBridge* raiz.

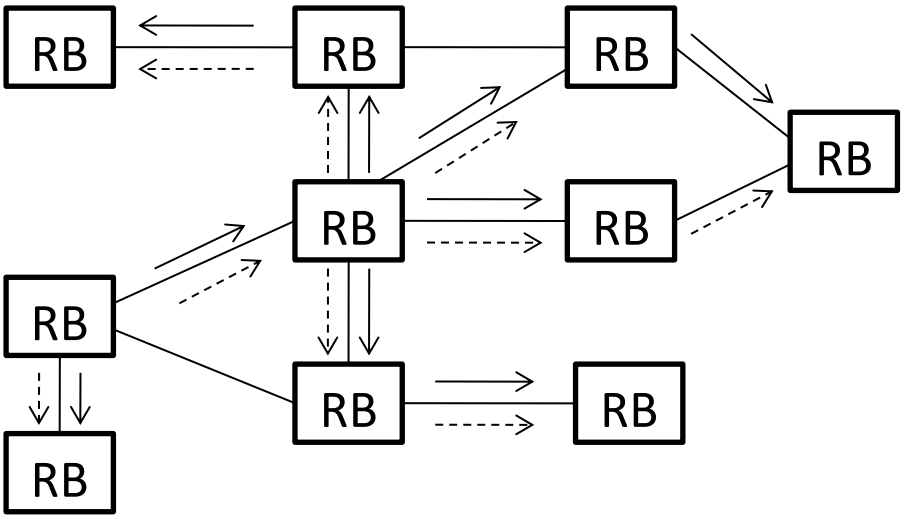


Figura 4-17 – Constituição de duas árvores de distribuição.

Observa-se que nessa figura não se encontram representadas as eventuais filtragens (*pruning*) ocorridas com base nos *data labels* dos quadros de dados.

As duas árvores de distribuição dessa figura são coincidentes, exceto nos *links* de acesso à RB5. Uma das árvores utiliza o *link* conectado à RB3, enquanto a outra árvore utiliza o *link* conectado à RB4.

Embora uma única árvore de distribuição seja logicamente suficiente para um *TRILL Campus* inteiro, a computação de árvores de distribuição adicionais justifica-se por duas razões.

Em primeiro lugar, essa computação possibilita a constituição de múltiplas árvores de distribuição com menores custos (*multipath shortest path distribution trees*) para tráfego multidestino, o que enseja melhor distribuição de tráfego pelos *links* da rede TRILL.

A segunda razão é que essa forma de computação possibilita a escolha, para uma *RBridge* de ingresso de tráfego, de uma *bridge* raiz mais próxima dessa *RBridge* de ingresso. Isso porque é possível a constituição de uma ou de múltiplas *bridges* raiz e, no limite, de um número de *RBridges* raiz igual ao número de *RBridges* de ingresso.

Na medida em que aumenta o número de *RBridges* raiz, aumenta também a possibilidade de escolha de uma *RBridge* raiz mais próxima de cada uma das *RBridges* de ingresso.

Há um valor de compromisso no número de *RBridges* raiz, que deve ser definido para cada rede TRILL. A escolha desse número deve levar em consideração, de um lado, a eficiência na transmissão de quadros e, do outro lado, o total de custos computacionais.

A eficiência máxima é obtida quando o número de *RBridges* raiz é igual ao número de *RBridges* de ingresso. Acresce-se o fato de que a probabilidade de ocorrência de transmissões fora de ordem (*out-of-order delivery*) é praticamente

eliminada na hipótese de coincidência total entre *RBridges* de ingresso e *RBridges* raiz.

O aumento do número de *RBridges* raiz aumenta também a proporção do número de *links* da rede TRILL utilizados. Os *chips* de silício para roteadores mais populares e mais efetivos em termos de custo, são capazes, pelo menos até o momento, de processar até 16 árvores de distribuição em uma *RBridge* raiz. Assim, quanto maior o número de *RBridges* raiz, maior será a proporção de *links* de menor custo utilizados em uma rede TRILL.

É possível que em uma rede TRILL de menor dimensão uma única *bridge* raiz, ou um número reduzido de *bridges* raiz, seja suficiente para utilização total dos *links* dessa rede.

Por outro lado, quando o número de *bridges* raiz aumenta, o custo computacional para a constituição de árvore de distribuição aumenta na mesma proporção, o que praticamente inviabiliza a solução que maximiza eficiência e elimina a ocorrência de transmissão de quadro fora de ordem, solução essa que, como visto anteriormente, consiste em igualar o número de *RBridges* de ingresso e de *RBbridges* raiz.

No TRILL, a *RBridge* com a maior prioridade para ser a *RBridge* raiz anuncia para as demais *RBridges* (por meio de LSPs) quantas e quais árvores de distribuição devem ser calculadas.

As árvores de distribuição centradas em uma *RBridge* raiz são calculadas com um *tie-breaker* determinístico, para que todas as *RBridges* da rede TRILL calculem as mesmas árvores de distribuição.

Uma raiz pode ser uma *RBridge* ou um pseudonó. Cada raiz adquire múltiplos *nicknames*, um para cada árvore de distribuição, e utiliza o *nickname* de cada árvore de distribuição no algoritmo *tie-breaker*. Dessa forma, as árvores de distribuição centradas na raiz, embora igualmente de menor custo, utilizam diferentes *links* da rede TRILL.

A transmissão de quadros de dados do TRILL multidestino em uma árvore de distribuição realiza-se com o *multidestination flag* M igual a 1, e com um *nickname* de egresso adquirido pela *bridge* raiz. Nesse *nickname* encontram-se mapeados um ou um grupo de VLAN IDs ou de FGLs.

No caso limite em que existe a coincidência total de *RBridges* de ingresso e de *RBridges* raiz, os *nicknames* de egresso e de ingresso em quadros de dados do TRILL multidestino utilizam o mesmo valor.

4.4.8. **Transmissão Unicast**

A Figura 4-18 será utilizada para ilustrar a forma em que ocorre a transmissão *unicast* no TRILL, isto é, a transmissão de quadros de dados do TRILL com endereços *unicast* conhecidos.

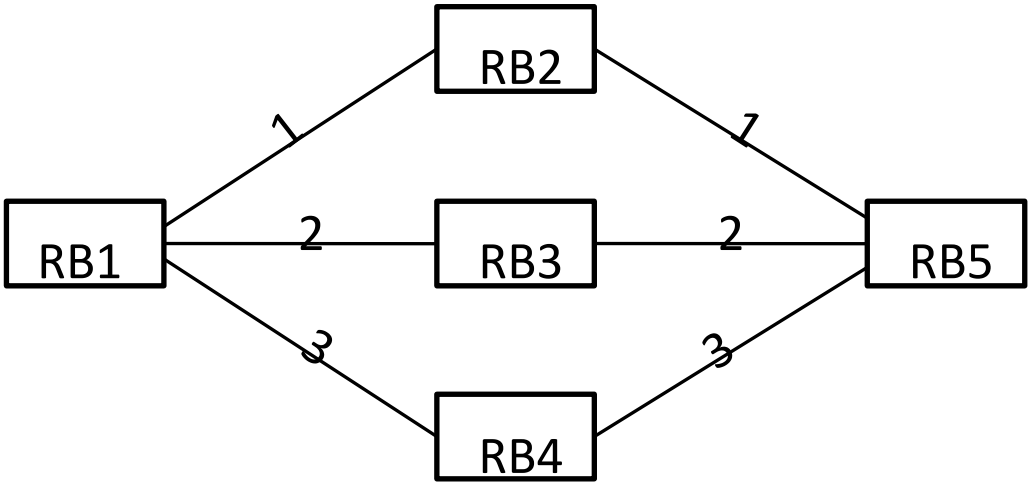


Figura 4-18 – Configuração de uma rede TRILL.

Suponha que RB5 adquiriu o *nickname* N51 e mapeou o *data label* (X1, Y1) nesse *nickname*, e que o protocolo *IS-IS for TRILL* atribuiu o *link* 1 para o *nickname* N51.

Suponha que RB1 aprendeu o *nickname* N51 associado ao endereço MAC de uma estação final conectada (certamente) à RB5, (endereço MAC8, por hipótese), no *data label* (X1,Y1).

Suponha que RB1 adquiriu o *nickname* N13, que o protocolo IS-IS for TRILL atribuiu esse *nickname* para o *link* 3 e que RB1 mapeou o *data label* (X1, Y1) nesse *nickname (*N13).

Suponha agora que RB1 vai transmitir um quadro de dados (enviado por uma estação final com endereço *unicast* MAC3, por hipótese) para o endereço MAC8 e que esse quadro possua o *data label* (X1,Y1).

Pelo endereço MAC8 e pelo *data label* (X1,Y1), RB1 obtém o *nickname* N51 (aprendido), que será o *nickname* de egresso do quadro de dados do TRILL a ser enviado. Esse *nickname* é a rota de egresso que indica o caminho a ser seguido (caminho 1, no caso).

O *nickname* de ingresso é o N13, obtido por RB1 a partir do *data label* (X1,Y1).

Com base nesses *nicknames*, RB1 transmite o quadro de dados direcionalmente para RB5.

Caso não houvesse ocorrido a aprendizagem do endereço MAC8, RB1 transmitiria o quadro de dados por uma árvore de distribuição.

A recepção de um quadro de dados por RB5 enseja que RB5 aprenda a associação entre o endereço MAC3, no *data label* (X1,Y1), associado ao *nickname* N13.

Caso a estação final MAC8 responda a esse quadro de dados, RB5 utilizará N51 como *nickname* de ingresso com base em (X1, Y1) e N13 como *nickname* de

egresso com base no endereço MAC3 e em (X1,Y1). Em consequência, essa resposta *unicast* será transmitida pelo caminho 3.

Verifica-se então o uso de diferentes caminhos para a transmissão e a recepção de um quadro por uma *RBridge*, conforme menção no início deste item.

Caso RB1 houvesse transmitido quadros iniciais por uma árvore de distribuição, após a recepção da primeira resposta enviada por RB5, RB1 faria uma transição, passando a transmitir (se tiver o que transmitir) de forma *unicast* (*known unicast*).

Em consequência dessa transição na forma de transmissão de tráfego por RB1 de *unknown* para *known*, RB5 poderá receber quadros de dados fora de ordem (*out-of-order*). Isso decorre do fato de que a transmissão de um quadro da *RBridge* de ingresso para uma árvore de distribuição e dessa árvore até alcançar o destino, pode seguir um caminho mais longo (e consequentemente com maior *delay* de trânsito) que o caminho de menor custo direto utilizado pela transmissão *unicast* que ocorre após o aprendizado.

4.4.9. **Filtragem de Tráfego**

Um quadro de dados do TRILL contém um VLAN ID nativo, seja de forma exclusiva (VL TRILL) seja na forma (X,Y) do FGL TRILL onde Y representa o VLAN ID nativo.

A presença desse VLAN ID em um quadro de dados do TRILL multidestino permite que esse quadro seja enviado apenas para *RBridges* que possuam pelo menos uma porta conectada a nós finais que sejam da VLAN indicada pelo VLAN ID nativo. Tais *RBridges* constituem as folhas de uma sub-árvore de distribuição dentro da árvore de distribuição utilizada.

Para isso, tais *RBridge*s anunciam por seus LSPs a quais VLANs estão conectadas. Em caso de acesso compartilhado, cabe à *RBridge*, atuando como *Appointed Forwarder,* realizar esse anúncio.

Esse mecanismo assemelha-se ao MVRP (*Multiple VLAN Registration Protocol*), visto no capítulo 3 anterior deste livro.

Para um quadro de dados com endereço de grupo, o TRILL possibilita a filtragem adicional baseada em endereços MAC derivados de endereços IP *multicast*. As *RBridge*s anunciam o conjunto de endereços MAC de grupo que elas desejam receber.

A filtragem é feita por "bisbilhotagem" (*spoofing*) em mensagens de controle IP *multicast*, principalmente em mensagens IGMP (*Internet Group Management Protocol*). Os endereços de grupo MAC são obtidos por *spoofing* nos endereços de grupo IP contidos nas mensagens de controle IP *multicast*.

4.4.10. **DRB, Appointed Forwarders e Pseudonode em Links de Acesso**

O presente subitem aborda a conceituação, em um *link* Ethernet compartilhado de acesso, de DRB (*Designated RBridge*), de *Appointed Forwarders* e de *pseudonode*.

4.4.10.1. DRB (Designated RBridge)

Quando existem múltiplas *RBridges* juntamente com nós finais em um mesmo *link* Ethernet compartilhado (um segmento de LAN, por exemplo), é necessário que apenas umas das *RBridges* encapsule um quadro MAC originário de um nó final em uma VLAN dentre um grupo de VLANs. As *RBridges* que assumem essa função são denominadas *Appointed Forwarders* para o grupo de VLANs considerado.

A Figura 4-19 ilustra essa configuração para o caso de acesso por um segmento de LAN.

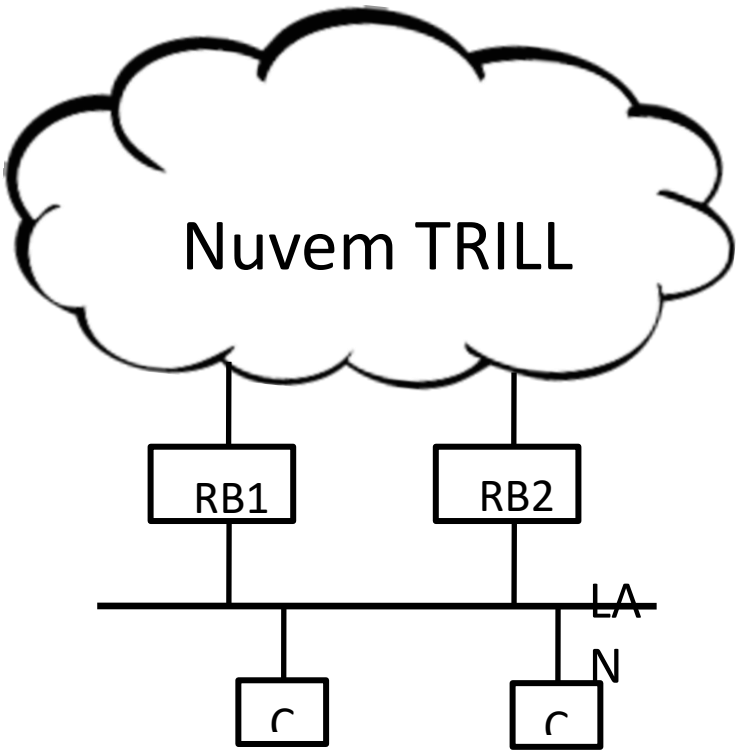


Figura 4-19 – Configuração de acesso por um segmento de LAN.

Para definir a forma de encaminhamento do quadro de dados pelas *RBridges* conectadas a um *link* Ethernet compartilhado de acesso, o IS-IS possui um protocolo de eleição que define uma dessas *RBridges* como sendo a *Designated RBridge* (DRB) desse *link*.

Diferentemente do L3 IS-IS, onde é possível a existência de múltiplas DRBs, o *IS-IS for TRILL* admite a existência de apenas uma DRB em cada *link* compartilhado de acesso.

Para possibilitar a distribuição de tráfego na rede, certamente em função dos VLAN IDs dos quadros, a DRB pode delegar a função de encapsulamento/decapsulamento (e de encaminhamento) de um quadro para si própria ou para uma das demais *RBridges* conectadas ao acesso compartilhado.

A *RBridge* à qual foi delegada a função de encaminhamento (transmissão e recepção) de quadros de dados em um determinado grupo de VLANs, é referida como *Appointed Forwarder RBridge*, ou simplesmente *Appointed Forwarder*, para os quadros de dados recebidos de um ou mais nós finais contidos nesse grupo de VLANs.

4.4.10.2. Appointed Forwarders

A RFC 6439 (*RBridges: Appointed Forwarders*) tem como propósito atualizar a RFC 6325 quanto a *Appointed Forwarders.*

O mecanismo *Appointed Forwarder* só se aplica em portas de uma *Appointed Forwarder RBridge* que estejam configuradas para o serviço de acesso compartilhado. São excluídas, consequentemente, portas configuradas para o entrocamento de *RBridges* (portas tronco) e portas configuradas para o serviço de acesso ponto-a-ponto.

Caso ocorra um engano de configuração, de forma tal que duas *RBridges* se considerem o *Appointed Forwarder* para um mesmo grupo de VLANs, aplica-se um mecanismo de inibição especificado na RFC 6439. Na realidade, essa RFC define um conjunto de mecanismos para a inibição de diferentes hipóteses de erro de funcionamento.

O mecanismo *Appointed Forwarder* do TRILL suporta correção de falhas no modo *active-standy*, evita a ocorrência de *loops* e, com configuração administrativa, possibilita a distribuição de tráfego (*load spreading*) com base em VLANs.

Esse mecanismo requer que todas as *RBridges* que participam do grupo de acesso intercambiem quadros do *TRILL for IS-IS Hello* através de suas portas de acesso.

A detecção da impossibilidade de acesso a uma outra *RBridge* do grupo de acesso no mecanismo *Appointed Forwarder*, baseada no tempo de expiração do temporizador do *TRILL Hello*, demanda um tempo da ordem de segundos, embora esse tempo possa ser reduzido com o uso do BFD (*Bidirectional Forwarding Detection*).

4.4.10.3. Pseudonode

Caso em um *link* Ethernet compartilhado de acesso existam N *RBridges* conectadas (incluindo a DRB), o *IS-IS for TRILL* permite à DRB modelar todo o *link*

como um *pseudonode*. A DRB atribui um nome para o *pseudonode*, de forma tal que essas N *RBridges* no *link* possam reportar conectividade apenas para o *pseudonode*.

Pseudonodes são ferramentas sintéticas de simulação utilizadas para simplificar o cálculo de caminhos mais curtos.

O IS-IS *for TRILL* possibilita à DRB, como alternativa, relatar a conectividade para cada *Appointed Forwarder* individualmente, nos casos em que essa opção seja justificável.

4.4.11. **Acesso Active-Active ao TRILL**

A RFC 7379 (*Problem Statement and Goal for Active-Active Connection at the TRILL Edge*) define um novo mecanismo para acesso compartilhado de equipamentos de usuário (CEs) a uma rede TRILL. Um CE, ou seja, um nó final, pode ser uma estação final ou uma *bridge*.

Esse novo mecanismo, referido como Acesso *Active-Active* ao *TRILL* (*Active-Active Access to TRILL*), destina-se à conexão de CEs que se encontram multiplamente (*multply)* conectados a um grupo de *RBridges*, conforme ilustração na Figura 4-20.

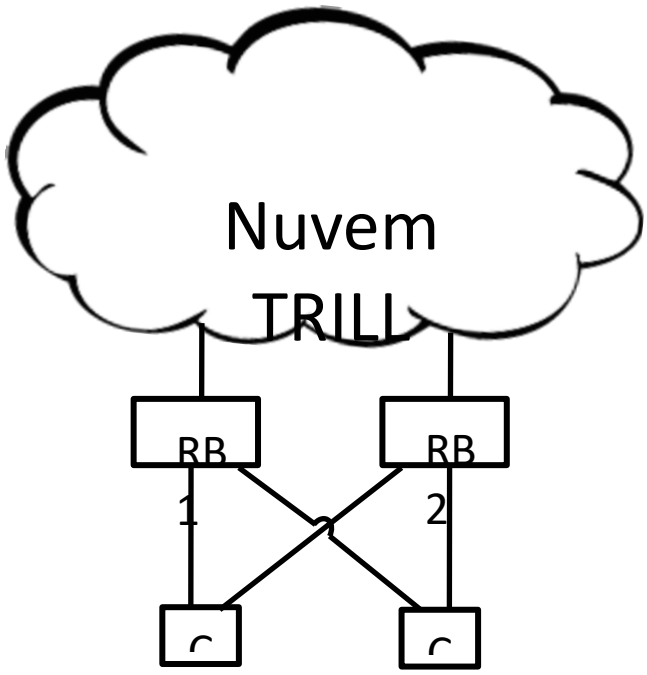


Figura 4-20 – Acesso Active-Active ao TRILL.

Comparativamente com o mecanismo *Appointed Forwarder*, a conexão de acesso *active-active* apresenta as seguintes principais características:

— Balanceamente de tráfego com base em fluxos de dados, e não em VLANs;

— Recuperação de falhas mais rápida.

Deve ser possível, exceto em condições de falha, que a conexão *active-active* realize distribuição de tráfego na granularidade de fluxos de dados entre os múltiplos *links* existentes. Assim, contrariamente ao que ocorre com os *Appointed Forwarders*, quadros de dados de uma mesma VLAN são distribuídos, por fluxos de dados, entre as *RBridges* do grupo de acesso.

O funcionamento do mecanismo de acesso *Active-Active* requer o uso de um protocolo LAALP (*Local Active-Active Link Protocol*). O LAALP é qualquer protocolo similar ao MC-LAG (*Multi-Chassis Link Aggregation*), que se aplica de forma distribuída em um CE, no grupo de *RBridges* de acesso e nos *links* entre esse CE e esse grupo de *RBridges*.

O LAALP é usualmente, na atualidade, um protocolo proprietário, variando conforme o fornecedor.

Um MC-LAG é uma extensão proprietária do padrão IEEE 802.1AX (*Link Aggregation*) que padroniza o comportamento do CE, deixando ainda proprietário o lado do grupo de *RBridges* de acesso. Se o LAALP se constituir em um MC-LAG, reduz-se a sua dependência de fornecedores.

Encontra-se em andamento, no IEEE, a definição de um adendo ao padrão IEEE 802.1AX, referido como IEEE 802.1X-REV, que remove totalmente as características proprietárias do MC-LAG, o que permitirá o uso de um LAALP totalmente aberto.

Complementando a RFC 7379, encontra-se em análise no IETF o *draft TRILL: Pseudo-Nickname for Active-Active Access*, que define os conceitos de *Virtual RBridge* (RBv) e do correspondente *pseudo-nickname*.

Virtual RBridge (RBv) representa, virtualmente, um grupo de *RBridges* de acesso *active-active* a uma rede TRILL. Uma RBv é também referida como um AAE (*Active-Active Edge RBridge Group*).

O conceito de RBv assemelha-se ao conceito de *pseudonode* visto anteriormente neste item. É atribuído um *pseudo-nickname* a uma RBv, utilizado para representá-la como um todo.

A DRB de uma RBv é referida como vDRB (*Virtual DRB*). A vDRB é responsável pela decisão quanto ao *pseudo-nickname* atribuído à RBv.

A seleção de um valor de *pseudo-nickname* ocorre na forma normal (RFC 6325), considerando-se a RBv como uma única *RBridge* de acesso. A vDRB seleciona um valor de *nickname* disponível, e anuncia esse valor (com prioridade máxima) para as demais *RBridges* da RBv por meio de LSPs do *IS-IS for TRILL*.

Quando uma *RBridge* de um AAE recebe um quadro de dados do TRILL, *unicast* ou multidestino, em uma de suas portas RBv, essa *RBridge* transmite esse quadro de dados, qualquer que seja o seu VLAN ID, utilizando o *pseudo-nickname* como *nickname* de ingresso.

A *RBridge* remota que recebe tais quadros de dados aprende o *pseudo-nickname*, de forma válida, para todas as VLANs constantes dos fluxos de dados

recebidos. Dessa forma, evita-se a alternância de aprendizagem de *nicknames* para diferentes *VLANs*. Essa alternância é referida como MAC *flip-flopping*.

4.4.12. **RBridge Channel**

A RFC 7178 (TRILL: *RBridge Channel Support*) define um mecanismo generalizado de constituição de canais para o envio de mensangens, tais como mensangens BFD (*Bidirectional Forwarding Detection*), entre *RBridges* e entre *RBridges* e estações finais. Esse mecanismo, que foi definido com base em extensões do protocolo TRILL, é denominado *RBridge Channel.*

O BFD é um processo de OAM largamente utilizado em redes IP e em redes MPLS, que será abordado com maior profundidade no próximo subitem dada a sua utilização no TRILL.

O *RBridge Channel* foi definido de forma generalizada para atender a múltiplos processos. Com essa forma generalizada, o *RBridge Channel* acomoda todos os modos regulares de transmissão de dados no TRILL, incluindo transmissões *unicast single-hop* e *multiple-hop*, assim como distribuição multidestino no escopo de uma VLAN.

As mensagens do *RBridge Channel* são estruturadas de modo a se assemelharem a quadros de dados do TRILL através de um *RBridge Campus*.

A comunicação no *RBridge Channel* entre *RBridges* e estações finais é bidirecional, e o formato das respectivas mensagens são, na maior parte, iguais àquelas trocadas entre *RBridges* no *RBridge Channel.*

O *RBridge Channel* é funcionalmente similar ao *MPLS Generic Associated Channel* definido na RFC 5586. Enquanto o *Generic Associated Channel* usa um label MPLS especial para multiplexação estatística de diferentes tipos de mensagens atendidas, uma mensagem do *RBridge Channel* é identificada por um *multicast inner.MacDA* especial e por um *inner Ethertype* (*Ethertype* interno).

4.4.12.1. Estrutura de um Quadro do RBridge Channel

A Figura 4-21 representa a estrutura de um quadro do *RBridge Channel.*

Figura 4-21 – Estrutura de um quadro do RBridge Channel.

Considerando-se o *Link Header* (ou *Outer Header*) como sendo um *link* Ethernet entre *RBridges*, o *outer.MacSA* é o endereço MAC da porta pela qual o quadro do *RBridge Channel* é enviado. O *outer.MacDA* é o endereço MAC da porta da *RBridge next-hop* para mensagens *unicast* do *RBridge Channel*, ou o endereço MAC *multicast All-RBridges* para mensagens multidestino do *RBridge Channel*.

Para o *TRILL Header* de quadros do *RBridge Channel*, a IANA (*Internet Assigned Numbers Authority*) atribuiu o valor de *nickname* 0xFFC0 (*Any-RBridge*), que se encontrava previamente disponível.

O *payload* é específico para o tipo de protocolo atendido, como, por exemplo, uma mensagem BFD.

4.4.12.2. RBridge Channel Header

A Figura 4-22 apresenta o formato do cabeçalho do *RBridge Channel*, que representa a parte mais específica do *RBridge Channel* na estrutura da figura anterior.

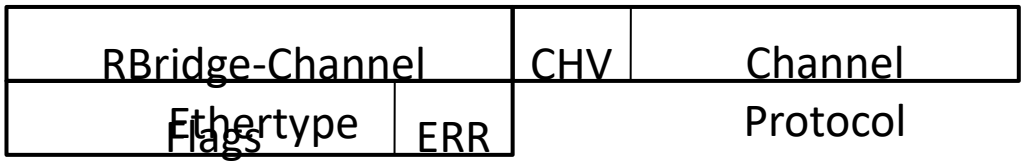


Figura 4-22 – Formato do cabeçalho do RBridge Channel.

Envelopado por esse cabeçalho encontram-se as informações específicas do *payload* do quadro do *RBridge Channel*.

Para o *RBridge Channel Ethertype* foi atribuído o valor 0x8946 pela *IEEE Registration Authority*, conforme menção na RFC 7178.

O campo CHV (*RBridge Channel Header Version*), com 4 bits, indica a versão atualizada para o cabeçalho, sendo o seu valor atual igual a zero.

O Campo *Channel Protocol*, com 12 bits, especifica um protocolo particular que utilize o *RBridge Channel*. A RCF 7178 menciona a atribuição de um único valor (0x001 para *RBridge Channel Error*).

O campo *Flags* provê a disponibilidade de 12 bits para a indicação de *flags*, conforme a Figura 4-23.

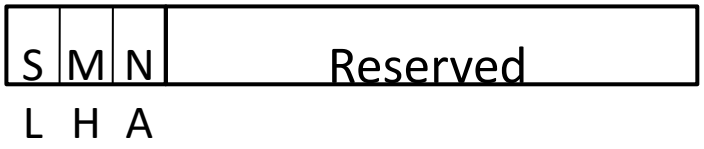


Figura 4-23 – Flags do cabeçalho do RBridge Channel.

Os três primeiros desses *flags* têm os seguintes significados:

— Bit SL (*Silent bit*): se igual a 1, as mensagens de erro do *RBridge Channel* foram suprimidas;
— Bit MH (*Multi-Hop bit*): é utilizado para informar ao destino que a mensagem pode ser *multi-hop* (MH=1) ou foi enviada com a intenção de ser apenas *one-hop* (MH=0);
— Bit NA (*Native bit*): se o valor do bit NA for igual a 1, indica-se que o *payload* é um quadro de protocolo nativo. Se setado para zero, indica-se que o *payload* é um quadro de protocolo não-nativo (ou seja, ocorreu o envelopamento prévio desse quadro pelo protocolo TRILL).

Finalmente, o campo ERR (*Error*s), com 4 bits, especifica os tipos de erro que podem ser indicados por quadros do *RBridge Channel*. A RFC 7178 especifica seis códigos de erros (códigos de 0 a 5), não atribuindo valores para os códigos de 6 a 14 e reservando o código 16.

Os códigos já atribuídos são os seguintes:

— Código 0: *No Error*;
— Código 1: *Frame too short (truncated Ethertype or Channel Header)*;
— Código 2: *Unrecognized Ethertype*;
— Código 3: *Unimplemented value of CHV*;
— Código 4: *Wrong value of NA flag*;
— Código 5: *Channel Protocol is reserved or unimplemented*.

4.4.13. **TRILL OAM e TRILL BFD**

OAM (*Operations, Administration and Maintance*) e a aplicação de BFD em redes TRILL têm como suporte cinco RFCs específicas:

— RFC 6905: *Requirements for OAM in TRILL*;
— RFC 7174: *TRILL OAM Framework*;
— RCF 7175: *TRILL: Bidirectional Forwarding Dectection (BFD)*;
— RFC 7455: *TRILL: Fault Management*;
— RFC 7456: *Loss and Delay Measurement in TRILL.*

Além dessas RFCs específicas, devem ser considerados, no TRILL, a recomendação ITU-T G.8013/Y.1731 (*OAM Functions and Mechanisms for Ethernet Based Networks*) e o padrão IEEE 802.1-2014 (Cláusulas 18 até 22), relativos a OAM em redes Ethernet.

A RFC 6905 define os requisitos para OAM no TRILL, que representam a base para as demais RFCs específicas.

Conforme a RFC 7174, a Figura 4-24 ilustra os diversos níveis de OAM e de BFD em uma rede TRILL.

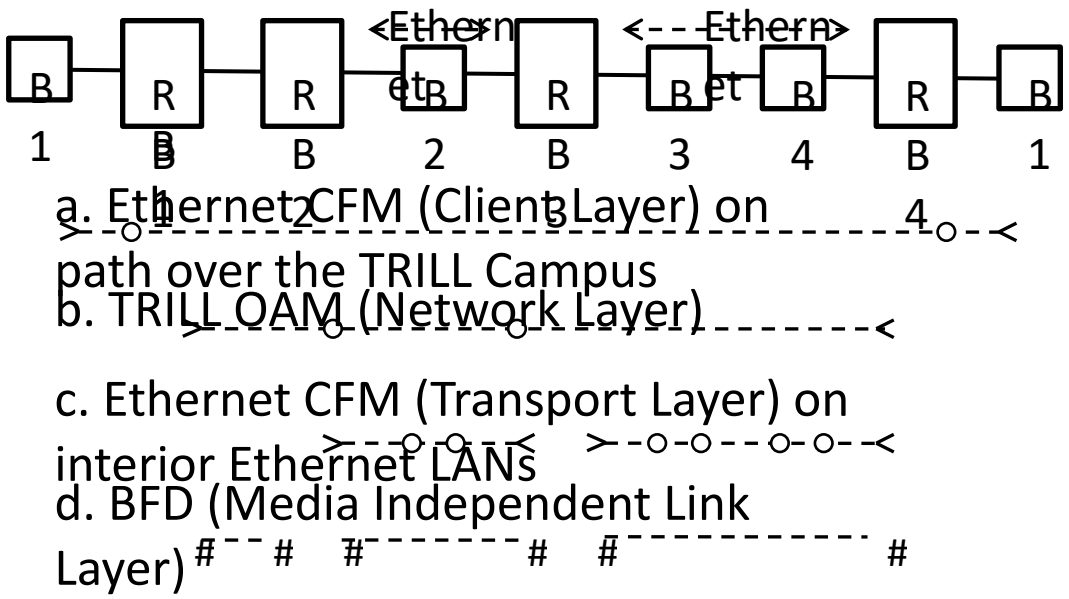


Figura 4-24 – OAM e BFD em uma rede TRILL.

Nessa figura, os símbolos > e < representam MEPs (*Maintance End Points*), o símbolo **o** representa MIPs (*Maintance Intermediate Points*), o símbolo # representa pontos de terminação do BFD e o símbolo * representa pontos de terminação de OAM de *link*.

4.4.13.1. Estrutura do TRILL OAM (RFC 7174)

A RFC 7174 adota, em essência, os conceitos e os procedimentos definidos para o *Ethernet* OAM na recomendação ITU-T G.8013/Y.1731 e no padrão IEEE 802.1Q-2014. Ressalva-se, no entanto, a definição de alguns aspectos próprios do TRILL OAM, particularmente no que tange aos quadros TRILL OAM.

A RFC 7455 e a RFC 7456, além de definir os aspectos específicos de gerenciamento de falhas e de medições de perdas e de *delays* respectivamente, complementam a RFC 7174 quanto a aspectos estruturais do TRILL OAM.

Para a acurada ralização de monitoração e/ou de diagnósticos, as mensagens do TRILL OAM devem seguir o mesmo caminho que os quadros de dados correspondentes.

A estrutura das mensagens do TRILL OAM encontra-se na Figura 4-25.

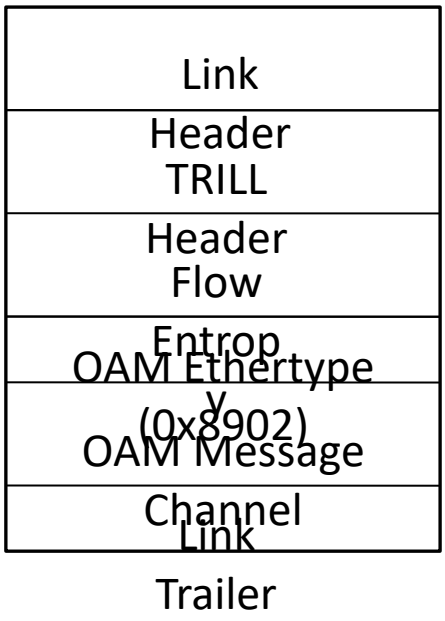


Figura 4-25 – Estrutura das mensagens do TRILL OAM.

Como se observa nessa figura, o *inner Ethernet* é referido como *Flow Entropy*. Esse campo, de tamanho fixo, contém pacotes de dados reais ou informações de dados sintéticos que realizam a emulação (*mimicking*) do fluxo de dados que se deseja monitorar.

Para a identificação do quadro OAM, utiliza-se um valor *EtherType* específico, que é o CFM *Ethertype* (valor 0x8902).

O *OAM Message Channel* disponibiliza métodos para comunicar detalhes específicos do TRILL OAM.

O *TRILL OAM Message Channel* é composto pelo *TRILL OAM Message Header* e por um ou mais *OAM Message TLVs*.

A Figura 4-26 apresenta o formato do *TRILL OAM Message Channel*.

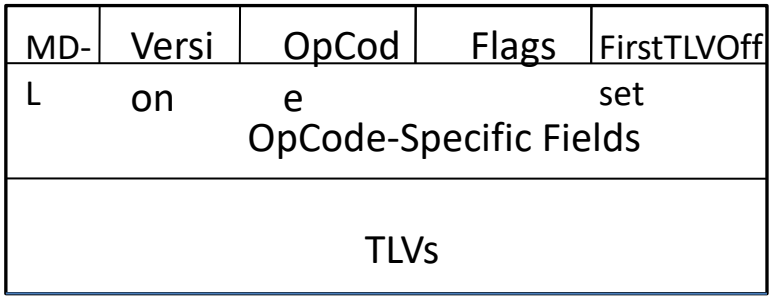


Figura 4-26 – Formato do TRILL OAM Message Channel.

O *TRILL OAM Message Header* é representado, nessa figura, pelos campos MD-L, *Version*, *OpCode* e *FirstTLVOffset*.

Esses campos têm os seguintes significados:

— *Maintenance Domain Level* (MD-L), com 3 bits, utiliza, para o TRILL em geral, o valor zero. Para o *TRILL Base Mode* (*Appendix B* da RFC 7455), o valor do MD-L é igual a 3, enquanto que para extensões do TRILL o MD-L pode assumir diferentes valores apropriados;
— *Version*, com 5 bits. A versão atual é a versão zero;
— *Operation Code* (*OpCode*), com 8 bits, especifica a operação realizada pela mensagem do TRILL OAM;
— *Flags*, com oito bits, incluem *flags*, operacionais e a sua definição é específica para cada *OpCode* utilizado;
— *FirstTLVOffset*, com 8 bits, define a localização do primeiro TLV, em bytes, a partir do final do próprio campo *FirstTLVOffset*;
— *OpCode-Specific Information*, que pode conter, por exemplo, o número de identificação da sessão, um *timestamp*, etc..

A RFC 7455 apresenta os *TRILL OAM TLVs.* São apresentados 10 TLVs definidos no padrão IEEE 802.1Q-2014 que são comuns ao CFM (*Connectivity Fault Management* e ao TRILL.

São apresentados também 11 TLVs específicos para o TRILL, com os respectivos códigos de identificação (entre 64 e 74). Esses TLVs foram definidos na própria RFC 7455.

Dentre os TLVs específicos para o TRILL destaca-se o *TRILL OAM Application Identifier TLV* (código 64), que é obrigatoriamente o primeiro TLV em qualquer mensagem do OAM TRILL. Esse TLV contém o campo *Return Code* (com valor zero para indicar mensagens *Request* e com o valor 1 para indicar mensagens *Reply*) e o campo *Return Sub-Code* que indica sub-códigos para cada um dos *Return Codes*.

O *End TLV* (código 0) deve ser obrigatoriamente incluído no final de qualquer mensagem de OAM.

Um quadro do TRILL OAM é identificado, em conjunto, pelo valor do CFM *Ethertype* 0x8902 e pelo *Alert flag*(A) com o seu valor setado para 1.

O *Alert flag*, indicado como A, representa, conforme a RFC 7455, a utilização do segundo bit de *flag* do *TRILL Header* que se encontrava reservado.

Uma *RBridge* pode ser:

— Incapacitada para OAM;
— Capacitada para OAM com as novas extensões;
— Capacitada para OAM com retrocompatibilidade (*backwards compatibility*).

Uma *RBridge* capacitada para OAM que pretende enviar uma mensagem OAM *Request*, deve conhecer a capacitação OAM da *RBridge* receptora para enviar uma mensagem apropriada.

A notificação de capacitação OAM realiza-se pelo protocolo *IS-IS for TRILL*, utilizando o sub-TLV *TRILL Version* (TRILL-VER) definido da RFC 7176. Essa informação é codificada no *flag O* (*OAM Capable*) e no *flag B* (*Backwards-Compatible OAM*) desse sub-TLV.

Os conceitos de ME (*Maintenance Entity*), de MA (*Maintenance Association*), de MD (*Maintenance Domain*), de MEP (*Maintenance End Point*) e de MIP (*Maintenance Intermediate Point*), são aqueles definidos no padrão IEEE 802.1Q-2014 e apresentados no Capitulo 7 deste livro.

A IANA atribuiu um endereço *unicast* (00-00-5E-90-01-00) e uma endereço *multicast* (01-00-5E-90-01-00) para a identificação de quadros do OAM TRILL.

4.4.13.2. Gerenciamento de Falhas (RFC 7455)

A RFC 7455 especifica o gerenciamento de falhas no TRILL OAM. Os métodos definidos acompanham aqueles do CFM (*Connectivity Fault Management*) especificados no padrão IEEE 802.1Q-2014 e reutiliza as suas ferramentas quando possível.

A RFC 7455 apresenta os seguintes tipos de mensagem:

— Mensagens *Loopback*;
— Mensagens *Path Trace*;
— Mensagens *Multi-Destination Tree Verification* (MTVMs);
— Mensagens CCM (*Continuity Check Messages*).

As mensagens *Loopback,* cujas funções encontram-se definidas no Capítulo 7 deste livro para o CFM, utiliza, para o TRILL OAM, o valor de *OpCode* 3, e o valor de *OpCode* 2 para as mensagens de resposta (*Reply messages*).

As *messages Lookpback* utilizam o *Out-of-Band Reply Address TLV* (código 65), o *Diagnostic Label TLV* (código 66) e o *Sender ID TLV* (código 1).

Para as mensagens *Path Trace* e as MTVMs, a RFC 7455 especificou os seguintes *OpCodes* (contidos no registro *CFM OAM IETF Opcodes*):

— Código 64 (*Path Trace Reply*);
— Código 65 (*Path Trace Message*);
— Código 66 (*Multi-Destination Tree Verification Reply*);
— Código 67 (*Multi-Destionation Tree Verification Message*).

O uso primário das mensagens *Path Trace* é para o isolamento de falhas. Essas mensagens podem ser também utilizadas para tracejar o caminho utilizado de uma *RBrigde* para outra *RBridge*.

O padrão IEEE 802.1Q-2014 atende os objetivos das mensagens *TRILL Path Trace* utilizando mensagens *Link Trace.* O uso das mensagens *Link Trace* no TRILL OAM não é possível pelo fato de que, no TRILL, os caminhos *unicast* e *multicast* não são congruentes.

As mensagens *Path Trace* adotam o mesmo formato que as mensagens *Link Trace*.

As mensagens *Multi-Destionation Tree Verification* (MTVMs) possibilitam a verificação da integridade das árvores de distribuição no TRILL, assim como a verificação da operacionalidade dos seus mecanismos de filtragem (*pruning*).

Adicionalmente, as MTVMs podem ser utilizadas para tracejar uma dada árvore de distribuição em um *TRILL Campus*.

O formato das MTVMs é o mesmo que os das mensagens *Loopback*, com exceção do valor do *OpCode* que passa a ser 67.

A RFC 7455 define a aplicação de CCMs no TRILL. De um modo geral, as OAM TRILL CCMs utilizam os mesmos critérios de aplicação das IEEE 802.1Q-2014 CCMs, exceto em alguns aspectos que são específicos para o TRILL OAM.

Enquanto que pelo padrão IEEE 802.1Q-2014 existe sempre um único caminho entre dois MEPs em um dado momento, no TRILL podem existir múltiplos caminhos.

As OAM TRILL CCMs, diferentemente das CCMs no padrão IEEE 802.1Q-2014, definem múltiplos fluxos em um objeto CCM, cada um correspondendo a um fluxo que um dado MEP deseja monitorar. Então, em ambientes *multipath* como o do TRILL, as CCMs monitoram a conectividade por fluxo.

Para identificar quais os fluxos, ou qual o fluxo, encontram-se em falha, a *RBridge* de origem utiliza uma combinação de parâmetros, dentre os quais certamente se encontra o *Flow Identifier TLV* (código 72).

4.4.13.3. Medições de Perdas e de Delays no TRILL (RFC 7456)

A Monitoração de Performance (PM) é um aspecto chave de OAM, que permite aos operadores de rede verificar a SLA (*Service Level Agreement*) assumida junto aos usuários e a detectar anomalias nas redes.

A RFC 7456 especifica dois mecanismos fundamentais da PM no TRILL OAM, relativos a medições de perdas de quadros e a *delays* de quadros em redes TRILL. Esses mecanismos são baseados na funcionalidade da Monitoração de Performance definida na recomendação ITU-T G.8013/G.8013/Y.1731.

O protocolo de medição de perdas de quadros entre duas *RBridges* pode se realizar mediante o envio de um conjunto de quadros sintéticos, procedendo-se então à contagem do número de quadros enviados e de quadros recebidos durante um período de testes.

As medições de perdas de quadros (*IPPM-Loss*) ocorre no modo *one-way Synthetic Loss Measurement* (1SL). A RFC 7456 faz distinção entre perdas de quadros *near-end* e de perdas de quadros *far-end* (entre uma *RBridge* e outra *RBridge* remota).

O protocolo de medição de *delays* objetiva a medição de *delays* e de variação de *delays* (*jitter*) de quadros entre duas *RBridges*. Esse protocolo utiliza, para isso, mensagens OAM contendo *timestamps*.

As medições de *delays* ocorrem no modo *one-way delay measurement* (IPPM-1DM) e no modo *two-way delay measurement* (IPPM-2DM).

Aplicam-se as seguintes siglas:

— IPPM (*Inter-Packet Performance Monitoring*);
— SLM (*Synthetic Loss Measuremente Message*);
— SLR (*Synthetic Loss Measurement Reply*);
— DMM (*Delay Measurement Message*);
— DMR (*Delay Measurement Reply)*.

As medições de PM no TRILL podem ser aplicadas por meio de um protocolo de medição ponto a ponto ou de um protocolo de medição ponto a multiponto.

No caso de protocolo ponto a multiponto, uma *RBridge* envia mensagens PM para múltiplas *RBridges* utilizando mensagens *multicast*. Em caso de *two-way PM,* as *RBridges* refletoras respondem utilizando mensagens *unicast*, após um período aleatório de espera de zero a 2 segundos para proteção contra tempestades de respostas (*reply storms*).

4.4.13.4. TRILL BFD (RFC 7175)

A RCF 7175 especifica o uso do protocol BFD em redes TRILL, com base na extensão *RBridge Channel* (RFC 7178) ao protocolo TRILL.

O BFD, especificado na RFC 5880 (*Bidirectional Forwarding Detection*), é um mecanismo de OAM amplamente utilizado em redes IP e em redes MPLS. Esse macanismo representa, na realidade, um simples exemplo de protocolo *Hello*, similar às versões desse protocolo em protocolos de roteamento existentes.

Um par de sistemas troca periodicamente quadros BFD. Se um dos sistemas para de receber esses quadros por um tempo consideravelmente longo, algum componente do caminho bidirecional é considerado como estando em falha.

Sob certas condições, a troca de quadros BFD pode não ocorrer de forma periódica.

O suporte ao *BFD-over-TRILL* é similar ao suporte ao *BFD-over-IP* (RFC 5881), exceto no que diz respeito a diferenças explicitamente mencionadas na RFC 7175.

Conforme a RFC 7175, uma rápida convergência em redes TRILL em caso de falhas na rede é considerada criticamente necessária. O mecanismo *TRILL IS-IS Hello*, definido na RFC 7177 (*TRILL: Adjacency*), não satisfaz a essa necessidade, convergindo em dezenas de segundos ou, na melhor das hipóteses, em um segundo.

Um quadro do TRILL BFD utiliza a estrutura de quadros do *RBridge Channel*, sendo também adotado, para o seu cabeçalho, o formato do *RBridge Channel Header*.

O TRILL suporta quadros do TRILL *BFD Control one-hop* e *multi-hop* e quadros do TRILL *BFD Echo* entre vizinhos, sobre o *RBridge Channel.*

Para o caso específico do TRILL BFD, foram definidos pela IANA (conforme a RFC 7175) os seguintes números para o protocolo *RBridge Channel*:

— Número 0x002, indicando o protocolo TRILL *BFD Control*;
— Número 0x003, indicando o protocolo TRILL *BFD Echo*.

Os quadros do *TRILL BFD Control* são quadros *unicast* do *TRILL RBridge Channel.* As informações específicas associadas ao protocolo *TRILL BFD Control* são aquelas definidas na Seção 4.1 da RFC 5880 para o *BFD Control*.

One-hop TRILL BFD Control é tipicamente utilizado para detectar rapidamente falhas em *links* e em *RBridges*. Os quadros do *TRILL BFD Control* utilizados com esse propósito são enviados com alta prioridade, ou seja, o *Inner.VLAN tag* deve utilizar a prioridade 7. Em caso de *Ethernet Outer Link,* o *Outer.VLAN tag* deve também utilizar a prioridade 7.

Nos quadros do *TRILL BFD Control One-hop*, o *flag* MH do *RBridge Channel Header* é setado para zero, enquanto que nos quadros do *TRILL BFD Control Multi-hop* esse *flag* é setado para 1.

A definição do *multi-destination BFD Control* encontra-se em progresso, conforme o IETF *draft BFD for Multipoint Networks*.

Um quadro do *TRILL BFD Echo Protocol*, da mesma forma, utiliza o formato dos quadros *unicast* do *RBridge Channel*. Um quadro do *TRILL BDF Echo* deveria ser enviado em retorno por um vizinho imediato que o recebeu, pelo fato de que ambos os *nicknames* neles contidos são setados para um único valor (valor do *nickname* de ingresso).

4.4.14. **Plano de Controle do TRILL**

O plano de controle do TRILL engloba, conforme menção anterior, o protocolo *TRILL Hello* (RFC 7177), o protocolo *IS-IS for TRILL* (RFC 7176) e o protocolo ESADI (RFC 7357).

Como o protocolo ESADI foi apresentado anteriormente neste item, o presente subitem abordará o protocolo *TRILL Hello* e o protocolo *IS-IS for TRILL.*

À exceção do ESADI, que possui formatação própria, os quadros de controle do TRILL são compostos pelos seguintes campos:

— *IS-IS Common Header*;
— *IS-IS PDU Specific Fields;*
— *IS-IS TLVs.*

4.4.14.1. Protocolo TRILL Hello

Utiliza-se no TRILL um protocolo *Hello* com ligeiras modificações, sendo que o resultado dessas modificações representa o protocolo *TRILL IS-IS Hello*, ou simplesmente protocolo *TRILL Hello*. O protocolo *TRILL Hello* foi definido

inicialmente na RFC 6325, sendo a sua versão definitiva especificada na RFC 7177 (*TRILL: Adjacency*).

O *IS-IS Hello* é um protocolo pelo qual roteadores (*RBridges* no caso do *TRILL Hello*) enviam mensagens *Hello* para seus vizinhos, dando conhecimento de sua existência. Além disso, o roteador que se anuncia inclui também em suas mensagens *Hello* uma lista dos vizinhos que já se anunciaram para ele. Um roteador somente considera um outro roteador como vizinho se, além de se anunciar para esse outro roteador, verificar também a sua presença na lista de vizinhos já enviada por esse outro roteador. Ou seja, a conectividade entre vizinhos constitui-se de forma *two-way*.

Para o *TRILL Hello* foi definido um tipo de mensagem específica, referida como *TRILL Hellos,* que podem ser *TRILL LAN Hellos* ou *TRILL P2P Hellos.*

A RFC 7177 define nove TLVs para as *TRILL Hellos,* indicando quando esses TLVs são opcionais ou obrigatórios para ambas as opções de mensagem.

Mensagens *TRILL Hello*, como todas as demais mensagens do *IS-IS for TRILL*, são primariamente distinguidas das mensagens *Layer 3 IS-IS* pela utilização do endereço MAC *multicast* de destino *All-IS-IS-RBridges* (01-80-C2-00-00-41) e pela utilização do valor *L2-IS-IS Ethertype* (0x22F4).

Uma *TRILL IIH PDU* é uma variação para o TRILL das *Layer 3 IIH PDUs.* Em uma *TRILL IIH PDU*, o *IS-IS Common Header* e o *PDU Header* fixo são os mesmos que os das *Layer 3 Level 1 IIH PDUs*.

Conforme menção anterior neste item, às mensagens *TRILL Hello* foi atribuído o valor de NPID (*Network Layer Protocolo Identifier*) igual a 0xC0.

Para um *link* Ethernet compartilhado de acesso a uma rede que utiliza o IS-IS, o protocolo *IS-IS Hello* elege um ou mais roteadores com a função DIS (*Designated Intermediate System*). No TRILL, o equivalente ao DIS é referido como DRB (*Designated RBridge*).

Quando da escolha de um DIS para um *link* compartilhado de acesso no caso de utilização do protocolo L3 *IS-IS Hello,* podem ocorrer diferentes agrupamentos (*cliques*) de roteadores que tenham obtido conectividade *two-way*. Para cada um desses agrupamentos é eleito um DIS, sem prejuízo na transmissão de quadros de dados.

Quando se utiliza um *link Ethernet* compartilhado de acesso ao TRILL, cabe à DRB a indicação dos *Appointed Forwarders* para grupos de VLANs que utilizam esse *link*. Em consequência, diferentemente do que ocorre no *L3 IS-IS Hello*, é importante que exista uma única DRB. Para que isso ocorra, a eleição da DRB baseia-se apenas nas prioridades das *RBridges* envolvidas, e não na verificação da conectividade *two-way* entre *RBridges*.

Em um *link Ethernet* compartilhado que interliga *RBridges*, pode existir equipamentos ou serviços que restrinjam a conectividade de VLANs. Para que isso não ocorra com quadros do TRILL (quadros de dados e quadros do TRILL IS-IS) que transitam nesse *link*, atribui-se esses quadros a uma VLAN especial,

denominada VLAN Designada (*Designated VLAN*), também referida como DVLAN.

Compete à DRB do *link* Ethernet compartilhado a indicação da VLAN Designada e a atribuição do *Outer.VLAN ID* para a VLAN Designada. É necessário que a DRB anuncie esse *Outer.VLAN ID* para as demais *RBridges* conectadas ao *link* Ethernet compartilhado.

Para o desempenho de suas funções, uma DRB possui uma de suas portas no estado DRB, porta essa referida como porta DRB.

A DVLAN de um *link* Ethernet entre *RBridges* é determinado pela *Desired DVLAN* da DRB nesse *link*. A *Desired DVLAN* de uma RBridge, inclusive da DRB do *link*, é usualmente configurada manualmente pelo operador da rede TRILL.

Outra diferença entre o *L3 IS-IS Hello* e o *TRILL IS-IS Hello* diz respeito ao tamanho das mensagens *Hello*. No *Layer 3 IS-IS Hello* essas mensagens devem ser preenchidas (*padded*) para o tamanho máximo. No protocolo *TRILL Hello* as mensagens *Hello* são limitadas em tamanho, não sendo preenchidas (*padded*), de modo a remover impedimentos artificiais para comunicação entre *RBridges* vizinhas.

Uma outra função desempenhada pelo protocolo *TRILL Hello* é a de determinar a MTU (*Maximum Transmission Unit*) com a qual é possível a comunicação segura entre *RBridges* vizinhas.

4.4.14.2. Protocolo IS-IS for TRILL

O procolo IS-IS for TRILL (ou IS-IS TRILL), foi especificado na RFC 7176 (*TRILL Use of IS-IS*).

O protocol *IS-IS for TRILL* consiste em extensões do protocol *Layer 3 IS-IS* destinadas ao TRILL, da mesma forma que a RFC 6329 estende o protocolo *Layer 3 IS-IS* para atendimento ao SPB.

A RFC 7176 define novos TLVs, sub-TLVs e PDUs para o *IS-IS for TRILL*, além de utilizar, com as devidas adequações, aqueles já existentes no *Layer 3 IS-IS.* A RFC 7176 especifica também formatos e *code points* para suporte ao TRILL.

A RFC 7176 obsoletou e substituiu a RFC 6326, e tem como base a RFC 6325, a RFC 1195 e o padrão ISO-10589.

- **TLVs Adicionais**

O protocolo IS-IS for TRILL utiliza, em adição ao *Layer 3 IS-IS*, os seguintes TLVs:

— *Group Address* (GADDR) *TLV* (código 142);
— *Multi-Topology-Aware Port Capability* (*MT-Port-Cap*) *TLV* (código 143);
— *MT_Capability TLV*(código 144);
— *Router Capability TLV* (código 242);
— *TRILL Neighbors TLV* (código 145);

— *MT-Intermediate Systems Neighbors* (MT-ISN) *TLV* (código 222);
— *Extended IS Reachability TLV* (código 22).

As informações sobre esses TLVs encontram-se sumarizados na Figura 4-27.

RFC	PDU	TLV	Tipo de TLV
7176	LSP	GADDR	142
6165	IIH	MT-Port-Cap-TLV	143
6329	LSP	MT-Capability	144
4971	LSP	Router Capability	242
7176	IIH	TRILL Neighbor	145
5120	LSP	MT-ISN	222
5305	LSP	Extended IS Reachability	22

Figura 4-27 – TLVs adicionais para os protocolos IS-IS for TRILL.

O *TRILL neighbor TLV* substitui, para o *IS-IS for TRILL*, o *IS-IS neighbor TLV* do L3 IS-IS.

- **Sub-TLVs Adicionais**

A Figura 4-28 relaciona os sub-TLVs definidos pela RFC 7176 respectivamente para cada um dos TLVs, sendo que o *TLV TRILL Neighbor* (código 145) não possui sub-TLVs.

TLV	SUB-TLV	
	Nome	**Tipo**
142	Group MAC Address (GMAC-ADDR)	1
	Group IPv4 Address (GIP-ADDR)	2
	Group IPv6 Address (GIPV6-ADDR)	3
	Group Labeled MAC Address (GLMAC-ADDR)	4
	Group Labeled IPv4 Address (GLIP-ADDR)	5
	Group Labeled IPv6 Address (GLIPV6-ADDR)	6
143	Special VLANs and Flags (VLAN-FLAGS)	1
	Enabled-VLANs (Enabled-VLANs)	2
	Appointed Forwarders (AppointedFwrdrs)	3
	Port TRILL Version (PORT-TRILL-VER)	7
	VLANs Appointed (VLANs-Appointed)	8
144 e 242	Nickname (NICKNAME)	6
	Trees (TREES)	7
	Tree Identifier (TREE-RT-IDs)	8
	Trees Used Identifiers (TREE-USE-IDs)	9
	Interested VLANs and Spanning Tree Roots (INT-VLAN)	10
	TRILL Version (TRILL-VER)	13
	VLAN Group (VLAN-GROUP)	14
	Interested Labels and Spanning Trees Roots (INT-LABEL)	15
	RBridge Channel Protocols (RBCHANNELS)	16
	Affinity (AFFINITY)	17
	Label Group (LABEL-GROUP)	18
222 e 22	MTU (MTU)	28

Figura 4-28 – Sub-TLVs para o protocolo IS-IS for TRILL.

- **Teste de MTUs**

Existem dois novos tipos de mensagens (PDUs) para uso entre pares de *RBridges* vizinhas, com o objetivo de testar, de forma bidirecional, o tamanho máximo (MTU) dos quadros transitados nas respectivas conexões.

Essas PDUs são as seguintes:

— *MTU-Probe*;
— *MTU-Ack*.

Essas PDUs são preenchidas (*padded*) para os tamanhos sendo testados.

O envio de *MTU-probes* é opcional, mas uma *RBridge* que recebe uma dessas PDUs deve respondê-la com uma *MTU-Ack* preenchida para o mesmo tamanho que é a *MTU-probe* recebida.

As *MTU-probes* podem ser *unicast* ou *multicast*. Uma *MTU-Ack* é normalmente *unicast* quando enviada para a *RBridge* que transmitiu a *MTU-probe*, mas pode ser *multicast* para o endereço *All-RBridges*.

Quadros nativos que após o encapsulamento pelo TRILL excedam a MTU de um *link* pelo qual foram enviados são geralmente descartados.

4.4.15. **Módulo RBridge MIB (RFC 6850)**

A RFC 6850 (*Definitions of Managed Objects for RBridges*) descreve um modelo para o gerenciamento de *RBridges*. Mais particularmente, essa RFC descreve os objetos específicos para o TRILL que constituiem o módulo *RBridge MIB* (RBRIDGE-MIB).

A RBridge-MIB foi definida com considerável ênfase na relação com outros módulos MIB, maximizando a reutilização dos objetos gerenciáveis desses módulos, sem todavia duplicá-los.

Dentre tais módulos MIB podem ser citadas a IF-MIB (*Interfaces MIB*), a BRIDGE-MIB, a P-BRIDGE-MIB, a Q-BRIDGE-MIB, a IEEE8021-BRIDGE-MIB, a IEEE8021-Q-BRIDGE-MIB e a ISIS-MIB.

A RFC 6850 apresenta a relação da RBRIDGE-MIB com cada uma dessas MIBs.

As *RBridges* são compatíveis com as IEEE 802.1D e IEEE 802.1Q *bridges*, assim como com roteadores e nós finais IPV4 e IPv6. No modelo para gerenciamento definido na RFC 6850, as *RBridges* são vistas primariamente como *bridges* de usuário.

O módulo P-BRIDGE-MIB, que é uma extensão do módulo BRIDGE-MIB, define objetos gerenciáveis para as classes de tráfego e para o aprimoramento de filtragem de tráfego *multicast*, conforme o padrão IEEE 802.1D e o padrão IEEE 802.1t (*MAC Bridges: Technical and Editorial Corrections*) que é um adendo ao padrão 802.1D.

O módulo MIB para *bridges* virtuais, referido como Q-BRIDGE-MIB, define objetos gerenciáveis para aprimoramentos de *virtual LAN bridging*, conforme o padrão IEEE 802.1Q-2014 e os padrões IEEE 802.1u (*Technical and Editorial Corrections*) e IEEE 802.1v (*VLAN Classification by Protocol and Port*) que são adendos ao padrão IEEE 802.1Q-2014.

Os objetos da RBRIDGE-MIB são acessados geralmente através do SNMP (*Simple Network Management Protocol*). Tais objetos são definidos utilizando-se os mecanismos definidos na SMI (*Structure of Managemente Informations*), que hoje se encontra na versão 2 (SMIv2), conforme o padrão STD 58 (correspondente às RFC 2578, RFC 2579 e RFC 2580).

A RFC 6850 estrutura os objetos gerenciáveis da RBRIDGE-MIB em sub-árvores e cada sub-árvore é organizada como um conjunto de objetos relacionados. Conforme menção anterior, esses objetos são suplementados pela utilização de objetos de outros módulos MIB.

Foram especificados 9 tipos de sub-árvore, que são as sub-árvores *rbridgeBase, rbridgeFdb, rbridgeVlan, rbridgeESADI, rbridgeCounters, rbridgeSnooping, rbridgeDtree, rbridgeTrill* e a sub-árvore *Notifications*.

CAPÍTULO 5

LAN BRIDGING EM REDES DE PROVEDORES

5.1. PREÂMBULO

Como já mencionado em capítulos anteriores deste livro, os órgãos de padronização vêm desenvolvendo esforços, em diferentes caminhos, no sentido da especificação de padrões visando a tornar *carrier-grade* a operação de *bridged* (*switched*) *LANs*, particularmente de *bridged Ethernet LANs*.

As preocupações englobam escalabilidade, qualidade de serviço, padrões de desempenho e gerência de redes, dentre outros quesitos, além da especificação de serviços com funcionalidades adequadas.

Este capítulo dedica-se à utilização de redes 802.1Q (LANs C-*tagged*, PB, PBB e PBB-TE) como redes de provedores de serviço. Observa-se que o termo provedores de serviço não se restringe a empresas prestadoras de serviços de telecomunicações, mas também aplica-se à prestação de serviço a clientes internos dentro de uma mesma organização.

Observa-se que o padrão IEEE 802.1Q-2014 incorpora redes *bridged Ethernet* de provedores de serviço, antes especificadas em emendas ao padrão IEEE 802.1Q-2014 (IEEE 802.1ad, IEEE 802.1ah e IEEE 802.1Qay, respectivamente para PB, PBB e PBB-TE).

A simples utilização do padrão IEEE 802.1D-2004 mostra-se inadequada para esse propósito, não lhe sendo então atribuída qualquer atenção no presente capítulo.

A utilização das LANs apenas C-*tagged* do padrão IEEE 802.1Q-2014 mostra-se também inadequada para o propósito deste capítulo, como será visto subsequentemente. Entretanto, VLANs C-*tagged* formam a base das demais tecnologias vistas neste capítulo, e é essencial entendê-las.

Desconsiderando-se então a utilização do padrão IEEE 802.1D-2004 como tecnologia de rede de provedor, serão consideradas as seguintes tecnologias:

— VLANs C-*tagged*;
— *Provider Bridges* (PB), conforme as cláusulas 15 e 16 do padrão IEEE 802.1Q-2014;
— *Provider Backbone Bridges* (PBB), conforme conforme as cláusulas 25 e 26 do padrão IEEE 802.1Q-2014;
— *Provider Backbone Bridges-Traffic Engineering* (PBB-TE), conforme as cláusulas 25 e 26 do padrão IEEE 802.1Q-2014.

5.2. VLANs C-TAGGED

O uso de VLANs C-*tagged* em redes *bridged Ethernet* de provedores de serviço consiste na simples extensão da aplicação dos princípios funcionais das VLANs C-*tagged* em redes privativas.

Sugere-se aos leitores que, neste ponto, procedam a uma revisão do capítulo 3 deste livro, particularmente do item 3.12, relativo à constituição de VLANs C-*tagged*.

Na Figura 3.19 do capítulo 3 anterior, encontra-se o formato de um quadro MAC contendo o campo C-VLAN *tag* (ou Q-*tag*), além do formato desse campo. O campo C-VLAN *tag* possui o (sub) campo VID, também designado por VLAN-ID, que identifica cada VLAN C-*tagged* da rede. O C-VLAN *tag* possui também o subcampo *User Priority* com 3 bits, o que possibilita a definição de uma escala com 8 níveis de prioridade, utilizada para fins de administração da qualidade de serviço diferenciado na rede.

A inadequabilidade da simples extensão da definição de VLANs C-*tagged* para sua utilização em redes de provedor de serviço é consequência de dois importantes aspectos:

— Compartilhamento de endereços MAC e de VLAN-IDs entre os clientes e o provedor de serviços;
— Escalabilidade inadequada em função do tamanho do campo VLAN ID.

Como não existe, nesse caso, separação entre VLAN *tags* (e entre VLAN IDs, em consequência), ocorre a superposição funcional da administração desses campos entre o usuário (que necessita manter o controle sobre suas VLANs), e o provedor de serviços (que requer também a administração de suas VLANs). Essa questão torna evidentemente problemática a simples extensão de VLANs C-*tagged* para redes de provedor.

Por outro lado, o campo VLAN ID possui apenas 12 bits, o que possibilita a existência teórica de 4096 VLANs (4094, na prática, devido ao uso de dois VLAN IDs para fins administrativos). Tal quantidade de VLANs é viável (precariamente) para redes contidas em ambientes corporativos, mas torna-se inviabilizadora para redes de provedor atendendo a diferentes usuários, cada um requerendo um número considerável de VLANs.

Esses aspectos praticamente inviabilizam essa solução. Como será visto a seguir, as novas tecnologias de *bridging* em redes de provedores buscam, dentre outras funcionalidades, contornar essas dificuldades.

5.3. PROVIDER BRIDGES (PB)

O padrão IEEE 802.1ad, referente a *Provider Bridges* (PBs), encontra-se hoje obsoletado e incorporado ao padrão IEEE 802.1Q-2014, nas seções 15 e 16 desse padrão.

O PB define a forma pela qual os provedores de serviço podem prestar serviços a seus usuários mediante a utilização de VLAN *tags* adicionais em suas redes *bridged LANs.* Tais VLAN *tags* adicionais são utilizados para a identificação de instâncias do serviço prestado aos usuários, operando como um mecanismo de envelopamento dos VLAN *tags* das redes dos usuários.

Em consequência do que foi dito no parágrafo anterior com relação ao envelopamento dos VLAN *tags* do usuário (referidos como *C-tags*, onde C significa *customer*) pelos VLAN *tags* do provedor de serviço (referidos como *S-tags*, onde S significa *service*), o PB é também denominado *Q-in-Q bridging*, ou simplesmente *Q-in-Q*.

O VLAN ID de um *C-tag* é referido como C-VID, e uma VLAN do usuário identificada por um C-VID é denominada C-VLAN. Do mesmo modo, um *S-tag* contém um S-VID, que identifica uma S-VLAN.

Uma rede PB é referida como uma PBN (*Provider Bridged Network*).

5.3.1. **Formato Básico de Quadros MAC PB**

A Figura 5-1 apresenta o formato básico de um quadro MAC PB.

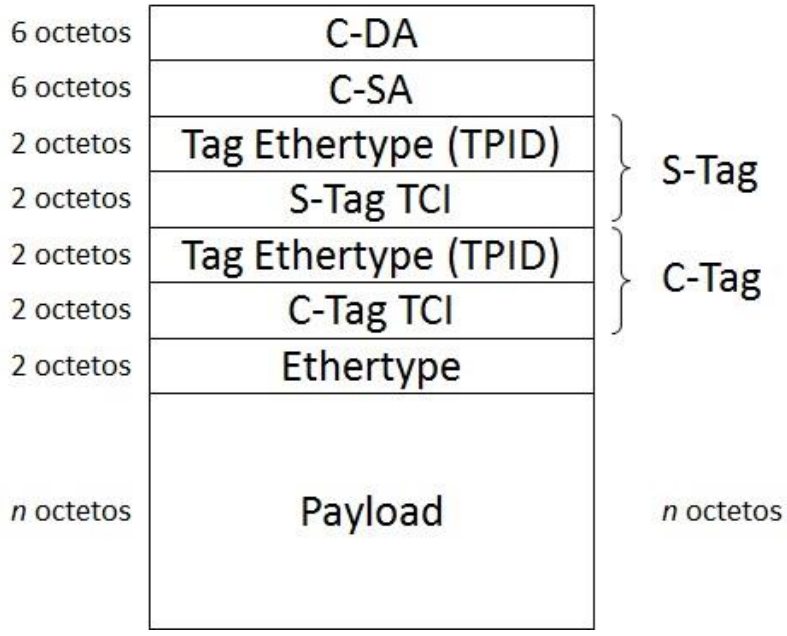


Figura 5-1 - Formato básico de um quadro MAC PB.

Nessa figura, TPID significa *Tag Protocol Identifier* e TCI *Tag Control Information*.

Como se observa nessa figura, o *S-tag* foi inserido entre o C-SA (endereço MAC de origem da rede global) e o *C-tag*, o que significa que o *C-tag* foi envelopado pelo *S-tag*.

Um quadro Q-*in*-Q é transmitido através da PBN contendo o *C-tag* do quadro de serviço enviado pelo usuário, sem alteração pelo provedor. Caso o

quadro de serviço enviado pelo usuário seja *untagged*, ele continua sem *C-tag* ao ser transmitido através da PBN.

As *bridges* da rede do provedor utilizam o *S-tag* para identificar uma instância de serviço PB, instância essa que representa um conjunto de *C-tags* a critério do usuário. O *S-tag* dessa instância de serviço envelopa os *C-tags* das C-VLANs que lhe correspondem.

Como um S-VID contém 12 bits, a exemplo de um C-VID, podem ser constituídas 4094 S-VLANs sobre a rede PB. Esse número representa uma limitação no número de instâncias de serviço oferecidas por uma rede PB de provedor.

Uma outra característica marcante e limitadora do PB é a de que a rede do provedor e as redes dos usuários utilizam os mesmos endereços MAC. Dessa forma, essas redes, em conjunto, aparecem como uma única grande rede global, o que pode tornar-se problemático.Por outro lado, um problema adicional de segurança se apresenta, pelo fato de que as informações de endereçamento passam a ser visíveis exteriormente ao domínio do usuário.

5.3.2. **Formato do S-tag**

O formato de um *C-tag* é aquele que representa um VLAN *tag* na Figura 3.19 do capítulo 3 deste livro. Os *S-tags* adotam esse mesmo formato, com duas alterações:

— O código *EtherType* que identifica o tipo de *tag* tem o seu valor alterado;
— O bit CFI (*Canonical Format Indicator*), que consta do campo TCI do C-*Tag*, perde a sua função, passando agora a significar DEI (*Drop Elegible Indicator*).

A Figura 5-2 representa o formato do *S-tag* no PB.

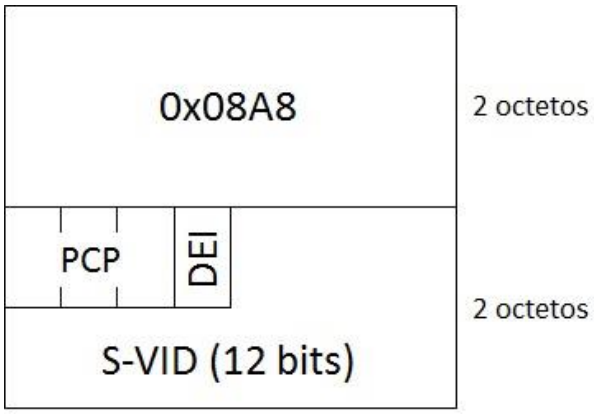


Figura 5-2- Formato do *S-tag* no PB.

Nessa figura, PCP significa *Priority Code Point* e DEI significa *Drop Eligible Indicator*.

Conforme o capítulo 3 anterior, um VLAN *tag* é identificado pelo código *EtherType* 0X8100. No caso do PB, esse código é utilizado pelos *C-tags*, enquanto os

S-tags são identificados pelo código *EtherType* 0x08A8. No PBB, os B-*tags* utilizam o mesmo código *EtherType* que os S-*tags* (0x88A8), enquanto os I-*tags* utilizam o código *EtherType* 0x88C7.

O bit CFI utilizado em VLAN *tags* de cliente (*C-tags*) é de utilidade, por exemplo, em configurações onde existem redes *Ethernet* operando juntamente com redes *Token Ring*. Esse bit passou a representar, no *S-tag*, a função DEI (*Drop Eligible Indicator*). O bit DEI tem função análoga à do bit DE (*Discard Elegibility*) no *Frame Relay*.

A codificação para elegibilidade de quadros para descarte no *S-tag* pode ocorrerutilizando-se exclusivamente o campo PCP (*Priority Code Point*), com 3 bits, e mantendo-se sem função o campo DEI. Como visto no subitem 3.12.9 do capítulo 3 deste livro, foram definidas quatro alternativas de codificação, designadas por 8P0D (*default*), 7P1D, 6P2D e 5P3D, onde P significa prioridade e D significa elegibilidade para descarte.

O bit DEI no VLAN *tag* pode ser utilizado, alternativamente, para indicar a elegibilidade para descarte. O uso do bit DEI permite que o VLAN *tag* codifique 8 prioridades distintas, cada uma delas com um indicador de elegibilidade para descarte.

Caso esse recurso seja suportado e esteja habilitado em uma determinada porta, nos quadros transmitidos por essa porta a elegibilidade para descarte será codificada no bit DEI. Por outro lado, um quadro recebido por essa porta será considerado elegível para descarte se o bit DEI estiver setado ou se a tabela de decodificação do VLAN *tag* configurada nessa porta denotar, pelo valor do PCP, que esse quadro é elegível para descarte.

Se o recurso não estiver habilitado, o bit DEI nos quadros transmitidos pela porta deverá ter o valor zero, e deverá ser ignorado nos quadros recebidos.

Por *default*, o recurso está desativado nas portas.

Como as *bridges* no interior de uma rede PB são transparentes aos *C-tags*, o domínio de *broadcasting* de uma S-VLAN é definida pelo S-VID.

5.3.3. **Conceitos Básicos no PB**

Aplicam-se à tecnologia PB alguns conceitos básicos, que dizem respeito aos seguintes aspectos:

— Componentes VLAN-*aware bridges*;
— Tipos de VLAN-*aware bridges*;
— Tipos de portas de VLAN-*aware bridges.*

5.3.3.1. Componentes VLAN-Aware Bridges

VLAN-*aware bridges* são *bridges* habilitadas para reconhecer, inserir ou remover VLAN-*tags*.

Para o caso de VLAN-*aware bridges* utilizadas no PB, consideram-se dois tipos de componentes VLAN-*aware bridges*:

— Componente C-VLAN: é um componente VLAN-*aware bridge* habilitado para reconhecer, inserir ou remover C-VLAN *tags*;
— Componente S-VLAN: é um componente VLAN-*aware bridge* habilitado para reconhecer, inserir ou remover S-VLAN *tags*.

5.3.3.2. Tipos de VLAN-Aware Bridges

Foram definidos os seguintes tipos de *bridges* na concepção do PB:

— C-VLAN *bridges* (ou simplesmente VLAN-*bridges*): são *bridges* contendo exclusivamente um único componente C-VLAN;
— S-VLAN *bridges*: são *bridges* contendo exclusivamente um único componente S-VLAN;
— *Bridges* de borda de provedor (*Provider Edge Bridges*): são *bridges* contendo um único componente S-VLAN e um ou mais componentes C-VLAN;
— *Bridges* de provedor (*Provider Bridges*): são S-VLANs *bridges* ou *bridges* de borda de provedor.

Observa-se que uma *bridge* de provedor encontra-se normalmente instalada em dependências do provedor, mas pode também se localizar em dependências do usuário. Nesse último caso, essa *bridge* de provedor pode ser opcionalmente operada e mantida pelo usuário.

A conexão entre as *bridged* LANs do provedor no PB e as redes do usuário ocorre por meio de uma *bridge* de provedor (S-VLAN *bridge* ou *bridge* de borda de provedor), localizada em dependências do provedor ou em dependências do usuário. É no componente S-VLAN dessa *bridge* que ocorre o reconhecimento, inserção e a remoção dos *S-tags*.

5.3.3.3. Tipos de Portas

Foram especificados os seguintes tipos de porta para o PB:

— *Customer Network Port* (porta de rede de cliente):uma CNP é uma porta de um componente S-VLAN de uma *bridge* de provedor ou uma porta interna do componente S-VLAN de uma *bridge* de borda de provedor, que recebe e transmite quadros de um único sistema de cliente, de um componente C-VLAN no interior de uma *bridge* de borda de provedor, ou de uma S-VLAN *bridge* controlada pelo cliente. Um sistema de cliente pode ser uma estação, uma C-VLAN *bridge* ou um roteador de cliente diretamente conectados à *Customer Network Port* de uma S-VLAN *bridge*;
— *Customer Edge Port* (porta de borda de cliente): uma CEP é uma porta de um componente C-VLAN de uma *bridge* de borda de provedor que o conecta ao equipamento CE do cliente;

— *Provider Network Port* (porta de rede de provedor): uma PNP é uma porta de um componente S-VLAN de uma *bridge* de provedor, que o conecta a outro componente S-VLAN de outra *bridge* de provedor, observando-se que uma S-VLAN *bridge* representa também um componente S-VLAN;
— *Provider Edge Port* (porta de borda de provedor): uma PEP é uma porta de um componente C-VLAN no interior de uma *bridge* de borda de provedor que o conecta ao componente S-VLAN também no interior dessa *bridge*.

5.3.4. Interfaces de Serviço para Clientes

O serviço PB pode oferecer a um cliente um ou mais tipos de interface, cada um provendo diferentes possibilidades de seleção e identificação de instâncias de serviço, de seleção de prioridade de serviço e de proteção de acesso de serviço. Em qualquer dos tipos, a segregação das diferentes instâncias de serviço é obtida em interfaces totalmente sob controle do provedor de serviço.

A proteção de acesso de serviço é obtida mediante o provisionamento de acessos de serviço sobre múltiplas LANs de acesso, o que será melhor abordado adiante neste item.

O acesso a uma determinada instância de serviço pode ser provido pelos seguintes tipos de interface:

— Interfaces de serviço baseadas em porta(*port-based)*;
— Interfaces de serviço *C-tagged*;
— Interfaces de serviço *S-tagged*.

5.3.4.1. Interface de Serviço Baseada em Porta (Port-Based)

Na interface de serviço baseada em porta, a CNP (*Customer Network Port*) utilizada provê uma única instância de serviço S-VLAN, conforme a configuração ilustrada na Figura 5-3.

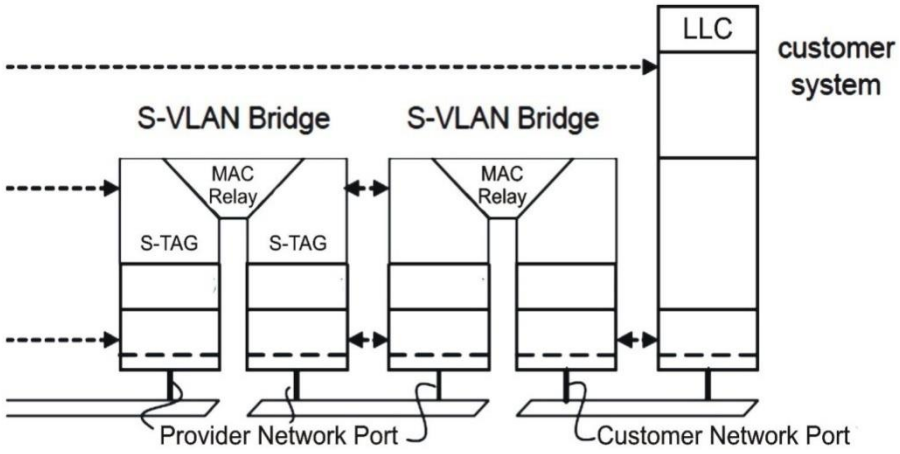


Figura 5-3 - Configuração de interface de serviço baseada em porta em PB.

A Figura 5-4 apresenta a constituição da interface de serviço baseada em porta correspondente à configuração ilustrada na figura anterior.

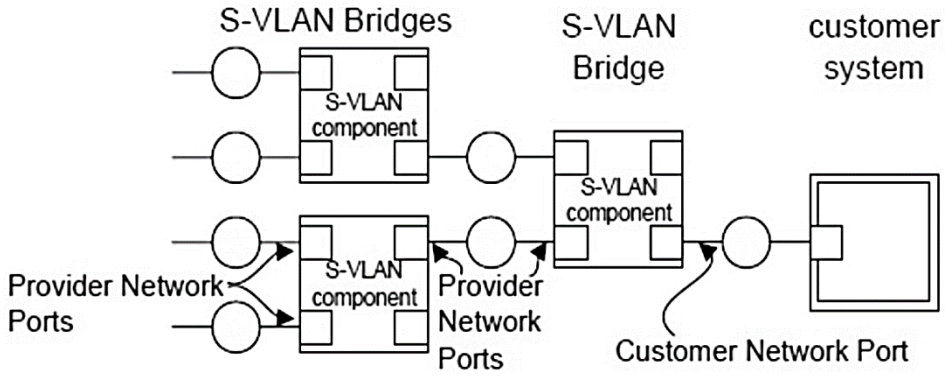


Figura 5-4- Constituição da interface correspondente à figura anterior.

Nessas figuras, o sistema de cliente conectado à *Customer Network Port* pode ser uma *bridge*, um roteador ou uma estação.

Essa interface é baseada em porta no sentido de que cada instância de serviço PB (identificada por um S-VID) é selecionada pela simples associação dessa instância de serviço a uma determinada *Customer Network Port* da *bridge* S-VLAN.

Cada instância de serviço (S-VID) utilizada está associada a uma *Customer Network Port* diferente. Quadros transmitidos para esse tipo de porta por um sistema de cliente C-VLAN-*aware* não incluem um S-VID, mas podem conter um S-*tag* sem o S-VID (ou seja, podem ser quadros *priority-tagged*) para permitir a identificação de sua prioridade de serviço.

A seleção de uma instância de serviço para interfaces de serviço baseadas em porta é feita mediante a configuração de uma *Customer Network Port* com um valor de PVID correspondente ao S-VID utilizado para identificar a instância de serviço. Na *Customer Network Port* utilizada, deve ser configurado o parâmetro *Acceptable Frame Types* com o valor A*dmit Only Untagged and Priority Tagged Frames*.

Como nessa forma de acesso à rede PB considera-se apenas a identificação da porta de entrada para a definição de uma instância de serviço (isto é, de uma S-VLAN), os *C-tags* dos quadros recebidos não são tomados em conta. Assim sendo, podem ser transmitidos inclusive quadros de cliente *untagged*. Isso significa que as redes PB podem ser conectadas, nessa forma de acesso, também a *bridges* IEEE 802.ID (*VLAN-unaware bridges*) de clientes.

5.3.4.2. Interface de Serviço C-tagged

As interfaces de serviço C-*tagged* possibilitam que sejam atribuídos diferentes S-VLAN IDs da rede PB aos quadros pertencentes a diferentes agrupamentos de C-VLANs de redes de usuário.

- **Configuração Básica**

Uma interface de serviço *C-tagged* é provida por uma *bridge* de borda de provedor contendo um ou mais componentes C-VLAN conectados a interfaces baseadas em portas internas em um único componente S-VLAN, como ilustrado na Figura 5-5 e na Figura 5-6.

A Figura 5-5 apresenta uma configuração de interface de serviço C-*tagged* em PB.

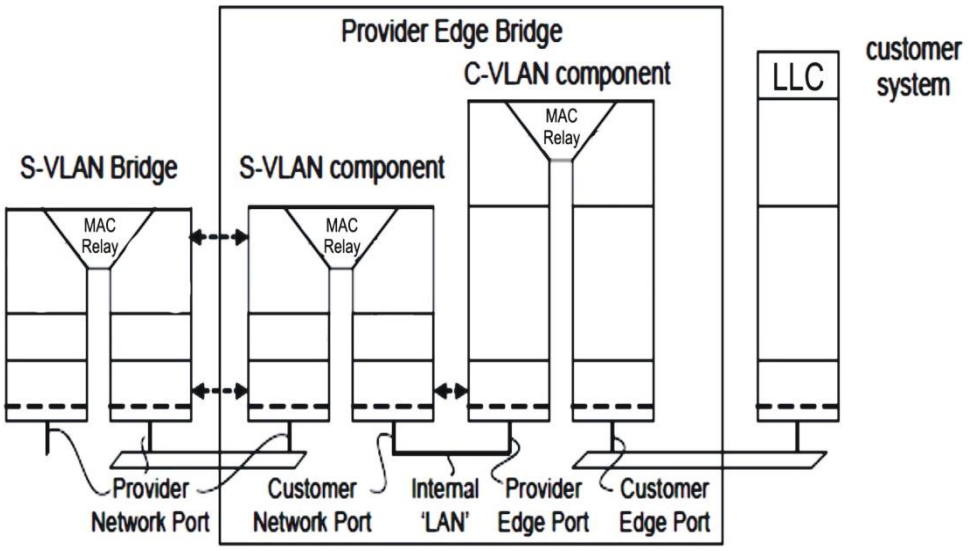


Figura 5-5- Configuração de interface de serviço C-tagged em PB.

A Figura 5-6 apresenta a constituição da interface de serviço *C-tagged* da configuração ilustrada na figura anterior.

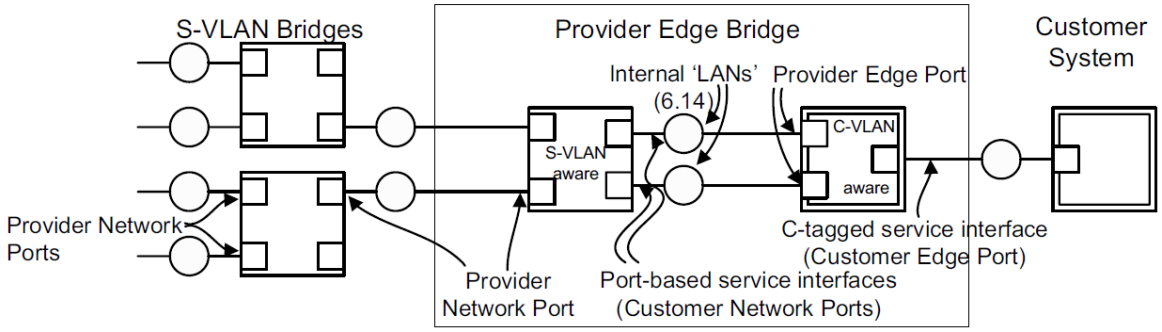


Figura 5-6 - Constituição da interface correspondente à Figura 5-5 anterior.

Essas figuras representam uma configuração em que a *bridge* de borda do provedor contém apenas um componente C-VLAN. Como mostra a segunda figura, os C-VLAN IDs são agrupados em dois subconjuntos, cada um deles correspondendo a uma das duas PEPs do componente C-VLAN. Observa-se que um subconjunto pode conter um ou mais C-VLAN IDs.

A Figura 5-7 apresenta a constituição de uma configuração de acesso por interfaces de serviço *C-tagged* com múltiplos componentes C-VLAN, que podem atender a um usuário ou a múltiplos usuários.

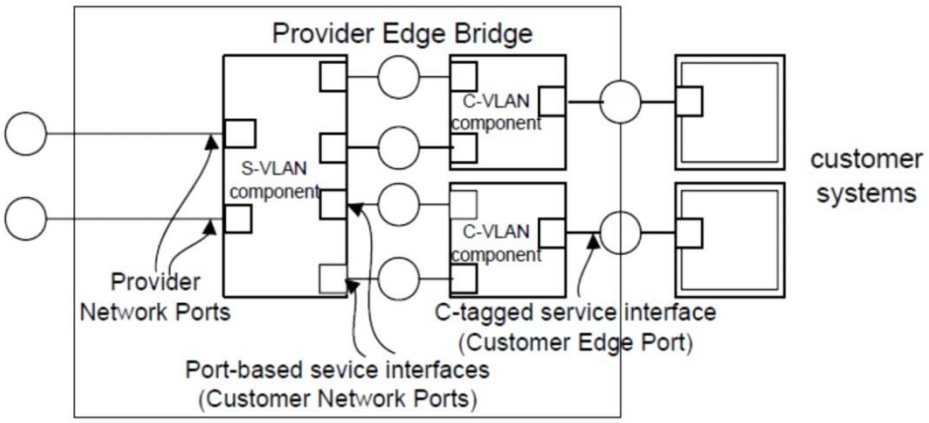


Figura 5-7 - Interface de serviço C-tagged com múltiplos componentes C-VLAN.

- **Funcionamento da Interface de Serviço C-tagged**

Para o funcionamento da interface de serviço C-*tagged* o cliente deve informar ao provedor a constituição de cada um dos agrupamentos de C-VLANs desejados, e quais *bridges* de borda de provedor irão atender a cada um desses agrupamentos.

A seleção de instâncias de serviço para interfaces de serviço C-*tagged* é provida por um componente C-VLAN em uma *bridge* de borda de provedor. O componente C-VLAN verifica o C-VID para direcionar um quadro entrante por uma *Customer Edge Port* (CEP) a uma *Provider Edge Port* (PEP) que suporta uma instância de serviço específica (ou seja, identificada por um valor de S-VID). Essa verificação ocorre mediante consulta à tabela de registro de C-VIDs (*C-VID Registration Table*) contida na CEP.

Os quadros de no máximo uma C-VLAN podem ser transportados *untagged* em uma instância de serviço.

Uma conexão interna entre uma CNP no componente S-VLAN e uma PEP na componente C-VLAN é criada para cada instância de serviço. A CNP é identificada pelo identificador da CEP mais o valor de S-VID usado como PVID da CNP. A PEP é identificada pelo identificador de CEP mais o valor de C-VID utilizado como PVID da PEP. Esses valores de PVID e a conexão entre a CNP e a PEP são estabelecidas mediante entradas (*entries*) recíprocas na tabela de registro de C-VID e na tabela de configuração de PEP, da maneira descrita a seguir:

— Cria-se uma entrada na tabela de registro de C-VID para cada C-VLAN suportada no componente C-VLAN associada a uma CEP. A combinação do identificador da CEP com um C-VID identifica uma PEP. A entrada contém, entre outros parâmetros, um valor de S-VID que corresponde ao

PVID da CNP, o qual associa a PEP com uma CNP em particular. Observa-se que a combinação CEP/C-VID identifica uma PEP única, porém múltiplas combinações CEP/C-VID podem identificar uma mesma PEP caso as entradas para essas combinações de CEP/C-VID contenham o mesmo valor de S-VID. Esse mapeamento "muitos em um" de CEP/C-VIDs em uma mesma PEP permite constituir uma agregação (*bundling*), ou seja, mapear diversos C-VLANs na mesma instância de serviço.

— Cria-se uma entrada na tabela de configuração da PEP para cada S-VID que corresponda a uma instância de serviço acessada pelo componente C-VLAN. A combinação do identificador da CEP com o S-VID identifica uma CNP. A entrada contém, entre outros parâmetros, um valor de C-VID que corresponde ao PVID da PEP, o qual associa a CNP com uma PEP em particular. Observa-se que a combinação CEP/S-VID identifica apenas uma CNP, porém múltiplas combinações CEP/S-VID podem identificar a mesma CNP caso as entradas correspondentes a essas combinações CEP/S-VID contenham o mesmo valor de C-VID. Esse mapeamento do tipo "muitos em um" de CEP/S-VIDs em uma mesma CNP permite configurar VLANs assimétricas e pode ser utilizado para constituir uma conectividade do tipo multiponto com raiz.

A tabela de registro de C-VIDs (*C-VID Registration Table*) reside na *Customer Edge Port* (CEP) desse componente C-VLAN. Nessa tabela, encontra-se o mapa que associa valores de C-VID a valores de S-VID.

A tabela de registro de C-VIDs provê a funcionalidade equivalente a:

— Configurar o PVID da *Customer Network Port* (CNP) do componente S-VLAN;
— Acrescentar a *Provider Edge Port* (PEP) do componente C-VLAN ao *Member Set* da C-VLAN;
— Acrescentar a *Provider Edge Port* (PEP) e/ou a *Customer Edge Port* (CEP) ao *untagged set* dessa C-VLAN (caso se deseje que os quadros enviados a essa porta sejam transmitidos *untagged* para essa C-VLAN).

O agrupamento "muitos em um" de C-VIDs para um S-VID mencionado acima pode ser configurado de forma tal que a totalidade dos C-VIDs atendidos pela CEP sejam agrupados em um único grupo. Esse agrupamento "todos em um" cria uma interface de serviço similar, mas não idêntica, à interface de serviço baseada em porta.

Em uma interface de serviço C-*tagged*, ao contrário do que ocorre na interface de serviço baseado em porta, o componente C-VLAN possibilita a modificação do C-*tag*. Essa modificação pode ocorrer pela inserção de um C-*tag* com o valor do PVID em quadros *untagged*, pela inserção do valor do PVID no *tag* de quadros *priority-tagged*, ou pela remoção do C-*tag* de quadros a serem enviados através da

Provider Edge Port(PEP). O componente C-VLAN, contudo, não envia quadros com o C-VID nulo.

Observa-se que uma interface de serviço baseada em porta pode ser implementada em uma *bridge* de borda de provedor mediante a configuração da porta de cliente como uma *Customer Network Port* (CNP) e não como uma *Customer Edge Port* (CEP). Neste caso, não há um componente C-VLAN associado.

Quando se deseja transmitir um quadro de uma C-VLAN após esse quadro alcançar o componente C-VLAN da *bridge* de borda de provedor de entrada, o valor de C-VID é verificado para indicar por qual *Provider Edge Port* deve o quadro ser transmitido. Essa verificação ocorre na tabela de registro de C-VIDs desse componente C-VLAN.

5.3.4.3. Interface de Serviço S-tagged

Uma interface de serviço *S-tagged* pode ser oferecida por uma *bridge* S-VLAN controlada pelo provedor a uma *bridge* S-VLAN operada pelo cliente, ou a uma *bridge* de borda de provedor também operada pelo cliente. Nessa última opção, a *bridge* de borda de provedor controlada pelo cliente oferece interfaces de serviço *C-tagged* dentro da própria rede do cliente.

A Figura 5-8 representa a configuração de uma interface de serviço *S-tagged* em uma rede PB, na hipótese de uso de uma *bridge* S-VLAN controlada pelo cliente.

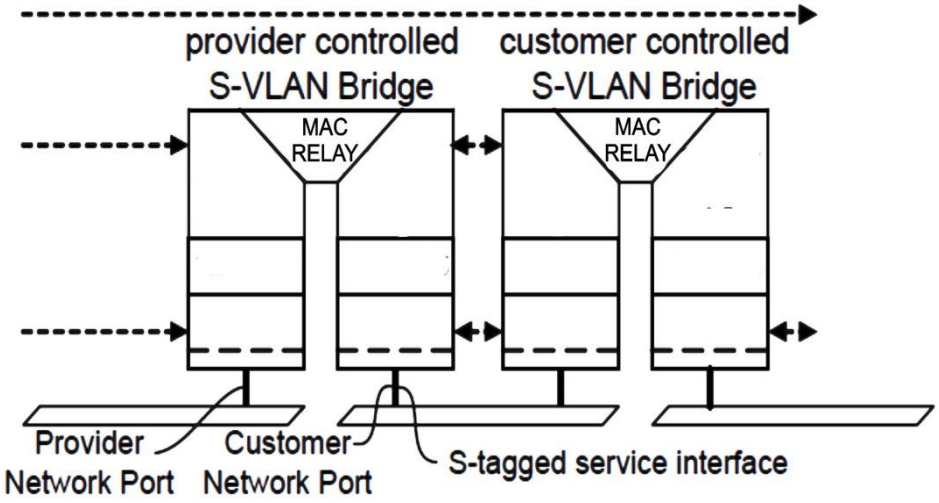


Figura 5-8- Configuração de interface de serviço S-tagged em PB.

A constituição de interface de serviço *S-tagged* em PB correspondente a essa figura encontra-se na Figura 5-9.

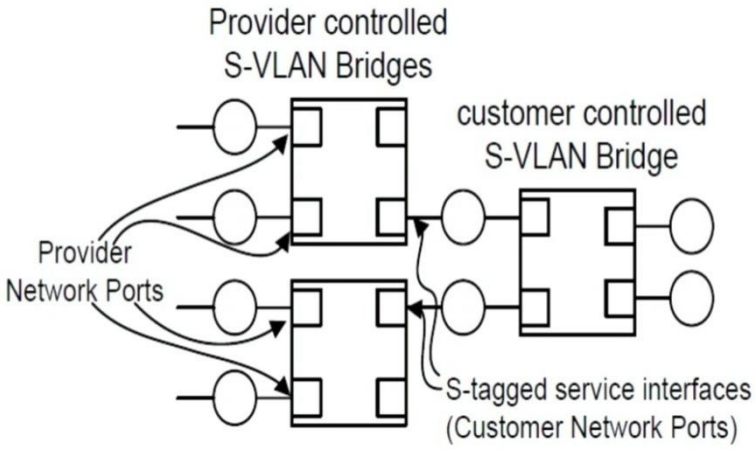


Figura 5-9 - Constituição de interface de serviço S-tagged em PB.

Nesse tipo de interface, a seleção de instâncias de serviços é feita pelo próprio cliente. O provedor recebe, por uma CNP (*Customer Network Port*), os quadros MAC oriundos do cliente já contendo *S-tags*.

O provedor de serviço pode atribuir S-VIDs de maneira independente dos S-VIDs atribuídos pelo cliente (ou por um outro provedor de serviço). As CNPs das *bridges* S-VLAN de borda controladas pelo provedor mapeiam diretamente cada valor de S-VID recebido em um valor de S-VID na rede PB mediante consulta à tabela de tradução de VIDs. Essas portas são incluídas como membros dos *Member Sets* das correspondentes S-VLANs.

5.3.4.4. Visão Global de Interfaces de Serviço em PB

A Figura 5-10 apresenta um exemplo de configuração de uma rede PB com duas interfaces de serviço *port-based* e uma interface de serviço *C-tagged*.

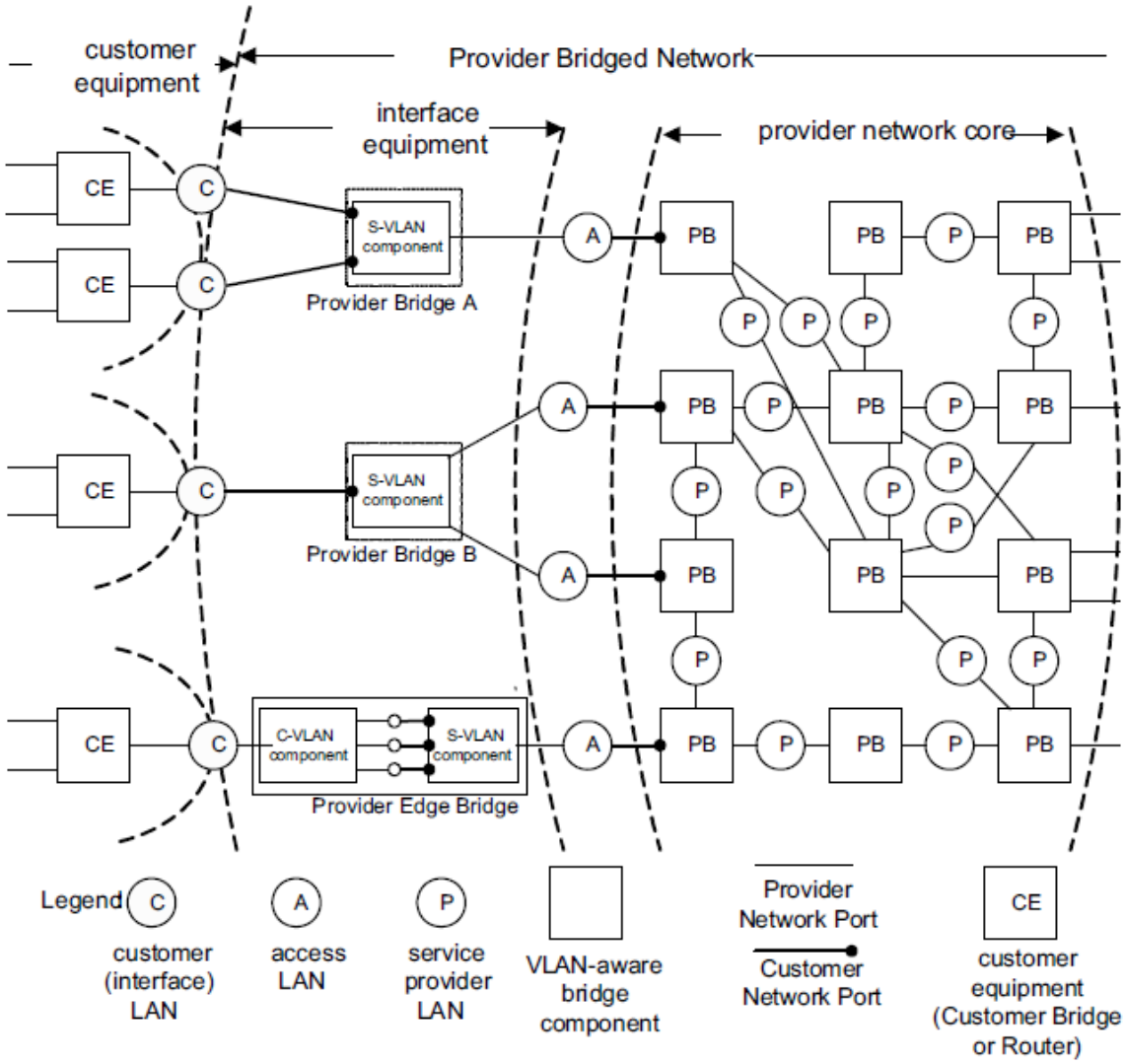


Figura 5-10 - Configuração de rede PB.

Nessa figura, equipamentos CE (*bridges*, roteadores ou estações) conectam-se à rede PB por meio de LANs de interface com o cliente. A figura apresenta a situação típica em que os equipamentos de interface da rede PB (componentes S-VLAN da *bridge* de provedor A, da *bridge* de borda de provedor B e da *bridge* de borda de provedor) se conectam ao núcleo (*core*) da rede PB utilizando uma ou mais LANs de acesso.

A topologia física arbitrária do interior da rede PB e a conectividade por ela oferecida para o suporte de distintas instâncias de serviço MAC PB é definida e gerenciada pelo provedor de serviços. A transmissão de quadros só pode ocorrer com a concordância do provedor, o que garante a segurança dessa transmissão.

O componente S-VLAN da *bridge* de provedor A da figura utiliza duas LANs de interface de cliente, o que possibilita o provimento de duas interfaces de serviço baseadas em porta distintas para dois equipamentos CE e utiliza *S-tags* para multiplexar as correspondentes instâncias de serviço em uma única LAN de acesso.

O componente S-VLAN da *bridge* de provedor B, por sua vez, atende a uma única LAN de interface com o cliente, através de uma única instância de serviço. A utilização de duas LANs de acesso objetiva o provimento de resiliência ao acesso. Quando múltiplas instâncias de serviço são providas em um único ponto de conexão de acesso ao cliente, os dois *links* de acesso podem ser utilizados.

A *bridge* de borda de provedor utiliza um único componente C-VLAN e provê uma única interface de serviço *C-tagged* para um CE, e utiliza *C-tags* nas respectivas LANs de interface com o cliente para selecionar entre múltiplas instâncias de serviço.

5.3.4.5. Prioridade de Serviço e Eligibilidade para Descarte

Para todas as interfaces de serviço a prioridade de serviço é selecionada com base na prioridade recebida para cada quadro, possivelmente regenerada com a utilização da tabela de regeneração de prioridades (consultar a Seção 6.5.9 do padrão IEEE 802.1Q-2014). O mecanismo para a determinação da prioridade recebida varia em função do tipo de interface de serviço empregado.

Para o caso de interface de serviço baseada em porta, a seleção de prioridade utiliza a prioridade sinalizada pelo método de acesso ao meio da porta. Em caso de inexistência dessa sinalização, utiliza-se um valor único de prioridade para todos os quadros.

Uma outra alternativa para as interfaces de serviço baseadas em porta é a sinalização, pelo equipamento CE, da prioridade a ser adotada pela rede PB para um quadro, mediante o envio desse quadro *S-tagged*, contendo a prioridade desejada e com o S-VID igual a zero (ou seja, um quadro do tipo *priority-tagged*). Como foi visto anteriormente, o *S-tag* desse quadro pode indicar, além da prioridade, também a elegibilidade para descarte do quadro que o contém.

As interfaces de serviço *C-tagged* utilizam, para os *S-tags*, as prioridades e a elegibilidade para descarte conduzidas nos *C-tags* dos quadros recebidos. A interface de serviço *C-tagged* pode utilizar um único valor de S-VID para um conjunto contendo todos os C-VIDs recebidos, e assim funcionar como uma interface de serviço baseada em porta. A diferença reside na forma de sinalização de prioridade e de elegibilidade para descarte, conforme os parágrafos anteriores. Reside também na possibilidade de modificação de *C-tags*, que existe na interface *C-tagged* e não existe em interfaces baseadas em porta.

Para as interfaces de serviço *S-tagged*, a prioridade de serviço e a elegibilidade para descarte dos S-*tags* de quadros na PBN são obtidas pelo mapeamento dos valores contidos no campo PCP (e, possivelmente, também no bit DEI) dos *S-tags* recebidos do cliente.

5.3.5. **Protocolos de Controle no PB**

Serão considerados, neste subitem, dois aspectos referentes à operacionalização de protocolos de controle de camada 2 em redes PB:

— Reserva e bloqueio de endereços MAC;
— Operacionalização de protocolos STP em redes PB.

Os protocolos de controle de camada 2 são identificados pelo campo Tipo, situado imediatamente após os campos de endereço MAC.

5.3.5.1. Reserva e Bloqueio de Endereços MAC

A Figura 3.21 do capítulo 3 deste livro apresenta a relação dos endereços MAC reservados para o componente *C-VLAN* de *bridges* de provedores.

A seção 8.6.3 do padrão IEEE 802.1Q-2014 reduziu a relação de endereços MAC reservados para o contexto de componentes S-VLAN de um total de 16 para um total de 11 endereços MAC. Dessa forma, foram desbloqueados alguns endereços MAC da relação da figura citada.

A relação de endereços MAC reservados para o contexto dos componentes S-VLAN das *bridges* de provedor encontra-se na Figura 3.22 do capítulo 3 deste livro.

Verifica-se que não consta daquela figura o endereço reservado *Bridged Group Address* (01-80-C2-00-00-00), que existe na tabela referente aos componentes C-VLAN. A conseqüência dessa inexistência é que a rede PB pode tornar-se transparente às BPDUs dos protocolos *spanning tree* das redes de cliente, uma vez que tais BPDUs não são bloqueadas pelas *bridges* de provedores (*bridges* de redes PB).

Observa-se que o endereço *Provider Bridge Group Address* (01-80-C2-00-00-08) foi mantido, sendo utilizado para o trânsito de BPDUs dos protocolos *spanning tree* no interior de PBNs.

5.3.5.2. Operação do MSTP em PB

A operação do MSTP no interior de uma rede PB independe da operação de qualquer protocolo *spanning tree* nas redes de cliente a ela conectadas.

Alcança-se essa independência mediante a utilização do endereço de grupo de *bridge* de provedor (*Provider Bridge Group Address*), igual a 01-80-C2-00-00-08, como endereço de destino de todas as MSTP PDUs transmitidas em todas as portas das *bridges* da rede PB. Os quadros recebidos pelas portas de rede de cliente e dirigidas ao endereço de grupo de *bridges* (01-80-C2-00-00-00) são submetidas ao mesmo tratamento de seleção de instância de serviço e despacho que os quadros de dados de cliente.

As BPDUs de cliente dirigidas ao endereço de grupo de *bridges* são transmitidas de maneira transparente pela rede PB, o que permite que o cliente utilize uma instância de MSTP ou RSTP totalmente independente da rede PB.

5.3.6. **Aprendizagem de Endereços MAC de Cliente**

Como nas redes PB os endereços MAC são únicos e visíveis para toda a rede, verifica-se uma grande carga de tráfego no processo de aprendizagem de endereços MAC. Qualquer mudança na rede do usuário poderá impactar o interior da rede do provedor. Esse fato constitui-se em uma limitação para as redes PB. Quando ocorre, por exemplo, a adição de um *host* na rede do usuário, o novo endereço MAC deve ser aprendido também por *bridges* da rede PB.

Quadros de serviço de cliente para uma dada instância de serviço MAC ficam restritos à parte da rede de provedor que suporta a topologia VLAN da S-VLAN associada, e, ademais, ficam restritos também ao ocorrer a aprendizagem dos endereços de origem dos quadros.

Em uma PBN que provê interfaces para cada cliente em uma pequena parcela do número total de interfaces de cliente, a necessidade de aprender endereços de estações finais de clientes pode ser significativamente reduzida mediante a aplicação de critérios de utilidade de filtragem aprimorada (*Enhanced Filtering Utility*) conforme a seção 8.7 do padrão IEEE 802.1Q-2014.

Essa aprendizagem pode ser também limitada pela aplicação do método SVL (*Shared VLAN Learning*), quando basta que um endereço MAC seja aprendido em uma S-VLAN de uma FID (*Filtering Identifiers*) para que essa aprendizagem possa ser utilizada nas demais S-VLANs da mesma FID.

5.4. **PROVIDER BACKBONE BRIDGES (PBB)**

A tecnologia referida como *Provider Backbone Bridges* (PBB), inicialmente especificada no padrão IEEE 802.1ah (2008), teve sua especificação incorporada no padrão IEEE 802.1Q-2014 em suas cláusulas 25 e 26.

Uma PBBN (*PBB Network*) é uma *bridged* LAN constituída por BEBs (*Backbone Edge Bridges*) na borda da PBBN, por BCBs (*Backbone Core Bridges*) no núcleo da PBBN e pelas LANs conectadas, sob o controle administrativo de um provedor único. É possível a interconexão de duas ou mais PBBNs de diferentes provedores de serviço.

A Figura 5-11 ilustra a configuração básica de uma PBBN.

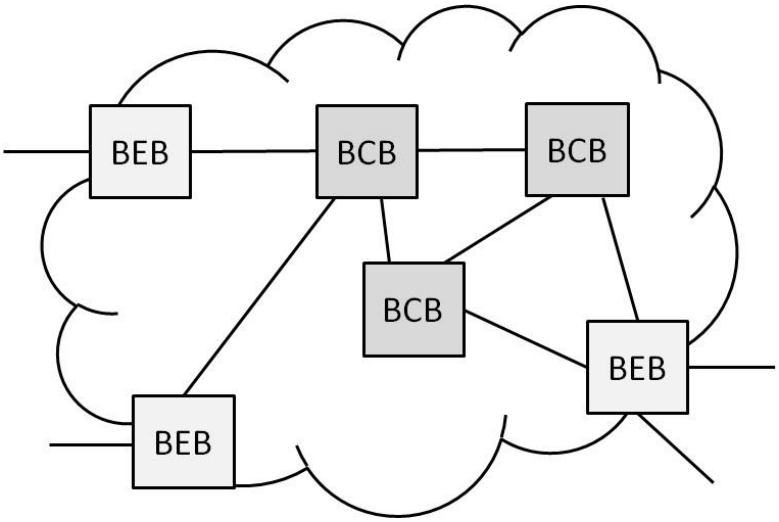


Figura 5-11 – Configuração básica de uma PBBN.

Os equipamentos de cliente podem estar conectadas às BEBs diretamente ou através de redes de acesso, como, por exemplo, uma rede PB. É possível também o transporte de tráfego S-VLAN não encapsulado no interior da PBBN, alocando-se parte do espaço de VIDs da PBBN para S-VLANs, ou seja, para a constituição de PBNs utilizando a infraestrutura de uma PBBN.

O PBB amplia a escalabidade de redes *bridged Ethernet* de provedores, no caso referidas como PBBNs. O PBB permite a constituição de até 2^{24} instâncias de serviço (praticamente 16 milhões de instâncias), em oposição às redes PB, que se limitam à constituição de até 2^{12}(até 4096) instâncias de serviço.

O funcionamento do PBB tem como base a adição de mais um cabeçalho MAC pelas BEBs da PBBN. Esse cabeçalho encapsula os quadros MAC que transitam pela PBBN, criando, assim, uma camada MAC adicional e independente. Por essa razão, o PBB é referido também como padrão *MAC-in-MAC.*

A rede PBB aplica internamente processos e protocolos de controle Ethernet de forma independente da rede Ethernet de cliente, com base nos valores de B-tag, MAC B-DA e MAC B-SA. Esse é o caso do algoritmo de aprendizagem de endereços MAC e do protocolo MSTP (ou RSTP), MMRP (Multiple MAC Registration Protocol) e MVRP (Multiple VLAN Registration Protocol), como visto no capítulo 5 anterior deste livro.

5.4.1. **Formatos de Quadros em PBB**

A Figura 5-12 apresenta o formato básico de um quadro MAC PBB.

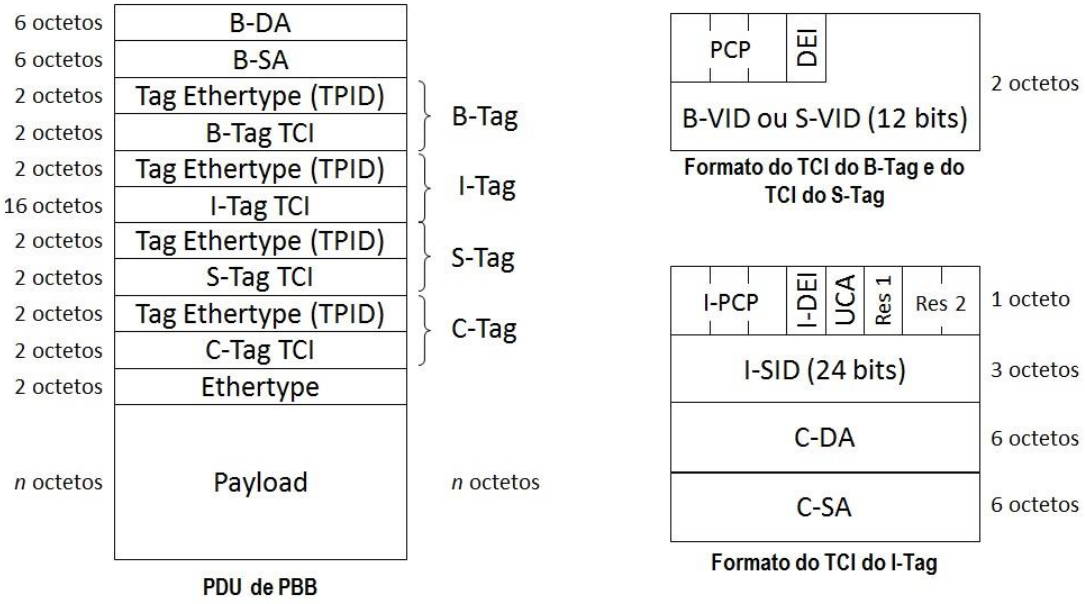


Figura 5-12 - Formato de um quadro PBB.

Nessa figura, TPID significa *Tag Protocol Identifier*,TCI significa *Tag Control Information*, PCP significa *Priority Code Point*, DEI significa *Drop Eligible Indicator* e UCA significa *Use Customer Addresses*.

5.4.1.1. PDU de PBB

Os endereços MAC B-DA e MAC B-SA representam, respectivamente, os endereços MAC de destino e de origem das PIPs (*Provider Instance Ports*). Para uma PIP que é conectada à BCB por uma LAN externa, esses endereços são efetivamente os endereços da PIP. Caso a PIP e a BCB pertençam a uma mesma BEB, esse endereço pode ser o endereço da própria *bridge*.

O PBB é constituído por componentes I (I-*comps*) e por componentes B (B-*comps*). O padrão IEEE 802.1Q-2014 especifica também a utilização de componentes T (T-*comps*).

Observa-se que, no PBB, *I* significa *instance* (instância de serviço), enquanto *B* significa *backbone* e *T* significa *transparent*.

Cada I-*comp* ou T-*comp* é responsável pelo encapsulamento dos quadros recebidos dos clientes e atribui cada quadro a uma instância de serviço *backbone* (*Backbone Service Instance* – BSI), identificada por um I-SID (*Backbone Service Instance Identifier*).

O quadro do cliente é encapsulado por um cabeçalho *MAC-in-MAC*, cabeçalho esse que engloba um I-*Tag*, um B-*Tag* e os endereços MAC *backbone* (B-SA e B-DA).Compete aos I-*comps* ou T-*comps* a inserção e a retirada dos I-*Tags*, que incluem o TPIDs (*Tag EtherType*) e o TCI do I-*Tag*, cujo formato e conteúdo se encontram na Figura 5-12 anterior.

Os I-SIDs, que possuem 24 bits de comprimento, possibilitam a criação de 2^{24} instâncias de serviço na PBBN,o que confere uma alta escalabilidade a esse tipo de rede.

Cabe aos B-*comps* a inserção e a retirada dos B-*Tags* (*Backbone VLAN tags*), que têm um papel análogo ao de S-*tags* utilizados em redes PB, e são utilizado sem conjunto com os endereços *backbone* MAC.

Os B-*Tags* possuem capacidade para a criação de 10^{12} B-VIDs (ou seja, de B-VLANs). Esse tipo de *tag* possui a mesma formatação dos *S-tags*.

Os quadros I-*tagged*, encapsulados pelos B-*Tags*, são transportados no interior das B-VLANs. São configuradas nas CBPs as *CBP Backbone Service Instance Tables* que mapeiam I-SIDs para B-VIDs e que são criadas como parte do serviço de provisionamento.

Somente as BEBs processam os *I-Tags* (e também os B-*Tags*), enquanto as BCBs processam exclusivamente os *B-Tags*, sendo transparentes aos I-*Tags*.

Um *B-comp* único é responsável pela comutação de quadros de cliente, encapsulados entre esse *B-comp* e *I-comps*/*T-comps*.

O número de *I-comps* (zero ou mais), *T-comps* (zero ou mais) e *B-comps* (zero ou um) dentro de uma BEB depende do tipo e do número de interfaces externas suportadas pela BEB. Os *I-comps*, *T-comps* e o *B-comp* podem estar na mesma BEB ou em BEBs diferentes.

5.4.1.2. Campo TCI do I-Tag

O campo I-*Tag* de um quadro MAC PBB utiliza o valor OX88E7 no seu subcampo *Type*/*Length*, que representa o IEEE 802.1Q *Backbone Service Instance Tag Type (*802.1Q *ITagType*) indicando o PBB. Lembra-se que esse campo tem o valor 0x8100 para o 802.1Q*TagType* indicando rede *C-tagged* e o valor 0x88A8 para o 802.1Q*STagType* (PB) indicando o PB.

Os subcampos I-PCP e I-DEI têm os mesmos significados que os desses subcampos no PB.

O bit UCA (*Use Customer Addresses*) constitui um *flag* que, quando setado para 1, sinaliza a uma Entidade de Mutiplexação de Instância de Serviço *Backbone* que ela deve utilizar os endereços contidos nos campos C-DA e C-SA.

O subcampo I-SID, conforme menção anterior, utiliza o comprimento de 24 bits, e representa o Identificador de Instância de Serviço *Backbone.*

Alguns valores de I-SID foram reservados para utilizações ou significados específicos. A Figura 5-13 representa tais valores.

I-SID value (hexadecimal)	Meaning/Use
0	Reserved for use by future amendments to this standard.
1	Default value—Unassigned ISID on a VIP.
2 through FF	Reserved for use by future amendments to this standard.
FFFFFF	Reserved for implementation use. This I-SID value shall not be configured as an identifier for a backbone service instance or transmitted in a Backbone Service Instance Tag header. This I-SID value may be used to indicate a wildcard match for the I-SID in management operations.

Figura 5-13 – Valores de I-SID reservados.

5.4.2. Componentes de PBNs

Neste subitem serão abordados os tipos de *bridges* em PBN e os tipos de portas de *bridge* em PBNs.

5.4.2.1. Tipos de Bridges em PBBNs

Os tipos de *bridges* em PBBNs encontram-se relacionados na Figura 5-14, onde constam os quatro tipos de BEB existentes.

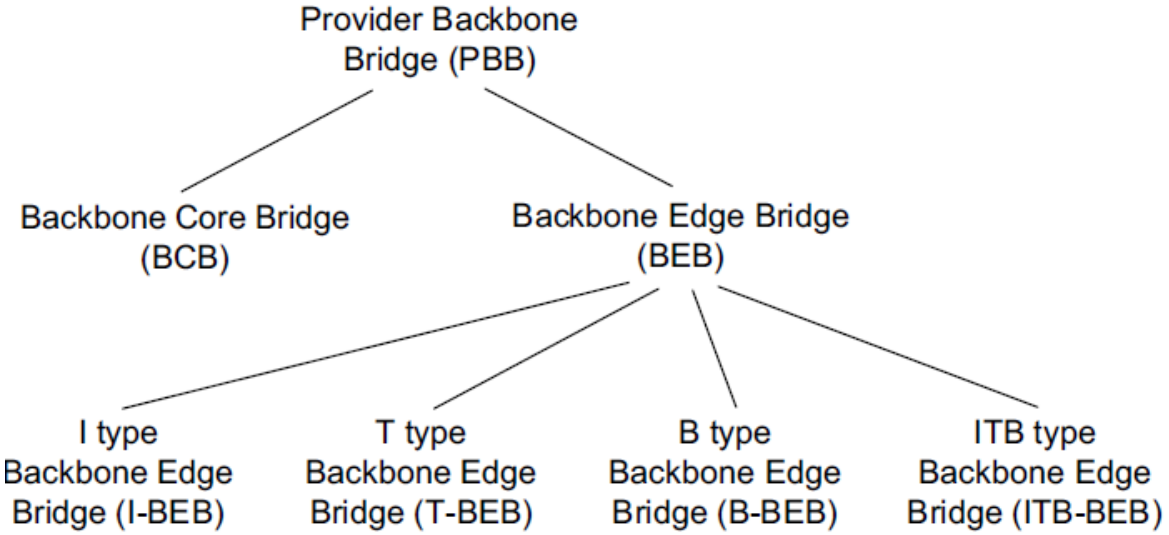


Figura 5-14 - Tipos de bridges em PBBNs.

Os tipos de BEBs podem classificar-se em I-BEB, T-BEB, B-BEB, and ITB-BEB, conforme a figura.

Uma BEB do tipo B (B-BEB) contém, apenas, um B-*comp*.

As I-BEBs e T-BEBs representam a interface efetiva da PBBN com os sistemas de cliente. Uma B-BEB pode conectar-se a uma ou mais I-BEBs e/ou T-BEBs, sendo que cada uma dessas conexões pode ocorrer por um ou mais *links* externos.

Uma T-BEB constitui um caso particular de uma *TPMR bridge*.*TPMR bridges*, como visto anteriormente neste livro, são *VLAN-unaware bridges* que possuem exatamente duas portas acessíveis externamente e que suportam um subconjunto de funcionalidades de uma MAC *bridge* com alto grau de transparência, como especificado nas seções 5.13 e 5.14 do padrão IEEE 802.1Q-2014.

Uma T-BEB possui um T-*comp*, sendo esse T-*comp* um componente TPMR (*Two-Port MAC Relay*) com uma porta PIP.

As BEBs do tipo ITB (ITB-BEBs) integram, em uma única unidade, as funções das B-BEBs e das I-BEBs e/ou T-BEBs. Uma ITB-BEB contém apenas um B-*comp* e um ou mais I-*comps* e/ou T-*comps*. O B-*comp* conecta-se a cada I-*comp* ou T-*comp* de uma ITB-BEB por um ou mais *links* internos.

Um caso particular de ITB-BEBs são as IB-BEBs, as quais contêm um B-*comp* e um ou mais I-*comps*.

5.4.2.2. Tipos de Portas de Bridges em PBBNs

As PBBNs utilizam os seguintes tipos de porta:

— Portas *Backbone* de Cliente (*Customer Backbone Ports* – CBPs) são portas de uma BEB que podem intercambiar (transmitir e receber) quadros I-*tagged* com múltiplos clientes e que podem atribuir B-VIDs e traduzir I-SIDs com base em I-SIDs recebidos;

— Portas de Instância Virtual (*Virtual instance Ports* – VIPs) são portas de *bridges* em um I-*comp* ou em um T-*comp* de uma BEB que provê acesso a uma única instância de serviço *backbone*;

— Portas de Instância de Provedor (*Provider Instance Ports* – PIPs) são conjuntos de VIPs que são suportadas por uma única instância de ISS (*Internal Sublayer Service*);

— Portas de Rede de Cliente (*Customer Network Ports* – CNPs) são portas de um I-*comp* ou de um T-*comp* de uma BEB que se conectam aos sistemas de clientes;

— Portas de Rede de Provedor (*Provider Network Ports* – PNPs) são portas de um B-*comp* de uma BEB que se conectam às BCBs, ou são portas de BCBs.

Observa-se que as denominações CNP e PNP são também utilizadas em PB.

5.4.3. Características do PBB

Além da capacidade de ampliar a escalabilidade da rede do provedor em termos de número de VLANs implementáveis (até cerca de 16 milhões de I-SIDs, como se viu), o PBB apresenta características adicionais, algumas representando vantagens funcionais sobre o PB.

A funcionalidade de tunelamento *MAC-in-MAC* do PBB significa a total separação entre os endereços do PBB e os endereços das redes C-*tagged* que se

conectam à PBBN. A conexão de redes C-*tagged* à PBN pode ocorrer diretamente ou através de uma PBN.

Os endereços MAC das redes de usuário atravessam o interior da PBBN transparentemente, envelopados nos quadros PBB.

Outra característica diz respeito ao uso dos 4094 B-VIDs da PBBN. O PBB possibilita a segregação de diferentes tecnologias e funcionalidades. Assim, por exemplo, parte da rede pode ser utilizada para o PBB-TE (o que será visto adiante neste capítulo), sendo o restante da rede dedicado a serviços no PBB (sem *Traffic Engineering*).

Considerando-se que o processo de tunelamento do PBB pode aplicar-se para qualquer tipo de quadro do usuário, os protocolos de controle das redes de cliente operam independentemente dos seus correspondentes na PBBN.

Um importante exemplo dessa independência de protocolos de controle é a aplicação do MSTP nas duas redes. A árvore MSTP de uma B-VLAN associada, via I-SID, a, por exemplo, uma S-VLAN, é independente da árvore MSTP dessa S-VLAN. As BPDUs da S-VLAN atravessam transparentemente a PBBN.

Os protocolos MMRP e MVRP aplicam-se normalmente na PBBN, inclusive em suas BCBs. Observa-se que o recurso de *snooping* ("bisbilhotagem") das tabelas IGMP não é factível no PBB, sendo então o MMRP a única alternativa para o envio eficiente de quadros *multicast*.

O MMRP permite que um quadro *multicast* endereçado a um endereço de grupo seja direcionado, na PBBN, exclusivamente pelos *links* onde foi declarada a necessidade da transmissão desse quadro.

O MVRP permite construir uma topologia B-VLAN que representa um subconjunto da topologia ativa (definida pelo MSTP) circunscrita às portas onde houve registro, pelo MVRP, de participação efetiva nessa B-VLAN. Dessa maneira, quadros *broadcast* ou quadros *unicast* sem endereço de destino conhecido inundam a rede dentro dos limites da topologia B-VLAN correspondente. Analogamente, a aprendizagem de endereços MAC *backbone* também ocorre dentro dos limites dessa topologia B-VLAN.

5.4.4. **Conectividade em Backbone VLANs**

A operação do MSTP no interior de uma PBBN é independente da operação de qualquer protocolo *spanning tree* em redes de provedor ou em redes de clientes conectadas a essa PBBN.

A aplicação do MVRP no contexto de cada topologia ativa de *spanning tree* em uma PBBN configura a extensão da topologia ativa de cada B-VLAN nessa PBBN. Essa configuração define os limites de conectividade entre os pontos de conexão de cliente que participam efetivamente de cada B-VLAN. Ainda que a uma *spanning tree* estejam mapeadas múltiplas VLANs (*C-VLANs*, *S-VLANs* ou *B-VLANs*), a filtragem do MVRP se aplica independentemente para cada uma dessas VLANs.

O MVRP pode reconfigurar essa conectividade em caso de alteração da topologia ativa.

No interior da PBBN, as MRPDUs do MVRP utilizam o endereço *Provider Bridge MVRP* (01-80-C2-00-00-0D). As MRPDUs do MVRP utilizando o endereço *Customer Bridge MVRP* (01-80-C2-00-00-21) são transportadas pela PBBN como se fossem quadros de serviço de clientes.

A operação do MVRP no interior de uma PBBN é independente da operação de qualquer protocolo de configuração no interior de redes de cliente ou de PBNs conectadas a essa PBBN.

5.4.5. **Interfaces de Serviço para Clientes**

Da mesma forma que no serviço PB, o PBB pode oferecer diferentes tipos de interface de serviço para os seus clientes. Também da mesma forma, a segregação dos diferentes instâncias de serviço é obtida em interfaces totalmente sob controle do provedor.

São possíveis os seguintes tipos de interface de serviço para clientes no PBB:

— Interface de serviço baseada em porta (*port-based*);
— Interface de serviço *S-tagged;*
— Interface de serviço I-*tagged*;
— Interface de serviço transparente (*Transparent Customer Service Interface*).

As interfaces de serviço para clientes oferecidos pelo PBB podem ser agrupadas conforme a Figura 5-15.

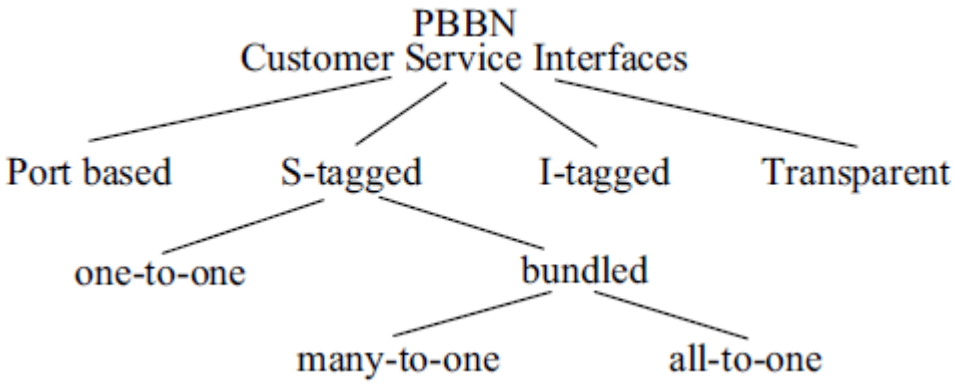


Figura 5-15 - Tipos de interface de serviço para cliente no PBB.

Os detalhes elucidativos dessa classificação de interfaces de serviço serão vistos nos próximos subitens.

5.4.5.1. Interface de Serviço Baseada em Porta

Uma interface de serviço baseada em porta (*port-based*) em uma PBBN provê o mesmo tipo de serviço daquele prestado por uma interface de serviço baseada em porta em PBNs, como visto anteriormente neste capítulo.

Isso significa que os quadros recebidos em portas de rede de cliente de I-*comps* de entrada podem ser quadros *C-tagged* ou quadros *untagged*, mas não quadros *S-tagged* (não existem PBNs conectadas nessas portas) com exceção dos quadros *priority tagged*.

Uma interface de serviço baseada em porta no PBB, a exemplo do que ocorre no PB, pode conectar-se a uma C-VLAN *bridge* (IEEE 802.1Q-2014), a uma IEEE 802.ID *bridge*, a um roteador ou a uma estação final.

Uma interface de serviço baseada em porta em uma PBBN é disponibilizada em uma CNP de uma BEB, conforme a Figura 5-16 e a Figura 5-17 adiante. O serviço provido por essa interface encaminha todos os quadros recebidos sem um S-*Tag* para o *backbone* em uma única instância de serviço *backbone*.

Todos os quadros recebidos com um S-*Tag* que não contenha um valor nulo de S-VID são descartados. Um S-*Tag* com valor nulo de S-VID serve apenas para indicar a prioridade do respectivo quadro.

Quando da utilização do serviço baseado em porta, todo o tráfego de uma rede de cliente que entra por uma CNP é transmitido na PBBN em uma única instância de serviço *backbone*, ou seja, utilizando um único valor de I-SID.

A Figura 5-16 representa uma configuração de interface de serviço baseada em porta em PBB.

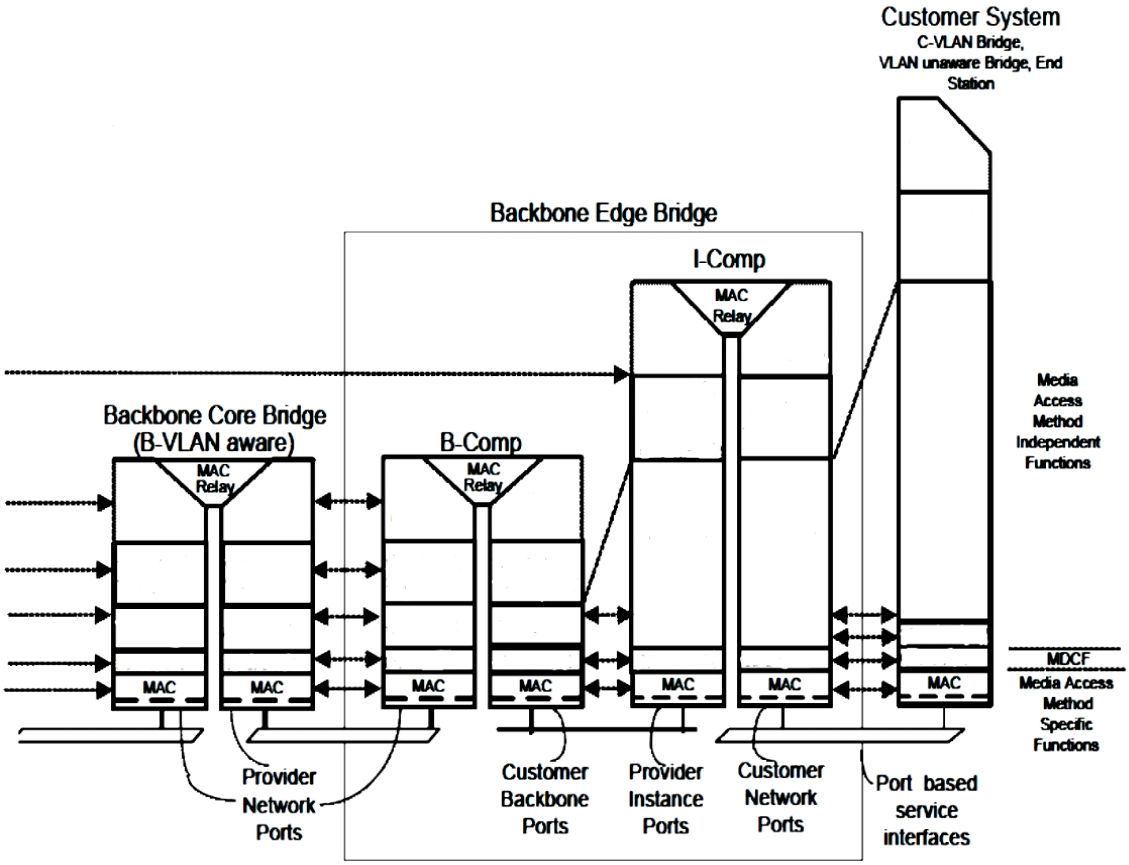


Figura 5-16 - Configuração de interface de serviço baseada em porta em PBB.

A Figura 5-17, por sua vez, apresenta a constituição de interface de serviço baseada em porta em PBB, contendo dois I-*comps* na respectiva BEB.

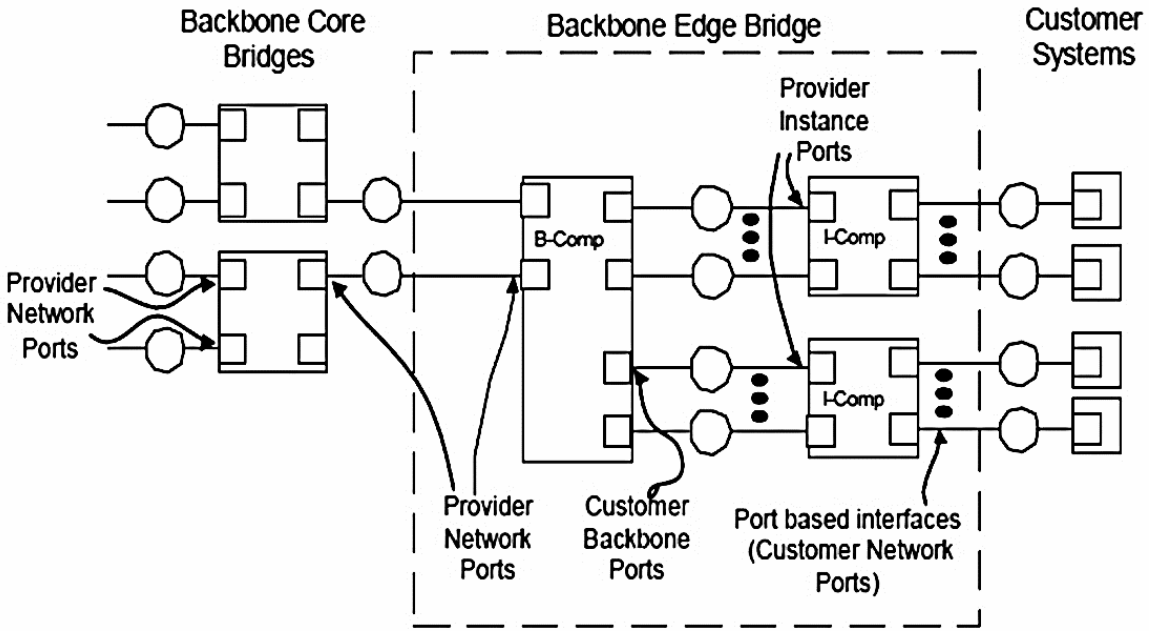


Figura 5-17- Constituição de interface de serviço baseada em porta em PBB.

Nessa figura estão representadas quatro interfaces de serviço baseadas em porta, sendo duas em cada um dos dois I-*comps* representados. Cada CNP (*Customer Network Port*) pode estar configurada como uma interface de serviço baseada em porta, sendo então mapeada em um I-SID escolhido.

Um valor de I-SID adotado deve ser mapeado em um B-VID, sendo que em um B-VID podem estar mapeados múltiplos I-SIDs.

5.4.5.2. Interface de serviço *S*-tagged

Uma interface de serviço S-*tagged* mapeia, numa porta CNP do I-*comp* da BEB, cada instância de serviço, identificada por um S-VID, em uma instância de serviço PBB, identificada por um I-SID. Existem dois tipos de interface de serviço S-*tagged* no PBB, uma realizando mapeamento um a um entre S-VIDs e I-SIDs, e a outra mapeando agrupamentos (*bundling)* de S-VIDs em um único I-SID.

No mapeamento um a um entre S-VIDs e I-SIDs, os S-*tags* não são transportados pelos quadros I-*tagged* no interior da PBBN. O valor de cada S-VID na saída da PBBN é obtido com base no valor de I-SID a ele mapeado.

Os valores do DEI e do PCP do quadro S-*tagged* entrante são transpostos para os correspondentes campos I-DEI e I-PCP do I-*Tag* do quadro I-*tagged*. No destino, ocorre operação contrária, sendo os valores do I-DEI e I-PCP transpostos para os campos DEI e PCP do quadro S-*tagged* de saída.

No mapeamento por agrupamento é indispensável o transporte dos S-*tags*. Os valores de DEI e PCP do quadro S-*tagged* entrante, além de serem mantidos nos *S-tags* envelopados nos quadros I-*tagged*, são transpostos para os correspondentes campos I-DEI e I-PCP dos I-*Tags.*

No mapeamento por agrupamento, o cliente (que administra a rede que se conecta à PBBN) define, para cada valor de I-SID, os agrupamentos de S-VIDs a serem mapeados. As correspondências entre I-SIDs e agrupamentos de S-VIDs são registradas no I-*comp* da BEB. Nas BEBs de saída, o valor de I-SID permite a identificação e a utilização da CNP de saída.

A Figura 5-18 apresenta uma configuração de interface de serviço *S-tagged* em PBB.

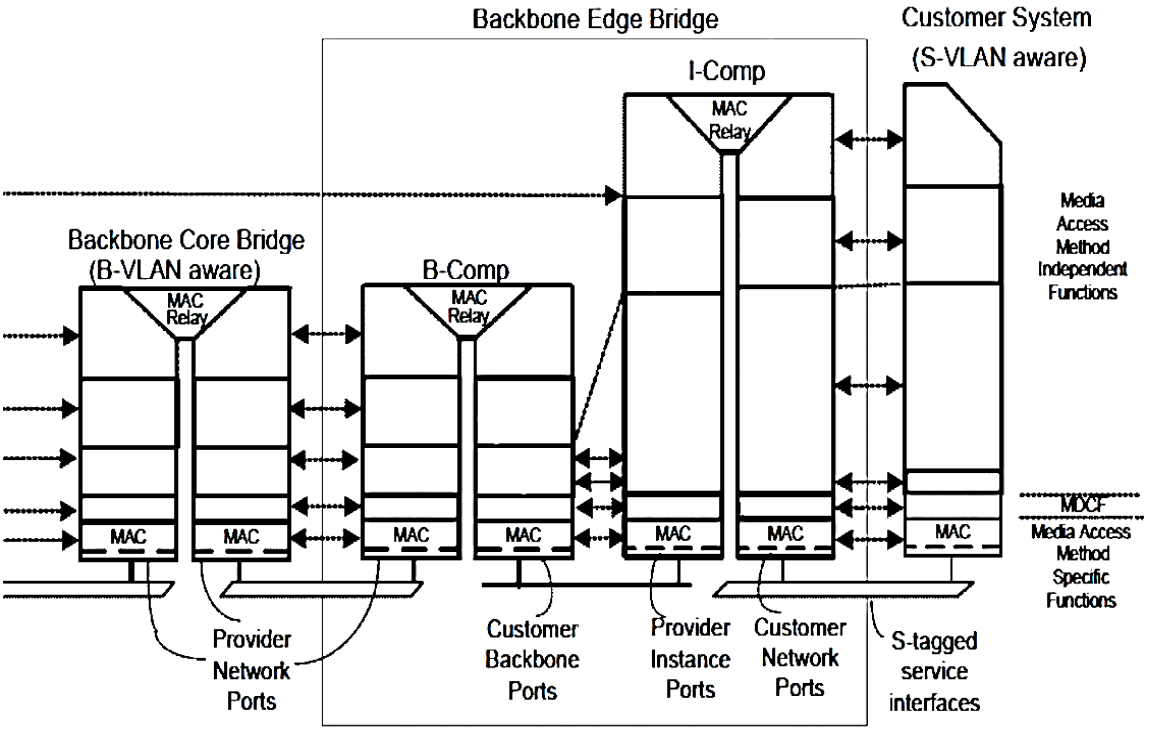


Figura 5-18 - Configuração de interface de serviço *S-tagged* em PBB.

A Figura 5-19 representa a constituição de interface *S-tagged* em PBB, contendo dois I-*comps*.

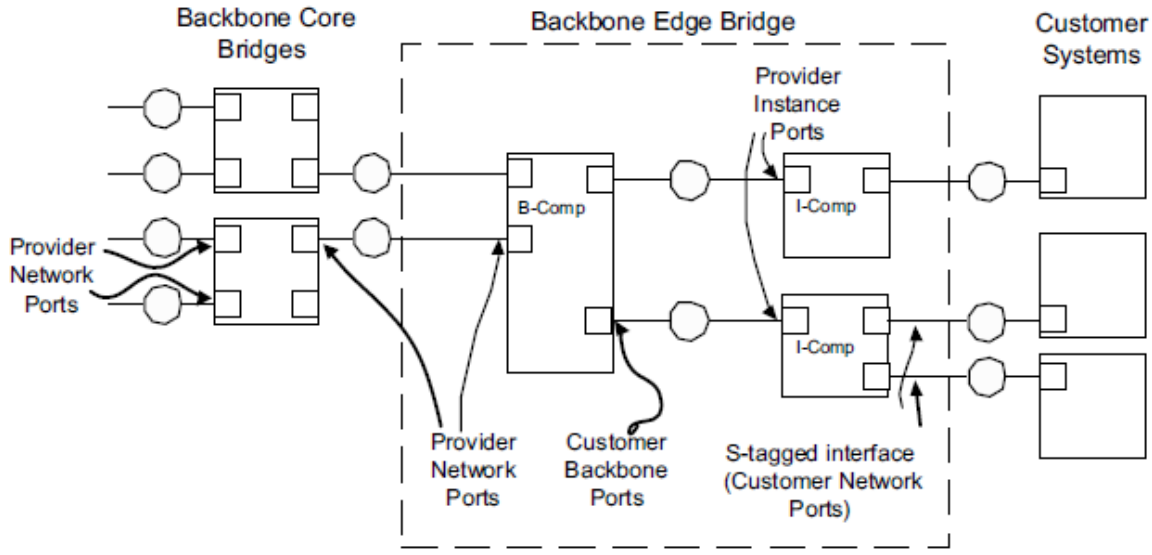


Figura 5-19 - Constituição de interface de serviço S-*tagged* em PBB.

5.4.5.3. Interface de Serviço C-tagged

O padrão IEEE 802.1Q-2014 não especifica interface de serviço C-*tagged* para o PBB, ocorrendo o acesso direto de redes *C-tagged* de usuários a cada PBBN por interfaces *port-based*. Esse padrão admite, contudo, a possibilidade de suporte a esse

tipo de interface por uma BEB, mediante a inclusão de um componente C-VLAN (ou seja, de um C-*comp*) nessa BEB.

Em particular, é possível que uma BEB suporte uma interface de serviço C-*tagged* para uma *bridged LAN*. O componente resultante é uma associação entre um B-*component* e um componente C-VLAN (sem um S-*component* explicito, como se observa).

Conforme fabricantes, essa possibilidade pode ser concretizada mediante a configuração de uma porta CEP (*Customer Edge Port*), própria das *Provider Edge Bridges*, em uma BEB. Essa solução possibilita o acesso direto de C-VLANs à CEP, inclusive com a possibilidade de *bundling* de C-VIDs em S-VIDs. No interior da BEB há uma relação de um para um entre um S-VID e um I-SID. Como essa relação é de um para um, o S-*tag* é desnecessário e é descartado, resultando, na PIP da BEB, quadros I-*tagged* envelopando diretamente C-*tags*.

5.4.5.4. Interface de Serviço I-tagged

Uma PBBN pode prover uma interface de serviço I-*tagged* para conexão a outra PBBN ou a uma porta PIP (*Provider Instance Port*) de um usuário. O acesso às instâncias de serviço *backbone* é controlado pela configuração de tabelas de instâncias da CBP (*Customer Backbone Port*) da BEB de provedor utilizada. Cada I-SID na interface de serviço I-*tagged* identifica uma instância de serviço *backbone*. O cliente deve prover o endereço B-DA para quadros entregues em uma interface de serviço I-*tagged*.

Observa-se que essa BEB de provedor é sempre uma B-BEB, isto é, possui sempre apenas um B-*comp*.

Do lado do usuário encontra-se uma BEB por ele controlada, que pode ser uma I-BEB (ou seja, possui um I-*comp*) ou uma B-BEB (ou seja, possui um B-*comp*).

A Figura 5-20 apresenta uma configuração de uso de interface de serviço I-*tagged*, quando a BEB controlada pelo usuário é uma B-BEB (figura superior) ou uma I-BEB (figura inferior).

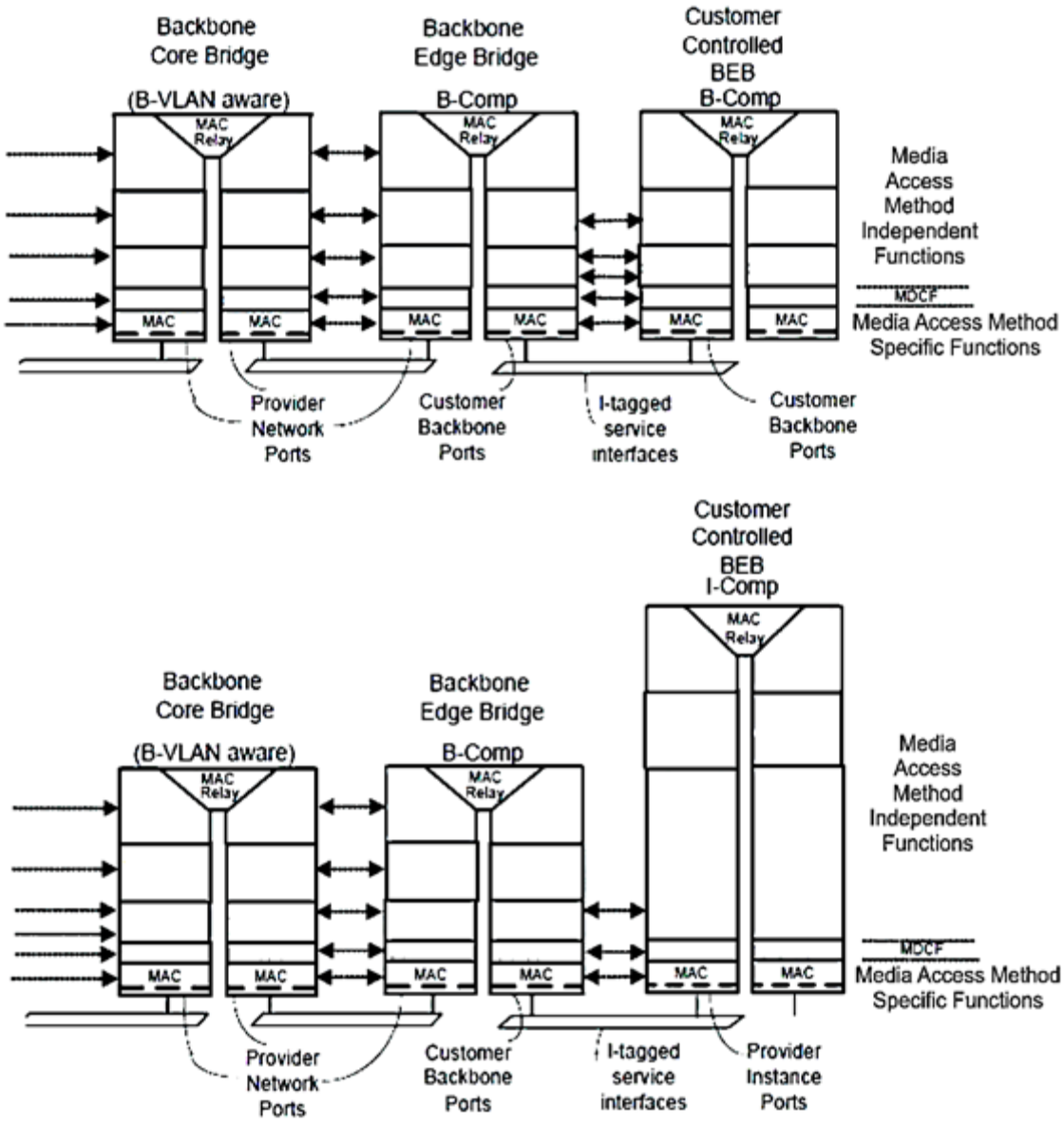


Figura 5-20 - Configurações de interface de serviço I-tagged em PBB.

A Figura 5-21 representa, de forma integrada, a constituição de uma configuração com interface I-*tagged*, conectada a uma BEB (B-BEB ou I-BEB) controlada pelo usuário.

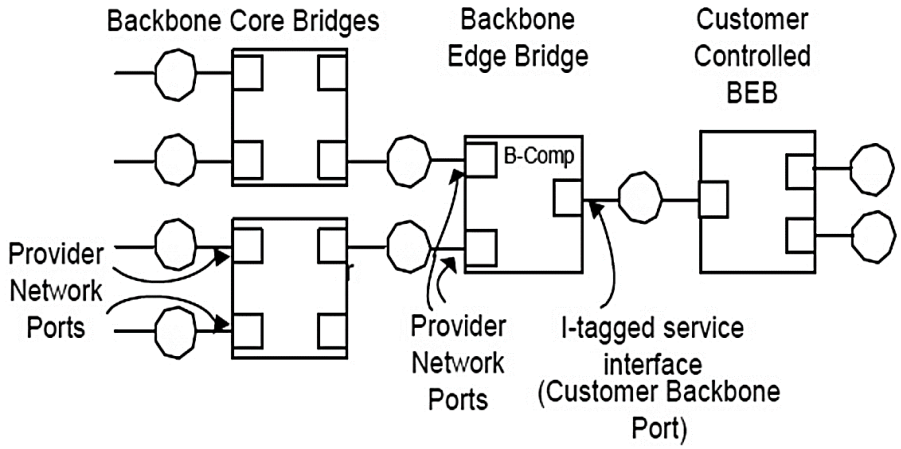


Figura 5-21 - Constituição de interface de serviço I-tagged.

Quando a BEB do usuário é uma I-BEB, o quadro é enviado pelo usuário já contendo um I-*Tag*, isto é, como um quadro I-*tagged*. O *I-Tag* enviado pelo usuário é mapeado em um I-*Tag* do provedor, sendo o quadro inserido na PBBN (B-VLAN) contendo um B-VID associado ao I-SID contido no I-*Tag* do provedor.

Na saída do quadro da PBBN ocorre uma operação inversa à da entrada. O quadro atinge a porta de rede de provedor de saída com base no seu MAC B-DA, e filtrado pela B-VLAN utilizada. O B-*comp* da BEB de saída mapeia o I-*Tag* do provedor do quadro no correspondente I-*Tag* do usuário, e o envia pela sua respectiva porta *backbone* de usuário (CBP).

No outro caso, no qual a BEB do usuário é uma B-BEB, o quadro enviado pelo usuário é do tipo B-*tagged*. Nesse caso, além do mapeamento do *I-Tag* do quadro oriundo do usuário (envelopado pelo respectivo B-*Tag*) em um *backbone* I-*Tag*, o B-BEB do provedor deve também mapear o B-*Tag* recebido do usuário em um *backbone* B-*Tag* (B-*Tag* do provedor).

5.4.5.5. Interfacede Serviço Transparente

O padrão IEEE 802.1Q-2014 especifica a interface de serviço transparente, disponibilizada em uma CNP de uma BEB. Essa BEB é constituída por um B-*comp* e um ou mais T-*comps*. Uma interface de serviço transparente transmite todos os quadros para o *backbone* em uma única instância de serviço *backbone*, identificada por um I-SID, da mesma forma que uma interface de serviço baseada em porta.

A transparência oferecida pela interface de serviço transparente, no entanto, é superior à da interface de serviço baseada em porta. Não ocorre a restrição imposta a quadros S-*tagged* de clientes na interface de serviço transparente. Assim, a esse tipo de interface podem ser conectados inclusive S-VLAN *bridges*, além das C-VLAN *bridges*, IEEE 802.1D *bridges*, roteadores e estações finais.

Uma interface de serviço transparente não reconhece, acrescenta, remove, ou modifica S-*tags* ou C-*tags* em quadros de cliente. Ela não aprende endereços de cliente nem filtra quadros com base em endereços MAC, com exceção dos

endereços MAC reservados para o componente TPMR relacionados na Figura 3.23 do Capítulo 3 deste livro e do endereço de gerenciamento do componente TPMR.

5.4.6. **Exemplo de PBBN**

A Figura 5-22 apresenta um exemplo de configuração de uma PBBN.

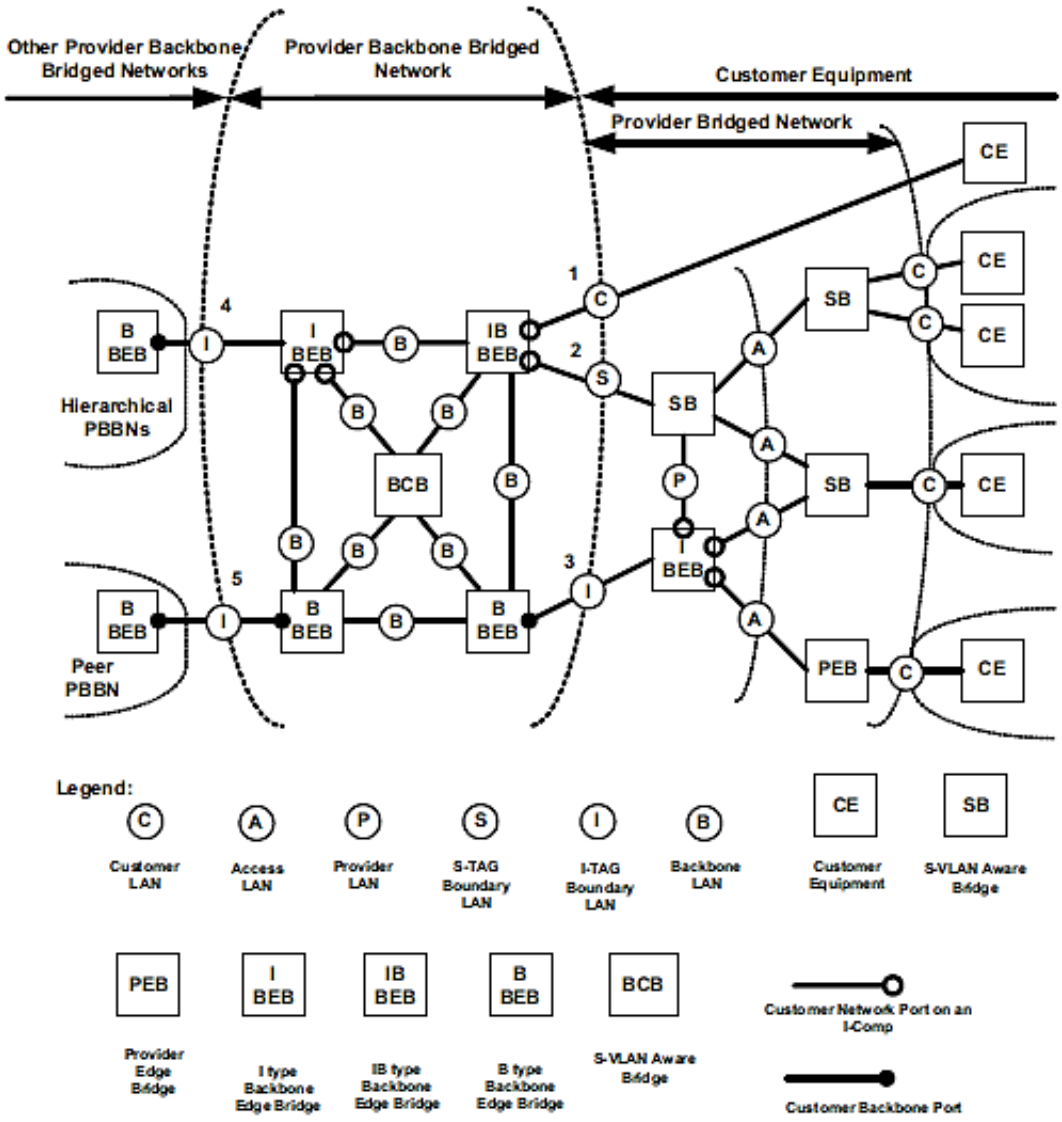


Figura 5-22 – Exemplo de configuração de uma PBBN.

A rede apresentada nesse exemplo é dividida em três áreas relevantes:

— A área de equipamentos de cliente, que possui interfaces com a PBBN;
— A área da PBBN, de propriedade do provedor do serviço *backbone* e que está interconectada com os equipamentos de cliente e com outras PBBNs;

— Uma área de outras PBBNs, de propriedade do mesmo ou de outros provedores de serviço *backbone*, permitindo que a rede atinja um nível de escalabilidade praticamente ilimitada.

Para um melhor entendimento desse exemplo, recomenda-se aos leitores uma revisão dos tipos de interfaces de serviço para clientes anteriormente apresentadas, ou recorrer à seção 26.2 do padrão IEEE 802.1Q-2014.

5.4.7. **Interconexão de PBBNs**

É importante a possibilidade de interconexão de PBBNs operadas por diferentes organizações. As interfaces entre PBBNs são exemplos da E-NNI (*External Network to Network Interface*) definida pelo MEF (*Metro Ethernet Forum*).

A interconexão de PBBNs por meio de E-NNIs pode ocorrer na forma de interconexão hierarquizada ou na forma de interconexão entre pares (*peer interconnection*).

A interconexão hierarquizada conecta as PBBNs em uma hierarquia por níveis. Nesse modelo de interconexão, uma PBBN de nível n+1 superpõe-se a uma PBBN de nível n. Dessa forma, amplia-se a escalabilidade da oferta de serviços. Cada nível de interconexão hierarquizada acrescenta uma camada de encapsulamento adicional, até o limite definido pelo tamanho máximo de quadro para o tipo de MAC utilizado.

A PBBN de nível n+1 transporta as B-VLANs da PBBN de nível n como serviços, de forma análoga à maneira como a PBBN de nível 1 transporta as S-VLANs de clientes como serviços.

A interconexão entre pares resulta em uma rede PBB global com um único nível hierárquico. Nesse tipo de interconexão, uma instância de serviço da PBBN A é conectada a uma instância de serviço da PBBN B pela E-NNI do tipo par que as conecta. A E-NNI do tipo par é formada por uma interface de serviço I-*tagged* entre duas CBPs (*Customer Backbone Ports*) situadas em cada uma das PBBNs pares. O valor de I-SID utilizado na E-NNI do tipo par deve ser acordado entre os provedores de serviço envolvidos.

Em cada terminação da E-NNI do tipo par, a CBP pode, opcionalmente, traduzir o valor de I-SID utilizado para um novo valor de I-SID para uso na PBBN correspondente. A CBP também mapeia o I-SID para uma B-VLAN no interior da respectiva PBBN, uma vez que não ocorre a transmissão de B-*tags* através de uma E-NNI do tipo par.

5.4.8. **Prioridade de Serviço e Elegibilidade para Descarte**

Os processos para a obtenção de valores de prioridade de serviço e de elegibilidade para descarte que devem constar dos I-*Tags* e B-*Tags* em PBBNs são, basicamente, os mesmos utilizados em PBNs, com os devidos ajustes.

Para a interface de serviço baseada em porta, aplicam-se os mesmos processos que em PBNs, inclusive com a possibilidade de envio, pelo cliente, de quadros incluindo artificialmente S-*Tags* com o S-VID para sinalizar a prioridade de serviço e a elegibilidade para descarte.

Na interface de serviço S-*tagged* utilizam-se informações contidas nos S-*Tags* enviados pelos clientes.

No caso da interface de serviço I-*tagged*, as informações relativas à prioridade de serviço e à elegibilidade para descarte estão contidas nos I-*Tags* enviados pelo cliente.

5.4.9. **Endereços MAC Backbone (B-MAC Addresses)**

Quando um quadro de cliente penetra em uma PBBN, ele é encapsulado pela adição de um endereço MAC *backbone* individual de origem (B-SA), pela adição de um endereço MAC *backbone* de destino (B-DA) e pela criação e adição de um I-*Tag*. Esse encapsulamento é o resultado de ações na PIP, sendo que a PIP representa um ponto de terminação de uma instância de serviço *backbone*, e o endereço MAC *backbone* identifica esse ponto de terminação (registra-se que, na CBP, o quadro ganha um B-*Tag*, o que completa a PDU do quadro PBB).

Mais especificamente, caso uma PIP possua mais de uma VIP (*Virtual Instance Port*), são as VIPs que, efetivamente, representam os pontos de terminação das instâncias de serviço *backbone* que terminam na PIP. Assim, um ponto de terminação de uma instância de serviço *backbone* é identificado pela combinação do endereço MAC *backbone* da PIP com o valor de I-SID atribuído a essa instância de serviço. As VIPs de uma PIP, correspondentes a cada I-SID dessa PIP, portanto, podem compartilhar o endereço MAC *backbone* da PIP.

O endereço MAC *backbone* de origem (B-SA) de um quadro encapsulado identifica a PIP que realizou o encapsulamento. Para uma PIP que se conecta a uma LAN externa, esse endereço é o endereço MAC da PIP. Para uma PIP com conexão interna a uma CBP (*Customer Backbone Port*) no interior de uma BEB, esse endereço pode ser qualquer endereço que resulte na entrega do quadro I-*tagged* para o devido I-*comp*, como, por exemplo, o endereço de *bridge* desse I-*comp*.

O endereço MAC *backbone* de destino (B-DA) de um quadro encapsulado identifica a PIP ou as PIPs às quais o quadro deverá ser entregue. O valor *default* para o B-DA é o endereço *Backbone Service Instance Group*. Esse endereço *default* é constituído pela concatenação dos três octetos do OUI (no caso, 00-1E-83) com os três octetos do I-SID, e configurando-se o bit I/G do primeiro octeto do valor resultante dessa concatenação para significar endereço MAC de grupo. Esse valor utilizado de OUI é referido como IEEE 802.1Q *Backbone Service Instance Group Address OUI*.

Quando o B-DA de um quadro é o *Backbone Service Instance Group Address*, o comportamento da PBBN é o de distribuir o quadro para todas as CBPs alcançáveis

dentro da B-VLAN na qual a instância de serviço é mapeada. Esse modo de operação pode resultar em um uso ineficiente dos recursos da rede.

Uma forma mais eficiente é obtida pela configuração de entradas estáticas nas *backbone bridges*, ou utilizando o MMRP (*Multiple MAC Registration Protocol*), para restringir a entrega aos pontos de terminação pertencentes efetivamente à instância de serviço *backbone*, representada pelo respectivo endereço de grupo.

Outra forma mais eficiente é obtida mediante autilização de endereços *backbone* MAC de destino (B-DAs) individuais. Neste caso, os endereços *backbone* MAC são aprendidos pelas PIPs a partir de quadros recebidos, os quais contêm tanto o endereço MAC de cliente quanto o endereço *backbone* MAC da PIP encapsuladora. A partir da aprendizagem, a PIP associa os endereços MAC de cliente aos endereços *backbone* MAC individuais.

No caso de uma instância de serviço *backbone* que inclua apenas duas VIPs (ou seja, ponto a ponto), o processo é ainda mais simples, uma vez que somente existe um *backbone* DA possível. Uma vez aprendido o endereço *backbone* MAC da PIP remota, esse endereço pode passar a ser usado como *default* e todos os quadros encaminhados nessa instância de serviço *backbone* pela PIP de origem passam a usar esse endereço como B-DA.

5.4.10. **Proteção do Acesso ao Serviço**

Foram definidas três classes de proteção do acesso ao serviço para o PBB: Classe I, Classe II e Classe III. O padrão IEEE 802.1Q-2014 apresenta, de forma ilustrada, a aplicação dessas três classes de proteção às interfaces de serviço baseadas em porta, às interfaces de serviço S-*tagged* e às interfaces de serviço I-*tagged*.

Na Classe I, não existe proteção. A conexão entre o sistema de usuário e o I-*comp* da PBBN (nos casos de interface de serviço baseada em porta e de interface de serviço S-*tagged*) ou a conexão entre o I-*comp* ou B-*comp* controlado pelo cliente e o B-*comp* da PBBN (no caso de interface de serviço I-*tagged*) é realizada por uma única LAN.

Na Classe II, as conexões de que trata o parágrafo anterior realizam-se por múltiplos *links* utilizando o mecanismo de agregação de *links* (padrão IEEE 802.1AX-2008).

Na Classe III, além da multiplicidade de LANs entre pares de nós, utiliza-se também multiplicidade de pares de nós da PBBN e do ambiente do cliente.

5.4.11. **Gerenciamento de PBB Bridges**

Como visto anteriormente, as PBB *bridges* podem ser BCBs e BEBs.

As BCBs são VLAN *bridges* comuns, que são gerenciadas utilizando objetos gerenciáveis definidos pelo padrão IEEE 802.1Q-2014 e especificados em detalhe nos itens de 12.4 a 12.14 desse padrão.

Cada I-*comp* e cada *B-comp* de uma BEB são gerenciados também utilizando objetos gerenciáveis sumarizados na seção 12.2 do padrão IEEE 802.1Q-2014, mas a seção 12.16 desse padrão acrescenta objetos gerenciáveis que, em conjunto, gerenciam a configuração geral de PBBNs, assim como as VIPs, PIPs e CBPs de uma BEB. A conjunção desses objetos, definidos nos citados padrões, possibilita o gerenciamento da configuração geral e o gerenciamento das VIPs, PIPs e CBPs de uma BEB.

O presente item especifica, em linhas gerais, os objetos gerenciáveis acrescidos pela seção 12.6 do padrão IEEE 802.1Q-2014 acima mencionada para o gerenciamento de BEBs. Tais objetos gerenciáveis, em linhas gerais, têm as seguintes funções:

— Suportam um objeto gerenciável de configuração de uma BEB, identificando todos os componentes dessa BEB;
— Identificam os tipos de porta da BEB;
— Provêem objetos gerenciáveis adicionais para o gerenciamento de VIPs, PIPs e CBPs.

Existe apenas uma única visão dos objetos gerenciáveis de uma BEB. Essa visão pode conter apenas os objetos necessários para o tipo de BEB sendo gerenciada. Se a BEB contém apenas um I-*comp*, a visão única englobará apenas os seguintes objetos gerenciáveis:

— Objeto de configuração de BEB;
— Objeto de configuração de portas de BEB;
— Objeto de configuração de VIPs;
— Objeto de configuração de PIPs.

Por outro lado, se a BEB contém apenas um B-*comp*, então a visão única contemplará apenas os seguintes objetos gerenciáveis:

— Objeto de configuração de BEB;
— Objeto de configuração de portas de BEB;
— Objeto de configuração de CBPs.

5.5. PROVIDER BACKBONE BRIDGES-TRAFFIC ENGINEERING (PBB-TE)

O PBB-TE, também referido inicialmente como PBT (*Provider Backbone Bridge*), é uma tecnologia de *Ethernet bridging*, definida inicialmente pelo padrão IEEE 802.Qay e incorporada ao novo padrão IEEE 802.1Q-2014 nas seções 25 e 26. O plano de controle do PBB-TE encontra-se especificado na RFC 6060 (*GMPLS Control of Ethernet Provider Backbone Traffic Engineering*), destinada ao controle do PBB-TE pelo GMPLS.

O PBB-TE é fruto de um esforço no sentido da definição de uma tecnologia *Ethernet* orientada a conexão para operar como rede de transporte.

Tradicionalmente, as redes de transporte consistiam em redes mutiplexadoras modo circuito, como as redes SONET/SDH (além das redes PDH) e redes OTN (*Optical Transport Networks*).

Mais recentemente, os órgãos internacionais de padronização vêm buscando a especificação de tecnologias de transporte modo pacote. As redes resultantes do uso dessas tecnologias modo pacotes são genericamente referidas como PTNs (*Packet Transport Networks).*

Para que uma tecnologia modo pacote seja considerada uma PTN, devem ser atendidos alguns requisitos operacionais típicos de redes de transporte, como, por exemplo:

- Orientação a conexão;
- Suporte a QoS;
- OAM hierarquizado e proteção nível *carrier class*;
- Supervisão fim a fim das conexões;

Atualmente, duas tecnologias modo pacote são utilizadas como PTNs, o MPLS-TP e o PBB-TE.

O T-MPLS (*Transport* MPLS) foi inicialmente desenvolvido pelo ITU-T, mas certas dificuldades em sua implementação conduziram a um esforço conjunto do ITU-T e do IETF, sob a coordenação do IETF, com o objetivo de definir uma concepção compatível com o MPLS. Essa concepção foi denominada MPLS-TP (MPLS *Transport Profile*).

5.5.1. **O PBB-TE como PTN**

Como será visto com maior profundidade adiante neste item, o PBB-TE, por operar com base em um subconjunto do PBB e em parte com base em uma nova concepção, oferece as seguintes características:

- Possibilita a constituição de circuitos virtuais orientados a conexão;
- Possibilita a implementação de *Traffic Engineering;*
- Possibilita a constituição de circuitos *backup*, o que melhora a proteção e a resiliência da rede;
- Provê uma plataforma ideal para a emulação de serviços TDM tradicionais (voz e vídeo);
- Elimina o uso de alguns protocolos de controle (MSTP e MVRP, por exemplo) e do tráfego por *flooding*;
- Oferece uma plataforma de OAM mais robusta.

5.5.2. **Funcionamento do PBB-TE**

A introdução de *Traffic Engineering* em redes *Ethernet* na definição PBB-TE permite a constituição individualizada de circuitos orientados a conexão ponto a ponto aos quais possam ser atribuídas diferentes classes de serviço, com reserva

de banda passante, o que representa a essência de *Traffic Engineering*. É possível também a constituição de circuitos virtuais ponto a multiponto no PBB-TE.

O PBB-TE aproveita o fato de que, desativando algumas funcionalidades do PBB, é possível a introdução de um novo comportamento funcional para as *Ethernet bridges*. Esse novo comportamento se resume na introdução do modo de transferência de quadros orientado a conexão, por circuitos virtuais.

Para obter esse resultado, o PBB-TE passa a adotar um processo administrativo no plano de gerenciamento ou um processo de sinalização no plano de controle na constituição de circuitos virtuais. Dessa forma, são preenchidas as tabelas de comutação em *bridges* de uma PBBN com entradas estáticas capazes de conduzir, por circuitos virtuais, os quadros a seus destinos.

O conjunto de entidades PBB-TE nas *bridges* da PBBN habilitadas para o PBB-TE é referido com uma Região PBB-TE.

No PBB, os B-VIDs são globalmente únicos na PBBN, e apresentam a funcionalidade comum de identificar as diferentes B-VLANs. Nas Regiões PBB-TE, os B-VIDs são utilizados para outra função. O B-VID, agora denominado ESP-VID, passa a compor o identificador de um circuito virtual em seu trajeto pela rede.

Em redes PBB-TE os circuitos virtuais são denominados ESPs (*Ethernet Switched Paths*) pelo padrão IEEE 802.1Q-2014.

Como a necessidade de ESP-VIDs no PBB-TE em decorrência da possibilidade de sua repetição em diferentes ESPs (*Ethernet Switched Paths*) é reduzida, grande parte dos 4094 possíveis valores de B-VID pode ser utilizada para o PBB ou para outras tecnologias em uma única rede. O PBB-TE pode coexistir então em uma única rede não só com o PBB, mas também, por exemplo, com futuras implementações.

Ressalta-se que a arquitetura e o formato dos quadros do PBB-TE são exatamente os mesmos que os do PBB. O que varia, basicamente, como visto anteriormente, é a funcionalidade atribuída aos B-VIDs dos quadros B-*tagged*.

5.5.3. **Suporte de Serviço MAC por uma Região PBB-TE**

Uma Região PBB-TE compreende um conjunto contínuo de IB-BEBs e BCBs com capacidade de prover TESIs e que alocam um conjunto comum de B-VIDs para serem controladas por um agente externo responsável pelo estabelecimento de ESPs (*Ethernet Switched Paths*, ou, no dizer do IETF, *Ethernet LSPs*).

Uma Região PBB-TE que provê serviço MAC às estações conectadas à rede segue o modelo de uma sequência de funções de comutação. A tabela de configuração de MST (*MST Configuration Table*) associada a cada BCB ou B-*component* de uma IB-BEB especifica um conjunto de VIDs, cada um dos quais pode ser utilizado de forma exclusiva em toda a Região PBB-TE como parte de um identificador de ESP. Um VID desse tipo denomina-se ESP-VID e está associado a um valor especial do MSTID que consta da tabela de configuração de MST, o TE-

MSTID, que indica que esse VID não se encontra sob o controle de um protocolo *spanning tree*.

Os I-*components* das IB-BEBs que suportam PBB-TE são responsáveis por encapsular quadros recebidos dos clientes e por atribuir cada quadro a uma instância de serviço *backbone*, assim como por desencapsular os quadros recebidos da rede através das instâncias de serviço *backbone*. A atribuição de quadros de clientes às instâncias de serviço *backbone* que serão transportadas pelas intâncias de serviço TE é realizada na PIP. Na PIP, o quadro de cliente é encapsulado por uma I-TAG, que inclui o I-SID, e um conjunto de endereços MAC *backbone* de origem e de destino. O endereço MAC *backbone* de origem corresponde ao endereço MAC da PIP, o qual é configurado para assumir o mesmo valor que o endereço MAC da CBP localizada no B-*component* conectada no interior da IB-BEB. Resulta daí que os endereços MAC *backbone* dos quadros associados a TESIs, os endereços ESP-MAC, sempre são os mesmos que os endereços MAC das CBPs.

Para um I-SID mapeado em uma TESI, a tabela de instâncias de serviço *backbone* contém o endereço MAC da CBP de terminação, referido como ESP-DA, e o ESP-VID do ESP associado. Os I-SIDs mapeados em uma B-VLAN não devem usar B-VIDs que pertençam ao espaço de ESP-VIDs. O comutador de quadros do B-*component* transmite os quadros pertencentes a um ESP entre as CBPs e PNPs ou entre PNPs usando os valores de ESP-DA e ESP-VID.

5.5.4. **ESPs, TESIs e Shared Forwarding**

A RFC 5828 (*GMPLS Ethernet Label Switching Architecture and Framework*) define uma moldura geral para a aplicação do GMPLS em redes *Ethernet* com comutação de *labels*. Assim, a RFC 5828 é aplicada ao PBB-TE.

A RFC 5828 utiliza a denominação *Eth-LSPs* para os ESPs, sendo que o IETF faz uso dessas duas denominações.

Existem dois tipos básicos de ESP:

— ESPs ponto a ponto (P2P ESPs);
— ESPs ponto a multiponto (P2MP ESPs).

O padrão IEEE 802.1 Q-2014 utiliza o conceito de *Traffic Engineering Service Instance* (TESI) suportada por um conjunto de ESPs, formando um serviço bidirecional. Assim, foram definidas P2P TESIs e P2MP TESIs. Uma TESI é coletivamente identificada por um único TE-SID, formando um serviço bidirecional.

Cada ESP é identificado pelo conjunto de três valores:

<ESP-DA, ESP-SA, ESP-VID>,

onde os três componentes têm os seguintes significados:

— O ESP-SA é o endereço da *Provider Instance Port* (PIP) que encapsula a instância de serviço de cliente, o qual é, por configuração, o mesmo que

o endereço da CBP de origem internamente conectada à PIP encapsuladora.
— O ESP-DA identifica o endereço ou os endereços da CBP de destino ou os endereços das CBPs de destino;
— O ESP-VID é um valor de VID que distingue as ESPs que possuem o mesmo par de endereços de origem e destino.

Embora um ESP seja identificado por uma 3-*tuple*, a comutação de quadros utiliza apenas a 2-*tuple* <ESP-DA, ESP-VID> em cada direção. Em PBB-TE, a combinação <ESP-DA, ESP-VID> é tratada como se fosse um único endereço de 60 bits, sendo 48 deles relativos ao endereço MAC B-DA (ESP-DA) e os 12 restantes correspondentes ao ESP-VID. O campo ESP-VID é usado como um seletor de caminho para a CBP de destino, e não como um B-VLAN-ID, permitindo a existência de até 2^{12} árvores de roteamento em direção a uma dada CBP.

5.5.4.1. P2P ESPs e P2P TESIs (P2P Eth-LSPs)

Como foi visto anteriormente, uma Região PBB-TE compreende um conjunto contínuo de IB-BEBs e de BCBs, com a capacidade de prover TESIs (*Traffic Engineering Service Instances*).

Um P2P ESP (*Eth-LSP*) é definido entre um par de CBPs situadas em diferentes BEBs. Um P2P ESP unidirecional é identificado por uma 3-*tuple* <ESP-SA, ESP-VID, ESP-DA>, sendo o ESP-SA e o ESP-DA, respectivamente, os endereços MAC individuais da CBP de início e da CBP de término do P2P ESP.

Uma P2P TESI é composta por dois P2P ESPs em sentidos reversos utilizando o mesmo par de endereços MAC-ESP individuais, mas cujos ESP-VIDs podem ser diferentes em um e outro sentido. O IETF utiliza o termo *Ethernet LSP* (*Eth-LSP*) bidirecional para designar uma P2P TESI.

Devido a razões relacionadas com a monitoração e diagnóstico de TESIs, simplicidade operacional e outros fatores, o padrão IEEE 802.1Q-2014 presume que os ESPs associados a TESIs P2P são simétricos, o que significa que as rotas nos dois sentidos são as mesmas. Adicionalmente, os dois sentidos são estabelecidos simultaneamente, em um mesmo processo de sinalização.

Diferentes ESPs dirigidos de uma BEB a outra BEB são identificados pelo uso de diferentes valores de ESP-VID, e podem utilizar diferentes sequências de TE *links* da rede de suporte ao PBB-TE.

A Figura 5-23 apresenta dois diferentes P2P ESPs entre as BEBs S1 e S2, com rotas diferentes.

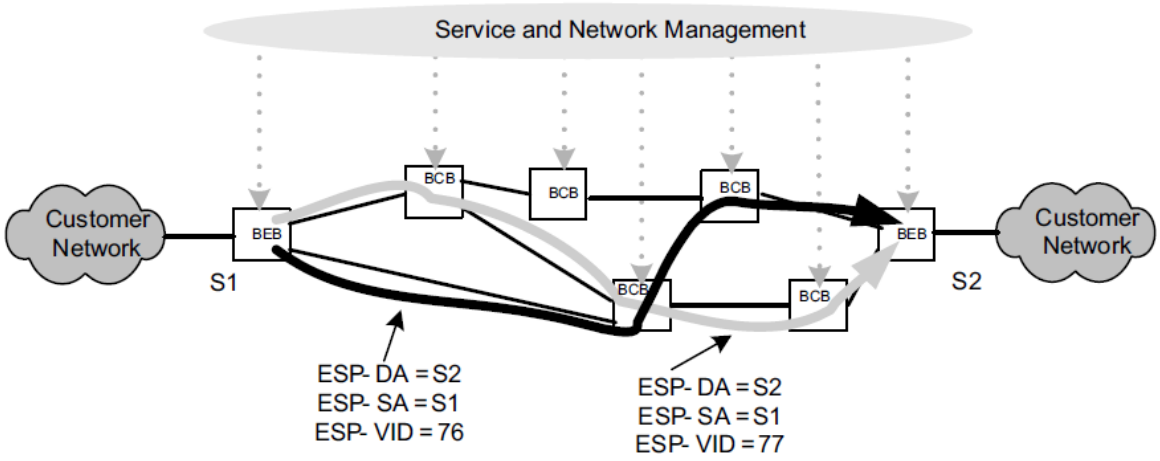


Figura 5-23 - ESPs entre um par de BEBs.

A configuração dessa figura pode justificar-se se considerarmos que cada um dos ESPs representados objetiva o atendimento a uma diferente classe de serviço ou a fins de proteção, por exemplo.

Uma mesma 2-*tuple* <ESP-VID, ESP-DA> pode ser utilizada em diferentes ESPs em direção a uma mesma CBP mas que possuam diferentes endereços ESP-SA. Esse modo de operação é denominado *shared forwarding*, sendo a configuração resultante referida como uma *rooted tree*.

O plano de controle e gerência de PBB-TE pode formar uma topologia ativa de árvores de roteamento com raiz em cada CBP que pertença a uma determinada região PBB-TE, definindo caminhos até a CBP raiz desde cada uma das demais CBPs da mesma região. Essas árvores com raiz e com engenharia de tráfego, definem os trajetos seguidos pelos quadros pertencentes a ESPs dentro da região PBB-TE. Cada uma das árvores com raiz é identificada pelo endereço MAC da CBP raiz, pelo ESP-VID e pelos endereços MAC das CBPs folhas.

O plano de controle e gerência de PBB-TE também gerencia a largura de banda de todos os ESPs ao longo de cada árvore de roteamento com raiz.

Para cada CBP de destino, o PBB-TE mantém árvores com caminhos simétricos desde a CBP raiz até cada CBP folha. O ESP-VID utilizado nos caminhos reversos não possuem necessariamente os mesmos ESP-VIDs usados nos caminhos diretos.

A Figura 5-24 representa os ESPs de uma *rooted* tree centrada na CBP-3.

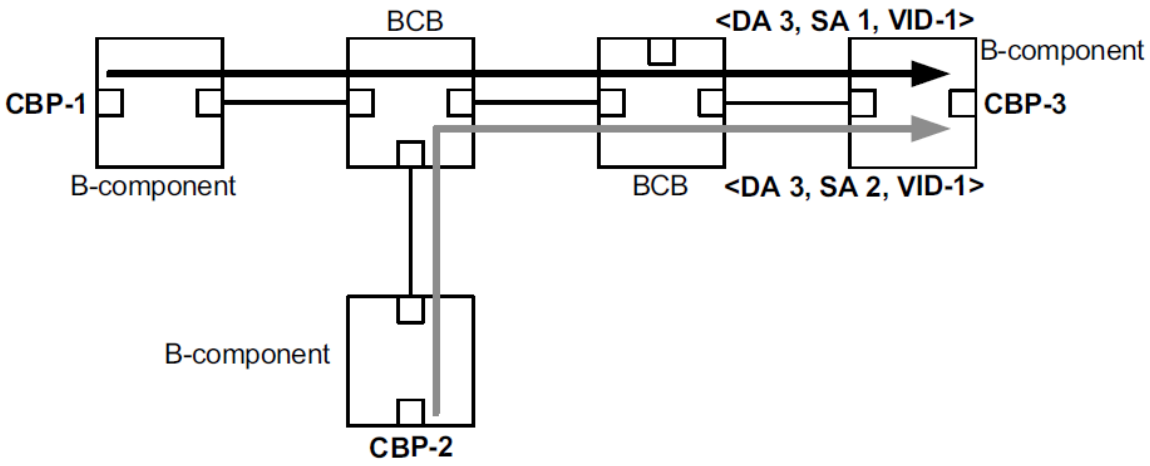


Figura 5-24 - ESPs de uma rooted tree.

Dentro de uma *rooted tree* os parâmetros de TE e de QoS são próprios (reserva de banda passante, por exemplo).

O uso de *rooted trees* se justifica por razões de escalabilidade (redução do número do ESP-VIDs e de entradas de filtragem estáticas nos FDBs).

Todos os ESPs de uma *rooted tree* podem suportar uma mesma classe de serviço, o que, adicionalmente, pode simplificar o processo de atribuição de valores de ESP-VID para esses ESPs.

O GMPLS possibilita a alternativa de constituição de P2P ESPs unidirecionais. O suporte a P2P ETSIs não congruentes (não corroteados), contudo, não é adotado nas extensões do GMPLS para o PBB-TE.

5.5.4.2. P2MP ESPs e P2MP TESIs

Um P2MP ESP é um ESP unidirecional constituído entre uma CBP raiz e múltiplas CBPs folha. Nesse caso, o *label* PBB-TE *Ethernet*(que será definido adiante quando da abordagem da RFC 6060, que trata da aplicação de GMPLS a PBB-TE) consiste em um VID e um endereço MAC de grupo que corresponde às múltiplasCBPs folha. O endereço ESP-SA é o endereço MAC individual da CBP raiz.

Cada um dos ramos da árvore P2MP unidirecional é associado a um P2P ESP simétrico e congruente no sentido reverso.

Uma P2MP TESI corresponde ao conjunto formado por um P2MP ESP e pelos múltiplos P2P ESPs entre as CBPs folha e a CBP raiz.

Os procedimentos definidos na RFC 3473 e na RFC 4875 podem ser adaptados para sinalizar P2MP Eth-LSPs no sentido da raiz para os ramos. Os procedimentos completos para sinalizar P2MP TESIs (P2MP Eth-LSPs bidirecionais) estão fora do escopo da RFC 6060.

5.5.5. **Constituição e Utilização de ESPs**

Os ESPs e os serviços no PBB-TE são sempre originados e terminados em IB-BEBs.

Como vimos, os B-*comps* das IB-BEBs dizem respeito à estrutura e aos mecanismos que suportam a comutação de quadros MAC na rede PBB-TE *backbone*. O I-*comps* dizem respeito à utilização dos I-*Tags*.

No PBB-TE, os I-*comps* inserem e retiram os I-*Tags* e os endereços ESP-SA e ESP-DA, enquanto os B-*comps* inserem e retiram os B-*Tags* (onde são contidos os ESP-VIDs).

A retirada desses *tags* ocorre na IB-BEB de destino, após o devido processamento, sendo os B-*Tags* os primeiros a serem processados e retirados.

As PIPs (nos I-*comps*) e as CBPs (nos B-*comps*) de uma IB-BEB são interconectados por meio de LANs internas, sendo que as PIPs e as CBPs não são visíveis externamente.

Os I-*comps* de origem recebem quadros dos usuários em suas CNPs, e inserem os I-SIDs e os endereços ESP-SA e ESP-DA, enquanto os B-*comps* correspondentes inserem os ESP-VIDs. Os I-SIDs e os ESP-VIDs são contidos, respectivamente, nos correspondentes I-*Tags* e B-*Tags*.

Os B-*comps* utilizam suas PNPs para se conectarem a BCBs no interior da rede PBB-TE *backbone*, por meio de TE *links*.

5.5.6. **Gerência de Falha de Conectividade (CFM) em Redes PBB-TE**

Os processos de OAM para redes *bridged Ethernet* definidos pelas seções 18, 19, 20, 21 e 22 do padrão IEEE 802.1Q-2014, referidos como CFM (*Connectivity Fault Management*), aplicam-se às redes PBB-TE, porém com a definição de algumas particularidades adicionais.

Sugere-se aos leitores reportarem-se, neste ponto, ao Capítulo 8 deste livro, e, para maior profundidade, às seções referidas acima do padrão IEEE 802.1Q-2014.

Conforme a cláusula 26.9 do padrão IEEE 802.1Q-2014, as particularidades adicionais relativas à aplicação do CFM ao PBB-TE dizem respeito aos seguintes aspectos:

— Endereçamento de PBB-TE MEPs (*PBB-TE Maintenance Association End Points*);
— Identificação de TESIs;
— Constituição de PBB-TE MEPs em portas de *bridges*;
— Constituição de PBB-TE MIPs (*PBB-TE Maintenance Domain Intermediate Points*) em portas de *bridges*;
— Domínios de Manutenção de TESIs;
— Melhorias dos protocolos CFM específicos para o PBB-TE.

As melhorias referidas no último quesito acima dizem respeito a todo o CFM, ou seja, aos protocolos *Continuity Check* (CC), *Loopback* (LB) e *Linktrace* (LT).

5.5.7. **Comutação de Proteção para P2P TESIs**

O PBB-TE provê um esquema de proteção linear fim a fim para P2P TESIs, onde uma TESI de proteção é constituída para proteção 1:1 de uma TESI de serviço, sendo essas duas TESIs iniciadas e terminadas nas mesmas BCPs.

A Figura 5-25 apresenta um exemplo de configuração com uma TESI de serviço e uma TESI de proteção 1:1.

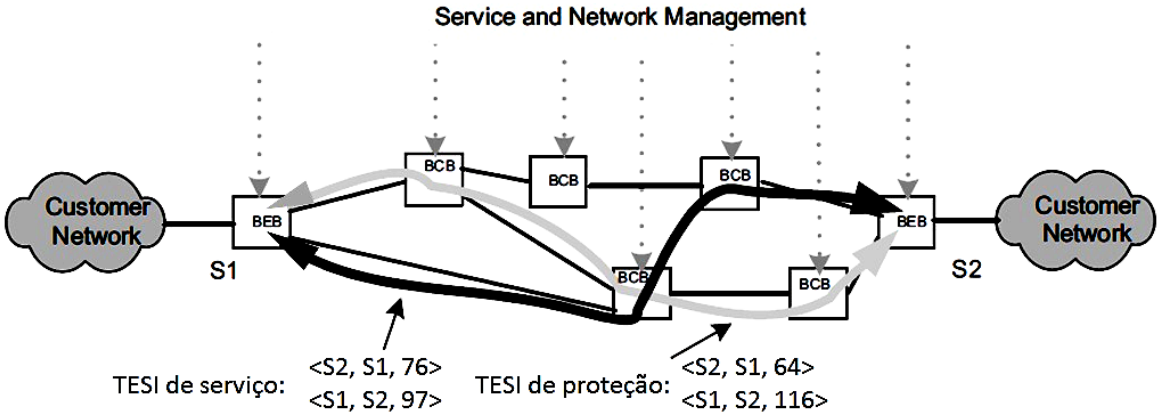


Figura 5-25 - TESI de serviço e TESI de proteção 1:1.

Considerando-se a utilização de TE, presume-se que a TESI de proteção atende os mesmos requisitos da classe de serviço suportada pela TESI de serviço.

Quando uma falha ocorre na TESI de serviço (*working TESI*), o tráfego é automaticamente comutado para a TESI de proteção. Esse processo é referido como comutação de proteção.

Como o ESP de proteção encontra-se preestabelecido, é assegurada a sua disponibilização em um tempo de comutação (*switchover*) inferior a 50 ms.

A Figura 5-26 representa os elementos essenciais do esquema de proteção de TESIs no PBB-TE. Como a proteção é do tipo 1:1, o tráfego é enviado apenas por uma das TESIs em um dado momento.

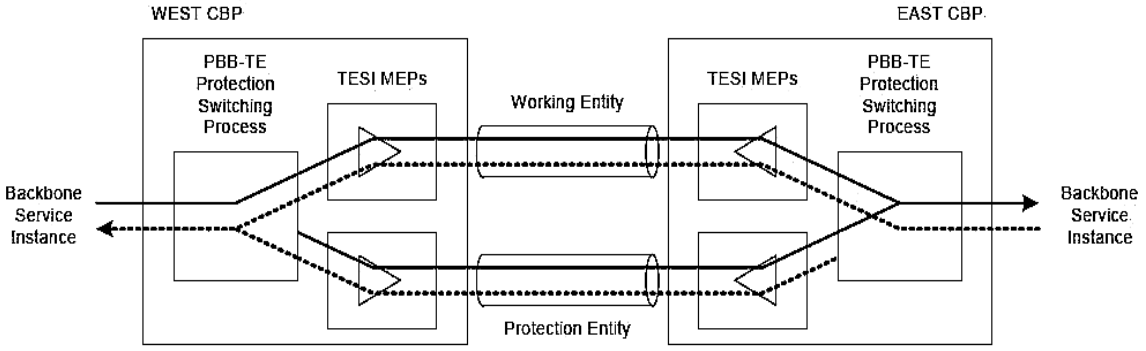


Figura 5-26 - Arquitetura de comutação de proteção.

A comutação de proteção nessa figura pode ser acionada por operação manual ou pelo CFM. As informações do CFM resultam da monitoração periódica das TESIs de serviço e de proteção, ou pela monitoração da camada física.

Para se obter um tempo de *switchover* inferior a 50 ms, o intervalo entre *Continuity Check Messages* deve ser programado nas respectivas MAs com um valor no máximo igual a 10 ms.

O esquema de comutação de proteção no PBB-TE pode ser do tipo reversível ou não-reversível após a correção da falha na TESI de serviço.

Para uma abordagem mais ampla, a descrição da comutação de proteção encontra-se no padrão IEEE 802.1Q-2014 e na recomendação ITU-T G.8031.

O padrão IEEE 802.1Q-2014 abstém-se da especificação de esquemas de comutação de proteção para P2MP ESPs.

5.5.8. **Plano de Controle do PBB-TE**

O plano de controle do PBB-TE tem como base a RFC 6060(*Generalized Multiprotocol Label Switching (GMPLS) Control of Ethernet Provider Backbone Traffic Engineering (PBB-TE)*), que define uma extensão do GMPLS especificamente voltada para o PBB-TE. A RFC 6060 complementa a RFC 5828 (*Generalized Multiprotocol Label Switching (GMPLS) Ethernet Label Switching Architecture and Framework*), que especifica a arquitetura e a visão geral do que se donominou *GMPLS Ethernet Label Switching* (GELS).

O plano de controle do PBB-TE, baseado no GMPLS, engloba os seguintes aspectos:

— Roteamento *constraint-based*;
— Definição de classes de serviço (CoS);
— Algoritmo para seleção e cálculo de rotas;
— Sinalização para a constituição de ESPs conforme as rotas selecionadas;
— Possibilidade de reserva de banda passante nos ESPs constituídos.

Uma visão geral do GMPLS e os processos de roteamento e de sinalização no GMPLS serão vistos a seguir.

5.5.9. **Generalized Multiprotocol Label Switching (GMPLS)**

O GMPLS tem a sua arquitetura definida na RFC 3945 (*Generalized Multiprotocol Label Switching (GMPLS) Architecture*).

O GMPLS é uma família de protocolos do plano de controle especificados para operar em redes orientadas a conexão e capacitadas para transporte com engenharia de tráfego (TE). O GMPLS é aplicável a uma faixa de tecnologias de rede, incluindo redes L2SC (*Layer 2 Switching Capable*), isto é, redes comutadas com base nos endereços de camada 2 (tais como redes Ethernet comutadas com base no

endereço MAC), e redes modo circuito (tais como SDH, OTN, WDM e redes de fibras ópticas).

O GMPLS, assim como o plano de controle do MPLS-TE, do qual o GMPLS é uma extensão, opera de forma centralizada tipicamente nos comutadores de borda das redes (IB-BEBs no caso de PBB-TE, por exemplo) que cumprem, como *head-ends*, as seguintes funções:

— Recebem as informações transmitidas pelo protocolo de roteamento (GMPLS OSPF-TE, por exemplo);
— Recebem informações relativas aos atributos de QoS das aplicações a serem atendidas;
— Definem, com base em um algoritmo *constraint-based* de cálculo de rotas, como o CSPF (*Constrained Shortest Path First*), os circuitos virtuais a serem sinalizados que melhor atendam às especificações de QoS;
— Sinalizam, pelo GMPLS RSVP-TE, mediante a explicitação da rota desejada, a constituição dos circuitos virtuais definidos.

O GMPLS introduz, relativamente ao MPLS-TE, restrições adicionais para o cálculo de rotas de LSPs GMPLS.A RFC 4655 (*A Path Computation Element (PCE)-Based Architecture*) especifica uma arquitetura para um modelo de PCE para atendimento a essa questão.

5.5.9.1. Gerenciamento de Links no GMPLS com LMP

O GMPLS utiliza o protocolo LMP, definido pela RFC 4204 (*Link Management Protocol – LMP*). O LMP desempenha as seguintes funções:

— Gerenciamento do canal de controle;
— Correlação de propriedades de *links*;
— Verificação de conectividade de *links*;
— Gerenciamento de falhas.

Um TE-*link* entre nós adjacentes, associado a um conjunto de parâmetros de TE (*Traffic Engineering*), é composto por um ou mais *data links* entre esses nós. Os *data links* são utilizados para transferir dados de usuários.

Os canais de controle usados para o intercâmbio de informações do plano de controle do GMPLS (tais como informações de roteamento, de sinalização, do próprio LMP e outras) existem de maneira independente dos *data links* gerenciados, isto é, são desacoplados dos *data links* que o GMPLS gerencia.

Em consequência, os canais de controle entre dois nós adjacentes não são obrigados a usar o mesmo meio físico que os *data links* entre esses nós. Por exemplo, um canal de controle pode utilizar um circuito virtual separado, um comprimento de onda, uma fibra óptica, um *link* Ethernet ou uma rede IP.

Os dois procedimentos básicos e obrigatórios do LMP são o gerenciamento do canal de controle e a correlação de propriedades de *links*.

O gerenciamento do canal de controle é utilizado para a estabelecer e manter os canais de controle entre nós adjacentes.

A correlação de propriedades de *links* é usada para assegurar que os extremos local e remoto de um determinado TE-*link* são do mesmo tipo (ou seja, IPv4, IPv6 ou não numerado). A correlação permite também alterar atributos TE dos *links* e para formar e modificar a agregação de *links*.

Os procedimentos de verificação de conectividade de *links* e de gerenciamento de falhas são opcionais.

O procedimento de verificação da conectividade de *links*, quando implementado, pode ser utilizado para verificar a conectividade física dos *data links* e aprender dinamicamente as associações entre um TE *link* e *Interface_Ids*. Esse procedimento ocorre quando do estabelecimento de TE-*links* e dos respectivos *data links* e, posteriormente, de forma periódica para os *data links* não alocados (livres) pertencentes ao TE *link.*

O procedimento de gerenciamento de falhas, quando implementado, ocorre mediante a notificação rápida do estado de um ou vários *data links* de um TE *link*. O gerenciamento de falhas inclui a detecção, a localização e a notificação de falhas de nós e de *links*. A detecção das falhas deve ser implementada na camada mais próxima da falha. Note-se que o mecanismo de notificação de falhas do LMP é independente do mecanismo usado para detectar falhas. O LMP apenas se vale do fato de que uma falha foi detectada.

A notificação da ocorrência de falhas possibilita a localização das falhas. As mensagens de notificação de falhas indicam se a falha afetou um ou vários *data links* ou o TE *link* inteiro. A correlação das falhas é efetuada localmente em cada nó quando da recepção da notificação, o que permite identificar a localização da falha.

5.5.10. **Roteamento no GMPLS**

Para o roteamento no GMPLS, foram definidas duas alternativas de protocolo: GMPLS OSPF-TE (RFC 4203) e GMPLS IS-IS TE (RFC 5307).Essas duas RFCs têm como base a RFC 4202 (*Routing Extensions in Support of GMPLS*)

O protocolo GMPLS OSPF-TE representa um conjunto de extensões do protocolo OSPF-TE (RFC 3630), enquanto o o protocolo GMPLS IS-IS TE representa um conjunto de extensões do protocolo IS-IS TE (RFC 5305).

É apresentada a seguir a evolução da família de protocolos OSPF até a definição do protocolo GMPLS OSPF-TE.

5.5.10.1. Protocolo OSPF

O protocolo OSPF Versão 2 foi especificado pela RFC 2328 (*Open Shortest Path First (OSPF) Version 2*). O OSPF Versão 2 é um protocolo de roteamento IP do tipo IGP (*Interior Gateway Protocol*) baseado em uma métrica *link state*, aplicável ao IPv4. Para o IPv6 foi especificado o protocolo OSPF Versão 3.

O domínio abrangido pelo OSPF é denominado sistema autônomo (*Autonomous System* – AS), e os sistemas autônomos são divididos em áreas OSPF.

O processo de roteamento é distribuído pelos roteadores, cada um possuindo a sua tabela de roteamento. As informações de roteamento são difundidas pelos roteadores mediante inundação (*flooding*). A inundação ocorre pela difusão de mensagens referidas como LSAs (*Link State Advertisements*).

As LSAs são compostas por um cabeçalho comum e por um *payload* constituído por TLVs e por sub-TLVs. Os TLVs e sub-TLVs, por sua vez, contêm um conjunto de atributos específicos.

5.5.10.2. Protocolo OSPF-TE

Para o uso do OSPF com engenharia de tráfego no MPLS-TE, a lógica de funcionamento foi completamente alterada. O processamento das informações de roteamento passa a ser centralizado nos roteadores (LSRs) de borda (PEs) da rede, que funcionam como *head-end LSRs*. Os roteadores intermediários (Ps) repassam as LSAs recebidas de outros roteadores, também por inundação, cujas informações, ao final, são recebidas centralmente pelos *head-end LSRs*.

Foi emitida então a RFC 3630 (*Traffic Engineering Extensions to OSPF Version 2*). A RFC 3630 especifica uma nova modalidade de LSA, referida como LSAs opacas, que se encontram especificadas na RFC 5250 (*The OSPF Opaque LSA Option*). As LSAs opacas, assim denominadas por não serem processadas por roteadores IP intermediários, provêem um mecanismo generalizado para possibilitar a criação de futuras extensões para o OSPF. As informações contidas nas LSAs opacas podem ser utilizadas diretamente pelo OSPF ou indiretamente por alguma aplicação que deseje distribuir informações através do domínio OSPF.

As LSAs opacas consistem em um cabeçalho LSA padronizado seguido de informações específicas.

Foram definidos três tipos de LSA opaca:

— LSA opaca tipo 9 (abrange um *link*);
— LSA opaca tipo 10 (abrange uma área OSPF);
— LSA opaca tipo 11 (abrange todo um sistema autônomo).

A RFC 3630 utiliza apenas a LSA opaca tipo 10. Dentro desse tipo, foi definida uma LSA, denominada TE LSA, identificada pelo código 1 (tipo 1).

A TE LSA utiliza alternativamente dois *top-level TLVs*, que são o *Router Address TLV* (tipo 1, ocupando 32 bits, correspondente a um endereço IPv4) ou o *Link TLV* (tipo 2, de comprimento variável).

O *Link TLV* descreve os atributos de cada *link* do sistema autônomo em seus sub-TLVs.

5.5.10.3. Protocolo GMPLS OSPF-TE

O protocolo GMPLS OSPF-TE, definido pela RFC 4203 (*OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)*), que tem como base a RFC 4202 (*Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)*), representa um conjunto de extensões ao protocolo OSPF-TE. Tais extensões consistem na criação de quatro novos sub-TLVs para a TE LSA e na criação de uma nova TE-LSA opaca do tipo 9 (abrangendo um *link*, portanto), referida como *TE Link Local LSA*.

A TE LSA, que é uma LSA opaca do tipo 10 (abrangendo uma área OSPF-TE), possui apenas um *top-level TLV* e possui um ou mais sub-TLVs aninhados. O *top-level TLV* pode assumir o valor (1) que representa o *Router Address TLV* ou o valor (2) que representa o *Link TLV*. Os quatro novos sub-TLVs introduzidos pela RFC 4203 dizem respeito ao *Link TLV* e definem novos *link attributes* para a TE LSA, conforme a Figura 5-27.

Tipo de sub-TLV	Comprimento	Nome do sub-TLV
11	8 octetos	*Link Local/Remote Identifiers*
14	4 octetos	*Link Protection Type*
15	variável	*Interface Switching Capability Descriptor*
16	variável	*Shared Risk Link Group (SRLG)*

Figura 5-27 – Novos link attributes para o GMPLS.

O sub-TLV *Link Local/Remote Identifier* (tipo 11) possui o comprimento igual a 8 octetos referentes à identificação local (4 octetos) e à identificação remota (4 octetos) dos *links* não numerados (*unnumbered links*) P2P que compõem a rede à qual o GMPLS se aplica. Opcionalmente, o identificador do *link* local pode ser enviado pela TE *Link Local LSA*.

O sub-TLV *Link ProtectionType* representa a capacidade de proteção que existe para um *link*. Os tipos especificados para o sub-TLV *Link Protection Type* foram os seguintes:

- 0x01 – *Extra Traffic*;
- 0x02 – *Unprotected*;
- 0x04 – *Shared* (1:N ou M:N onde $1 \leq M \leq N$);
- 0x08 – *Dedicated 1:1*;
- 0x10 – *Dedicated 1+1*:
- 0x20 – *Enhanced*;
- 0x40 e 0x80 – *Reserved*.

O campo *Value* do sub-TLV *Interface Switching Capability (ISC) Descriptor* (tipo 15) é composto por uma sequência de grupos de 4 octetos. O primeiro grupo de 4 octetos contém o campo *Switching Capability* e o campo *Encoding*, além de bits reservados. Os oito campos seguintes contêm a largura de banda reservada para cada uma das prioridades (*Max LSP Bandwidth at priority 0* até *Max LSP Bandwidth*

at priority 7). Segue-se um campo, de comprimento variável, contendo informações específicas (*Switching Capability-specific information*).

Para o campo *Switching Capability* (*Switching Cap*), foram especificados os seguintes valores de *switching capabilities*, um dos quais deve estar contido nesse campo:

— 1 *Packet-Switch Capable-1 (PSC-1)*
— 2 *Packet-Switch Capable-2 (PSC-2)*
— 3 *Packet-Switch Capable-3 (PSC-3)*
— 4 *Packet-Switch Capable-4 (PSC-4)*
— 51 *Layer-2 Switch Capable (L2SC)*
— 100 *Time-Division-Multiplex Capable (TDM)*
— 150 *Lambda-Switch Capable (LSC)*
— 200 *Fiber-Switch Capable (FSC)*

O valor do parâmetro Largura de Faixa do LSP em uma determinada prioridade (*Max LSP Bandwidth at priority x*) é codificado em uma lista de oito campos de 4 octetos no formato IEEE de ponto flutuante, indicando a largura de banda em bytes (e não em bits) por segundo.

O sub-TLV *Shared Risk Link Group* (SRLG) transporta uma lista não ordenada de grupos de *links* com idêntico risco (um conjunto de comprimentos de onda em uma mesma fibra óptica, por exemplo) aos quais o *link* sinalizado pertence.

A nova TE LSA opaca criada (*TE Link Local LSA*) é do tipo 9, com o campo *Opaque Type* com o valor 1 (o que significa que se trata de uma TE-LSA) e o campo *Opaque ID* (*instance*) com o valor 0.

Essa nova TE-LSA objetiva o intercâmbio dos *Link Local Identifiers* referentes a interfaces não numeradas entre nós adjacentes. Esse intercâmbio ocorre por meio de uma PDU específica para a *TE Link Local LSA*, que utiliza exclusivamente um TLV (TLV *Link Local Identifier*, tipo 1, com comprimento igual a 32 bits).

5.5.11. **Sinalização no GMPLS**

A sinalização no GMPLS consiste no uso do protocolo GMPLS RSVP-TE, definido na RFC 3473 (*GMPLS Signaling RSVP-TE Extensions*) e nas respectivas atualizações. A RFC 3473 tem como suporte a RFC 3471 (*GMPLS Signaling Functional Description*).

Além de possibilitar o estabelecimento (*setup*) de LSPs, a sinalização do GMPLS pode ser utilizada, também, para a modificação e para a liberação de LSPs.

Apresenta-se a seguir a evolução da família de protocolos RSVP até a definição do protocolo GMPLS RSVP-TE.

5.5.11.1. Protocolo RSVP Versão 1

O protocolo RSVP versão 1 foi especificado na RFC 2205 (*RSVP Version 1 Functional Specification*).

Os parâmetros transportados pelas mensagens do RSVP são codificados em diferentes objetos, em conformidade com as funções desempenhadas. Um objeto, identificado por um *Class-Number* (*Class-Num*), pode apresentar variações em suas funções e em seus campos, sendo cada uma dessas variações identificada por um código C-*Type*.

O RSVP Versão 1 especificado na RFC 2205 utiliza os tipos de mensagens relacionados na Figura 5-28.

TIPO DE MENSAGEM	CÓDIGO	SENTIDO
RSVP Path	1	Downstream
RSVP Resv (Reservation)	2	Upstream
RSVP PathErr (Error)	3	Upstream
RSVP ResvErr	4	Downstream
RSVP PathTear	5	Downstream
RSVP ResvTear	6	Upstream
RSVP ResvConf	7	Downstream

Figura 5-28 - Tipos de mensagens do protocolo RSVP versão 1.

5.5.11.2. Protocolo RSVP-TE

A RFC 3209 (*RSVP-TE: Extensions to RSVP for LSP Tunnels*) ampliou o protocolo RSVP Versão 1 para permitir a sinalização no processo de estabelecimento de LSPs no MPLS-TE, com extensões que definem o protocolo RSVP-TE.

Foi especificado, na RFC 3209, um novo tipo de mensagem, denominada *RSVP Hello*. Além de utilizar os objetos definidos para o RSVP Versão 1, a RFC 3209 introduziu um conjunto adicional de objetos:

- Objeto LABEL (*Class-Num* 16), com um *C-Type* (*C-Type* 1);
- Objeto LABEL_REQUEST (*Class-Num* 19), com três *C-Types* (*C-Type* 1, *C-Type*2 e *C-Type* 3);
- Objeto EXPLICIT_ROUTE (*Class-Num* 20), com um *C-Type* (*C-Type* 1), objeto esse também referido como ERO (*Explicit Route Object*);
- Objeto RECORD_ROUTE (*Class-Num* 21), com um *C-Type* (*C-Type* 1), objeto esse também referido como RRO (*Record Route Object*);
- Objeto SESSION_ATTRIBUTE (*Class-Num* 207), com dois *C-Types* (*C-Type* 1 e *C-Type*7);
- Objeto HELLO (*Class-Num* 22), com dois *C-Types* (*C-Type* 1 para HELLO REQUEST e *C-Type* 2 para HELLO ACK).

Além disso, foram definidos três sub-objetos para o ERO (*EXPLICIT_ROUTE Object*) e três sub-objetos para o RRO (*RECORD_ROUTE Object*).

Para cada um dos objetos SESSION (*Class-Num* 1), FILTER_SPEC (*Class-Num* 10) e SENDER_TEMPLATE (*Class-Num* 11), a RFC 3209 definiu dois novos *C-Types*, que são o *C-Type* 7 e o *C-Type* 8.

O objeto SESSION_ATTRIBUTE (*Class-Num* 207), definido na RFC 3209, apresenta limitações quanto ao número máximo de condições binárias que pode indicar (igual a 8) e quanto à impossibilidade de explicitação de atributos. Por essa razão, foi emitida a RFC 5420 (*Encoding Attributes for MPLS LSP Establishment Using RSVP-TE*), que, em complementação à RFC 3209, especificou os seguintes tipos de objeto:

— Objeto LSP_ATTRIBUTES (*Class-Num* 197), com um *C-Type* (*C-Type 1*);
— Objeto LSP_ REQUIRED_ATTRIBUTES (*Class-Num* 67), com um *C-Type* (*C-Type 1*).

5.5.11.3. Protocolo GMPLS RSVP-TE

A RFC 3473 (*GMPLS Signaling RSVP-TE Extensions*), com base na RFC 3471 (*GMPLS Signaling Functional Description*), define o protocolo GMPLS RSVP-TE para a constituição de GMPLS LSPs. A RFC 4974 (*GMPLS RSVP-TE Signaling Extensions in Support of Calls*) provê as extensões do GMPLS RSVP-TE para a constituição de chamadas (*calls*).

Em determinadas topologias de rede, pode ser vantajoso manter associações entre pontos de terminação e pontos principais de trânsito para suportar uma instância de serviço. Tais associações são denominadas *Calls* (chamadas). Uma chamada não provê a conectividade efetiva para a transmissão de tráfego de usuário, senão estabelece apenas uma relação mediante a qual uma ou múltiplas conexões podem então ser construídas. No GMPLS tais conexões denominam-se GMPLS LSPs.

A RFC 3473 objetiva ampliar a funcionalidade do protocolo RSVP-TE para atender, particularmente, as seguintes características apresentadas pelo GMPLS:

— Atendimento a múltiplas tecnologias de rede, incluindo redes modo circuito;
— Estabelecimento de GMPLSLSPs bidirecionais simétricos e congruentes em um único processo sinalização.

A RFC 3473 introduziu, em complemento às mensagens já definidas para os protocolos RSVP e RSVP-TE já apresentados neste capítulo, dois novos tipos de mensagens de sinalização, que são a mensagem *RSVP Notify* e a mensagem *RSVP Notify Ack* (*Acknowledgement*). A mensagem *RSVP Notify* foi definida para permitir a um LSR informar a LSRs não adjacentes falhas relacionadas com LSPs, uma vez que as mensagens RSVP de erros limitam-se ao alcance de LSRs adjacentes.

A RFC 3473 introduziu também oito novas classes de objeto:

— Objeto UPSTREAM_LABEL (*Class-Num* 35), com o *C-Type* do *label* utilizado;
— Objeto LABEL_SET (*Class-Num* 36), com um *C-Type*(*C-Type* 1);
— Objeto PROTECTION (*Class-Num* 37), com um *C-Type* (*C-Type* 1);
— Objeto SUGGESTED_LABEL (*Class-Num*129), com o *C-Type* do *label* utilizado;
— Objeto ACCEPTABLE_LABEL_SET (*Class-Num*130), com um *C-Type* (*C-Type* 1);
— Objeto RESTART_CAP (*Class-Num*131), com um *C-Type* (*C-Type* 1);
— Objeto NOTIFY_REQUEST (*Class-Num*195), com dois *C-Types* (*C-Type* 1 para IPv4 e *C-Type* 2 para IPv6);
— Objeto ADMIN_STATUS (*Class-Num*196), com um *C-Type* (*C-Type* 1).

A RFC 3473 define um novo *C-Type* para o objeto LABEL_REQUEST (*Class-Num* 19), denominado GENERALIZED LABEL_REQUEST *(C-Type 4*), que contém os seguintes campos fundamentais:

— *LSP Encoding Type*;
— *Switching Type;*
— *Generalized Payload Identifier* (GPID).

O parâmetro *LSP Encoding Type* indica o formato de codificação de LSPs (por exemplo: SONET/SDH/GigE, etc.) que será usado para os dados no LSP específico. Esse parâmetro representa a natureza do LSP, e não a natureza dos enlaces atravessados pelo LSP. Um enlace pode suportar um conjunto de formatos de codificação, o que significa que um enlace é capaz de transportar e comutar um sinal em um ou mais desses formatos de codificação, a depender da disponibilidade de recursos e da capacidade do enlace.

Os formatos permitidos para a codificação de LSPs encontram-se relacionados na Figura 5-29:

Valor	Tipo
1	Packet
2	Ethernet
3	ANSI/ETSI PDH
5	SDH ITU-T G.707/SONET ANSI T1.105
7	Digital Wrapper
8	Lambda (photonic)
9	Fiber
11	FiberChannel

Figura 5-29 – Formatos para a codificação de LSPs.

Os valores 4, 6 e 10 encontram-se reservados.

Os valores para o campo *Switching Type* foram já apresentados neste item, quando da referência ao sub-TLV *Interface Switching Capability Descriptor* (*Sub-TLV Type*15) do protocolo GMPLS OSPF-TE.

O parâmetro *Generalized Payload Identifier* (G-PID) é o identificador do *payload* transportado por um GMPLS LSP, ou seja, identifica a camada cliente desse GMPLS LSP. O G-PID é utilizado pelos nós localizados nos pontos de terminação do GMPLS LSP ou, em certos casos, no *penultimate hop*. Para *payloads* de Ethernet e de pacotes usam-se os valores *EtherType* padronizados. Para outros tipos de *payload*, a RFC 3471 relaciona 47 outros valores para o G-PID.

A RFC 3473 define também dois novos *C-Types* para o objeto LABEL (*Class-Num* 16):

— *C-Type* 2 (objeto GENERALIZED LABEL), que opera no sentido *upstream*;
— *C-Type 3* (objeto WAVEBAND SWITCHING).

O objeto UPSTREAM_LABEL (*Class-Num* 35) é enviado pelo nó *upstream* para o nó *downstream* com o propósito de possibilitar a constituição de GMPLS LSPspara tráfego de serviço no sentido *upstream*, para permitir, em um único processo de sinalização, a comunicação bidirecional. Esse *label* deve ser utilizado pelo nó *downstream* quando do envio de quadros de serviço no sentido *upstream*.

Os objetos LABEL_SET (*Class-Num* 36) e SUGGESTED_LABEL (*Class-Num* 129), que são transmitidos no sentido *downstream*, objetivam possibilitar ao nó *upstream* influenciar o nó *downstream* na escolha do *label* a ser por este enviado para uso do nó *upstream* quando do envio de quadros de serviço no sentido *downstream*.

A RFC 3473 foi complementada por outras RFCs que definem novas classes de objeto, novos *C-Types* e novos subobjetos:

— RFC 4872 (*RSVP-TE Extensions in Support of End-to-End GMPLS Recovery*);
— RFC 4873 (*GMPLS Segment Recovery*);
— RFC 4974 (*GMPLS RSVP-TE Signaling Extensions in Support of Calls*);
— RFC 5420 (*Encoding of Attributes for MPLS LSP Establishment Using RSVP-TE*)
— RFC 6002 (*GMPLS Data Channel Switching Capable (DCSC) and Channel Set Label Extensions*);
— RFC 6510 (*RSVP Message Formats for LSP Attributes*).

A RFC 5420, anteriormente citada neste capítulo e que complementa o protocolo RSVP-TE, complementa igualmente o protocolo GMPLS RSVP-TE.

5.5.11.4. Constituição de P2MP TE LSPs (RFC 4875)

A RFC 4875 (*Extensions to RSVP-TE for P2MP TE LSPs*) define extensões do protocolo RSVP TE para a constituição de P2MP TE LSPs unidirecionais, no MPLS TE e em redes que utilizam o GMPLS no plano de controle, no sentido de um LSR de ingresso (*Source LSR*) para múltiplos LSRs de egresso (*Leaf LSRs*). A

RFC 4875 tem como suporte a RFC 4461 (*Signaling Requirements for Point-to-Multipoint TE MPLS LSPs*).

Um P2PM TE LSP unidirecional é composto pos múltiplos S2L (*Source-to-Leaf*) sub-LSPs. Uma mensagem *RSVP Path* pode sinalizar a constituição de um ou de múltiplos S2L sub-LSPs de um único P2PM TE LSP unidirecional em um único processo de sinalização. Um P2PM TE LSP unidirecional pode ser constituído em um ou em múltiplos processos de sinalização.

A RFC 4875 especifica um novo objeto e amplia cinco objetos já existentes. O novo objeto definido é o objeto S2L_SUB_LSP (*Class-Num* 50), com 2 *C-Types* (*C-Type* 1 para o IPv4 e *C-Type* 2 para o IPv6).

Foram ampliados o objeto SESSION, (*Class-Num*1), o objeto FILTER_SPEC (*Class-Num* 10), o objeto SENDER_TEMPLATE (*Class-Num* 11), além do objeto SECONDARY_EXPLICIT_ROUTE, também referido como SERO (*Class-Num* 200), e o objeto SECONDARY_RECORD_ROUTE, também referido como SRRO (*Class-Num* 201).

Os objetos SERO e SRRO já haviam sido definidos na RFC 4873 (*GMPLS Segment Recovery*).

A sinalização para a constituição de P2MP TE LSPs bidirecionais no GMPLS não foi ainda especificada nem na RFC 4875 nem na RFC 6060. O assunto encontra-se no *draft A Framework for Point-to-Multipoint MPLS in Transport Networks*.

5.5.11.5. P2P GMPLS TE LSPs Assimétricos

A RFC 3473 limita-se à constituição de P2P GMPLS LSPs bidirecionais congruentes e simétricos (com reserva da mesma largura de banda nos dois sentidos). Foi emitida a RFC 6387 (*GMPLS Asymmetric Bandwidth Bidirectional LSPs*) com o objetivo de ampliar os termos da RFC 3473 para permitir a constituição de P2P GMPLS LSPs congruentes e assimétricos.

A RFC 6387 definiu três novos objetos com esse propósito:

- Objeto UPSTREAM_FLOWSPEC (*Class-Num* 120), com os mesmos C-*Types* que o objeto FLOWSPEC (*Class-Num* 9);
- Objeto UPSTREAM_TSPEC (*Class-Num* 121), com os mesmos C-*Types* que o objeto SENDER_TSPEC (*Class-Num* 12);
- Objeto UPSTREAM_ADSPEC (*Class-Num* 122), com os mesmos C-*Types* que o objeto ADSPEC (*Class-Num* 13).

5.5.12. GMPLS em Redes Ethernet

A RFC 5828 (*Generalized Multiprotocol Label Switching (GMPLS) Ethernet Label Switching Architecture and Framework*) especifica a arquitetura e a visão geral do que se denominou *GMPLS Ethernet Label Switching* (GELS), ou seja, da aplicação do GMPLS em redes Ethernet.

O papel das redes Ethernet está se expandindo rapidamente no sentido de que elas se tornem redes de transporte modo pacote orientadas a conexão, ou seja, que se tornem PTNs (*Packet Transport Networks*).

A RFC 5828 objetiva definir um plano de controle baseado no GMPLS para redes Ethernet orientadas a conexão e capacitadas como PTNs. O GMPLS é utilizado com propósito similar também para outras tecnologias, como, por exemplo, o MPLS-TP.

Para uso específico em Ethernet como rede de transporte, são necessárias algumas extensões do plano de controle GMPLS. A RFC 5828 tem como objetivo prover um quadro geral e uma arquitetura para essas extensões.

Provavelmente, GELS irá requerer mais de um tipo de comutação para suportar os diversos modelos que surgirão. Uma vez que as extensões necessárias irão depender do tipo de comutação, serão emitidos documentos específicos para cada tecnologia.

O PBB-TE constitui um exemplo de rede Ethernet orientada a conexão, ou seja, um exemplo de rede CO-Ethernet (*Connection-Oriented Ethernet*) e orientada a transporte, enquadrando-se, portanto, na aplicação do GELS.

Os trajetos virtuais em redes Ethernet de transporte definidos conforme a RFC 5828, estabelecidos com base no GMPLS, são referidos como *Eth-LSPs* (*Ethernet Label Switched Paths*), que podem ser unidirecionais ou bidirecionais.

Uma P2P TESI é suportada por um único *P2P Eth-LSP* bidirecional ou, alternativamente, por dois *P2P Eth-LSP* unidirecionais congruentes em sentidos opostos.

Uma P2MP TESI é suportada por um conjunto formado por um *P2MP Eth-LSP* da raiz para as folhas e um *P2P Eth-LSP* congruente de cada uma das folhas para a raiz.

Conforme menção anterior, o GMPLS utiliza diferentes descritores de ISC (*Interface Switching Capabilities*). A RFC 5828 estabelece que as L2SC ISCs não serão utilizadas para representar interfaces capacitadas para suportar *Eth-LSPs* definidos nessa RFC ou em outros documentos normativos do IETF em suporte a paradigmas de comutação Ethernet de transporte.

Devem ser definidas, na medida da necessidade, uma ou mais novas ISCs para suportar *Eth-LSPs* em cada tipo de serviço de comutação Ethernet específico que venha a se constituir em exemplo de aplicação do GELS. Para o PBB-TE foi definida, pela RFC 6060, a ISC denominada 802_1 PBB-TE, com o código 40.

5.5.13. **GMPLS no PBB-TE (RFC 6060)**

A RFC 6060 (*GMPLS Control of Ethernet PBB-TE*) define uma especificação complementar à RFC 5828 abordada no item anterior e que especifica o GELS.

A RFC 6060 descreve aspectos tecnológicos específicos da aplicação do plano de controle GMPLS para o PBB-TE. São descritos as extensões e os mecanismos do

GMPLS para o estabelecimento de conexões ponto a ponto, e menciona-se a possiblidade de constituição de conexões ponto a multiponto no PBB-TE.

A RFC 6060 suporta, mas não modifica, o plano de dados do PBB-TE especificado no padrão IEEE 802.1Q-2014 anteriormente abordado neste capítulo.

A rigor, a RFC 6060 restringe-se a aspectos concernentes à sinalização no PBB-TE, que ocorre com base no protocolo GMPLS RSVP-TE com pequenas extensões.

Outra restrição observada na RFC 6060 refere-se à não especificação da sinalização para a constituição de *P2MP Eth-LSPs* bidirecionais, embora tenha sido mencionada a viabilidade de adaptação das RFC 3473 e RFC 4875 para sinalizar o estabelecimento de *P2MP Eth-LSPs* unidirecionais da raiz (ponto) para as folhas (multiponto). Consequentemente, este subitem restringe-se à sinalização para a constituição de *P2P Eth-LSPs*.

O PBB-TE, conforme menção anterior, é projetado para ser bidirecional e roteado de forma simétrica, como a Ethernet. Ou seja, a funcionalidade completa dos protocolos Ethernet só é garantida para *Eth-LSPs* bidirecionais. Por conseguinte, a RFC 6060 contempla somente o estabelecimento de *P2P Eth-LSPs* bidirecionais.

No entanto, ressalva-se que é possível de uso de extensões do GMPLS RSVP-TE para configurar ESPs unidirecionais. Neste caso, o objeto UPSTREAM_LABEL não é incluído na mensagem PATH.

O roteamento no PBB-TE não é tratado na RFC 6060. Entretanto, assim como nas demais aplicações utilizando GELS, esse roteamento deve se basear no processo de roteamento do GMPLS. Isso significa que o roteamento no PBB-TE baseia-se no protocolo GMPLS OSPF-TE, definido pela RFC 4203 (*OSPF Extensions in Support of MPLS*), ou, alternativamente, no protocolo GMPLS IS-IS TE, definido pela RFC 5307 (*IS-IS Extensions for GMPLS*).

5.5.13.1. Procedimentos Básicos

Quando do envio de mensagens RSVP *Path* para inicializar o processo de constituição de um *P2P Eth-LSP* bidirecional, o iniciador deve utilizar os procedimentos definidos na RFC 3473, com as seguintes especificidades:

- Deve setar o LSP *Encoding Type* para *Ethernet* (código 2, conforme a RFC 3471);
- Deve setar o LSP *Switching Type* para 802_1 PBB-TE (código 40);
- Deve setar o G-PID para *Ethernet* (valor 33, conforme a RFC 3471);
- Deve setar o objeto *UPSTREAM_LABEL* para o valor de *label* por ele (ou seja, pelo iniciador) definido, que é identificado pela *tuple* <ESP-MAC1, ESP-VID1>, onde o ESP-VID1 é administrado localmente para o ESP-MAC1, que representa o endereço MAC do iniciador;

— Deveria setar o objeto *LABEL_SET* ou o objeto *SUGGESTED_LABEL* caso o iniciador deseje influenciar a escolha do *label* a ser utilizado no tráfego *downstream* (<ESP-MAC DA, ESP-VID>).
— Pode transportar um valor de I-SID via *Call/Connection* ID (conforme a RFC 4974).

O processamento nas *bridges* intermediárias (BCBs) e nas *bridges* de borda de egresso (IB-BEBs de egresso) não é alterado com relação à RFC 3473. Entretanto, as BCBs que suportam o LSP *Switching Type* setado para 802_1 PBB-TE não devem alterar os valores de *label* recebidos.

A *tuple* <ESP-MAC1, ESP-VID1> contida no UPSTREAM_LABEL é usada para criar uma entrada estática na base de dados de filtragem (FDB) na *bridge* presente em cada um dos *hops* no sentido *upstream*. A necessidade de inclusão das entradas estáticas para o sentido *upstream* é inferida do LSP *Switching Type* (correspondente a 802_1 PBB-TE) no objeto *GENERALIZED LABEL_REQUEST* contido na mensagem *RSVP Path*.

No destino, o ESP-VID2 é administrado localmente para o ESP-MAC2. A *tuple* <ESP-MAC2, ESP-VID2> é informada pelo objeto LABEL contido na mensagem *RSVP Resv*. Essa *tuple* será incluída em entradas estáticas nas FDBs das *bridges* ao longo do caminho no sentido oposto.

Esse procedimento tem como resultado a criação de um *Eth-LSP* bidirecional, uma vez que as mensagens *RSVP Path* e *RSVP Resv* seguem caminhos totalmente congruentes.

5.5.13.2. Encaminhamento Compartilhado (Shared Forwarding)

Quando do estabelecimento de conexões P2P PBB-TE entre múltiplas IB-BEBs e uma outra IB-BEB e se utiliza um único ESP-VID para os *Eth-LSPs* no sentido das múltiplas IB-BEBs para essa IB-BEB (ou seja, *Eth-LSPs* utilizando uma mesma *tuple* <ESP-MAC DA, ESP-VID>), cria-se a possibilidade de utilização do que se denomina encaminhamento compartilhado (*shared forwarding*).

Caso esses *Eth-LSPs* convirjam, total ou parcialmente, para uma determinada BCB, existirá um único caminho dessa BCB para a IB-BEB. Assim, basta uma única entrada estática de encaminhamento (*forwarding entry*) nas FDBs a partir do ponto de convergência para o encaminhamento de quadros desses *Eth-LSPs*.

A despeito do encaminhamento compartilhado, os *Eth-LSPs* que compartilham o caminho comum mantêm independentes os seus parâmetros de qualidade de serviço.

5.5.13.3. PBB-TE Ethernet Label

A RFC 6060 definiu um novo tipo de *generalized label* para uso em cada sentido no PBB-TE GMPLS RSVP-TE, denominado *PBB-TE Ethernet Label*.

O *PBB-TE Ethernet label* possui o formato representado na 5.30.

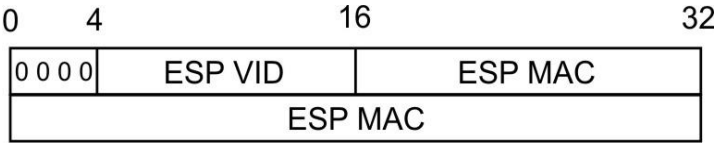


Figura 5-30 - Formato do PBB-TE Ethernet label.

Esse formato deve ser utilizado também para os *P2MP Eth-LSPs*, com a ressalva de que o ESP MAC, que é o endereço MAC de destino dos quadros a serem transmitidos, corresponde,no sentido da raiz para as folhas,a um endereço MAC de grupo.

Em qualquer hipótese, os *PBB-TE Ethernet labels* devem ser globalmente únicos na PBBN-TE de suporte, o que é obtido com base na unicidade global dos endereços ESP-MAC.

Ressalva-se que os valores do *PBB-TE Ethernet label* não são alterados ao longo dos *Eth-LSPs* a que dizem respeito.

5.5.13.4. Identificação de Instâncias de Serviço

Valores de I-SID são utilizados para identificar serviços em uma PBBN, inclusive em uma nuvem PBB-TE existente em uma PBBN. Para eliminar ambiguidade, os valores de I-SID devem ser globalmente únicos na rede ou, no mínimo, entre um par qualquer de *bridges* de borda.

O mapeamento da associação entre I-SIDs e ESPs é registrado em tabelas de instâncias de serviço *backbone* (*Backbone Service Instance Tables*) contidas em IB-BEBs. Essa associação pode ocorrer de forma manual ou de forma semiautomática por sinalização conforme a RFC 6060.

O processo de sinalização para o estabelecimento de mapeamento entre I-SIDs e ESPs ocorre mediante o envio de um novo TLV, denominado *Service* ID TLV, definido na RFC 6060, quando do estabelecimento de uma chamada (*Call Signaling*) conforme a RFC 4974 ou, alternativamente, quando do estabelecimento de um *Eth-LSP*. Esse envio ocorre em mensagens *RSVP Path*.

No primeiro caso, o novo *Service ID TLV* pode ser enviado no objeto CALL_ATTRIBUTES definido na RFC 4974, para indicar o mapeamento I-SID / ESP para o *Eth-LSP* que será estabelecido em associação com a chamada.

Na opção alternativa, o novo *Service ID TLV* pode ser transportado no objeto LSP_ATTRIBUTES definido na RFC 5420 no processo de sinalização para o estabelecimento do *Eth-LSP*.

O formato do *Service* ID TLV definido encontra-se na 5.31.

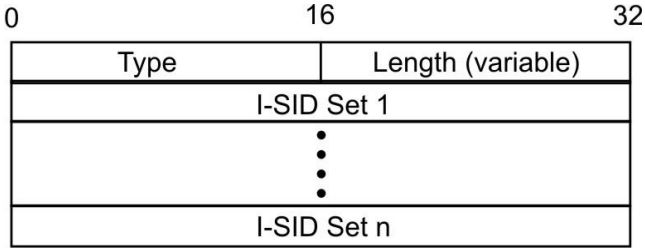


Figura 5-31 - Formato do Service ID TLV.

O campo *Type* possui o valor 3 quando o *Service ID TLV* é transportado no objeto CALL_ATTRIBUTES e o valor 2 quando transportado no objeto LSP_ATTRIBUTES.

O objeto I-SID *Set* é utilizado para informar uma lista ou uma faixa de valores de I-SID.

CAPÍTULO 6

ETHERNET SOBRE SDH E OTN

6.1. PREÂMBULO

Quadros Ethernet são transportados sobre redes Ethernet e por diversas outras tecnologias de rede. O transporte de Ethernet sobre redes Ethernet foi detalhadamente apresentado anteriormente no capítulo 4 (*LAN Bridging* com IS-IS) e capítulo 5 (*Bridging* em Redes de Provedores).

Adicionalmente, podem ser citadas as seguintes opções de transporte de Ethernet:

- Ethernet sobre MPLS (VPLS e VPWS);
- Ethernet sobre SDH;
- Ethernet sobre OTN;
- Ethernet sobre redes de WDM;
- Ethernet sobre fibras ópticas;
- Ethernet sobre RPR (*Resilient Packet Ring*);
- Ethernet sobre ATM;
- Ethernet sobre *Frame Relay*.

Serão abordados, neste capítulo, apenas o transporte de Ethernet sobre SDH e sobre OTN.

O capítulo 7 deste livro é dedicado à utilização do MPLS para transporte de Ethernet.

6.2. HIERARQUIA DIGITAL SÍNCRONA (SDH)

O SDH é a hierarquia de multiplexação hoje predominante em sistemas de transmissão digital síncrona. Ela é composta por agregados de linha, denominados STM-N (*Synchronous Transport Modules* de ordem N), onde N pode assumir os valores 1,4,16,64 e 256.

O agregado de linha básico é o STM-1, que representa a transmissão de 2.430 octetos (9 x 270) em um quadro, a cada 125 μs. Isso significa que esse quadro é transmitido 8.000 vezes por segundo, resultando a velocidade de transmissão de 155,520 Mbit/s.

Como será visto adiante neste subitem, os quadros STM-N possuem *payloads* básicos onde são mapeados sinais tributários Esses *payloads* são referidos como *containers* (C-12 e C-4, por exemplo). Os sinais tributários podem corresponder aos sinais digitais do PDH no padrão americano ou no padrão europeu. Outros tipos de sinais tributários foram especificados, como células ATM, por exemplo.

Adota-se no Brasil o padrão europeu, definido pelo ETSI (*European Telecommunications Institute*).

A arquitetura SDH é composta por três camadas:

— Seção de regeneração, ou camada de seção;
— Seção de multiplexação, ou camada de linha;
— Camada de trajeto.

A cada uma dessas camadas corresponde um tipo de *overhead*, que é processado, respectivamente, pelos regeneradores, equipamentos terminais de linha (LTEs) e equipamentos terminais de trajeto (PTEs). Tais equipamentos são partes integrantes de um circuito SDH.

Os equipamentos terminais do SDH apresentam-se nos seguintes tipos básicos:

— DXC (*Digital Cross Connect*);
— ADM (*Add-Drop Multiplexer*);
— RDXC (*Reconfigurable DXC*);
— RADM (*Reconfigurable ADM*);
— TM (*Terminal Multiplexer*).

Os ADMs caracterizam-se pela possibilidade de derivar sinais tributários do SDH ao longo de topologias em barramento ou anel.

Embora não sendo parte integrante do SDH, utiliza-se um agregado de linha que opera à taxa de 51,840 Mbit/s, designado por STM-0, que corresponde ao STS-1 (Synchronous Transport Signal de ordem 1) da hierarquia SONET. A hierarquia SONET antecedeu o SDH, sendo de uso disseminado nos Estados Unidos.

O SDH encontra-se especificado nas recomendações do ITU-T G.707, G.708 e outras.

6.2.1. **O Quadro STM-N**

O quadro STM-N tem a estrutura apresentada na Figura 6.1.

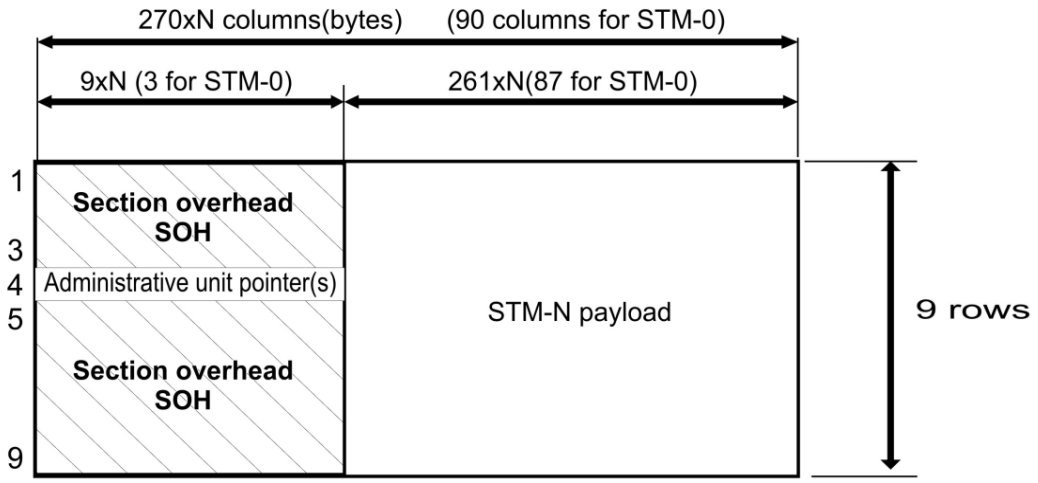


Figura 6-1 – Estrutura do quadro STM-N.

Observa-se nessa figura a presença do *overhead* de seção (SOH) e do(s) ponteiro(s) de unidade administrativa. O SOH é composto pelo RSOH (*Regenerator Section Overhead*), ocupando Nx27 octetos (NX3x9), e pelo MSOH (*Multiplex Section Overhead*), ocupando Nx45 octetos (Nx5X9).

Entre o RSOH e MSOH encontram-se o(s) ponteiro(s) de unidade administrativa, cuja função será vista adiante neste item.

Essa figura não exibe o POH (*Path Overhead),* componente incluído no *STM-N payload* (com Nx261 octetos) ao longo da montagem do quadro STM-N. O POH é composto em diferentes níveis.

Um quadro STM-N, como se observa na figura, é produto da multiplexação de N quadros STM-1 no período de 125 μs, do que resulta as velocidades de Nx 155,520 Mbit/s para um STM-N.

6.2.2. **Elementos de Sinal do SDH**

Serão abordados, a seguir, os elementos de sinal da estrutura de multiplexação do SDH.

- **Containers (Cs)**

Os *containers* são as unidades mais elementares da estrutura STM-N. Todos os tributários externos, inclusive os tributários PDH, são mapeados nos respectivos *containers*.

Existem os seguintes tipos de *containers*, na estrutura de multiplexação do SDH no padrão europeu: C-12, C-3, C-4, C-4-4c, C-4-16c, C-4-64c e C-4-256c. Essa relação não considera os tributários PDH do padrão ANSI (*American National Standarts Institute*).

- **Virtual Containers (VCs)**

Um VC corresponde à carga transportada pela rede SDH entre terminações de seção de trajeto na rede. Um VC é composto por um *container* (VC *payload*) e pelo respectivo POH (*Path Overhead)*.

Na hierarquia ETSI, adotada no mundo inteiro exceto nos Estados Unidos, os trajetos de baixa ordem (LPs, *low-order paths*) são constituídos por VCs-12 ou, alternativamente, por VCs-3. Estes são mapeados em VCs-4, que constituem os trajetos de alta ordem (HPs, *high-order paths*).

Um VC-4 pode estar estruturado como um espaço único de carga destinado a transportar uma carga externa de velocidade alta, ou estar subdividido em unidades tributárias, as quais recebem os VCs de baixa ordem.

- **Unidades Tributárias (TUs)**

Uma TU tem como objetivo o provimento de adaptabilidade entre camadas de trajeto de alta ordem e camadas de trajeto de baixa ordem. Uma TU é constituída pela adição de um ponteiro de TU (TU PTR) a um VC de baixa ordem, sendo o ponteiro utilizado para indicar o posicionamento do VC de baixa ordem com relação à posição de início do quadro do VC de alta ordem.

No padrão ETSI existem a TU-12 e a TU-3.

- **Grupos de Unidades Tributárias (TUGs)**

O papel de um TUG é coletar um ou mais TUs e colocá-los em uma posição fixa no payload de um VC de mais alta ordem. Não há acréscimo de *overhead* quando um TUG é formado por TUs.

Existem duas classes de TUG, que são o TUG-2 e o TUG-3.

- **Unidades Administrativas (AUs)**

Uma AU funciona como um adaptador entre a camada de trajeto de ordem mais elevada e a camada de seção de multiplexação. Uma AU consiste em um *payload* e um ponteiro de AU (AU PRT).

O *payload* transporta um VC de alta ordem e o AU PTR indica a posição relativa do VC de alta ordem ao longo da multiplexação em cascata proporcionada pelos AUGs.

- **Grupos de Unidades Administrativas (AUGs)**

A área de carga do sinal de linha STM-N constitui um grupo de unidades administrativas (AUG). Um AUG agrega uma ou mais AUs inseridas em um quadro STM-N. Um AUG-4, por exemplo, corresponde à área de carga de um sinal STM-4 e, na hierarquia ETSI, é composto por quatro AUs-4.

6.2.3. **Mapeamento de Tributários PDH**

Os tributários PDH (padrão europeu) são mapeados no SDH conforme a figura 6.2.

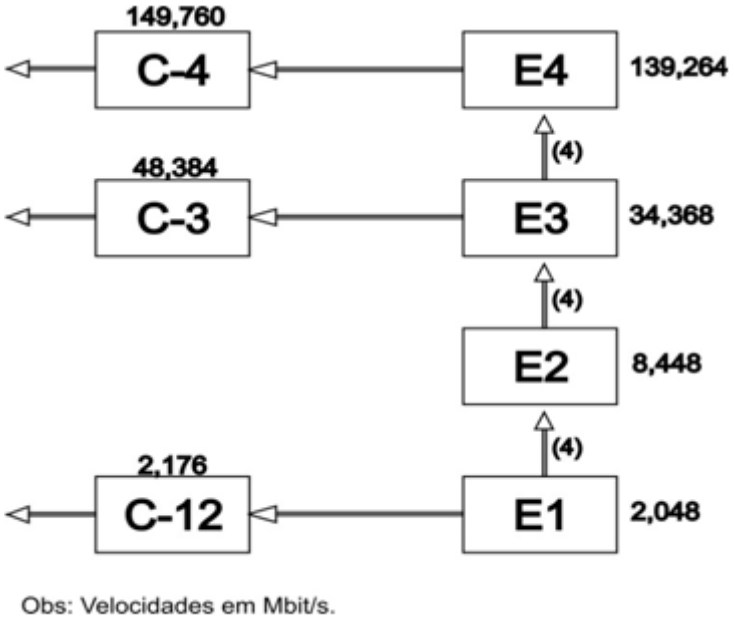


Figura 6-2 – Estrutura do quadro STM-N.

Como se observa nessa figura, o sinal E2 (8,448 Mbit/s) não é mapeado diretamente no SDH.

6.2.4. **Resumo de Velocidades**

A Figura 6.3 apresenta um resumo de velocidades utilizadas pelo SDH e seus tributários PDH no padrão europeu.

STM-N		VCs			Tributários PDH	
Tipo	Velocidade	Tipo	Velocidade de Quadro	Velocidade de *Payload*	Tipo	Velocidade
—	—	VC-12	2,240	2,176	E1	2,048
STM-0	51,840	VC-3	48,960	48,384	E3	34,368
STM-1	155,520	VC-4	150,336	149,760	E4	139,264
STM-4	622,080	VC-4-4C	601,344	599,040	—	—
STM-16	2.488,320	VC-4-16C	2.405,376	2.396,160	—	—
STM-64	9.953,280	VC-4-64C	9.621,504	9.584, 640	—	—
STM-256	39.813,120	VC-4-256C	38.486,016	38.338,560	—	—

Obs: Velocidades em Mbit/s.

Figura 6-3 – Estrutura do quadro STM-N.

6.2.5. **Next Generation SDH (NG-SDH)**

Para atenuar a ineficiência inerente ao SDH, que opera com tributários discretos com valores de capacidade muito afastados, o ITU-T especificou uma técnica fim a fim complementar ao SDH, referida como VCAT (*Virtual Concatenation*).

O uso da VCAT (que será abordado com maior profundidade adiante neste item) aumenta consideravelmente a eficiência de uso do SDH como suporte a diferentes tecnologias de rede, dentre as quais se encontra a *Ethernet*.

Como as necessidades de vazão de tráfego por parte dos usuários podem variar ao longo do tempo, tornou-se necessária a definição de um mecanismo de redimensionamento automático dos agrupamentos obtidos pela VCAT, sem a interrupção do tráfego que por eles cursa.

Esse objetivo foi alcançado mediante a especificação do mecanismo referido como LCAS (*Link Capacity Adjustment Scheme*), pela recomendação ITU-T G. 7042/Y. 1304.

A VCAT e o LCAS serão abordados adiante neste item.

Para a intermediação do mapeamento de quadros de diferentes tecnologias no SDH foram especificados, pelo ITU-T, o GFP (*Generic Framing Procedure*) e, como opção alternativa, o LAPS (*Link Access Procedure-SDH*), que serão vistos adiante neste capítulo para o caso particular de mapeamento de quadros Ethernet.

O uso combinado do GFP (ou mesmo do LAPS), da VCAT e do LCAS representa a otimização da funcionalidade e a maximização da eficiência do SDH como tecnologia de transporte de pacotes. A utilização desse conjunto para o transporte de diferentes tecnologias sobre o SDH denomina-se *Next Generation SDH* (NG-SDH).

No contexto do NG-SDH, o uso do GFP ou do LAPS independe da utilização da VCAT e do LCAS. O GFP pode ser mapeado em SDH com ou sem o uso da VCAT. A maximização de benefícios ocorre, contudo, com uso conjunto desses procedimentos.

A VCAT pode ser utilizada independentemente do LCAS, mas o uso do LCAS está condicionado ao uso da VCAT.

6.2.6. **Concatenação de VCs SDH**

Os VCs especificados para o SDH básico, que são os VC-12, VC-3 e VC-4 se for considerado apenas o acesso de tributários PDH no padrão europeu, possuem, como visto anteriormente, as velocidades dos respectivos *payloads* iguais a 2,176 Mbit/s, a 48,384 Mbit/s e a 149,760 Mbit/s.

Essas alternativas de capacidade apresentam as seguintes limitações:

— Alcançam limites reduzidos de velocidade, impossibilitando o atendimento de aplicações de alta vazão de tráfego;
— Apresentam elevado afastamento entre as suas velocidades, tornando ineficiente o atendimento de aplicações com velocidades intermediárias.

Para solucionar a primeira limitação, foi definida a tecnologia denominada concatenação contígua de VCs SDH. A concatenação contígua de VCs possibilita a elevação das velocidades do SDH, mas amplia ainda mais os afastamentos entre velocidades.

Para atenuar a limitação causada pelo afastamento de velocidades entre os VCs básicos e entre VCs obtidos por concatenação contígua de VCs, o ITU-T especificou uma nova tecnologia, referida com concatenação virtual (VCAT) de VCs SDH.

Essas tecnologias foram definidas na recomendação G.707 para o SDH, e pelo padrão ANSI T.105 para o SONET.

6.2.6.1. Concatenação Contígua de VCs

A concatenação contígua de VCs é obtida pela aglutinação de VCs ao longo de todo um trajeto SDH. Em todos os elementos de rede desse trajeto ocorre a concatenação de VCs.

Um conjunto de VCs concatenados de forma contígua é representado pela forma VC-n-Xc, onde **n** representa os VCs concatenados (VCs-11, VCs-12, VCs-2, VCs-3 ou VCs-4), e Xc representa o número de VCs concatenados de forma contígua (daí a letra ***c***).

Para o padrão europeu foram especificados apenas os agrupamentos VC-4-Xc (X igual a 1,4,16,64 ou 256).

Os agrupamentos VC-4-Xc objetivam exclusivamente a ampliação de velocidades. Os STMs (STM-4, STM-16, STM-64 e STM-256) podem resultar do mapeamento de um dos respectivos agrupamentos VC-4-Xc.

6.2.6.2. Concatenação Virtual (VCAT)

Na tecnologia VCAT, especificada na recomendação ITU-T G.707/Y.1322 (*Network Node Interface for the Synchronous Digital Hierarchy – SDH*), ao contrário do que ocorre na concatenação contígua, apenas as terminações do agrupamento de VCs participam da concatenação. Os VCs concatenados mantêm as suas individualidades ao longo dos respectivos trajetos (que podem ser diferentes), sendo recombinados nas terminações comuns desses trajetos.

Um conjunto de VCs concatenados em conformidade com a VCAT é representado pela forma VC-n-Xv, de modo assemelhado à concatenação contígua, sendo que a letra **v** indica concatenação virtual.

Os agrupamentos VC-n-Xv (ou agrupamentos VC-m-Xv) são também referidos como VC-n-Xv (ou VC-m-Xv) VCGs (*Virtual Concatenation Groups*), ou simplesmente como VCGs.

Ao contrário do que ocorre na concatenação contígua de VCs, onde os valores de X na expressão VC-n-Xc assumem valores definidos, o VCAT possibilita a concatenação de números arbitrários de VCs, respeitado o limite máximo de 256 VCs. Dessa forma, a capilaridade, e a conseqüente eficiência de uso da rede, é amplamente incrementada.

A recomendação ITU-T G.7043/Y.1343 (*Virtual Concatenation of Plesiochronous Digital Hierarchy – PDH*), complementando a definição da VCAT, define a concatenação virtual de sinais PDH (*Plesiochronous Digital Hierarchy*).

- **Velocidades SDH com VCAT**

A Figura 6.4 representa as possíveis concatenações virtuais de VCs SDH, com as velocidades dos respectivos *payloads*.

VC SDH	VC-n-Xv (Velocidades de *Payload*)			
	Valor de X	X=1	X=64,X=256	Degraus
VC-11	1 a 64	1,600 Mbit/s	102,400 Mbit/s	1,600 Mbit/s
VC-12	1 a 64	2,176 Mbit/s	139,264 Mbit/s	2,176 Mbit/s
VC-2	1 a 64	6,784 Mbit/s	434,176 Mbit/s	6,784 Mbit/s
VC-3	1 a 256	49,536 Mbit/s	12,681 Gbit/s	49,536 Mbit/s
VC-4	1 a 256	149,760 Mbit/s	38,339 Gbit/s	149,760 Mbit/s

Figura 6-4 – Velocidades SDH com VCAT.

- **Vantagens e Restrições da VCAT**

Conforme menção anterior, a VCAT tem como principal razão o ganho de eficiência no uso dos VCs SDH em algumas aplicações, como no caso de Ethernet a 10 Mbit/s, 100 Mbit/s (*Fast Ethernet*) e 1 Gbits/s (*Gigabit Ethernet*).

O atendimento de 10 Mbit/s *Ethernet* por um VC-3 (49,536 Mbit/s) utiliza apenas 20% da banda disponível. O uso do VCG VC-12-5v (10,880 Mbit/s) permite a utilização de cerca de 92% dessa banda.

Para o atendimento do *Fast Ethernet* pelo SDH é utilizado um VC-4 (149,760 Mbit/s), o que representa o aproveitamento de 67% da banda. O uso do VCG VC-12-46v (100,096 Mbit/s) representa a utilização de praticamente 100% da banda.

No caso da *Gibabit Ethernet*, a utilização de um VC-4-16c (2,396 Gbit/s) resulta na utilização de cerca de 42% da banda. O uso do VCG VC-4-7v (1,048 Gbit/s) representa a utilização de cerca de 95% da banda.

Uma outra vantagem da VCAT é a possibilidade de utilização de diferentes trajetos SDH para a constituição de VCGs. Em caso de topologia em anel com ADMs, um VCG pode utilizar trajetos nos dois sentidos do anel.

Observa-se, por outro lado, que o uso de diferentes trajetos em um VCG ocasiona retardos diferenciais entre VCs concatenados. Para compensar os retardos diferenciais, a terminação de recepção do VCG utiliza um *buffer*, dimensionado de forma tal que o valor máximo de retardo diferencial seja de 256 ms.

Os VCs de um VCG recebem um mesmo valor de *label* para possibilitar a concantenação na recepção. Esse *label* é referido como *Multi-Frame Indicator* (MFI).

A par dessas vantagens, a VCAT apresenta duas restrições, que resultam da inflexibilidade no número de VCs que constituem um VCG.

Se um ou mais VCs falham, todo o VCG deixa de operar. Por outro lado, caso a necessidade de banda passante aumente ou diminua durante a utilização do VCG, o ajustamento do número de VCs só pode se efetivar com a interrupção da utilização.

O LCAS (*Link Capacity Adjustment Scheme*) objetiva eliminar essas restrições.

6.2.6.3. Link Capacity Adjustment Scheme (LCAS)

O LCAS, definido na recomendação ITU-T G.7042/Y.1305 (*Link Capacity Adjustment Scheme* (*LCAS*) *for Virtual Concatenated Signals*), consiste em uma técnica destinada a aumentar ou reduzir a capacidade de um sinal concatenado virtualmente em uma rede SDH ou em uma OTN. O LCAS, que opera sobre a VCAT, pode possibilitar a continuidade operacional de um VCG quando ocorre aumento ou redução do número de seus membros.

Ademais, o LCAS possibilita o ajuste automático da capacidade de um VCG em caso de falha de um ou mais de seus membros. Possibilita também o retorno dos membros que falharam à operação normal após o reparo das falhas.

O LCAS é um mecanismo unidirecional, sendo necessários dois processos em sentidos inversos para possibilitar uma alteração de capacidade de modo bidirecional. Para cada um desses processos existem um lado fonte, ou de origem, referido como *So,* e um lado de destino (*sink*) referido como *Sk*.

A recomendação ITU-T G.7042/Y.1305 define os requeridos estados na terminação de origem e na terminação de destino de um grupo VCAT (VCG). Define também as informações de controle intercambiadas entre essas terminações, necessárias para o redimensionamento flexível do sinal concatenado virtualmente.

Para maiores detalhes relativos aos campos que conduzem essas informações de controle, recomenda-se aos leitores consultar também as recomendações ITU-T G.707/Y.1322 e ITU-T G.783 (*Characteristcs of Synchronous Digital Hierarchy* (SDH) *Equipment Functional Blocks*), ambas para o SDH, e as

recomendações ITU-T G.709/Y.1331 (*Interfaces for the Optical Transport Networks*) e ITU-T G. 798 (*Characteristics of Optical Network Hierarchy Equipment Functional Blocks*), ambas para OTN.

- **Pacotes de Controle**

A sincronização entre as terminações *So* e *Sk* quando ocorrem alterações de capacidade de transmissão em um VCG é obtida pelo intercâmbio de pacotes de controle. Cada pacote de controle descreve o estado do *link* para o próximo pacote de controle. Dessa forma, o receptor de um pacote de controle pode mudar para a nova configuração imediatamente após a sua recepção.

Pacotes de controle transportam informações da VCAT e informações do LCAS. Tais pacotes contêm informações enviadas pela *So* e informações enviadas pela *Sk*.

A Figura 6.5 apresenta os diversos tipos de informações de controle em pacotes de controle VCAT/LCAS, com os respectivos sentidos de transmissão.

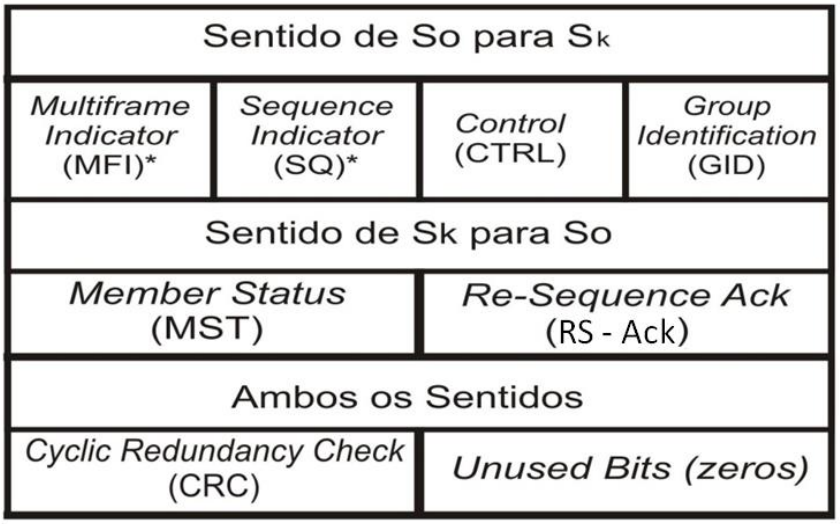


Figura 6-5 – Formato de um pacote de controle VCAT/LCAS.

A Figura 6.6 representa um resumo dos sentidos de transmissão de informações em um pacote de controle VCAT/LCAS.

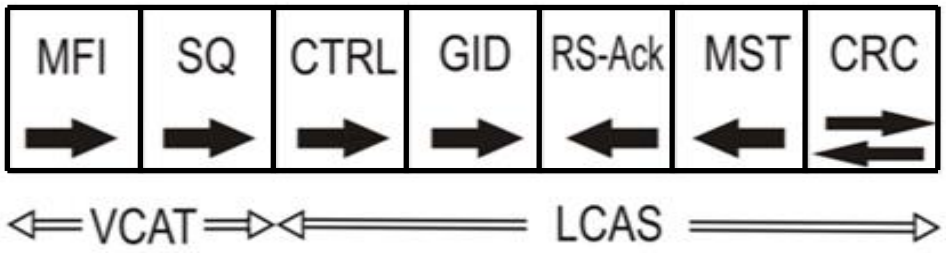


Figura 6-6 – Sentidos de transmissão de informações de controle.

- **Multiframe Indicator (MFI)**

Esse campo contém um valor que identifica o conjunto de todos os membros de um VCG. No lado *Sk*, o MFI deveria ser utilizado para realinhar os *payloads* de todos os VCs concatenados em um VCG, e para possibilitar a determinação de retardos diferenciais entre os VCs membros desse VCG.

- **Sequence Indicator (SQ)**

Esse campo representa o número sequencial de cada VC dentro de um VCG. O valor máximo possível para esse valor (X_M) é igual a 63 (para VCs de ordem inferior) e a 255 (para VCs de ordem superior). Observa-se que o valor inicial de SQ é zero.

Dentro de um VCG em funcionamento normal, o valor máximo de SQ (Xp) é o valor definido para o VCG. Por exemplo, para um VC-4-8v, Xp é igual a 8.

Quando ocorre uma redução temporária do número de VCs de um VCG, o valor máximo (X_A) é o valor real do SQ máximo. Por exemplo, se um VC-4-8v opera temporariamente como um VC-4-7v, X_A é igual a 7.

- **Control (CTRL)**

O campo *control* é usado para transferir informações da fonte para o destino, e contém os comandos de controle (palavras de controle) necessários para sincronizar os processos LCAS utilizados para iniciar alterações de faixas passantes. Esse campo indica o *status* individual de cada membro do VCG.

A Figura 6.7 representa as palavras de controle do LCAS (*LCAS CTRL Words*).

Valor	Comando	Observações
0000	FIXED	Indica o uso de banda fixa(modo não-LCAS)
0001	ADD	Esse membro está pronto para ser adicionado ao VCG
0010	NORM	Transmissão normal(transportando carga útil)
0011	EOS	Indicação de fim de sequência e de transmissão normal
0101	IDLE	Esse membro não é parte ou está pronto para remoção do VCG
1111	DNU	*Do Not Use* (o *payload*) porque Sk informou *Status Fail*

Figura 6-7 – LCAS Control Words (CTRL).

Quando da iniciação da fonte de um VCG, todos os membros devem enviar CTRL=IDLE até que eles sejam adicionados ao VCG (quando passam a enviar CTRL = ADD).

- **Group Identification (GID)**

Esse bit complementa o MFI na identificação de VCGs. O *GID bit* de um VCG tem o mesmo valor em pacotes de controle que possuam o mesmo valor de MFI.

O *GID bit* permite ao *Sk* verificar se todos os membros por ele recebidos originaram de um mesmo transmissor.

- **Re-Sequence Acknowledge (RS-Ack)**

Esse bit indica ao *So* uma renumeração dos valores de sequência de membros de um VCG que enviaram um campo CTRL indicando NORM, DNU ou EOS. Indica também que uma mudança dos valores de sequência desses membros foi detectada pelo Sk.

A notificação pelo *RS-Ack bit* realiza-se pela mudança de valor desse bit (de zero para 1 ou de 1 para zero).

- **Member Status (MST)**

O campo MST é utilizado pela *Sk* para transmitir informações para a So, a respeito do estado de todos os membros de um VCG para os quais a*So*tenha sinalizado com a palavra de controle ADD, NORM, EOS ou DNU.

Um pacote de controle que conduz o campo MST possui um bit para cada membro do VCG. Cada um desses bits pode reportar o estado OK (valor zero) ou o estado não OK (valor 1). Como um pacote de controle pode não conter bits

suficientes para todos os membros do VCG, podem ser utilizados múltiplos pacotes de controle para esse fim (o que se denomina MST multiframe).

O *Sk* utiliza, em cada informação MST, o mesmo valor SQ que ela recebeu da *So* para o respectivo membro do VCG.

- **Cyclic Redundancy Check (CRC)**

Para garantir as informações de controle do LCAS, cada pacote de controle é protegido por um CRC, aplicado após a completa recepção do pacote. Caso o CRC acuse a presença de erros, o pacote de controle deve ser rejeitado.

6.2.6.4. Operação do Protocolo LCAS

Conforme menção anterior, a operação do LCAS é unidirecional. Para adicionar ou remover membros bidirecionalmente, os procedimentos devem ser repetidos na direção oposta. Como esses procedimentos são independentes entre si, não há necessidade de sincronização entre eles.

A adição ou a remoção de membros de VCGs pode ocorrer de duas formas:

— Sob controle do sistema de gerência;
— De forma autônoma em caso de falhas ou reparo dessas falhas.

- **Procedimentos sob Controle do Sistema de Gerência**

A adição ou remoção de membros de um VCG ocorre sob controle do sistema de gerência da rede quando a operação é intencional. Nesse caso, os procedimentos ocorrem sem interrupção do tráfego (*hitless procedures*).

A adição de membros é acionada pelo comando ADD, sendo que cada membro adicionado recebe um valor de SQ. Esse processo consiste na adição de cada membro, seguida da adição do respectivo *payload*, processo esse que ocorre após a conclusão dos procedimentos necessários.

De forma similiar, ocorre a remoção planejada de membros de VCGs, também sob controle da gerência da rede.

Para evitar a ocorrência de interrupções em ambos os casos, foi desenvolvido um mecanismo de sincronização entre a *So* e a Sk.

- **Procedimentos Autônomos**

Quando em um membro de um VCG que envia um NORM ou um EOS ocorre uma falha de rede, essa condição de falha é detectada pela Sk, que envia um MST=FAIL para a *So*. A *So* altera a condição NORM ou EOS para a condição DNU e remove temporariamente o membro do VCG afetado pela falha.

Essa remoção ocorre de modo autônomo (automático), sem a interveniência da gerência da rede. De um modo geral ocorre interrupção de tráfego nesse procedimento (*non-hitless procedure*).

Quando o defeito que ocasiona a remoção é reparado, a *Sk* detecta esse fato. A *Sk* envia então um MST=OK para a *So* relativo ao membro afetado. A *So* altera a condição DNU para NORM ou EOS, e passa a utilizar novamente esse membro.

6.3. OPTICAL TRANSPORT NETWORK

A arquitetura para OTNs foi definida na recomendação ITU-T G.872 (*Architecture of Optical Transport Networks*). Essa recomendação foi conplementada pela recomendação ITU-T G.709/Y.1331 e pela recomendação ITU-T G.798 (*Characteristics of Optical Network Hierarchy Equipment Functional Blocks*) .

As OTNs podem ser utilizadas para o mapeamento nativo de sinais de cliente, para o mapeamento de canais SDH e para o mapeamento de quadros do GFP.

Embora o SDH possa alcançar velocidades de transmissão elevadas (foi especificado o STM-256 como limite máximo, correspondente a 40 Gbit/s), o desenvolvimento das OTNs se justificou por diferentes razões:

- Possibilidade de acréscimo de novos canais de *overhead* que acrescentariam as funcionalidades requeridas para a realização eficiente de funções de OAM&P na rede WDM;
- Provimento de uma forma mais poderosa de realização da função FEC (*Forward Error Correction*);
- Provimento de um envelope de *payload* menos granular para o transporte de sinais de cliente com larguras de banda mais elevadas.

A OTN foi inicialmente definida de forma limitada a 40 Gbit/s. A atual versão da recomendação ITU-T G.709/Y.1331 (2012), referida como OTN versão 3, suporta o mapeamento de sinais a 100 Gbit/s, através da ODU4. A OTN versão 3 suporta também a transmissão de sinais de 1 Gbit/s, pela ODU0.

A arquitetura de OTNs encontra-se resumida na Figura 6.8.

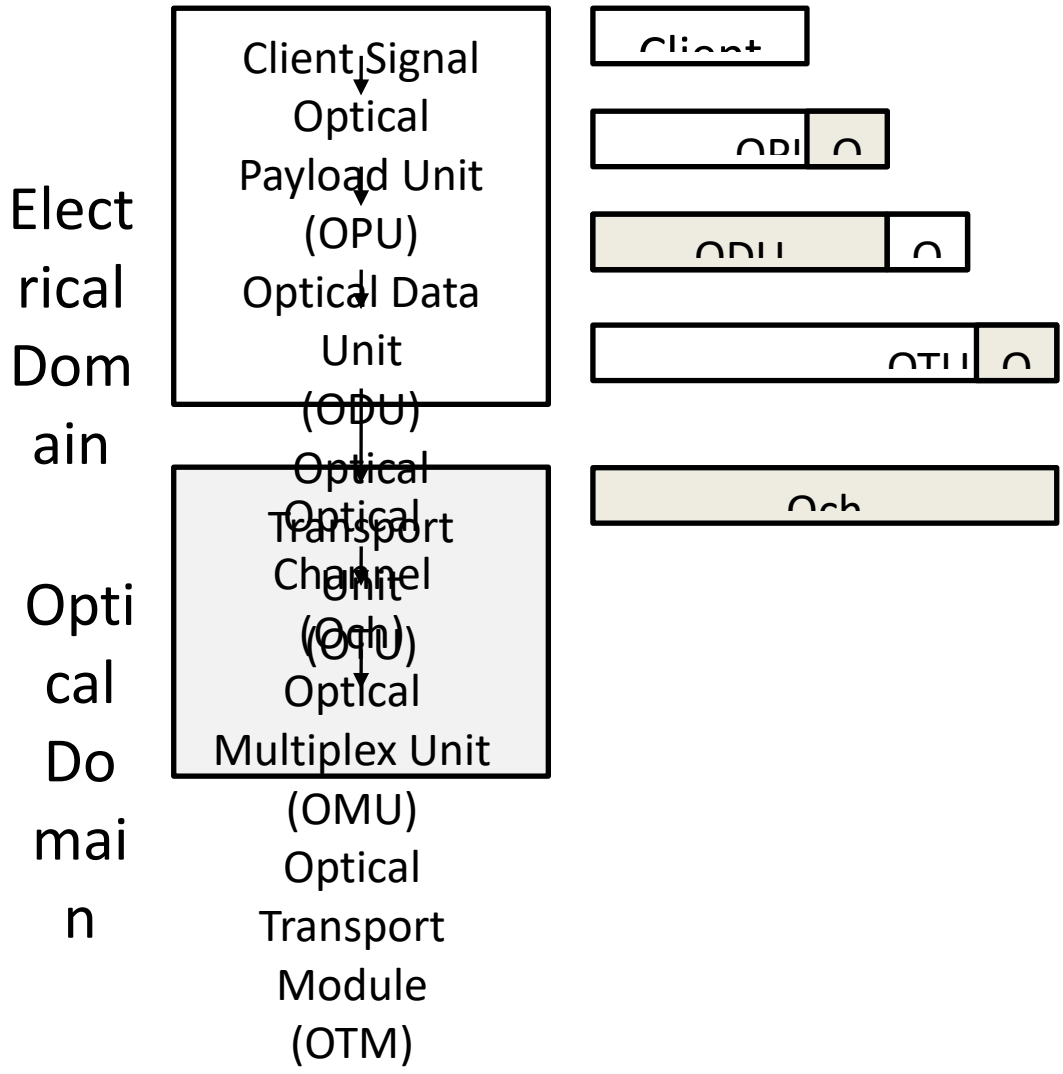


Figura 6-8 – Arquitetura de OTNs.

Como se observa nessa figura,um quadro OTU, ao penetrar no domínio óptico, passa a constituit um *Och* (*Optical Channel*). Múltiplos canais ópticos assim constituídos são multiplexados em frequência (em comprimento de onda, na verdade) pela OMU (*Optical Multiplex Unit*). Um sistema WDM (*Wavelenght Division Multiplex*) representa uma OMU.

O sinal óptico agregado na saída da OMU constitui o OTM (*Optical Transport Module*), que alimenta unidirecionalmente uma fibra óptica.

Estão especificados atualmente os seguintes tipos de ODU: ODU0, ODU1, ODU2, ODU2e, ODU3, ODU4 e ODUflex.

A Figura 6.9 apresenta as diferentes ODUs de alta ordem utilizadas como servidoras, assim como as ODUs de baixa ordem utilizadas como clientes de cada uma das ODUs servidoras. Nessa figura encontram-se relacionadas as velocidades de transmissão de todas as ODUs especificadas.

ODU Cliente	ODU Servidora
ODU0 (1,25 Gbit/s)	ODU1 (2,5Gbit/s)
ODU0, ODU1, ODUflex	ODU2 (10Gbit/s)
ODU0, ODU1, ODU2, ODU2e (10,3125 Gbit/s), ODUflex	ODU3 (40 Gbit/s)
ODU0, ODU1, ODU2, ODU2e, ODU3, ODUflex	ODU4 (100 Gbit/s)

Obs: Na ODUflex podem ser mapeados pacotes de clientes de 1,25 Gbit/s a 100 Gbit/s.

Figura 6-9 – ODUs servidoras e respectivas ODUs clientes.

Conforme a recomendação ITU-T G.709, as ODUs clientes são referidas como ODUj (j=0, 1, 2, 2e, 3 e flex), enquanto as ODUs servidoras são referidas como ODUk (k=1, 2, 3 e 4).

As OTUk são servidoras das ODUs de mesma classe (ou seja, das ODUk), como mostra a figura 6.10.

ODU Cliente	OTU Servidora
ODU0	-
ODU1	OTU1
ODU2	OTU2
ODU2e	-
ODU3	OTU3
ODU4	OTU4
ODUflex	-

Figura 6-10 – OTUk como servidoras de ODUs clientes (ODUk).

Verifica-se, nessa figura que as ODU0, ODU2e e ODUflex não possuem OTUs servidoras próprias, utilizando OTUs servidoras de outras ODUs. Assim a ODU2e, por exemplo, é cliente da ODU3 ou da ODU4, que por sua vez são clientes da OTU3 e da OTU4, respectivamente.

A ODU2 não é capaz de suportar um *bit stream* transparente à velocidade de 10 Gbit/s, necessário, por exemplo, para o transporte da velocidade de linha Ethernet 10GBASE-R nativamente (ou seja, sem intermediação do GFP ou do SDH).

Com o propósito de viabilizar esse transporte, foi especificado o contêiner virtual ODU2e, de baixa ordem, inicialmente na recomendação ITU-T G.sup43 (*Transport of IEEE 10GBASE-R in Optical Transport Networks (OTN)*). Essa

especificação foi incorporada à versão atual da recomendação ITU-T G.709 (de 2012). A ODU2e suporta a velocidade de 10,3125 Gbit/s, superior àquela suportada pela ODU2.

Uma ODUflex é similar à VCAT, por permitir a concatenação virtual de múltiplos sinais, evitando, contudo, *delays* diferenciais por restringir a transmissão de todo o ODUflex a um mesmo ODUk de ordem superior e por prover uma única entidade de transporte gerenciável por serviço.

A ODUflex foi definida de modo a suportar qualquer velocidade, desde que menor que a capacidade total do contêiner virtual de alta ordem utilizado como servidor. Entretanto, para a máxima eficiência do transporte de *payloads*, a ODUflex deve concatenar um número inteiro do menor tamanho de *time slot* ao longo do caminho percorrido (que pode ser igual a 1,25 Gbit/s ou a 2,5 Gbit/s).

Como visto anteriormente, para a ODUflex não existe uma OTU servidora própria, podendo a ODUflex ser transportada em uma ODU2, em uma ODU3 ou em uma ODU4.

As diferentes ODUk são constituídas por TSs (*time slots*) de 1,25 Gbit/s ou de 2,5 Gbit/s, conforme a Figura 6.11.

ODUk	**1,25 Gbit/s**	**2,5 Gbit/s**
ODU1	2	1
ODU2	8	4
ODU3	32	16
ODU4	80	-

Figura 6-11 – Número de TSs para cada ODUk.

Como se observa, a ODU4 é constituída exclusivamente por 80 TSs de 1,25 Gbit/s.

6.4. ETHERNET SOBRE SDH E OTN

A despeito da crescente atenção dispensada ao uso de redes de transporte modo pacote, ou seja, de PTNs (*Packet Transport Networking*) tais como o MPLS-TP e o PBB-TE, a utilização de redes SDH como suporte a *Ethernet* mantém-se em um elevado patamar de importância.

A tecnologia referida como EoS (*Ethernet over SONET / SDH*) tem como base a intermediação do protocolo GFP (*Generic Framing Procedure*), definido na recomendação ITU-T G. 7041/Y. 1303 (*Generic Framing Procedure*).

O GFP, além de ser utilizado no SDH (vide recomendações ITU-T G.707 e G.783), é aplicado também em OTN (vide recomendações ITU-T G.709 e G.798) e em outras redes de transporte.

Como uma alternativa ao GFP, o ITU-T definiu a tecnologia *Ethernet* sobre LAPS (*Link Access Procedure – SDH*), na recomendação ITU-T X.86/Y.1323 (*Ethernet over LAPS*), para o transporte de *Ethernet* sobre SDH. O LAPS é um protocolo assemelhado ao HDLC (*High-Level Data Link Control*).

Ethernet sobre LAPS é uma extensão da tecnologia IP sobre LAPS definida na recomendação ITU-T X.85/Y.1321 (*IP over SDH using LAPS*).

Com a definição do GFP, o LAPS vem sendo preterido em alguns países. O GFP pode suportar outras tecnologias além do IP e *Ethernet* (tais como RPR, PPP e *Fiber Channel*).

6.4.1. **Mapeamento Nativo de Ethernet em SDH**

O anexo F da recomendação ITU-T G.707 especifica uma forma de transporte de 10 Gbit/s *Ethernet* em um VC-4-64c, alternativa ao uso do protocolo GFP.

O padrão IEEE 802.3ae define uma interface de 10 Gbit/s *Ethernet* WAN, que é basicamente um STM-64 com um VC-4-64c, que contempla o mapeamento do *Ethernet* MAC no VC-4-64c usando uma codificação 64B/66B.

A codificação 64/66B, definida no padrão IEEE 802.3ae para a subcamada de codificação física PCS (*Physical Coding Sublayer*) a 10 Gbit/s, foi criada para reduzir o *overhead* da estrutura da PCS 8B/10B original, que é de 25%.

A nova PCS opera com blocos de 8 bytes, precedidos de dois bits que permitem o alinhamento dos blocos e o reconhecimento de seu tipo. Os quadros de dados seguem o formato original dos quadros MAC *Ethernet*, iniciados com o preâmbulo de 7 bytes.

6.4.2. **Transporte de Ethernet sobre OTN**

Na Figura 6.12 encontra-se representada uma tabela contendo as diferentes formas de mapeamento de sinais Ethernet em OTN.

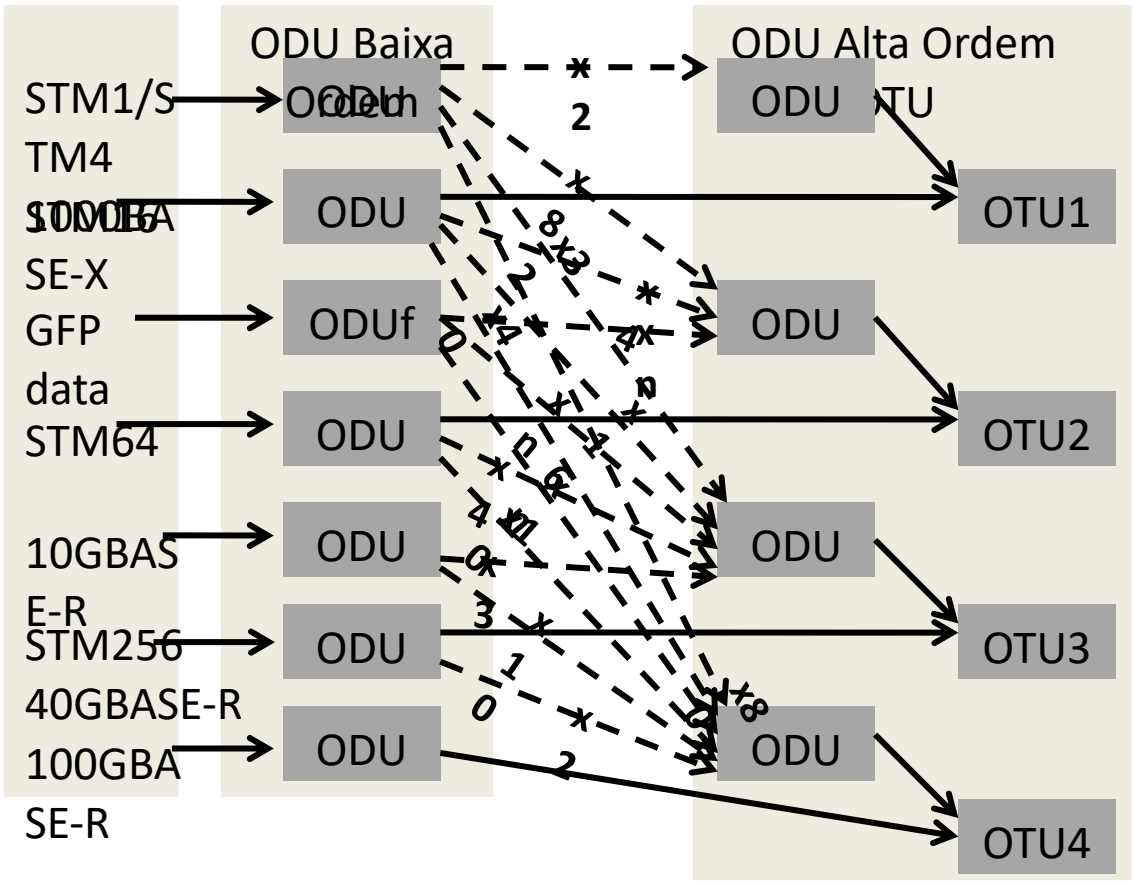


Figura 6-12 – Mapeamento de Ethernet em OTN.

Como se observa nessa figura, quadros Ethernet podem ser mapeados nativamente em uma rede OTN, ou, alternativamente, podem ser mapeados por via de quadros GFP ou através do SDH.

Conforme essa tabela, o mapeamento nativo de Ethernet em OTN ocorre das seguintes formas:

— O sistema de sinais 1000BASE-X é mapeado em uma ODU0;
— O sistema de sinais 10GBASE-R é mapeado em uma ODU2e;
— O sistema de sinais 40GBASE-R é mapeado em uma ODU3;
— O sistema de sinais 100GBASE-R é mapeado em uma ODU4.

Os sistemas de sinais Ethernet atendidos via quadros do GFP-F são mapeados em uma ODUflex, enquanto que aqueles atendidos via SDH são mapeados na ODU0 (STM-1), na ODU1 (STM-16), na ODU2 (STM-64) ou na ODU3 (STM-256).

6.4.3. **EoSDH pelo LAPS**

O protocolo LAPS (*Link Access Procedure-SDH*) foi inicialmente definido pela recomendação ITU-T X.85/Y.1321, com o propósito de permitir a transmissão do protocolo IP sobre SDH.

Foi posteriormente emitida a recomendação ITU-T X-86/Y.1323, com o propósito de estender a aplicação do LAPS para suporte a *Ethernet*, o que se denomina *Ethernet* sobre LAPS.

O LAPS é um protocolo assemelhado ao HDLC, possuindo as mesmas características, tais como simplicidade, baixa latência e controle remoto de erros e de fluxo. Esse protocolo admite a armazenagem (*buffering*) de um ou mais quadros MAC antes de seu envio.

O protocolo LAPS provê uma técnica simples para a constituição de *switched Ethernet* LANs e para a extensão dessas redes sobre MANs e WANs.

6.4.3.1. Arquitetura e Formatação do LAPS

A Figura 6.13 representa a arquitetura típica de redes utilizando *Ethernet* sobre LAPS.

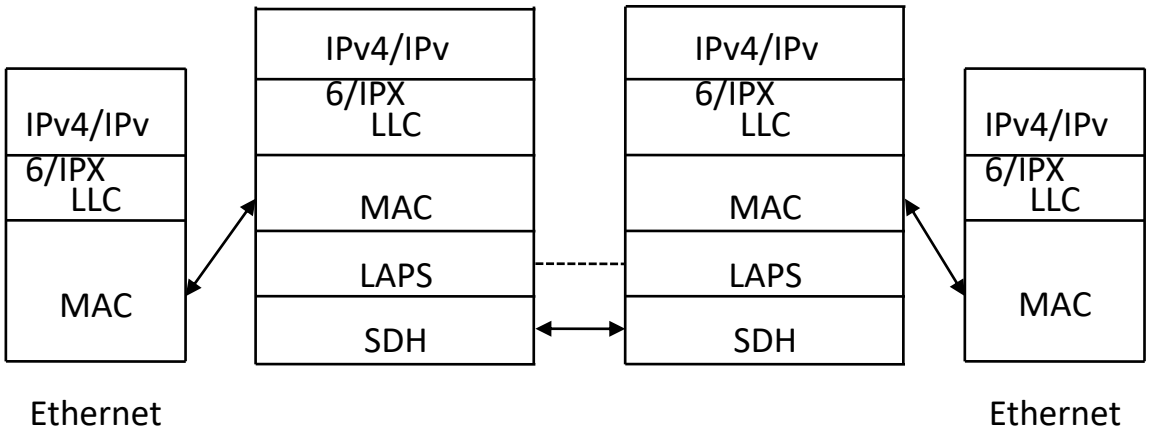


Figura 6-13 – Arquitetura de Ethernet sobre LAPS.

Como se observa nessa figura, um quadro MAC contendo um datagrama IP ou IPX é envelopado em um quadro LAPS, que por sua vez é mapeado em um VC do SDH.

O LAPS possui um cabeçalho LAPS, e, ao final do quadro, um *trailler* constituído pelo FCS do LAPS e por um *flag* (0x7E).

O cabeçalho LAPS, iniciado também por um *flag* (0x7E), possui um campo de endereço (igual a 0x04), um campo *control* (igual a 0x03) que representa o comando *Unumbered Information*, com o bit *Final*/*Poll* (P/F) setado para zero.

O SAPI (*Service Access Point Indentifier*), constituído por dois octetos, representa o ponto em que o LAPS, como camada provedora, presta serviço à camada cliente (camada MAC, no caso da recomendação X.86/Y.1323).

O valor do SAPI, que é igual a 0x0021 para o IPv4 e 0x0057 para o IPv6 conforme a recomendação X.85/Y.1321, passa a ser igual a 0xFE01 para o Ethernet de acordo com a recomendação X.86/Y.1323.

6.4.4. **EoSDH e EoOTN pelo GFP**

O GFP (*Generic Framing Procedure*) é um novo protocolo, definido na recomendação ITU-T G.7041/ Y.1303 (cuja última versão foi publicada em 2011) e pelo padrão ANSI T1.105.2.

O propósito do GFP é suportar o transporte de pacotes de tamanho fixo ou variável sobre um canal de alta velocidade, com alinhamento de bits ou de bytes.

São exemplos de tecnologias de rede transportadas pelo GFP o *Ethernet*, o PPP, o RPR e as tecnologias utilizadas em *Storage Area Networking* (SAN). Dentre as SANs podem ser citadas o *Fibre Channel*, a FICON (*Fibre Connection*), ESCON (*Enterprise Systems Connection*) e a SBCON (*Single-Byte Command Code Sets Connection*).

Por outro lado, o GFP pode ser aplicado sobre SDH, OTN, WDM ou diretamente sobre fibras ópticas (*dark fibers*).

A figura 6.14 apresenta a pilha de protocolos correspondente à utilização do GFP.

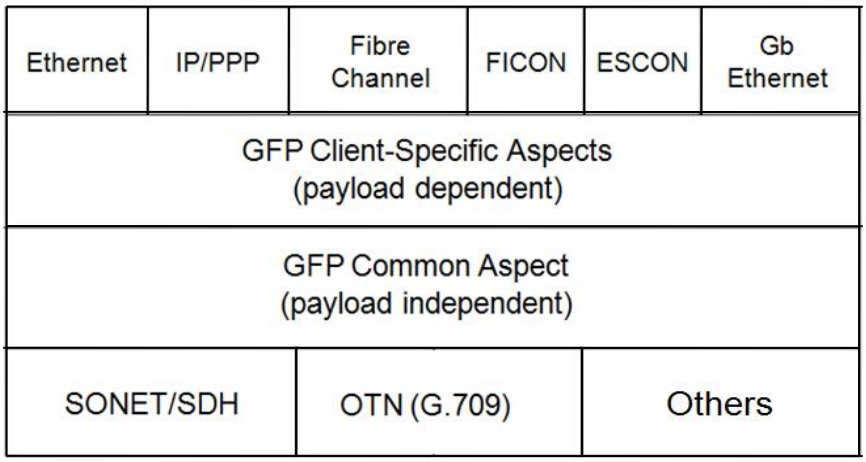


Figura 6-14 –Pilha de protocolos para o GFP.

Como se observa nessa figura, a especificação do GFP consiste em aspectos comuns e aspectos específicos para as tecnologias cliente. Os aspectos comuns aplicam-se a todos os tráfegos adaptados ao GFP.

Para os aspectos específicos do GFP foram definidos dois modos de adaptação de sinais de tecnologias cliente:

— Um modo de adaptação orientado a PDU, referido como *Frame-Mapped GFP* (GFP-F);
— Um modo de adaptação orientado a código de bloco (*block-code oriented*), referido como *Transparent GFP* (GFP-T).

O presente subitem enfatiza o uso do GFP como suporte para transmissão de *Ethernet* sobre SDH. Serão abordados, para isso, alguns aspectos básicos do GFP.

6.4.4.1. Quadros do GFP

A recomendação ITU-T G.7041/Y.1303 define dois tipos de quadro do GFP:

— Quadros de Cliente do GFP;
— Quadros de Controle do GFP.

6.4.4.2. Quadros de Cliente do GFP

Os Quadros de Cliente do GFP podem ser Quadros de Dados de Cliente (ou Quadros de Dados de Usuário), Quadros de Gerência de Cliente ou Quadros de Comunicações de Gerência.

O formato geral de Quadros de Cliente do GFP encontra-se na Figura 6.15.

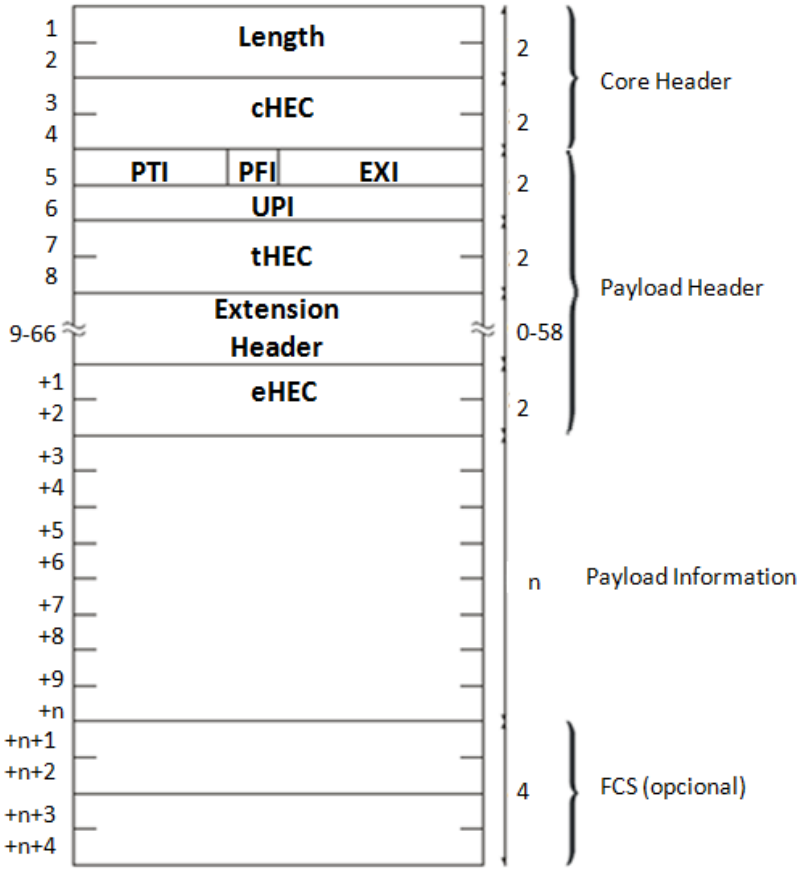


Figura 6-15 – Formato geral de Quadros de Cliente do GFP.

Como se observa nessa figura, os Quadros de Cliente do GFP consistem em um *GFP Core Header* (com 4 octetos) e em uma Área de *Payload* do GFP (*GFP Payload* Area).

- **GFP Core Header**

O *GFP Core Header* de Quadros de Cliente do GFP tem como propósito indicar o número de octetos contidos na Área de *Payload* do GFP. Essa informação é contida no campo PLI (*PDU Lenght Indicator*), com dois octetos.

Os dois últimos octetos do *GFP Core Header* representam o campo *cHEC* (*Core HEC*), que projege a integridade do próprio *Core Header*.

- **Área de Payload do GFP**

A Área de *Payload* do GFP, de tamanho variável, é utilizada para transportar informações específicas de protocolos de camada superior. Ela diz respeito aos aspectos específicos das tecnologias transportadas pelo GFP.

Essa área é composta pelo campo *Payload Header*, pelas informações nela contidas (campo *Payload Information*) e pelo campo *Payload FCS* do GFP, como mostra a figura 6.16.

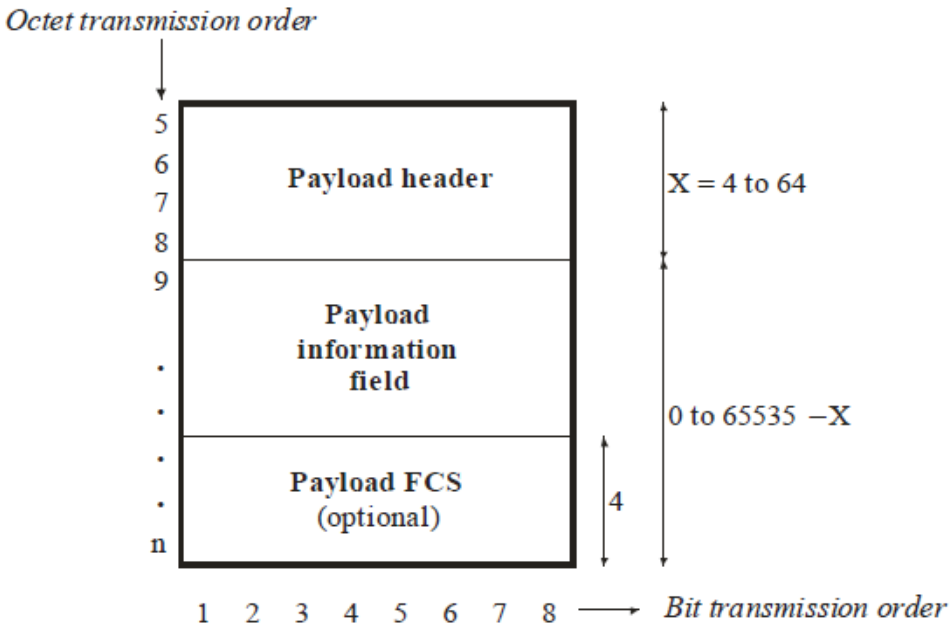


Figura 6-16 – Formato da Área de Payload do GFP.

O *Payload Header* é composto pelo campo *GFP Type* com dois octetos, pelo campo *Type HEC* (*tHEC*) com dois octetos, pelo campo *Extension Header* com zero a 60 octetos e pelo campo *Extension HEC* (*eHEC*) com dois octetos.

O campo *GFP Type* indica o conteúdo e o formato do campo de informações da Área de *Payload* do GFP (ou seja, do campo *Payload Information*). O campo GFP *Type* é composto pelos seguintes sub-campos:

— PTI (*Payload Type Indentifier*);
— PFI (*Payload FCS Indicator*);
— EXI (*Extension Header Identifier*);
— UPI (*User Payload Indentifier*).

O sub-campo PTI (*Payload Type Identifier*) identifica o tipo de Quadro de Cliente do GFP, conforme a Figura 6.17.

Type bits <15:13>	Usage
000	User data
100	Client management
101	Management communications
Others	Reserved

Figura 6-17 –Indentificadores de Tipo de Payload do GFP.

O sub-campo PFI (*Payload FCS Indicator*), com um bit, indica a presença (PFI=1) ou a ausência (PFI=0) do campo FCS na Área de *Payload* do GFP.

O sub-campo EXI (*Extension Header Identifier*) indentifica o tipo de *Extension Header* do GFP eventualmente existente no quadro do GFP. Os tipos definidos encontram-se na Figura 6.18.

Type bits <11:8>	Usage
0000	Null extension header
0001	Linear frame
0010	Ring frame
Others	Reserved

Figura 6-18 – Identificadores de tipo de Extension Header.

O sub-campo UPI (*User Payload Identifier*) identifica o tipo de informação de *payload* contido no campo *Payload Information* do quadro de cliente do GFP. Os valores de UPI são específicos para cada tipo de quadro de cliente (indicado pelo sub-campo PTI), e serão apresentados adiante neste item.

O campo *tHEC* (*Type HEC*) contém um código de controle de erros CRC-16, e protege a integridade do campo GFP *type*.

O campo *GFP Extension Headers* é composto por um campo que representa um ou mais cabeçalhos de extensão do GFP (de zero a 58 octetos) e por um campo *eHEC* (*Extension HEC*). O tipo de cabeçalho de extensão é indicado pelo sub-campo EXI como visto anteriormente.

O Cabeçalho de Extensão referente a um quadro linear (ponto a ponto), destina-se a cenários onde existem múltiplos *links* independentes que requerem agregação em um único caminho de transporte.

A recomendação ITU-T G.7041/Y.1303 postergou a definição do cabeçalho em anel (*ring frame*) para futuras emissões.

O campo *eHEC* (*Extension HEC*) contém um código de controle de erros CRC-16, e protege a integridade do conteúdo dos cabeçalhos de extensão (*Extension Headers*).

Por fim, o campo *pFCS* (*Payload Frame Check Sequence*), com quatro octetos, contém um código de controle de erros CRC-32, e protege o conteúdo do campo *GFP Payload Information*.

Como o próprio conteúdo do campo *GFP Payload Information* pode conter a sua própria proteção (FCS de um quadro Ethernet, por exemplo), o pFCS é de uso opcional.

6.4.4.3. Quadros de Dados de Cliente

Quadros de dados nativos enviados por clientes são envelopados para transporte em Quadros de Dados de Cliente do GFP.

O subcampo PTI do campo *GFP Type* para Quadros de Dados de Cliente possui o valor 000 (PTI = 000). Os demais subcampos desse campo adotam valores específicos para cada *payload* envelopado.

O subcampo PFI será setado de acordo com a habilitação do FCS. O subcampo EXI será setado de acordo com os requisitos da conexão GFP em termos de multiplexação de quadros e de topologia.

Os valores do subcampo UPI em Quadros de Dados de Cliente (PTI = 000), assim como em Quadros de Comunicações de Gerência (PTI = 101), encontram-se na Figura 6.19.

Observa-se que no caso específico de Quadros de Comunicações de Gerência (PTI = 101), nem todos os valores de UPI dessa figura são aplicáveis.

Como se observa, existem nessa figura valores de UPI relativos ao GFP-F e valores relativos ao GFP-T.

PTI = 000 and 101	
Type bits<7:0>	**GFP frame payload area**
0000 0000 1111 1111	Reserved and not available
0000 0001	Frame-mapped Ethernet
0000 0010	Frame-mapped PPP
0000 0011	Transparent fibre channel
0000 0100	Transparent FICON
0000 0101	Transparent ESCON
0000 0110	Transparent Gbit Ethernet
0000 0111	Reserved for future use
0000 1000	Frame-mapped multiple access protocol over SDH (MAPOS)
0000 1001	Transparent DVB ASI
0000 1010	Framed-mapped IEEE 802.17 resilient packet ring
0000 1011	Frame-mapped fibre Channel FC-BBW
0000 1100	Asynchronous transparent fibre channel
0000 1101	Frame-mapped MPLS (unicast) and frame-mapped T-MPLS
0000 1110	Frame-mapped MPLS (multicast)
0000 1111	Frame-mapped OSI network layer protocols (IS-IS, ES-IS, CLNP)
0001 0000	Frame-mapped IPv4
0001 0001	Frame-mapped IPv6
0001 0010	Frame-mapped DVB-ASI
0001 0011	Frame-mapped 64B/66B encoded Ethernet, including the Ethernet frame preamble
0001 0100	Frame-mapped 64B/66B encoded Ethernet ordered set information
0001 0101	Transparent transcoded FC-1200
0001 0110 to 1110 1111	Reserved for future standardization
1111 0000 to 1111 1110	Reserved for proprietary use

Figura 6-19 –Valores de UPI para Quadros de Cliente com PTI = 000 e PTI = 101.

6.4.4.4. Quadros de Gerência de Cliente

O GFP provê a possibilidade de definição de Quadros de Gerência de Cliente, que operam fim a fim, com o propósito de permitir funções tais como monitoração de desempenho e OAM.

O subcampo PTI do campo *GFP Type* apresenta o valor 100 (PTI = 100) para esse tipo de quadro, enquanto os demais subcampos desse campo possuem valores específicos para cada tipo de *payload* envelopado.

Os valores do sub-campo UPI em Quadros de Gerência de Cliente do GFP encontra-se na Figura 6.20.

PTI = 100	
UPI value	**Usage**
0000 0000 1111 1111	Reserved
0000 0001	Client signal fail (loss of client signal)
0000 0010	Client signal fail (loss of character synchronization)
0000 0011	Client defect clear indication (DCI)
0000 0100	Client forward defect indication (FDI)
0000 0101	Client reverse defect indication (RDI)
0000 0110 through 1101 1111	Reserved for future use
1110 0000 through 1111 1110	Reserved for proprietary use (Note)
NOTE – The use of proprietary code values is described in Annex A of [ITU-T G.806].	

Figura 6-20 –Valores de UPI para Quadros de Gerência de Cliente (PTI = 100).

O formato dos Quadros de Gerência de Cliente apresenta algumas particularidades. O campo *Payload Header,* com quatro octetos, se resume aos dois octetos do campo *GFP Type* e aos dois octetos do *tHEC*, separadamente dos campos *Extension Header* (opcional) e do campo que contém o *payload*.

Como o *payload* de quadros de gerência de clientes não possui dados nativos de cliente, o campo que o contém passa a se denominar *Payload Area*, e não *Payload Information* como no formato geral dos Quadros de Cliente.

6.4.4.5. Quadros de Comunicações de Gerência (MCFs)

Os Quadros de Comunicações de Gerência (MCFs), com PTI igual a 101, são utilizados para o provimento da possibilidade de existência de canais *in-band* para comunicações de gerência entre o início (fonte) do GFP e o término (*sink)* do GFP.

A especificação desse tipo de quadro não consta da atual versão da recomendação ITU-T G.7041/Y.1303.

6.4.4.6. Quadros de Controle do GFP

Os Quadros de Controle do GFP são utilizados para o gerenciamento de conexões do GFP.

Encontram-se em definição pelo ITU-T diferentes tipos de Quadros de Controle do GFP. Tais quadros deverão ser formatados diferentemente dos Quadros de Dados de Cliente, mas os respectivos *payloads* deverão ser protegidos por um HEC similar ao *cHEC*.

Foram especificados apenas os quadros de controle do GFP denominados quadros *GFP Idle*, contendo quatro octetos. Tais quadros consistem em um *Core Header*, com os campos PLI e *cHEC* iguais a zero. O objetivo desses quadros é operar como um mecanismo de adaptação dos quadros GFP para os diferentes meios de transporte, pela função *stuffing*, quando o meio de transporte tem capacidade superior à requerida pelo sinal do cliente.

Outros Quadros de Controle do GFP encontram-se em fase de estudo no ITU-T.

6.4.4.7. Modos de Adaptação de Sinais de Clientes

Conforme menção anterior neste item, foram especificados pelo ITU-T dois modos de adaptação de sinais de cliente em quadros do GFP:

— *Frame-Mapped Mode* (GFP-F);
— *Transparent-Mapped Mode* (GFP-T).

- **Modo Frame-Mapped (GFP-F)**

O GFP-F é um modo mais flexível de adaptação de sinais de cliente, apropriado para suportar aplicações modo pacote ponto a ponto ou ponto a multiponto. São exemplos de PDUs apropriadas para suporte pelo GFP-F os quadros Ethernet IEEE 802.3, RPR, MPLS e PPP.

Ocorre no GFP-F o envelopamento de uma PDU de camada superior (Ethernet, por exemplo) em uma GFP PDU, sendo alinhadas as fronteiras entre essas PDUs.

A Figura 6.21 ilustra o envelopamento de um quadro Ethernet em uma GFP-F PDU.

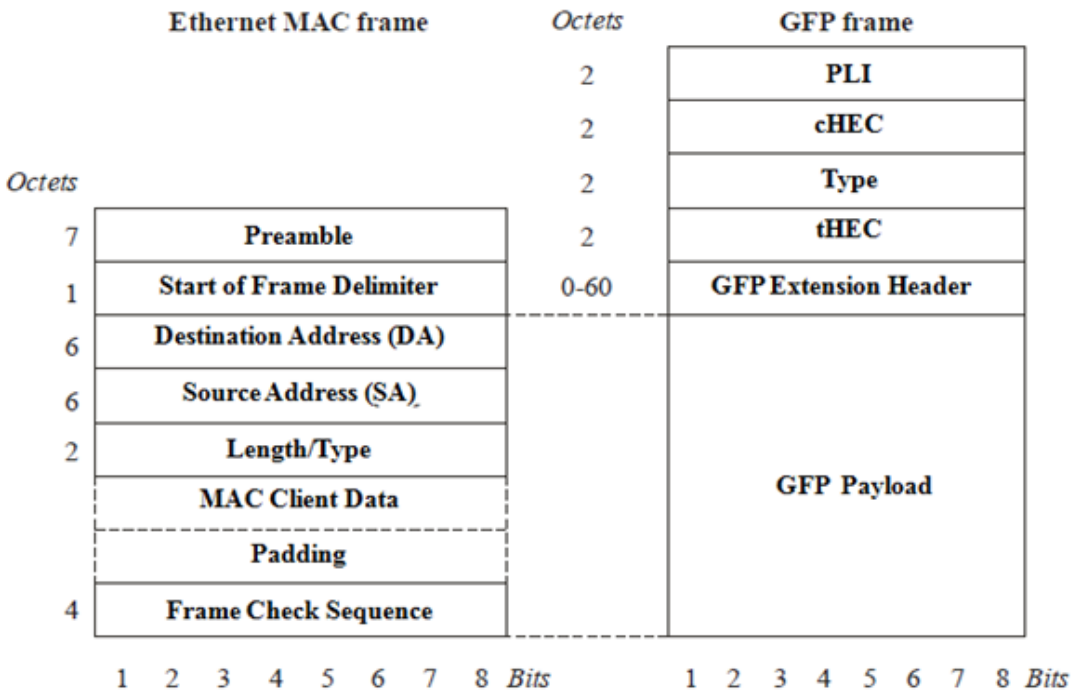


Figura 6-21 – Envelopamento de Ethernet em GFP-F.

A recomendação ITU-T G.7041/Y.1303 especifica critérios especiais para o transporte de quadros Ethernet 10GBASE-R sobre o GFP-F. Tais critérios especiais dizem respeito a *payloads* com transparência de preâmbulo e de SFD (ou seja, o preâmbulo e o SFD são também envelopados) e a *payloads* com informação de conjunto ordenado.

- **Modo Transparent-Mapped (GFP-T)**

O GFP-T é apropriado para aplicações ponto a ponto que requerem valores muito baixos de *delay* e de *jitter*. A versão atual do GFP-T foi especificada para operar com sinais utilizando a codificação 8B/10B transmitidos transparentemente. São exemplos de tecnologias suportadas pelo GFP-T o *Fibre Channel*, a FICON (*Fibre Connection*), a ESCON (*Enterprise Systems Connection*) e o Gigabit Ethernet.

Os quadros do GFP-T utilizam a mesma estrutura que os quadros do GFP-F, inclusive o *Payload Header*. Contudo, diferentemente do GFP-F que mapeia um quadro de dados nativo de cliente em seu formato original, no GFP-T caracteres individuais desse quadro são separados e mapeados em quadros de dados do GFP-T periódicos e de tamanho fixo.

CAPÍTULO 7

ETHERNET SOBRE MPLS

7.1. PREÂMBULO

O presente capítulo tem como objetivo definir a utilização de redes MPLS para transporte de Ethernet. Foram especificadas, pelo IETF, duas aplicações do MPLS com esse propósito.

Uma dessas aplicações, o VPWS (*Virtual Private Wire Service*), especificado na RFC 4447 (*Pseudowire Setup and Maintenance Using the Label Distribution Protocol-LDP*), possibilita o atendimento de diferentes tecnologias de rede em uma configuração ponto a ponto.

A RFC 4448 (*Encapsulation Methods for Transport of Ethernet over MPLS Networks*) define a utilização do VPWS para o transporte de Ethernet. Essa utilização do VPWS é referida como EoMPLS (*Ethernet over MPLS*) ou ETHoMPLS.

A segunda aplicação do MPLS para o transporte de Ethernet é o VPLS (*Virtual Private LAN Service*), cuja fundamentação encontra-se na RFC 4664 (*Framework for L2VPNs*).

Ao contrário do VPWS, o VPLS destina-se exclusivamente ao atendimento de redes Ethernet, operando nos moldes de uma rede *bridged* Ethernet, admitindo, portanto, tráfego multiponto a multiponto.

O VPWS como suporte a tráfego Ethernet e o VPLS serão abordados neste capítulo.

7.2. MPLS (MULTIPROTOCOL LABEL SWITCHING)

O MPLS é uma tecnologia de camada 2 orientada a conexão que utiliza diferentes formas de roteamento IP. Utiliza também diferentes protocolos de sinalização para fundamentar a constituição de redes cujo plano de dados opera por comutação baseada em *labels*. As centrais MPLS são referidas como LSRs (*Label Switching Routers*), enquanto os circuitos virtuais constituídos são denominados LSPs (*Label Switching Paths*).

A arquitetura do plano de dados do MPLS encontra-se representada na Figura 7.1.

Figura 7-1 – Arquitetura do plano de dados do MPLS.

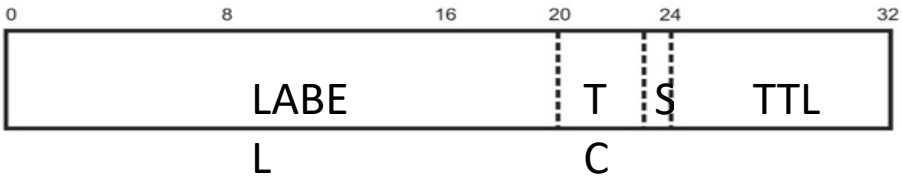
O MPLS é uma tecnologia de camada 2 que possui a singularidade de não possuir interface física. Assim, para ser utilizado sobre a camada física, o MPLS requer o uso da interface física de uma outra tecnologia de camada 2 (PPP, Ethernet, ATM e *Frame Relay*, por exemplo), como se verifica na figura.

Acima da interface física encontra-se o *label* de LSP, responsável pela definição do caminho virtual a ser seguido pelos quadros de dados que trafegam por um LSP.

Observa-se que quando da utilização do MPLS com suporte sobre ATM e *Frame Relay*, os *labels* de LSP são os próprios indentificadores de canal lógico dessas tecnologias. Para o ATM são utilizados os valores dos campos VPI/VCI (*Virtual Path Identifier/Virtual Channel Identifier*), enquanto que para o *Frame Relay* utiliza-se o campo DLCI (*Data Link Connection Identifier*).

Quando da utilização do PPP ou do Ethernet como suporte ao MPLS, torna-se necessária a introdução do um *label* de LSP específico, referido com *label* genérico ou *label* no formato genérico. A Figura 7.2 apresenta o formato de um *label* genérico.

Figura 7-2 – Formato do label genérico para o MPLS.



Os campos dessa figura têm os seguintes significados:

— *Label* no formato genérico, com 20 bits;
— Campo TC (*Traffic Class*), com 3 bits, onde é codificado o CoS ID e a precedência para descarte de quadros;
— *S bit*, que indica a posição do *label* na respectiva pilha de *labels* (*label stack*);

— Campo TTL (*Time to Live*), com 8 bits, que é utilizado para mitigação de *loops*.

O campo TC era anteriomente referido como EXP (*Experimental*).

Superiormente ao *label* de LSP encontra-se o *label* de aplicação. Esse *label*, ao contrário do *label* de LSP, é transparente para os LSRs intermediários, sendo processados apenas pelos LSRs de borda. Os LSRs de borda são também referidos como PEs (*Providers Edge Equipments*) ou como LERs (*Label Edge Routers*). Os LSRs intermediários, por sua vez, são referidos como Ps (*Provider Equipment*).

Observa-se que os *labels* de aplicação encontram-se exclusivamente no formato genérico.

7.2.1. **Tipos de MPLS**

O MPLS apresenta-se em um dos seguintes tipos:

— MPLS (ou MPLS básico);
— MPLS-TE (*MPLS Traffic Enginnering*);
— MPLS-TP (*MPLS Transport Profile*).

O MPLS (básico) utiliza uma lógica simples, em que não é possível a utilização de Engenharia de Tráfego (TE), ou seja, em que não é possível funções tais como reserva de banda e consideração de restrições (*constraints*) relativas a tráfego na constituição de LSPs.

O roteamento utiliza o protocolo OSPF ou o protocolo IS-IS. A sinalização tem como base o protocolo LDP (*Label Distribution Protocol*).

A partir do MPLS-TE foi introduzida a Engenharia de Tráfego, possibilitando um considerável *upgrade* funcional no MPLS.

O MPLS-TE utiliza o protocolo de roteamento OSPF-TE (*OSPF Traffic Engineering*) ou o protocolo de roteamento ISIS-TE (*IS-IS Traffic Engineering*). Para a sinalização, o protocolo utilizado é o RSVP-TE (*Resource Reservation Protocol Traffic Engineering*).

O MPLS-TP representa uma envolução do MPLS-TE pelas seguintes razões:

— Fortalecimento da operação orientada a conexão, pela eliminação das funções *label merging*, *ECMP* (*Equal Cost Multi-Path*) e PHP (*Penultimate Hop Poping*);
— Adoção do GMPLS (*Generalized MPLS*) como plano de controle;
— Possibilidade de uso de rede de controle independente;
— Possibilidade de constituição de LSP bidirecionais corroteados em um único processo de sinalização.
— Utilização de mecanismos de OAM e de proteção robustos.

O MPLS-TP utiliza o protocolo GMPLS OSPF-TE ou o protocolo GMPLS ISIS-TE para roteamento . Na sinalização, o protocolo utilizado é o GMPLS RSVP-TE.

7.2.2. **Aplicações do MPLS**

O MPLS, para qualquer um de seus tipos, opera nas seguintes aplicações básicas:

— Serviço MPLS publico;
— Serviço BGP/MPLS IP VPN, também referido como MPLS L3VPN;
— VPLS;
— VPWS.

O serviço MPLS público destina-se, basicamente, à utilização do MPLS como uma simples extensão das redes IP. Essa aplicação não possui *labels* próprios, consistindo apenas nos LSPs constituídos.

O serviço BGP/MPLS IP VPN consiste na emulação, pela rede MPLS, de um conjunto de L3VPNs baseadas em endereços IP da rede externa (a rede externa é necessariamente uma rede IP). Cada L3VPN constituída consiste em um conjunto de VRFs (*Virtual Routing Functions*) situadas nos respectivos LSRs de borda (PEs).

Como base para roteamento e sinalização entre PEs para a constituição de IPv4 L3VPNs sobre MPLS, utiliza-se o protocolo BGP-4 (definido na RFC 4271) estendido pela RFC 4760 (*Multiprotocol Extensions for BGP-4*), pela RFC 4364 (*BGP/MPLS IP Virtual Private Network*s) e pela RFC 3107 (*Carrying Label Information in BGP-4*). O protocolo BGP-4 estendido pela RFC 4760 é referido como protocolo MP-BGP (*Multiprotocol BGP-4*). A RFC 4364 introduziu a família de endereços VPN-IPv4.

Para a constituição de IPv6 VPNs sobre MPLS foi emitida a RFC 4659 (*BGP/MPLS IP Virtual Private Network Extension for IPv6 VPN*).

7.3. **ARQUITETURA PWE3**

A constituição de PWs sobre PSNs tem como base a arquitetura PWE3, definida na RFC 3985 (*Pseudo Wire Emulation Edge-to-Edge* (*PWE3*) *Architecture*). Essa RFC descreve, com base no conceito de PW, a emulação de serviços tais como ATM, *Frame Relay*, Ethernet e PPP sobre PSNs (*Packet Switching Networks*), utilizando como PSN redes IP ou redes MPLS.

7.3.1. **Aspectos Iniciais**

A arquitetura PWE3, que fundamenta a constituição de PWs emulando diferentes tecnologias sobre diferentes tipos de PSN, foi especificada na RFC 3985, como já mencionado.

A arquitetura PWE3 é um mecanismo que emula os atributos essenciais de um serviço de telecomunicações ponto a ponto sobre uma PSN, de modo a prover a funcionalidade mínima desse serviço, e de forma tal que o serviço emulado corresponda a um fio (*wire*) interligando dois pontos extremos da PSN, com o grau de serviço desejado. A conexão virtual assim obtida é denominada *pseudowire*.

Esses pontos extremos encontram-se nos equipamentos PE da PSN, que se conectam aos equipamentos CE externos por meio de configurações de rede de acesso, que de alguma forma transportam o protocolo sendo emulado.

Os PWs não são visíveis para os equipamentos Ps (interior da PSN), sendo processados apenas pelos PEs da PSN. Os PWs podem ser estabelecidos manualmente ou através de uso de protocolos de sinalização.

Pode-se afirmar, em suma, que a arquitetura PWE3 é uma concepção que objetiva a multiplexação estatística de um túnel em uma PSN de suporte, de modo a transportar, nesse túnel, diferentes instâncias de uma tecnologia de rede ou de diferentes tecnologias de rede. A cada uma dessas instâncias corresponde um PW.

Qualquer translação ou operação envolvendo a semântica do payload transportado é realizada por uma função incorporada aos PEs, denominada Native Service Processing (NSP). A NSP se situa externamente às terminações dos PWs, funcionando possivelmente no ingresso e no egresso de quadros, células ou bit streams (estruturados ou não estruturados). A RFC 3985 não define a funcionalidade de NSPs.

7.3.2. **Modelo de Referência para a Arquitetura PWE3**

A Figura 7.3 ilustra o modelo de referência para a implementação da arquitetura PWE3.

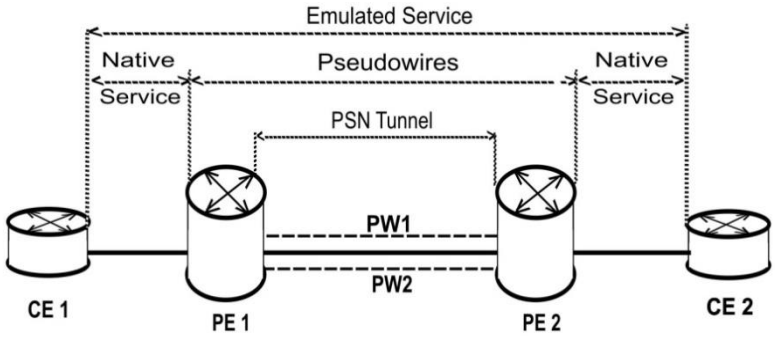


Figura 7-3 – Modelo de referência da arquitetura PWE3.

Os dois PEs (PE 1 e PE 2) provêm um ou mais PWs para possibilitar a comunicação entre os dois CEs (CE 1 e CE 2) através da PSN, de forma tal que o tráfego nos PWs, que emulam um tipo de serviço de telecomunicações, seja invisível para o interior da PSN. Unidades de dados nativas que atingem um PW são encapsuladas em uma PW-PDU, sendo essa PW-PDU transportada por via de um túnel na PSN.

Os PEs realizam o necessário encapsulamento e desencapsulamento de PW-PDUs, além de desempenhar outras funções requeridas pelo serviço PW, como sequenciamento e *timing*. Os PEs são roteadores de borda (*edge routers*) da PSN de suporte (IP ou MPLS) aos PWs.

Algumas aplicações requerem a efetivação de determinadas operações prévias nas unidades de dados nativas entrantes ou saintes de PEs. São exemplos dessas operações a translação de endereços ou de identificadores, em *Ethernet Bridging* e em SONET/SDH *Cross-Connect.* Essas operações podem ser realizadas no interior dos PEs ou entre as terminações dos PWs e os respectivos CEs. Algumas dessas operações constituem o denominado *Native Service Processing* (NSP).

7.3.3. **Camadas Constituintes da Arquitetura PWE3**

A figura 7.4 apresenta as camadas constituintes da arquitetura PWE3.

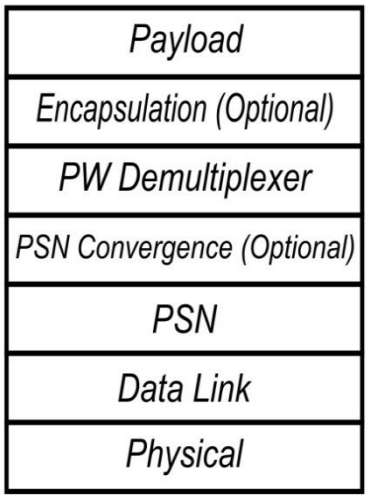


Figura 7-4 – Camadas da arquitetura PWE3.

A camada *Payload* contém quadros (PDUs) de camada 2, células ou bit *streams.*

A camada *Encapsulation* contém informações adicionais ao *payload*, se necessárias ao seu encapsulamento para fins de atendimento à funcionalidade do PW. Essa camada pode prover também suporte para processamento em tempo real e para sequenciamento de informações.

Os PWs são contidos em LSPs ou em túneis IP, ou em outros tipos de túnel, sendo que diferentes PWs podem ser multiplexados em um mesmo túnel. A camada *PW Demultiplexer* tem como função identificar os diferentes PWs dentro de um túnel que os contém.

A camada *PSN Convergence* tem como função acrescentar valores à PSN se necessários para o devido ajustamento dessa rede ao PW, tornando-os entidades independentes entre si. Se a PSN satisfaz naturalmente essa necessidade, a camada *PSN Convergence* é desnecessária e se torna vazia.

A camada PSN, como dito anteriormente, pode consistir em redes totalmente IP (IPv4 ou IPv6) ou em redes MPLS.

No caso de redes IP, o mecanismo normalmente utilizado para a constituição de PWs é o L2TPv3. No caso do MPLS, o serviço a isso destinado é o VPWS (serviço ponto a ponto) ou o VPLS (serviço multiponto a multiponto).

7.4. VIRTUAL PRIVATE WIRE SERVICE (VPWS)

O VPWS é um serviço que utiliza PWs constituídos em um túnel de uma PSN, cujas terminações (*PW endpoints*) são conectadas diretamente a redes externas por *Attachment Circuits* (ACs).

Para o caso particular em que a PSN utilizada represente uma rede MPLS, foram emitidos inicialmente pelo IETF dois *drafts*, denominados *drafts Martini* em decorrência do seu autor, L. Martini. Esses *drafts* dizem respeito ao transporte de PDUs de camada 2 sobre MPLS e aos métodos de encapsulamento para esse transporte. Posteriormente tais *drafts* foram oficializados pelo IETF mediante a emissão, respectivamente, das RFC 4906 e RFC 4905.

Nesse ínterim, o IETF emitiu também a RFC 4477, aplicando a arquitetura PWE3 ao MLPS, que se tornou o padrão de referência para o VPWS. Emitiu a seguir uma série de outras RFCs (RFC 4448, RFC 4618, RFC 4619 e RFC 4717), sendo cada uma delas específica para cada tecnologia de rede de camada 2 suportada pelo VPWS.

No VPWS, o PW *Demultiplexer* da Figura 7.4 é um *label* (PW *label*), atribuído e distribuído no modo UD (*Unsolicited Downstream*) pelo PW *endpoint* localizado no PE que se encontra a juzante no tráfego.

Registra-se que um PW *label* só tem significado em um sentido de transmissão. Para a comunicação P2P bi-direcional entre dois PW *endpoints* são necessários dois *PW labels*, cada um para um sentido de transmissão.

O túnel no interior da rede *backbone* que suporta o VPWS pode ser um LSP (MPLS básico, MPLS-TE ou MPLS-TP), ou pode basear-se em outras tecnologias. Um exemplo são os túneis IP nativo, L2TPv3 (*Level 2 Tunneling Protocol version 3*) ou GRE (*Generic Routing Encapsulation*), quando os equipamentos P (Ps) são roteadores IP.

Qualquer que seja o tipo de túnel de suporte, a funcionalidade VPWS restringe-se aos PEs que suportam esse serviço.

A RFC 4447 considera apenas a utilização de túneis MPLS para a condução e a agregação de PWs constituídos nos moldes da facilidade *hierarchy of routing knowledge,* de forma a isolar logicamente a rede MPLS *backbone* do processo de roteamento que ocorre no exterior dessa rede.

Os LSPs podem ser constituídos pelos processos do MPLS básico, utilizando protocolos IGP convencionais e o LDP. Os LSPs podem também consistir em LSPs MPLS-TE ou em LSPs MPLS-TP, incorporando assim as facilidades atinentes a *Traffic Engineering.*

7.4.1. **VPWS PWs**

A RFC 4447, em seu próprio título (*Pseudowire Setup and Maintenance Using* LDP) especifica o uso do protocolo LDP na sinalização para a constituição de VPWS PWs sobre MPLS LSPs, sendo definida uma extensão do LDP com esse propósito.

As RFCs complementares, citadas anteriormente, definem as formas de encapsulamento de PDUs de diferentes protocolos de camada 2 para a transmissão ponto a ponto por PWs constituídos em conformidade com a RFC 4447. Essas RFCs têm os seguintes títulos:

- RFC 4448: *Encapsulation Methods for Transport of Ethernet over MPLS Networks*;
- RFC 4618: *Encapsulation Methods for Transport of PPP/HDLC over MPLS Networks;*
- RFC 4619: *Encapsulation Methods for Transport of Frame Relay over MPLS Networks;*
- RFC 4717*: Encapsulation Methods for Transport of ATM over MPLS Networks.*

Tais RFCs estipulam que um *PW header* contendo um campo PW *demultiplexer*, ou seja, um *PW label*, seja acrescentado na origem e retirado no destino de uma PDU de camada 2 para cada uma das tecnologias atendidas, após cumprir o seu papel na identificação do PW no PE de egresso.

7.4.1.1. Aspectos Abordados

Serão abordados, neste item, os aspectos concernentes ao transporte ponto a ponto de serviços de rede de camada 2 sobre MPLS com base na arquitetura PWE3. Essa forma de transporte de serviços, formatados em quadros, células ou *bit streams* estruturados ou não estruturados, é definida na RFC 4447, onde serão fundamentados os termos deste item.

Será utilizada neste livro a denominação MPLS VPWS ou MPLS PWE3 com esse propósito. Fica convencionado, porém, que a simples denominação VPWS será suficiente para designar o MPLS VPWS.

7.4.1.2. Constituição de VPWS PWs

Para a constituição de PWs no VPWS seá admitida, neste item, a existência de um LSP entre o par de PEs considerados, sendo os PE 1 e PE 2 da Figura 7.3, por hipótese, os PEs de ingresso e de egresso desse LSP, respectivamente. Para isso foram distribuídos previamente valores de *Tunnel Labels* (*labels* de LSP) pelo LDP (MPLS básico), pelo RSVP-TE (MPLS-TE) ou pelo GMPLS RSVP-TE (MPLS-TP), podendo esse LSP ser comum a todos os PWs constituídos (e mesmo comum a outras aplicações do MPLS) sobre esse par de PEs.

O LSP entre os PE 1 e PE 2 é composto por esses PEs e por uma sucessão de equipamentos P no interior da rede MPLS. Esses equipamentos P são transparentes com relação aos PWs que utilizam o LSP. Somente os PE 1 e PE 2 participam da constituição de PWs e de suas funcionalidades.

Previamente à sinalização para a constituição de LSPs ocorre o processamento de um protocolo de roteamento do tipo IGP (*Interior Gateway Protocol*) na rede MPLS. Tais protocolos podem ser o OSPF ou o IS-IS para o MPLS básico, o OSPF-TE ou o ISIS-TE para o MPLS-TE ou o GMPLS OSPF-TE ou o GMPLS ISIS-TE para o MPLS-TP.

Cada terminação de PW em um PE é associada a um *Attachment Circuit* (AC) que o conecta ao CE correspondente à rede de camada 2 transportada.

A indispensável identificação da terminação de cada PW no PE de egresso ocorre pela consulta ao *PW label* (também denominado VC *label*), cujo valor se encontra no *Bottom Shim Label Header* do pacote MPLS recebido. Esse valor de *label* é associado ao PW correspondente na fase de sinalização, que consiste na utilização de mensagens do LDP estendido entre os PEs pares (PE 1 e PE 2, na hipótese considerada). A utilização a ser dada ao pacote MPLS recebido pelo PE 2 é baseada no *PW label* (atribuído previamente pelo PE 2 e enviado ao PE 1).

Os *PW labels* encontram-se obrigatoriamente na formatação genérica, o que significa que seus valores são atribuídos utilizando-se *Per-Platform Label Spaces* contidos nos PEs de egresso e transportados no interior de *Shim Label Headers* como *Bottom Shim Labels*. Os *labels* de LSP, ao contrário, podem ser formatados em conformidade com os LSRs que compõem a rede *backbone* MPLS, tais como ATM-LSRs, FR-LSRs ou *Frame-Based* LSRs. Os *labels de LSP* representam os *top labels* no quadro MPLS.

A atribuição e a distribuição de *PW labels* podem se realizar de modo estático ou mediante a utilização de extensões do LDP. Observa-se que alguns tipos de PW requerem o uso de *ControI Words*.

A distribuição de *PW labels* no protocolo LDP estendido para o VPWS processa-se no modo UD (*Unsolicited Downstream*), o que significa a utilização exclusiva de mensagens *Label Mapping* na distribuição de *PW labels*.

Uma outra particularidade que diz respeito ao LDP estendido para o VPWS é o fato de que podem ser utilizadas sessões *targeted* LDP, uma vez que os PEs não são necessariamente vizinhos adjacentes.

O *Label* TLV, que no LDP convencional é representado em diversas formas, dentre quais a formatação em VPIs/VCIs e DLCIs e a formatação genérica, quando aplicado para a constituição de PWs deve se restringir à formatação genérica.

7.4.2. **Transmissão de PDUs de Camada 2**

Suponha-se que o CE 1 da figura 7.3 (a terminação IP de origem, na realidade) tenha recebido um datagrama IP, cujo endereço IP de destino indica, na tabela de roteamento, o CE 2 (a terminação IP de destino, na realidade) como *next-*

hop. Um registro complementar indica o endereço de destino de camada 2 que o CE 1 deve inserir no quadro para o seu envio pelo devido *Attachment Circuit* (esse endereço de destino de camada 2 é obtido como resultado do processo de resolução de endereços previamente efetivado).

O CE 1 envelopa então o datagrama IP em uma PDU de camada 2, e a envia pelo devido *Attachment Circuit* ao PE 1. Ao receber essa PDU, o PE 1 identifica, a partir do *Attachment Circuit* de entrada, o PW a ser utilizado. À PDU de camada 2 são acrescentados então, sucessivamente, o *label* de PW e o *label* de LSP. O PE 1 envia então o pacote MPLS assim montado pelo LSP que conduz ao PE 2.

Ao receber o pacote MPLS, o PE 2 identifica o PW com base no *PW Label*, e em conseqüência identifica também o *Attachment Circuit* de saída. A função NSP realiza, se necessária, a translação de endereços de camada 2. A *Control Word*, se existir, é também processada.

Estando tudo correto, a PDU de camada 2 é então transmitida pelo PE 2 para o CE 2 pelo *Attachment Circuit* indicado. O CE 2 processa o *overhead* da PDU de camada 2 e obtém o datagrama IP de origem. A partir desse momento, o datagrama IP é enviado ao seu destino conforme os procedimentos convencionais do TCP/IP.

7.4.3. **Sub-Camada Control Word**

Na pilha de camadas da arquitetura PWE3 apresentada na figura 7.4 anterior neste capítulo, verifica-se a existência de uma subcamada opcional denominada *Encapsulation*, situada entre o *PW Demultiplexer* (*PW* l*abel*, no caso do VPWS), e o *payload* do pacote VPWS. A *Control Word* é um exemplo dessa subcamada, utilizada obrigatória ou opcionalmente, conforme o caso, no envelopamento de PDUs de camada 2 (*payload* do pacote VPWS na figura 7.4) durante o seu transporte através de um PW. O objetivo da *Control Word* é possibilitar a realização de algumas funções externamente às PDUs de camada 2 transportadas no PW.

Como visto anteriormente neste capítulo, as aplicações de camada 2 que utilizam a *Control Word* devem indicar essa utilização pelo LDP estendido quando da constituição dos respectivos PWs.

Se um PE é capaz de processar de forma não obrigatória a *Control Word*, ele deve ser configurado com um parâmetro que especifique se o uso da *Control Word* é *preferred* ou *not preferred*. Para cada PW deve existir um valor *default* para esse parâmetro. Se um PE não é capaz de processar a *Control Word*, ele deve ser configurado com o uso *not preferred* da *Control Word*.

A *Control Word* pode desempenhar as seguintes funções:

— Controle de sequenciamento de pacotes;
— Indicação e remoção de *padding*;
— Funções de OAM;
— Codificação de parâmetros dos protocolos de camada 2;
— Auxílio na fragmentação de *payloads*;

— Impossibilitação de ECMP (*Equal-Cost Multi-Path Load-Balancing*);
— Outras funções de carater geral.

A codificação de parâmetros dos protocolos de camada 2 será vista nos itens finais deste capítulo, que dizem respeito ao encapsulamento das diferentes opções de protocolo de camada 2.

Será aqui abordado o uso da *Control Word* como mecanismo que impossibilita a ocorrência de ECMP em PW PDUs. O mecanismo ECMP, que não deve ser utilizado em PW PDUs, tem como procedimento inicial a inspeção do *header* da MPLS PDU para verificar o protocolo de camada de rede utilizado. Se for o IPv4, o primeiro *nibble* (4 primeiros bits de um octeto) tem o valor 4 (0x0100), e se for o IPv6, esse *nibble* tem o valor 6 (0x0110).

É preciso evitar a possibilidade de se confundir o primeiro *nibble* das *Control Words* com esses valores. Para que isso aconteça, utiliza-se o valor zero (0x0000) ou o valor 1 (0x0001) para o primeiro *nibble* das *Control words*. Caso não fosse dessa forma, poderia acidentalmente ocorrer a interpretação errônea da um PW PDU *Header* como um pacote lP, o que agravaria a questão de sequenciamento de pacotes no VPWS.

7.5. ETHERNET SOBRE VPWS

A RFC 4448 especifica o encapsulamento de *Ethernet* PDUs para transporte sobre o VPWS. Essa RFC define os procedimentos para a utilização de PWs para a emulação de serviços *Ethernet* ponto-a-ponto sobre o MPLS.

O serviço consistindo no envelopamento de *Ethernet* PDUs para transporte no VPWS recebe a denominação EoMPLS (*Ethernet over* MPLS) também referido como ETHoMPLS.

Em uma configuração de EoMPLS, a função NSP é utilizada para o processamento de funções requeridas para as *Ethernet* PDUs, antes e após as terminações dos PWs. Por exemplo, no NSP de ingresso os campos *Preâmbulo* e FCS são retirados, enquanto no NSP de egresso esses campos, particularmente o FCS, são reinseridos. Outro exemplo são as possíveis funções de adição, retirada e alteração de valores de VLAN *tag* contidos em quadros Ethernet.

Cabem às terminações dos PWs as operações necessárias para estabelecimento e para manutenção dos PWs, e também aquelas necessárias para o encapsulamento e desencapsulamento das *Ethernet* PDUs para a sua transmissão pelo VPWS.

A RFC 4448 especificou ainda os seguintes aspectos para o serviço *Ethernet* sobre o VPWS:

— As Ethernet PDUs devem ser homogêneas em ambos os lados da rede MPLS, isto é, devem ser ambas Ethernet PDUs ou Ethernet VLAN PDUs.

É possível, embora não especificada, a transmissão de PDUs heterogêneas por meio de funcionalidades contidas nos NSPs;
— Na definição de MTUs devem ser considerados os overheads introduzidos pelo MPLS, pelo VPWS e pela eventual control word;
— Devem ser considerados procedimentos de QoS para o correto funcionamento do serviço.

7.5.1. **Modos de Operação de PWs**

As *Ethernet PDUs* encapsuladas podem conter ou não *VLAN tags* definidos pelo padrão lEEE 802.1Q-2014 para a constituição de VLANs. Para a transmissão dessas PDUs, os PWs podem operar de duas formas.

— *Raw mode;*
— *Tagged mode*.

7.5.1.1. Raw Mode

No *raw mode,* também denominado *Ethernet Port Mode*, a premissa básica é a de que o PW representa uma conexão virtual entre duas portas *Ethernet*. As *Ethernet* PDUs que entram no PW, contendo ou não IEEE 802.1Q *tags,* são transmitidas transparentemente até a terminação PW de egresso. Os *Attachment Circuits* são simplesmente as portas físicas no *raw mode.*

O PW *Type* 0x0005 identifica o encapsulamento de *Ethernet* PDUs no *raw mode.*

7.5.1.2. Tagged Mode

No *tagged mode,* também denominado *Ethernet* VLAN *Mode,* o PW, identificado pelo *Attachment Circuit* (porta física e valor de VLAN *tag),* representa uma conexão virtual entre duas terminações de uma VLAN. Opcionamente a operação *no tagged mode* possibilita a alteração do valor de VLAN ID pelo PE de egresso (por meio da função NSP), permitindo assim a comunicação entre duas VLANs com diferentes VLAN IDs através de um PW.

O PW *Type* Ox0004 identifica o encapsulamento de *Ethernet* PDUs no *tagged mode*.

O *tagged mode* possibilita duas formas de constituição de PWs:

— PW dedicado exclusivamente a uma VLAN;
— PW dedicado a um subconjunto de VLANs (por um único AC).

A Figura 7.5 ilustra essas duas possibilidades de constituição de PWs.

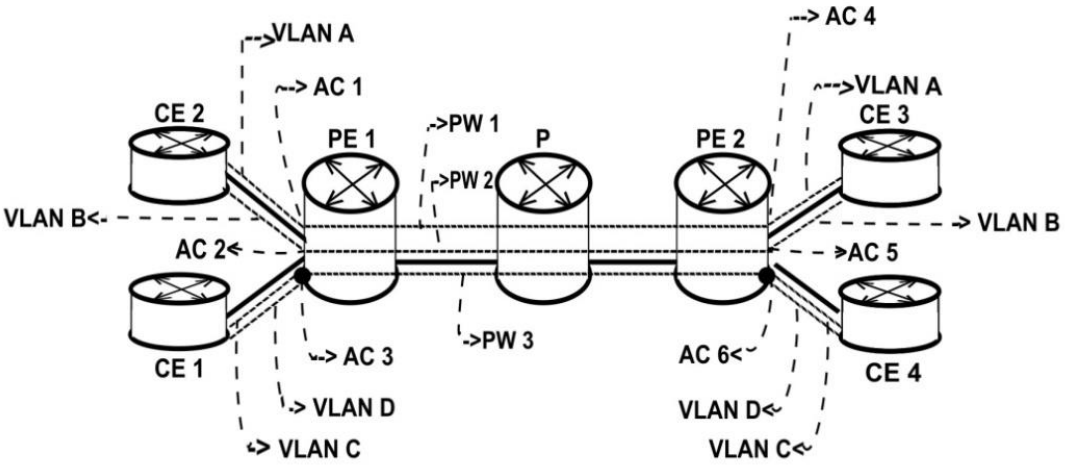


Figura 7-5 – Constituição de VPWS PWs no tagged mode.

Na parte superior dessa figura, cada VLAN (VLAN A e VLAN B) utiliza um PW próprio (PW 1 e PW 2, respectivamente). Na parte inferior da figura, duas VLANs (VLAN C e VLAN D) utilizam, por multiplexação estatística, um único PW (PW 3).

7.5.2. **Tipos de VLAN Tag**

Os *VLAN tags* podem ser de dois tipos:

— *Service-Delimiting Tags;*
— *Non-Service-Delimiting Tags.*

A distinção entre os tipos de *tag* recebidos é determinada por configuração local no PE. Um quadro *Ethernet* pode conter mais de um *tag*, quando no máximo um deles é um *Service-Delimiting-Tag*, o qual, se existir, é o *tag* que se encontra mais externamente no quadro.

Service-Delimiting Tag é um *tag* inserido nas *Ethernet PDUs* por um equipamento do SP e que é utilizado pelo SP para distinguir diferentes opções de tráfego.

Non-Service-Delimiting Tags são *tags* inseridos no quadro *Ethernet* por algum equipamento do usuário, e não tem qualquer significado para os PEs.

7.5.3. **Procedimentos Gerais**

Quando da transmissão de uma *Ethernet PDU*, o PE de ingresso realiza os seguintes procedimentos gerais:

— O NSP retira o FCS e o preâmbulo (se existir);
— Com base no *Attachment Circuit* de entrada, o PE identifica o PW a ser utilizado;

— A *Control Word,* se necessária, é inserida para o envelopamento da PDU;
— O *PW label* é inserido, envelopando a *Control Word* PDU;
— O *label* de LSP é inserido, envelopando a PW PDU;
— O pacote MPLS é montado sobre a rede de camada 2 que suporta o MPLS;
— O pacote MPLS é transmitido.

Quando o pacote atinge o PE de egresso, todo o *overhead* do quadro é processado e então retirado. O NSP regenera o FCS, realiza o necessário processamento de VLAN *tags*, e transmite a *Ethernet PDU* resultante para o respectivo CE pelo devido *Attachment Circuit*.

7.5.4. **Aspectos Operacionais**

Os subitens subsequentes neste item apresentam alguns aspectos operacionais relativos à transmissão de quadros Ethernet sobre o VPWS.

7.5.4.1. Gerência de MTU em Links PE-CE

Um PW só pode ser estabelecido se os valores de MTU em ambos os *links* PE-CE a ele associados são iguais. Se, apesar de medidas preventivas um PE de egresso receber uma *Ethernet* PDU cujo tamanho exceda a MTU do *link* PE-CE de egresso, essa PDU deve ser descartada.

7.5.4.2. Sequenciamento de Quadros

Se houver necessidade de controle do ordenamento sequencial de quadros, como o MPLS por si próprio não realiza essa função, é necessária a utilização da *Control Word* com esse propósito. Nesse caso, quadros recebidos fora de ordem devem ser descartados ou reordenados pela terminação PW de egresso.

7.5.4.3. Processamento de Erros em Quadros

Um pacote MPLS contendo uma PDU pode ser perdido, corrompido ou ter a sua ordem alterada no interior do PW utilizado. Os mecanismos nativos de processamento de erros em redes *Ethernet* devem ser ativados, competindo aos NSPs as medidas corretivas adequadas.

7.5.4.4. Controle de Fluxo

Para o controle de fluxo em uma rede *Ethernet* convencional foi emitido o padrão IEEE 802.3x, cujo uso é opcional. Quando utilizado, essa utilização ocorre em um *link* entre dois nós *Ethernet*. No caso de *Ethernet* sobre VPWS, essa função de controle pode existir apenas nos *links* PE- CE, não sendo aplicável nos PWs.

Essa função de controle de fluxo é exercida, basicamente, pelo envio de um quadro *Ethernet* especial, denominado *pause frame*, pelo nó que detecta uma condição de congestionamento. Se um PE detectar essa condição, ele pode enviar um *pause frame* para o CE, podendo encerrar essa função assim que o desejar.

7.5.4.5. Quality of Service (QoS)

Os *PRI bits* (*priority bits*), contidos no campo TCI (*Tag Control Information*) dos *headers* de VLANs, podem ser transpostos para o campo TC (*Traffic Class*) dos *Shim Label Headers* dos PW *Labels* e dos *labels* de LSP (registra-se que, mesmo em caso de MPLS sobre ATM ou *Frame Relay*, o *Top Shim Label Header* pode ser transmitido com o valor de *label* igual a zero).

No PE de egresso, valores contidos no campo TC podem ser re-inseridos nos *PRI bits* do campo TCl dos quadros de VLANs transmitidos para o respectivo CE.

Registra-se que o campo TC era referido anteriormente como campo EXP (*Experimental*).

7.6. VIRTUAL PRIVATE LAN SERVICE (VPLS)

O objetivo deste item é apresentar os diversos aspectos relativos ao VPLS sobre MPLS. O VPLS é um tipo de serviço L2VPN MP2MP para suporte exclusivo a *Ethernet*, que pode ser provido sobre o MPLS ou outras tecnologias PSN (como o IP, por exemplo). O VPLS é também conhecido pelas denominações TLS (*Transparent LAN Service*) e VPSN (*Virtual Private Switched Network).*

Os conceitos básicos e a arquitetura do VPLS encontram-se definidos na RFC 4664 (*Framework for Layer 2 Virtual Private Networks – L2VPNs*) e na RFC 4665 (*Service Requirements for Layer 2 Provider – Provisioned Virtual Private Networks*), referentes à constituição de L2VPNs. No que concerne ao plano de controle, mais precisamente aos mecanismos de sinalização para o estabelecimento e encerramento de conexões virtuais entre os PEs de uma VPLS VPN, foram definidas duas RFCS, com diferentes alternativas de sinalização:

— RFC 4761: VPLS *Using BGP for Auto-Discovery and Signaling;*
— RFC 4762: VPLS *Using LDP Signaling*.

Os fabricantes de sistemas de rede dividiram-se entre essas duas opções.

7.6.1. Configuração de VPLS VPNs

A configuração básica de uma rede MPLS dando suporte a uma única VPLS VPN encontra-se na Figura 7.6.

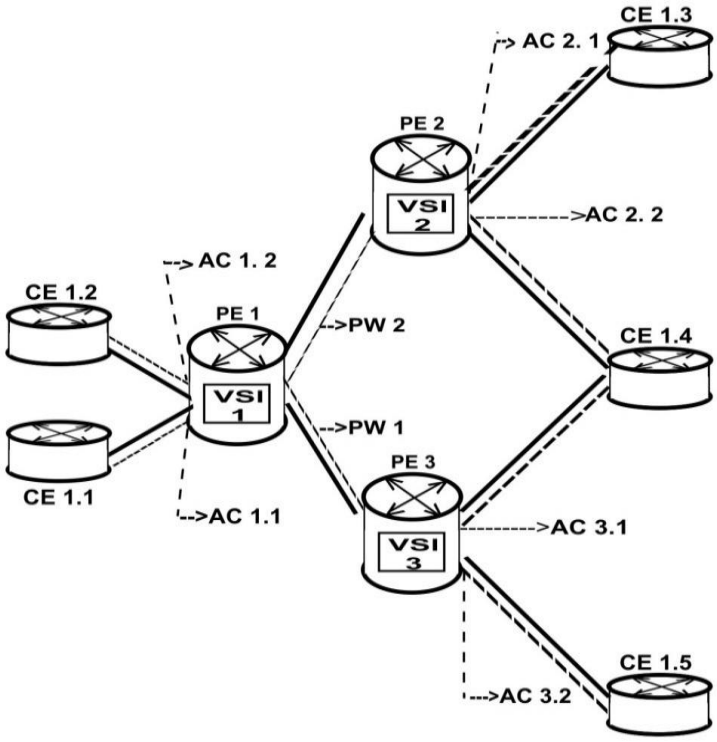


Figura 7-6 – Configuração básica de uma VPLS VPN.

Observa-se que nessa figura simplificada foram omitidos os equipamentos P da rede MPLS, que são totalmente transparentes ao VPLS. Registra-se também que os CEs podem ser roteadores ou *hosts* com interfaces VPLS, ou podem ser *Ethernet switches* aos quais se conectam, de alguma forma, roteadores ou *hosts*. Em ambos os casos os CEs conectam-se diretamente a terminações de *Attachment Circuits* (ACs).

Uma VSI (*Virtual Switching Instance*) é uma instância de entidade que emula a operação de um *Ethernet switch*, ou seja, que opera como um *Ethernet switch*. Uma VPLS VPN é composta por um conjunto de VSIs interconectadas por PWs. Esses PWs são constituídos de forma associada com as VSIs da respectiva VPLS VPN.

Um PE possui uma VSI por VPLS VPN. De um lado de uma VSI terminam os *Attachment Circuits*, e do outro lado os *pseudowires*.

É importante observar que, a exemplo das redes *switched Ethernet* tradicionais e do próprio VPWS, o VPLS suporta, além da opção *Port-Ethernet* (*raw mode*), a opção que considera a existência de VLANs IEEE 802.1Q em VPLS VPNs (ou seja, o *tagged mode*).

Uma VPLS VPN pode se destinar ao tráfego no *raw mode,* ou se destinar ao tráfego em uma VLAN específica ou ao tráfego em um subconjunto de VLANs no *tagged mode*.

As VSIs podem transmitir quadros *unicast*, *multicast* ou *broadcast*. A exemplo dos *Ethernet switches*, as VSIs inundam a rede no caso de recebimento de quadros *Ethernet* com endereço MAC de destino desconhecido, o que se constitui na forma

de inicio do processo de aprendizagem de endereços MAC (MAC *Address Learning*).

Como se observa, o VPLS, ao contrário do VPWS, utiliza os endereços MAC para fins de roteamento. O VPWS, por ser um serviço ponto a ponto entre CEs, não utiliza o processo de aprendizagem de endereços MAC.

Os PWs podem ser ponto a ponto ou ponto a multiponto. Os PWs ponto a ponto são sempre bidirecionais, enquanto os PWs ponto a multiponto são unidirecionais. Os PWs ponto a multiponto são adequados para o tráfego *multicast* ou *broadcast*, embora seja possível a transmissão *multicast* ou *broadcast* pela replicação de quadros em PWs ponto a ponto a partir de um PE.

Uma VSI inclui uma *Forwarding Information Base*, que registra o conjunto de informações necessárias para o envio de quadros *Ethernet* entre ACs e PWs, nos dois sentidos. Uma *Forwarding Information Base* pode ser preenchida dinamicamente (por exemplo, por aprendizagem com base em endereços MAC de origem) ou por configuração estática.

As portas no PE que se destinam aos CEs são configuradas como *Ethernet access ports* ou como 802.1Q *trunks*, conforme a sua natureza.

7.6.2. **Aspectos Gerais do VPLS**

Neste item serão abordados os seguintes aspectos do VPLS:

— Aprendizagem dinâmica de endereços MAC;
— Prevenção de *loops*.

7.6.2.1. Aprendizagem Dinâmica de Endereços MAC

Como existe roteamento MAC nas VSIs, tornam-se necessários mecanismos de aprendizagem de endereços MAC. Será aqui abordada a aprendizagem em sua forma dinâmica. Caso existam *loops* físicos na rede *Ethernet*, o processo de aprendizagem deve ser precedido por um processo de prevenção de *loops*, como será visto no próximo subitem.

- **Processo de Aprendizagem**

A Figura 7.7 representa uma configuração parcial de uma VPLS VPN, que será utilizada para ilustrar a aprendizagem dinâmica de endereços MAC.

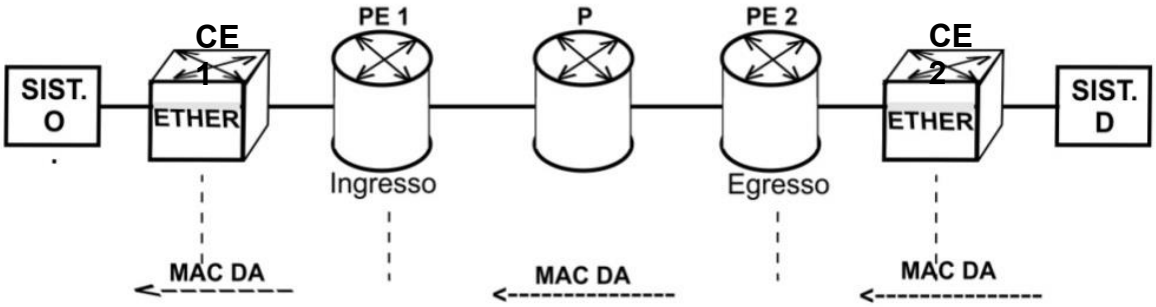


Figura 7-7 – Aprendizagem dinâmica de endereços MAC.

Foram admitidas nessa figura as seguintes hipóteses:

— Os quadros MAC transportam datagramas IP;
— Sist. O (sistema de origem): é a terminação IP mais próxima do PE na rede de acesso de origem;
— Sist. D (sistema de destino): é a terminação IP mais próxima do PE na rede de acesso de destino;
— O endereço MAC DA é o endereço MAC do sist. D;
— O endereço IP de destino é o endereço do terminal lP na rede de acesso de egresso (que pode estar além do sist.D);
— Os CE 1 e CE 2 são *Ethernet switches*.

Como se verifica na figura, o sist.D, por meio de quadros de serviço, possibilita a terminações Ethernet que se situam à esquerda na figura, a aprendizagem do endereço MAC do próprio Sist. D. Esse endereço é aprendido com as seguintes finalidades:

— O CE 1 utiliza o MAC DA para definir a porta que conduz ao devido *Attachment Circuit* de ingresso (esse *Attachment Circuit* conduz à VSI da VPLS VPN à qual pertence o quadro de serviço recebido);
— O PE 1 utiliza o MAC DA para definir o devido LSP para alcançar o PE 2;
— O PE 2 utiliza o MAC DA para definir o devido *Attachment Circuit* de egresso.

Quando o PE 1 recebe um quadro do PE 2, ele verifica se o valor do *label* de VPLS VPN recebido confere com o valor de *label* VPLS PW por ele sinalizado para o PE 2. Dessa forma, ele assegura a identificação do PE 2, e confere o *label* VPLS PW (*label* anteriormente sinalizado pelo PE 2) a ser utilizado no encapsulamento de quadros MAC que virão a ser destinados, na VPLS VPN, ao endereço MAC DA (do Sist.D).

O PE 1 aprende, nesse ato, o endereço MAC da terminação que gerou o quadro MAC (endereço MAC do Sist. D), o que permite a definição do LSP a ser utilizado.

Os endereços MAC aprendidos pelos PEs de uma VPLS VPN são registrados e mantidos temporariamente em tabelas de roteamento MAC. Esses registros são refrescados e retemporizados pelos quadros MAC de tráfego eventualmente recebidos. Se não houver tráfego de dados durante a temporização, os respectivos registros são apagados.

Vale registrar que, para a operacionalização dessa forma de aprendizagem, os endereços MAC têm de ser globalmente únicos em cada VPLS VPN. Entre diferentes VPLS VPNs, porém, pode haver a superposição de endereços MAC.

Os procedimentos de aprendizagem MAC podem ocorrer no âmbito de uma VLAN existente em uma VPLS VPN. Os valores de VLAN ID para uma VLAN 802.1Q são, de um modo geral, globalmente únicos dentro de cada VPLS VPN. Existem situações, no entanto, em que esses valores de VLAN ID assumem funções de identificação de serviços, o que requer que os VLAN IDs sejam globalmente únicos em todo o conjunto de VPLS VPNs constituídas sobre a rede MPLS.

O sist. O da Figura 7.7 requer a utilização de alguma forma de resolução do endereço MAC DA do Sist. B a partir do endereço IP do terminal IP de destino (a partir do prefixo IP da subrede onde se localiza o terminal IP de destino, na realidade) para a montagem dos quadros MAC de dados destinados ao Sist.D. A resolução de endereços MAC deve suceder o processo de roteamento IP na rede IP global, processo esse que se desenvolve transparentemente à rede MPLS (ou seja, aos equipamentos PE e P).

- **Aprendizagem Não-Qualificada e Qualificada**

Quando da existência de VLANs em uma VPLS VPN, existem duas formas de tratar a aprendizagem de endereços MAC, que são a aprendizagem não-qualificada e a aprendizagem qualificada.

Na aprendizagem não-qualificada, a existência das VLANs não é considerada pelo VPLS. Assim, uma distribuição de quadros em *broadcast* engloba todos os participantes da VPN, em todas as VLANs. Esse tipo de aprendizagem é utilizada no *raw mode*.

Na aprendizagem qualificada, a cada VLAN, isoladamente ou dentro de um subconjunto de VLANs, corresponde uma VPLS VPN, que possui um domínio de *broadcast* separado. A vantagem dessa forma é a maior eficiência no uso da rede MPLS. Essa é a forma de apredizagem utilizada no *tagged mode*.

7.6.2.2. Prevenção de Loops

Caso existam *loops* físicos na topologia de uma VPLS VPN, torna-se necessária a utilização de algum mecanismo de prevenção de *loops*, dada a frequente inundação nesse tipo de rede.

Os protocolos RSTP e MSTP podem ser utilizados no VPLS em determinadas circunstâncias. Para a topologia *full mesh* entre VSIs, contudo, que é amplamente

utilizada, foi definido um mecanismo *split horizon*, de maior simplicidade que esses protocolos.

O *split horizon*, nesse caso, consiste simplesmente na impossibilidade imposta às VSIs de retornar para o interior da rede quadros MAC recebidos em um PW. Isso significa que as VSIs não podem reenviar quadros em outros PWs da rede VPLS.

No caso de adoção de outras topologias de rede, como no exemplo de *hub-and-spoke*, torna-se inviável o uso do *split horizon*, sendo então necessário algum outro mecanismo de prevenção de *loops.* A operação da rede na topologia *hub-and-spoke* requer o trânsito de quadros por PWs entre VSIs, o que torna inviável a aplicação do *split horizon.* Um outro exemplo são as topologias que utilizam *multi-homing*, o que representa uma forma de *looping* físico.

A adoção da topologia *full mesh*, se por um lado é desejável, por outro torna-se problemática quando os PEs da VPLS VPN se afastam entre si e se tornam numerosos. Para atender a esse problema, foi definida uma forma hierarquizada de atendimento, denominada HVPLS (*Hierarchical VPLS*), que será vista adiante neste item.

7.6.3. **Procedimentos no VPLS**

Como no exemplo dos VPWS PWs, os VPLS PWs podem corresponder à operação do tipo *Ethernet raw mode* ou do tipo *Ethernet tagged mode.* Registra-se, contudo, que enquanto cada VPWS PW representa uma VPN ponto a ponto, os VPLS PWs participam de configurações de VPNs multiponto a multiponto. Isso porque no VPLS os pontos de terminação de um PW são conectados diretamente a VSIs, de onde serão alcançadas as redes Ethernet externas.

Os VPLS PWs do tipo *Ethernet raw mode*, denominados também *Ethernet* PWs, caracterizam-se pela sua transparência às VLANs eventualmente constituídas por um usuário. Assim, todo o tráfego desse usuário entrante por um *Attachment Circuit* em um PE, contendo ou não VLAN *tags*, diz respeito a uma VPLS VPN, sendo conduzido a uma única VSI nesse PE.

A parte desse tráfego que se destina a um outro PE de egresso será conduzida por um único PW entre o PE de ingresso e esse PE de egresso. Registra-se que os endereços MAC devem ser únicos na VPLS VPN no *Ethernet raw mode.*

Nos VPLS PWs do tipo *Ethernet tagged mode,* denominados também *Ethernet* VLAN PWs, o tráfego de uma determinada VLAN (que participa do *tagged mode*) é tratado separadamente, utilizando PWs próprios para essa VLAN ou para um subconjunto de VLANs ao qual essa VLAN pertence. O tráfego não-VLAN (quadros *untagged* e *priority tagged*) e o tráfego de VLANs que não participam do *Ethernet tagged mode* são processados no *Ethernet raw mode*, o que pode ocorrer em um mesmo PE.

Observa-se também que os endereços MAC devem ser únicos em cada uma das VPLS VPNs que operam no *Ethernet tagged mode,* assim como em uma VPLS VPN operando no *Ethernet raw mode.*

O encapsulamento de quadros MAC para o acesso a uma VPLS VPN, operando em qualquer dos modos acima citados, pode ocorrer de três formas:

— O encapsulamento é parte do quadro do usuário (como no exemplo de um VLAN ID da rede do usuário);
— O encapsulamento é unicamente um delimitador de serviço;
— O encapsulamento atende as duas condições anteriores.

Um delimitador de serviço é uma subcamada, ou simplesmente um campo, utilizado para identificar um usuário ou um serviço (uma VPLS VPN, no caso do VPLS). Um delimitador de serviço pode ser introduzido com esse objetivo específico, quando tem significado local, ou pode consistir de parte do encapsulamento inerente à própria PDU gerada pelo usuário em seu serviço local.

7.6.4. **Topologias em VPLS VPNs**

O conjunto de PWs pertinentes a uma VPLS VPN podem constituir diferentes topologias. Como os PWs são ponto a ponto ou ponto a multiponto, existe sempre uma única VSI em cada um dos *endpoints* remotos de um PW. Como os PWs ponto a multiponto são unidirecionais no sentido da raiz para os multipontos, o tráfego para a raiz desses PWs deve ocorrer por outros PWs ponto a ponto que terminam nessa raiz.

Podem ser citados os seguintes tipos de topologia adotados:

— *Full mesh*;
— *Tree Structered*;
— *Tree with Meshed Highest LeveI*.

7.6.4.1. Topologia Full Mesh

Nessa topologia, cada VSI em uma dada VPLS VPN tem um único PW para cada outra VSI dessa VPN, sendo esses PWs ponto a ponto ou ponto a multiponto.

A prevenção de *loops* se realiza por *split horizon*, conforme menção anterior neste capítulo.

7.6.4.2. Topologia Tree Structured

Nesse tipo de topologia, cada VSI ocupa uma posição em uma árvore estruturada que corresponde a um determinado nível hierárquico. Uma VSI tem apenas um PW no sentido superior dessa árvore. A raiz da árvore topológica possui o nível hierárquico máximo.

7.6.4.3. Topologia Tree with Meshed Highest Level

Essa topologia é uma variação do tipo de topologia *Tree Structured*, que permite a existência de múltiplas VSIs no nível hierárquico máximo. Essas VSIs devem apresentar uma topologia *full mesh* entre si. A Figura 7.8 ilustra esse tipo de topologia.

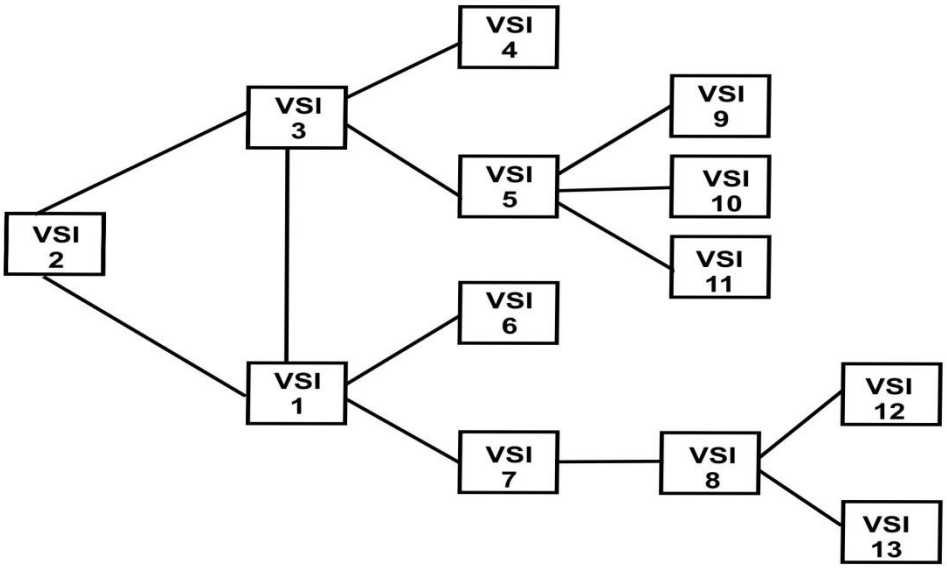


Figura 7-8 – Topologia Tree with Meshed Highest Level.

7.6.5. VPLS Hierarquizado (HVPLS)

A forma básica de operação do VPLS consiste em uma configuração *full mesh* de túneis LSP e de PWs entre os PEs participantes do serviço. Isso resulta obviamente em uma carga de *overhead* de sinalização e de replicação de quadros *multicast* e *broadcast* na rede. Tal condição se torna mais crítica quando da existência de uma maior quantidade de PEs, particularmente se um conjunto de PEs se concentra em um mesmo prédio.

Um prédio contendo diferentes empresas usuárias de serviços de telecomunicações é denominado uma MTU (*multitenant unit*). Torna-se conveniente a implementação de alguma forma de concentração do tráfego relativo aos PEs existentes em uma MTU, mesmo que esses PEs pertençam a VPLS VPNs de diferentes empresas.

A topologia de rede VPLS contendo equipamentos de agregação de tráfego configura o que se designa por HVPLS. Nessa topologia, a rede *backbone* MPLS é referida como *hub* (*hub* PWs, por exemplo), enquanto a rede que conecta o *hub* aos equipamentos de agregação de tráfego são referidos como *spoke*.

A Figura 7.9 representa as vantagens da utilização do HVPLS.

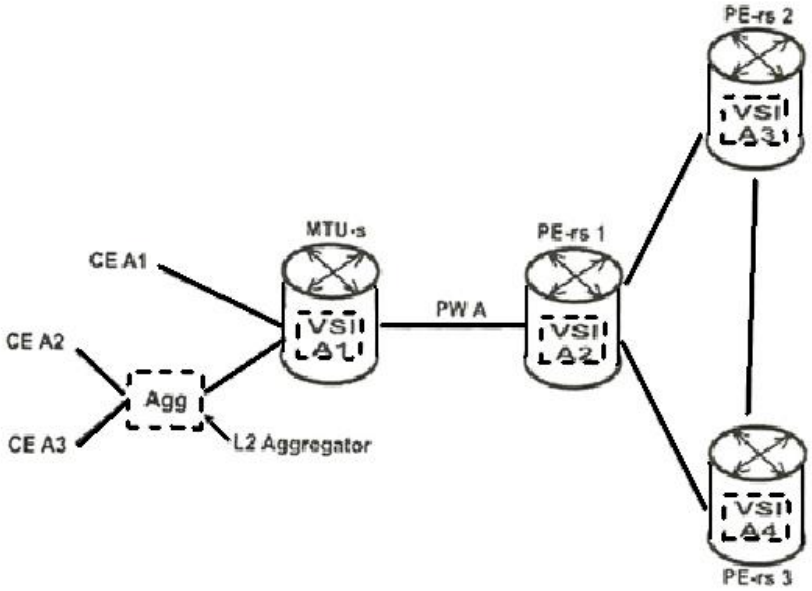


Figura 7-9 – Exemplo de configuração de rede com HVPLS.

MTU-s significa *MTU-switching*, enquanto PE-rs significa *PE-routing/switching*.

Nessa figura, considera-se a existência de uma única VPLS VPN cujo tráfego *spoke* cursa por um único *spoke PW* (PW A, na figura). Caso existissem múltiplas VPLS VPNs, a cada uma dessas VPNs corresponderia um *spoke PW*. Como se observa, os *spoke PWs* são do tipo VPWS PW, uma vez que constituem associações ponto a ponto.

O PW A termina, como se verifica na figura, na VSI A1 e na VSI A2. A VSI A2, diferentemente do que ocorre no VPLS convencional, é capaz de comutar tráfego entre o PW A e os PWs VSI A2/VSI A3 e VSI A2/VSI A4. Esses últimos PWs são *hub PWs* na concepção HVPLS.

Observa-se que o mecanismo *split horizon* não pode ser utilizado no HVPLS, uma vez que existe a possibilidade de tráfego entre *spoke* PWs e *hub PWs*, o que contraria o princípio de funcionamento desse mecanismo. Deve ser utilizado então um outro mecanismo de prevenção de *loops*, a exemplo do RSTP ou do MSTP. A necessidade de uso desses protocolos é realçada caso se utilize *dual-homing* no acesso *spoke*.

Existem duas terminologias referentes ao HVPLS, definidas respectivamente pelas RFC 4664 e RFC 4762.

CAPÍTULO 8

OAM, OPERAÇÃO E GERÊNCIA DE BRIDGES

8.1. PREÂMBULO

O presente capítulo objetiva a abordagem, em redes Ethernet, de OAM (*Operations, Administration and Maintance*), de alguns protocolos de operação e do processo de gerenciamento.

No que concerne OAM, serão apresentados os padrão IEEE 802.1Q-2014 (Cláusulas 18, 19, 20, 21, 22 e 29), a recomendação ITU-T G.8013/Y.1731 e a Cláusula 57 do padrão IEEE 802.3.

Os processos operacionais abordados são os seguintes:

— Mecanismos referentes a comutação automatica de proteção;
— *Link Layer Discovery Protocol* (LLDP), definido no padrão IEEE 802.1AB-2009;
— *Link Aggregation* (*Link Agg.*), definido no padrão IEEE 802.1AX-2008.

No que concerne os processos referentes à comutação automática de proteção, foram contemplados os seguintes padrões:

— *Ethernet Protection Switching*, definido na recomendação G.8031/Y.1342;
— *Ethernet Ring Protection Switching* (ERPS), definido na recomendação G.8032/Y.1344;
— *Extreme Networks' Ethernet Automatic Protection* (EAPS), definido na RFC 3619.
— Padrão IEC 62439.

Finalmente, serão expostas ideias genéricas sobre gerenciamento de *bridges*, dada a amplitude de informçaões cobertas por esse tema, com base na Cláusula 12 e na Cláusula 17 do padrão IEEE 802.1Q-2014.

8.2. OAM EM REDES ETHERNET (EOAM)

Serão abordados neste capítulo dois tipos de EOAM: OAM de Serviço (SOAM) e OAM de *Link*.

O SOAM engloba o SOAM CFM (*Connectivity Fault Management*), também referido como SOAM FM, e o SOAM PM (*Performance Monitoring*). Tais tipos de OAM são definidos nas Cláusulas 18 a 22 do padrão IEEE 802.1Q-2014 e na recomendação ITU-T G.8013/Y.1731.

Enquanto o padrão IEEE 802.1Q-2014 restringe-se ao SOAM CFM, a recomendação ITU-T G.8013/Y.1731, com maior amplitude, engloba também o SOAM PM.

O padrão IEEE 802.1Q-2014 apresenta, adicionalmente, em sua Cláusula 29, uma extensão do SOAM CFM referida como DDCFM (*Data-Driven and Data-Dependent CFM*).

Além do SOAM, o EOAM engloba também OAM de *Link* (*Link* OAM), objeto da Cláusula 57 do padrão IEEE 802.3.

O SOAM CFM em Ethernet engloba as funções necessárias para a detecção, verificação e localização (*isolation*) de falhas de conectividade em *bridged* LANs virtuais. Tais funções podem ser utilizadas em redes operadas por múltiplas organizações independentes, cada qual com acesso restrito de gerenciamento aos equipamentos das outras organizações.

As funções de SOAM PM em redes Ethernet permitem a medição de diferentes parâmetros de desempenho para conexões Ethernet ponto a ponto, parâmetros esses referentes a *delays* de trânsito de quadros, à perda de quadros, à vazão de tráfego (*throughput*) e à disponibilidade da rede.

Os mecanismos referentes a OAM de *Link* em redes Ethernet foram inicialmente definidos no padrão IEEE 802.3ah (*Ethernet in the First Mile*). Tais definições foram posteriormente incorporadas à Cláusula 57 do padrão IEEE 802.3-2008.

8.2.1. **SOAM CFM: Padrão IEEE 802.1Q-2014**

Na versão do IEEE, o SOAM CFM foi especificado inicialmente no padrão IEEE 802.1ag (*Connectivity Fault Management*). Tal padrão foi obsoletado e seus termos incorporados às Cláusulas 18, 19, 20, 21 e 22 do padrão IEEE 802.1Q-2014 (*Bridges and Bridged Networks*).

Tais cláusulas do padrão IEEE 802.1Q-2014 especificam protocolos, procedimentos e objetos gerenciáveis necessários ao suporte do gerenciamento de falhas de conectividade em redes constituídas por *bridges*, particularmente por VLAN-*aware bridges*. São providos recursos para a detecção, verificação, isolamento e comunicação de falhas de conectividades nesse tipo de rede, além da descoberta e da verificação de caminhos a serem seguidos por quadros de serviço originados e destinados a usuários de serviços específicos.

O padrão IEEE 802.1Q-2014 destina-se ao provimento de funções de OAM em variadas configurações de redes *Ethernet*, desde simples conexões ponto a ponto a complexas redes envolvendo uma multiplicidade de LANs e de *bridges*, incluindo LANs não IEEE 802.3.

As funções definidas no contexto do padrão IEEE 802.1Q-2014 e da recomendação ITU-T G.8013/Y.1731 não englobam aquelas referentes à recuperação de falhas de conectividade. Tais funções são de responsabilidade de outros mecanismos de controle de camada 2, tais como *spanning trees* e APS (*Automatic Protection Switching*).

Em consonância com o padrão IEEE 802.1Q-2014, os mecanismos definidos para a operação do CFM englobam os seguintes aspectos:

— Capacidade de configurar múltiplos Domínios de Manutenção (MDs) hierarquizados (*nested*) em uma *bridged* LAN ou em uma rede de *bridged* LANs interconectadas, podendo envolver diferentes organizações administrativas;
— Capacidade de configurar Associações de Manutenção (MAs) em MDs;
— Definição de protocolos, procedimentos e formatos de quadros requeridos pela CFM;
— Capacidade de configurar e gerenciar Pontos de Manutenção (MPs) dentro de MAs;
— Capacidade de comandar os MPs para realizar certas operações de isolamento de falhas e inspecionar os resultados.

As características do CFM podem ser utilizadas em redes operadas por um ou por múltiplos provedores de serviço independentes, cada um restrito ao gerenciamento de sua planta.

O CFM apresenta o seguinte escopo:

— Definição de Domínios de Manutenção (MDs), de Níveis de Domínio de Manuntenção, de Associações de Manutenção (MAs), dos respectivos Pontos de Manutenção (MPs) e dos objetos gerenciáveis necessários;
— Descrição dos protocolos e procedimentos utilizados pelos Pontos de Manutenção para detectar e diagnosticar falhas de conectividade dentro de um Domínio de Manutenção.

8.2.1.1. Conceitos Operacionais do CFM

O CFM baseia-se nos seguintes conceitos operacionais:

— Domínios de Manutenção (MDs) e Pontos de Acesso de Serviço de Domínios (DoSAPs);
— Pontos de Acesso de Serviços Intermediários (ISAPs);
— Instâncias de Serviço e Associações de Manutenção (MAs);
— Pontos de Manutenção (MPs);
— Pontos Terminais de MAs (MEPs);
— Entidades de Manutenção (MEs);
— Pontos Intermediários de MDs (MIPs);
— Níveis de MD;
— Endereçamento de MPs;

- **MDs e DoSAPs**

Um Domínio de Manutenção (MD) é uma rede, ou parte de uma rede, controlada por um único operador de serviço, na qual se efetiva o gerenciamento de falhas de conectividade.

Um MD é capaz de oferecer conectividade para sistemas a ele exteriores. Essa conectividade ocorre através dos Pontos de Acesso de Serviço de Domínios (DoSAPs) do MD. O MD deve ser capaz de suportar serviços entre quaisquer de seus DoSAPs.

Cada DoSAP é uma instância localizada em uma *bridge*, seja na EISS (*Enhanced Internal Service Sublayer*) ou na ISS dessa *bridge*.

O CFM admite variadas opções de redes que constituem um MD, desde um simples *link* ponto a ponto à totalidade de uma *bridged* LAN virtual.

Os MDs podem ser administrados separadamente, sendo que a cada MD é atribuído um MD *Name*. Os MD *Names* devem ser globalmente únicos em uma rede que engloba diferentes MDs.

A Figura 8-1 ilustra a estrutura de um MD sob o ponto de vista do operador de serviços.

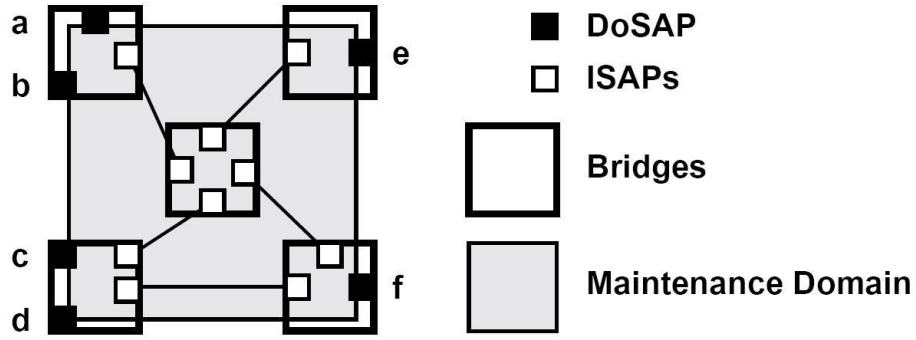


Figura 8-1 - Estrutura de um MD.

- **Instâncias de Serviço, MAs, MPs, MEPs, MEs e MIPs**

Quando se configura uma instância de serviço para um usuário, o administrador da rede configura um certo número de DoSAPs para dar acesso a esse usuário. Devem ser atribuídos identificadores a essa instância de serviço, de forma a segregá-la de outras instâncias de serviço eventualmente suportadas pelo respectivo MD (os VIDs são um exemplo desses identificadores).

A criação de uma instância de serviço estabelece uma conectividade sem conexão (*connectionless*) entre os DoSAPs configurados.

Um MEP (*Maintenance Association End Point*) é um tipo de *Maintenance Point* (MP) que pode gerar e receber CFM PDUs e acompanhar as respectivas respostas. Um MEP mantém uma ME (isto é, uma relação ponto a ponto entre MEPs) com cada outro MEP de sua MA.

MEPs podem ser criados em uma porta de *bridge* ou em um ponto de conexão de uma estação final a uma LAN.

Um MIP (*Maintenance Domain Intermediate Point*) é uma entidade CFM (ou seja, um MP) intermediária que pode reagir quando da passagem de determinados

tipos de CFM PDU e que pode gerar certas CFM PDUs. Essa geração ocorre somente em resposta a outras CFM PDUs recebidas.

O padrão IEEE 802.1Q divide o MIP em duas MHFs (MIP *Half Functions*) em uma porta de *bridge*, que são a *Up* MHF e a *Down* MHF.

Uma Up MHF é um *Up* MP, assim como um *Up* MEP, por residir em um MP de uma *bridge* que troca CFM PDUs diretamente com a *Bridge Relay Entity* dessa *bridge*. Uma *Down* MHF, ao contrário, troca CFM PDUs diretamente através de um *link*.

Uma MHF pode manter uma MIP CCM *Database* separada da MEP CCM *Database*.

Quando são configurados MEPs em cada um dos DoSAPs de uma instância de serviço, é criada uma MA para monitorar a conectividade (ou a conexão) nessa instância de serviço. Em uma conectividade (ou conexão) ponto a ponto, a MA consiste em apenas uma ME, enquanto que em uma conectividade multiponto a MA consiste em múltiplas MEs.

A cada MA corresponde um *Short* MA *Name*, que é único no respectivo MD.

Para a identificação de uma MA é criado um identificador, referido como MAID (MA *Identifier*). É necessário identificar também o MD que a contém. A plena identificação de uma MA (ou seja, uma MAID) é obtida então pela associação do *Maintenance Domain Name* com o *Short* MA *Name*.

Como será visto adiante neste item, os valores de MAID são transportados nas CFM PDUs para identificar inadvertidas conexões entre MEPs.

A cada MEP corresponde, em uma MA, um valor de *Maintenance Association End Point Identifier* (MEPID). Assim, a plena identificação de um MEP é obtida pela associação do MAID com o MEPID. Observa-se que o MEPID é representado por um número inteiro na faixa de 1 a 8191.

- **Níveis de MD**

Um MD pode prover instâncias de serviço a outro MD que o envolve ou pode utilizar instâncias de serviço providas por um terceiro MD por ele envolvido.

Para isso, são definidos níveis hierárquicos para esses *nested* MDs, referidos como níveis de MD. Os níveis de MD são incluídos nas CFM PDUs, evitando-se assim a necessidade de acrescentar qualquer informação aos quadros de dados para utilizar essa organização em camadas.

Diferentes níveis de MD são tipicamente atribuídos a usuários, a provedores de serviço e a operadores de serviço.

A Figura 8-2 representa o uso de níveis de MD para possibilitar a utilização de CFM sobre a conectividade provida por duas *bridged* LANs de operadores de serviço e pelo provedor de serviço que as interligam.

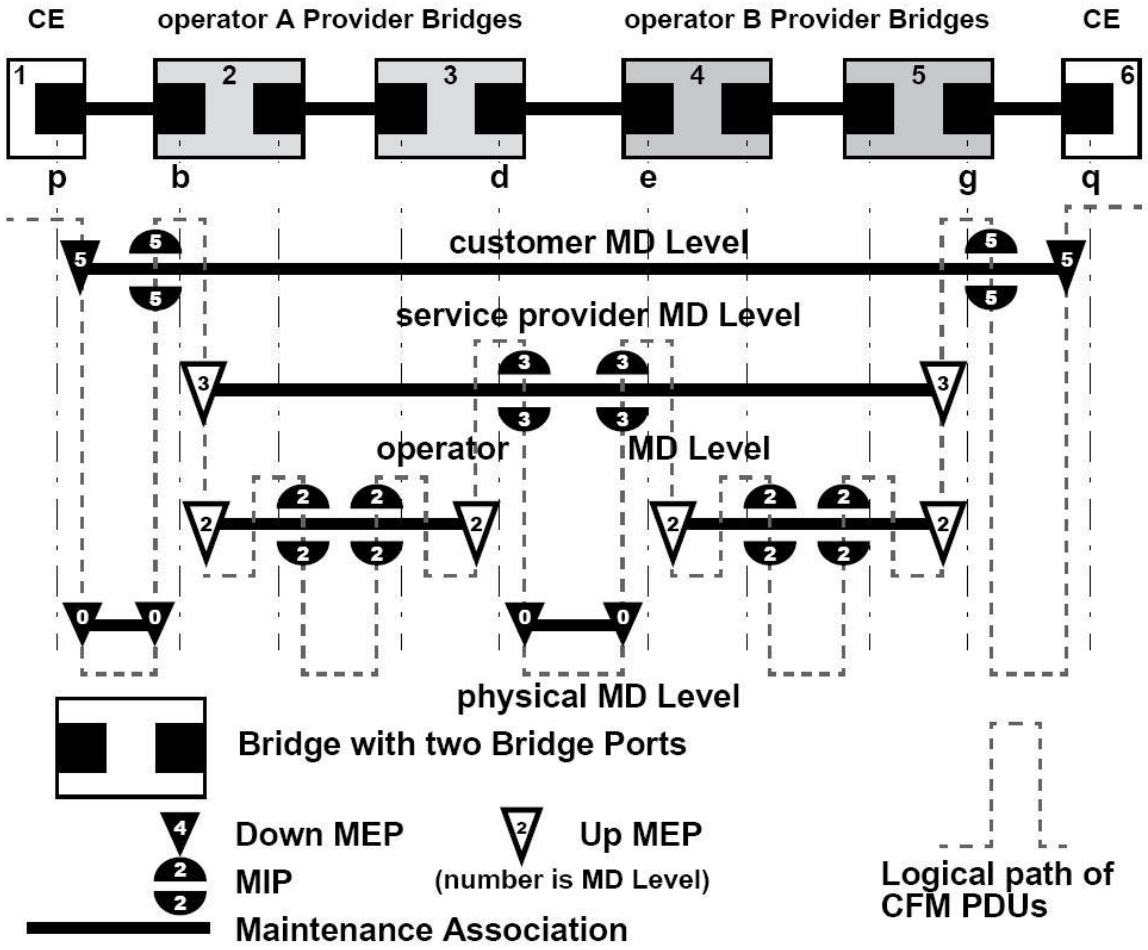


Figura 8-2 - Níveis de MD com MEPs e MIPs.

O uso *default* dos 8 possiveis níveis de MD em uma única porta de *bridge* encontra-se representado na Figura 8-3.

MD Level	7	6	5	4	3	2	1	0
Use	customer			service provider		operator		
	Highest	<- Higher lower ->						Lowest

Figura 8-3 - Uso default dos níveis de MD.

Devido à possibilidade de mascaramento de níveis de MD por encapsulamentos tais como aqueles providos por VLAN *tags*, o que torna desnecessária a aplicação do método acima exposto, o padrão IEEE 802.1Q-2014 define um método alternativo para alocação de níveis de MD.

8.2.1.2. Funções, Protocolos e Mensagens do CFM

As funções básicas do CFM são as seguintes:

— Descobrimento de caminhos;
— Detecção de falhas;
— Verificação e isolamento de falhas;
— Notificação de falhas;
— Recuperação de falhas.

Descobrimento de caminhos utiliza o protocolo *Linktrace* para determinar o caminho tomado para atingir um endereço MAC alvo.

Detecção de falhas utiliza o protocolo *Continuity Check* (Verificação de Continuidade) para detectar tanto falhas de conectividade quanto a conectividade não intencional entre instâncias de serviço.

Verificação e isolamento de falhas são processos administrativos utilizados tipicamente após a detecção de falhas.

Verificação de falhas pode também confirmar a iniciação ou a restauração de conectividade bem sucedidas. Para a verificação e o isolamento de falhas utiliza-se o protocolo *Loopback*.

Notificação de falhas é provida por um MEP que detectou falha de conectividade em sua MA, seja porque CCM PDUs esperadas não foram recebidas, porque CCM PDUs inválidas ou não esperadas foram recebidas, ou porque uma CCM PDU recebida continha uma notificação de falha na porta de *bridge* a ela associada. Uma única falha pode ser detectada por diferentes MEPs e pode resultar na geração de múltiplos alarmes de falha.

Para *bridged* LANs, a recuperação de falhas é provida tipicamente pelos protocolos *Spanning Tree* utilizados, enquanto que a recuperação de falhas em outras condições de rede é provida por mecanismos de comutação de proteção a serem vistos adiante neste item.

O padrão IEEE 802.1Q-2014 não especifica ações corretivas por parte do administrador de rede, tais como aquelas voltadas para erros de configuração ou para a substituição de componentes em falha.

Para a consecução dessas funções são utilizados os protocolos relacionados na Figura 8-4, com as respectivas mensagens.

Protocolos	Mensagens
Protocolo de Conferência de Continuidade	Mensagens de Conferência de Continuidade (CCMs)
Protocolo Loopback	>Mensagens Loopbak (LBMs) >Mensagens Loopback Replay (LBRs)
Protocolo Linktrace	>Mensagens Linktrace (LTMs) >Mensagens Linktrace Replay (LTRs)

Figura 8-4 - Protocolos e mensagens do CFM.

- **Protocolo de Conferência de Continuidade (CC Protocol)**

O processo de deteção de falhas de continuidade em uma MA utiliza o Protocolo de Conferência de Continuidade (*CC Protocol*). Esse protocolo permite também detectar falhas referentes a problemas verificados na recepção de CCMs (*Continuity Check Messages*).

Um MEP pode transmitir periodicamente CCM PDUs unidirecionais com endereço MAC *multicast* de destino, anunciando nessas mensagens também o seu endereço MAC e o respectivo MAID. O período de transmissão é configurável, podendo ser igual a 3,3 ms, 10 ms, 100 ms, 1s, 10s, 1 min ou 10 min.

O Protocolo de Conferência de continuidade tem também como função analisar as CCM PDUs recebidas dos demais MEPs da MA.

Uma falha de conectividade é detectada quando um MEP deixa de receber três CCMs consecutivas do seu MEP par. O tempo correspondente à não recepção de três CCMs consecutivas pode ser tão curto quanto 10,8 ms.

As discrepâncias entre as CCM PDUs recebidas e as expectativas configuradas no MEP receptor também apontam para falhas de conectividade.

Quando um MEP detecta uma condição de falha na recepção de CCM PDUs, ele deve setar para 1 o valor do *Remote Defect Indicator* (*RDI bit*), contido na própria CCM PDU, a partir da próxima CCM PDU por ele enviada. Dessa forma, o MEP que enviou as CCM PDUs não recebidas ou recebidas com falhas será notificado.

As CCM PDUs trocadas são armazenadas para exame pelo administrador da rede.

- **Protocolo Loopback**

O administrador da rede utiliza o protocolo *Loopback* para realizar a verificação de falhas. Um MEP pode transmitir uma mensagem LBM (*Loopback Message*) *unicast* para outro MP (MEP ou MIP) da respectiva MA, cujo endereço

MAC pode ser encontrado em mensagens CCM ou mensagens LTM/LTR. O MP receptor responde com a transformação da mensagem LBM em uma mensagem LBR (*Loopback Reply*) *unicast* e com o seu envio para o MEP de origem. O conteúdo dessas mensagens LBR deve ser analisado pelo administrador da rede.

A Figura 8-5 apresenta um exemplo de aplicação do protocolo *Loopback* com transmissão destinada a um MIP nesse exemplo.

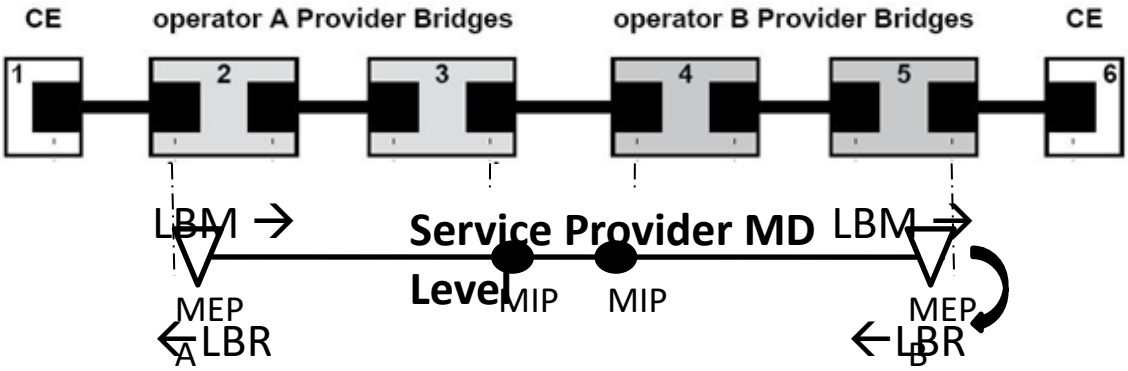


Figura 8-5 – Exemplo de aplicação do protocolo Loopback.

Protocolo Linktrace

O processo de descobrimento de caminhos ocorre pelo protocolo *Linktrace* (similar ao *IP Traceroute*), que determina o caminho utilizado para atingir um determinado endereço MAC alvo *(target*), MP a MP, desde o MEP de origem até o MP (MEP ou MIP) de destino no interior de uma MA. O endereço MAC de destino de uma LTM PDU é um endereço *multicast*, como será visto adiante neste item.

O protocolo *Linktrace,* além de possibilitar o descobrimento de caminhos, é também utilizado para o isolamento de falhas.

Embora destinada a um endereço MAC alvo individual (indentificado no interior da LTM PDU), uma LTM é enviada pelo MEP de origem utilizando um endereço MAC *multicast* específico para o protocolo *Linktrace*, que é função do nível de MD utilizado.

A transmissão da LTM é repetida a cada MIP da MA utilizada ao longo da *bridged* LAN, até atingir o MEP de destino (alvo). Cada MIP intermediário reenvia a LTM utilizando o seu próprio endereço MAC com endereço de origem, mas mantendo o endereço MAC do MEP de origem no interior da LTM PDU.

Cada MIP ao longo do caminho, assim como o MEP de destino, retorna uma mensagem LTR (*Linktrace Reply)*, com endereço de destino *unicast,* para o MEP de origem da mensagem LTM (*Linktrace Message*).

As LTMs são acionadas por ações administrativas. As respostas são catalogadas pelo MEP de origem das LTMs para análise pelo administrador da rede.

A Figura 8-6 apresenta um exemplo de aplicação do protocolo *Linktrace*.

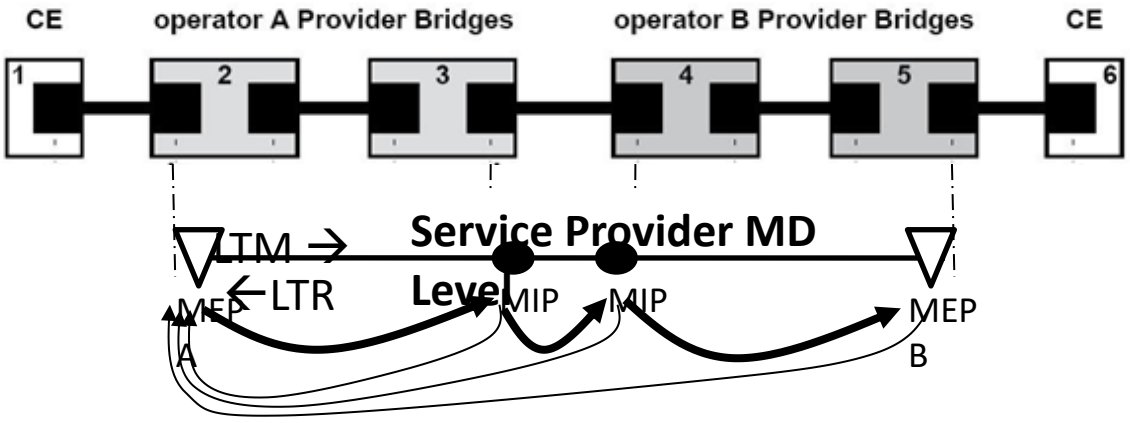


Figura 8-6 – Exemplo de aplicação do protocolo Linktrace.

8.2.1.3. Aspectos Comuns para a Codificação de CFM PDUs

A cláusula 21 do padrão IEEE 802.1Q-2014 especifica aspectos comuns para a codificação das CFM PDUs, que incluem os seguintes:

— Encapsulamento de CFM PDUs;
— Endereçamento de entidades CFM;
— Formato do cabeçalho comum de CFM PDUs;
— Tipos de TLV utilizados por CFM PDUs.

- **Encapsulamento de CFM PDUs**

As mensagens de OAM do padrão IEEE 802.1Q-2014 utilizam quadros MAC *Ethernet* convencionais, ressalvando-se o uso eventual de endereços MAC especiais atribuídos a esses tipos de mensagem.

O modo de identificação das CFM PDUS em quadros *Ethernet* depende do tipo de rede.

Para redes que utilizam o campo *Type/Length* (como redes IEEE 802.3, por exemplo, a indicação ocorre diretamente pelo código 0x8902.

Redes que requerem o uso de encapsulamento pelo LLC (como redes IEEE 802.11, por exemplo) utilizam o código *EtherType* 0x8902 via LSAP/SNAP (códigos AA-AA e OUI 00-00-00, respectivamente).

- **Endereçamento de Entidades CFM**

Existem três classes de CFM PDUs com respeito ao tipo de entidade CFM de destino:

— Aquelas endereçadas a todos os MEPs da MA (caso das CCM PDUs);

— Aquelas endereçadas ao conjunto de MPs imediatamente adjacentes ao MP transmissor, em um certo nível de MD, dentro da MA (caso das LTM PDUs);
— Aquelas endereçadas a um único e específico MP (caso das LBM PDUs, LBR PDUs e LTR PDUs).

A decisão quanto ao tratamento a ser dado a um quadro recebido por uma entidade CFM é dependente do endereço de destino desse quadro. Uma entidade CFM deve descartar qualquer quadro, a ela direcionada devido ao valor de VID e ao nível de MD, cujo endereço de destino não corresponda a um endereço MAC reconhecido por essa entidade.

Qualquer entidade CFM é configurada para reconhecer um ou mais endereços MAC individuais , ou um ou mais endereços MAC de grupo, em quadros por ela recebidos. É também configurada para utilizar um ou mais endereços MAC individuais, ou um ou mais endereços MAC de grupo, em quadros por ela transmitidas.

Os endereços MAC de destino utilizados dependem do tipo de CFM PDU e do nível de MD. Para as LBM PDUs, as LBR PDUs e LTR PDUs são transportados em quadros com endereços *unicast*.

AS CCM PDUs e as LTM PDUs, por outro lado, são transportadas em quadros com endereços *multicast* de destino, conforme a Figura 8.7.

Endereço: 01-80-C2-00-00-3*y*		
Nível de MD	**CCM PDUs**	**LTM PDUs**
	Valor de *y*	**Valor de *y***
7	7	F
6	6	E
5	5	D
4	4	C
3	3	B
2	2	A
1	1	9
0	0	8

Figura 8-7 – Endereços MAC de CCM PDUs e de LTM PDUs.

- **Formato do Cabeçalho Comum das CFM PDUs**

O formato do cabeçalho comum das CFM PDUs encontra-se representado na Figura 8-8.

	Octet
MD Level	1 (high-order 3 bits)
Version	1 (low-order 5 bits)
OpCode	2
Flags	3
First TLV Offset	4
Varies with value of OpCode	5
End TLV (0)	First TLV Offset + 5

Figura 8-8 - Formato do cabeçalho comum das CFM PDUs.

Para o campo MD Level o leitor deve se reportar à Figura 8-2 e à Figura 8-3 anteriores deste capítulo.

A versão atual do CFM PDU é a versão zero, sendo que esse valor é ignorado na recepção.

O campo *Opcode* especifica o formato e o significado do restante da CFM PDU. Os valores para esse campo encontram-se relacionados na Figura 8-9.

CFM PDU or organization	OpCode range
Reserved for IEEE 802.1	0
Continuity Check Message (CCM)	1
Loopback Reply (LBR)	2
Loopback Message (LBM)	3
Linktrace Reply (LTR)	4
Linktrace Message (LTM)	5
Reserved for IEEE 802.1	6 – 31
Defined by ITU-T Y.1731	32 – 63
Reserved for IEEE 802.1.	64 – 255

Figura 8-9 – Valores e faixas do campo OpCode.

A utilização do campo *Flags* é definida separadamente para cada tipo de mensagem identificado pelo respectivo *OpCode.*

O campo *First TLV Offset* indica a posição do primeiro TLV da CFM PDU a partir do início do primeiro octeto após o próprio campo *First TLV Offset*, até o início desse primeiro TLV na CFM PDU.

O campo *End TLV (0)* finaliza todas as CFM PDUs. Nesse TLV, o tipo é igual a zero e os campos Comprimento (*Lenght*) e Valor (*Value*) não se encontram presentes.

- **Tipos de TLV Utilizados**

Os tipos de TLV utilizados pelas CFM PDUs, com os respectivos códigos, encontram-se relacionados na Figura 8-10.

TLV or organization	Type field
End TLV	0
Sender ID TLV	1
Port Status TLV	2
Data TLV	3
Interface Status TLV	4
Reply Ingress TLV	5
Reply Egress TLV	6
LTM Egress Identifier TLV	7
LTR Egress Identifier TLV	8
Reserved for IEEE 802.1	9 – 30
Organization-Specific TLV	31
Defined by ITU-T Y.1731	32 – 63
Reserved for IEEE 802.1	64 – 255

Figura 8-10 - TLVs e respectivos códigos.

8.2.1.4. Formato de CFM CCM PDUs

A Figura 8-11 apresenta o formato das CFM CCM PDUs.

	Octet
Common CFM Header	1 – 4
Sequence Number	5 – 8
Maintenance association End Point Identifier	9 – 10
Maintenance Association Identifier (MAID)	11 – 58
Defined by ITU-T Y.1731	59 – 74
Reserved for definition in future versions of the protocol	
Optional CCM TLVs	First TLV Offset + 5
End TLV (0)	First TLV Offset + 5, if no Optional CCM TLVs are present

Figura 8-11- Formato de CFM CCM PDUs.

Como foi dito anteriormente, o campo *Flags* do cabeçalho comum é utilizado de forma específica para cada *OpCode*, ou seja, para cada tipo de mensagem CFM. Para o caso particular das CFM CCM PDUs, o campo *Flags* possui os seguintes significados:

- Primeiro *MSB bit*: representa o valor do RDI (*Remote Deffect Indicator*) *bit* , que, quando setado para 1, informa o MEP par a ocorrência de falhas no sentido das mensagens enviadas por esse MEP par;
- Segundo *MSB bit*: representa o *Traffic bit,* utilizado apenas em PBB-TE MAs;
- Últimos três *LSP bits*: constituem o campo de intervalo CCM (*CCM Interval*), que é condificado conforme a tabela 21-16 da Cláusula 21 do padrão IEEE 802.1Q-2014;
- Demais bits encontram-se reservados.

No corpo das CFM CCM PDUs, vale destacar o subcampo *Maintance Association Identifier* (MAID), já descrito anteriormente neste item. Ressalva-se, contudo, que os 48 octetos desse subcampo apresentam duas diferentes estruturas, dependendo da existência ou não do *Maintenance Domains Name* (*MD Name*) na composição do MAID.

O formato do campo MAID na hipótese da presença do *Maintenance Domain Name* (valor do primeiro octeto diferente de 1), encontra-se representado na Figura 8-12.

	Octet
Maintenance Domain Name Format (not 1)	1
Maintenance Domain Name Length	2
Maintenance Domain Name	3 – (mdnl [a] + 2)
Short MA Name Format	(mdnl + 3)
Short MA Name Length	(mdnl + 4)
Short MA Name	(mdnl + 5) – (mdnl + smanl [b] + 4)
0 pad, if necessary	(mdnl + smanl + 5) – 48

[a] Maintenance Domain Name Length.
[b] Short MA Name Length.

Figura 8-12 – Formato do campo MAID (MD Name presente).

Na hipótese de ausência do MD *Name* (valor do primeiro octeto igual a 1), o campo MAID assume o formato da Figura 8-13.

	Octet
Maintenance Domain Name Format (1)	1
Short MA Name Format	2
Short MA Name Length	3
Short MA Name	4 – (smanl [a] + 3)
0 pad, if necessary	(smanl + 4) – 48

[a] Short MA Name Length

Figura 8-13 – Formato do campo MAID (MD Name ausente).

Dentre os CCM TLVs opcionais das CFM CCM PDUs poderia constar o *Organization Specific TLV*. Além disso deveriam constar os seguintes TLVs:

— *Sender ID TLV;*
— *Port Status TLV;*
— *Interface Status TLV.*

8.2.1.5. Formato das CFM LBM PDUs e das CFM LBR PDUs

A Figura 8-14 apresenta o formato comum às CFM LBM PDUs e às CFM LBR PDUs.

	Octet
Common CFM Header	1 – 4
Loopback Transaction Identifier	5 – 8
Reserved for definition in future versions of the protocol	
Optional LBM/LBR TLVs	First TLV Offset + 5
End TLV (0)	First TLV Offset + 5, if no Optional LBM/LBR TLVs are present

Figura 8-14 - Formato das CFM LBM PDUs e das CFM LBR PDUs.

Ressalta-se a presença, nessa figura, do campo *Loopback Transaction Identifier*, necessário à associação entre uma LBM PDU e a respectiva LBR PDU, que ocorre pela replicação do valor desse campo contido na LBM PDU.

8.2.1.6. Formatos das CFM LTM PDUs e das CFM LTR PDUs

A Figura 8-15 apresenta o formato das CFM LTM PDUs.

	Octet
Common CFM Header	1 – 4
LTM Transaction Identifier	5 – 8
LTM TTL	9
Original MAC Address	10 – 15
Target MAC Address	16 – 21
Reserved for definition in future versions of the protocol	
Additional LTM TLVs	First TLV Offset + 5
End TLV (0)	First TLV Offset + 5, if no Additional LTM TLVs are present

Figura 8-15 - Formato das CFM LTMs.

O campo *Original MAC Address* representa um endereço MAC do MEP que originou a LTM PDU. Esse endereço é normalmente diferente do endereço MAC de origem, indicado no quadro Ethernet que contém a LTM PDU, pois esse endereço MAC de origem representa o endereço MAC de cada MP que retransmite a LTM PDU ao longo de seu trajeto.

Embora o endereço MAC de destino do quadro Ethernet que contém uma LTM PDU seja um endereço *multicast*, o valor do campo *Target MAC Address* pode conter qualquer valor de endereço MAC individual.

Finalmente, a Figura 8-16 apresenta o formato das CFM LTR PDUs.

	Octet
Common CFM Header	1 – 4
LTR Transaction Identifier	5 – 8
Reply TTL	9
Relay Action	10
Reserved for definition in future versions of the protocol	
Additional LTR TLVs	First TLV Offset + 5
End TLV (0)	First TLV Offset + 5, if no Additional LTR TLVs are present

Figura 8-16 - Formato das CFM LTR PDUs.

8.2.2. **DDCMF (Data-Driven and Data-Dependent CFM)**

A Cláusula 29 do padrão IEEE 802.1Q-2014 define uma extensão do SOAM CFM descrito nos subitens anteriores deste item, referida como DDCFM (*Data-Driven and Data-Dependent Conectivity Fault Management*). Essa cláusula descreve as funções do DDCFM e como o DDCFM pode ser operado e gerenciado.

As falhas *data-dependent* e as falhas *data-driven* (em conjunto referidos como DDFs) afetam exclusivamente determinados quadros ou determinadas sequências de quadros, não podendo então ser detectadas pelo CFM básico.

As falhas *data-dependent* resultam na perda repetitiva, ou no encaminhamento errôneo repetitivo, exclusivamente dos quadros de dados afetados. As falhas *data-driven*, mais complexas, ocorrem quando a perda de determinados quadros de dados é causada pela presença ou pela ausência de determinados outros quadros.

A forma básica para diagnosticar DDFs é isolá-las em segmentos reduzidos da rede, tais como uma porta de *bridge*, uma porta membro de um *LAG (Link Aggregation Group*) em uma porta de *bridge*, ou mesmo em uma LAN. Determina-se, então, se determinados quadros podem atravessar esses segmentos corretamente.

Uma DDF pode não ser aparente na ausência de tráfego real. Assim, o DDCFM é definido de forma a realizar diagnósticos enquanto o tráfego real é cursado.

Foram especificadas duas técnicas para diagnosticar e isolar DDFs:

— *Forward Path Test* (FPT), também referido como opção *Continue*, que se processa de forma transparente para o tráfego real;

— *Return Path Test* (RPT), que envolve o envelopamento dos quadros de dados.

8.2.3. **SOAM: Recomendação ITU-T G.8013/Y.1731**

A recomendação ITU-T G.8013/Y.1731 (*OAM Functions and Mechanisms for Ethernet Networks*), emitida em 11/2013, define as seguintes funções de OAM para o plano de dados de redes Ethernet;

— Funções de OAM para Gerenciamento de Falhas (*Fault Management*);
— Funções de OAM para Monitoração de Desempenho (*Perfomance Monitoring*).

Essas funções dizem respeito aos mecanismos requeridos para operar e manter os aspectos de rede e de serviço de redes Ethernet.

Tais mecanismos são aplicáveis, conforme a atual versão da recomendação ITU-T G.8013/Y.1731, a conexões *Ethernet* ponto a ponto e a conectividade *Ethernet* multiponto. As funções referentes a gerenciamento de falhas da recomendação ITU-T G.8013/Y.1731 são em grande parte alinhadas com os termos do padrão IEEE 802.1Q-2014.

A funcionalidade especificada por essa recomendação tem como fundamento os requisitos e os princípios estabelecidos na recomendação ITU-T Y.1730 (*Requirements for OAM Functions in Ethernet-Based Networks and EthernetServices*). A sua base arquitetural no que diz respeito ao suporte a conexões ponto a ponto e a conectividade multiponto encontra-se delineada na recomendação ITU-T G.8010/Y.1306 (*Architecture of Ethernet Layer Networks*).

8.2.3.1. Diferenças de Terminologia entre o Padrão IEEE 802.1Q-2014 e a Recomendação ITU-T G.8013/Y.1731

Existem algumas diferenças de terminologia entre os dois padrões, o que se encontra resumido da Figura 8-17.

IEEE Std 802.1ag-2007	ITU-T Y.1731 (2006)
Maintenance Association (MA)	Maintenance Entity Group (MEG)
Maintenance Association Identifier (MAID)	Maintenance Entity Group Identifier (MEGID)
Maintenance Domain (MD)	(No such construct present)
Maintenance Domain Level (MD Level)	Maintenance Entity Group Level (MEG Level)

Figura 8-17 - Diferenças de terminologia entre IEEE e ITU-T.

Observa-se que o nível de MD (*MD Level*) do padrão IEEE 802.1Q-2014 passa a ser referido, na recomendação G.8013/Y.1731, como *MEG Level* (MEL).

Adicionalmente, registram-se as seguintes diferenças entre esses padrões:

— A recomendação ITU-T G.8013/Y.1731 inclui um número adicional de *OpCodes*;
— A recomendação ITU-T G.8013/Y.1731 não considera um MIP como sendo constituído por duas MHFs como no padrão IEEE 802.1Q-2014;
— A recomendação ITU-T G.8013/Y.1731 não inclui o *LTM Egress Identifier TLV* nem o *LTR Egress Identifier TLV*, sendo que o padrão IEEE 802.1Q-2014 definiu medidas para possibilitar a compatibilidade entre implementações com os dois padrões;
— Existem diferenças no tratamento do campo TTL entre os dois padrões;
— O CFM conforme o padrão IEEE 802.1Q-2014 inclui dois bits extras no campo *Flags* não utilizados na recomendação ITU-T G.8013/Y.1731.

Finalmente, ressalta-se que a recomendação G.8013/Y.1731 especifica que as LBM PDUs podem possuir também endereços MAC de grupo como destino, enquanto o padrão IEEE 802.1Q-2014 limita-se a endereços MAC individuais como destino para essas mensagens.

Para compensar essa limitação, o padrão IEEE 802.1Q-2014 implementou a função *MP Loopback Responder*, que possibilita o envio de respostas a mensagens LBM enviadas com o endereço de grupo de CCM PDUs.

8.2.3.2. Conceitos Básicos

A recomendação ITU-T G.8013/Y.1731 apresenta os seguintes conceitos básicos:

— Entidade de Manutenção (ME);
— Grupo de MEs (MEG);
— Ponto Final de MEG (MEP);
— Ponto Intermediário de MEG (MIP);
— Ponto de Condicionamento de Tráfego (TrCP);
— Nível de MEG.

Grande parte desses conceitos são os mesmos que aqueles aplicáveis ao padrão IEEE 802.1Q-2014 já apresentados anteriormente neste capítulo, ressalvando-se as diferenças de nomenclatura mencionadas no subitem anterior.

É necessário, contudo, a apresentação do conceito de Ponto de Condicionamento de Tráfego (TrCP) e das particularidades do conceito de Nível de MEG.

- **Ponto de Condicionamento de Tráfego (TrCP)**

São pontos em um domínio de fluxo ETH capazes de realizar funções de condicionamento nos respectivos fluxos ETH, conforme especificação na recomendação ITU-T G.8010/Y.1306. Tais funções realizam os seguintes processos em quadros *Ethernet* desses fluxos ETH:

— Classificação: esse processo classifica cada quadro *Ethernet*;
— Medição: esse processo registra cada quadro *Ethernet* dentro de sua classe com o propósito de determinar a sua elegibilidade e a sua eventual ordem de procedência para descarte;
— Policiamento: esse processo policia os quadros de dados de acordo com o registro resultante de sua medição. Existem apenas duas ações, que são a transmissão ou o descarte de quadros.

- **Níveis de MEG**

Pode ocorrer a definição de MEGs hierarquizados (*nested*) em configurações de redes *Ethernet*.

É possível a identificação dos diferentes níveis hierárquicos dos MEGs definidos com base no próprio encapsulamento da ETH-*Layer*.

Quando isso não é aplicável, é necessária a atribuição de um nível de MEG a cada MEG hierarquizado, para explicitar o seu nível hierárquico.

O nível de MEG de cada MEG hierarquizado consta de todas as respectivas OAM PDUs. Com esse propósito, foi especificado um campo MEL (MEG *Level*) para essas PDUs, com três bits, do que resulta a possibilidade de atribuição de até 8 valores de nível de MEG.

Os 8 valores de nível de MEG podem ser compartilhados entre fluxos de OAM de usuários, de provedores de serviço e de operadores de serviço de rede.

A distribuição *default* desses valores é a seguinte:

— Três níveis para usuários: níveis 7,6 e 5;
— Dois níveis para provedores de serviço: níveis 4 e 3;
— Três níveis para operadores de serviço: níveis 2, 1, e zero.

Essa distribuição *default* pode ser alterada mediante acordo entre as partes.

Quadros de OAM referentes a um MEG de um determinado nível podem ser transportados transparentemente através de MEGs com níveis de MEG inferiores.

8.2.3.3. Formato do Cabeçalho Comum das OAM PDUs

O formato do cabeçalho comum a todas as OAM (SOAM) PDUs definidas na Recomendação ITU-T G.8013/Y.1731 encontra-se na Figura 8-18.

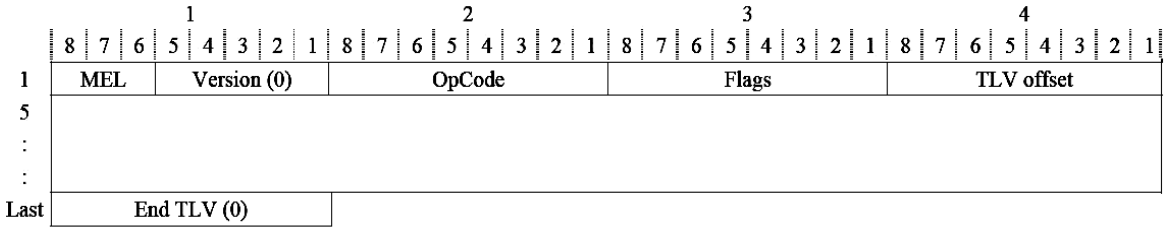


Figura 8-18 - Formato do cabeçalho comum das OAM PDUs.

Como se observa, essa figura corresponde à Figura 8-8 anterior referente ao cabeçalho comum das CFM PDUs de acordo com o padrão IEEE 802.1Q-2014.

Os campos do cabeçalho comum das OAM PDUs têm os seguintes significados:

— Campo MEL (MEG *level*): representa o nível de MEG correspondente à PDU;
— Campo *Version*: representa a versão do protocolo de OAM utilizado, sendo seu valor atual igual a zero;
— Campo *OpCode*: esse campo identifica cada OAM PDU, conforme a Figura 8-19 a seguir;
— Campo *Flags*: formado por 8 bits, cujos significados dependem do tipo de OAM PDU;
— Campo TLV *Offset*: esse campo indica a posição do primeiro TLV da OAM PDU com relação à posição desse próprio campo;
— Campo *End* TLV (0): representa um octeto com valores de bit zerados.

A Figura 8-19 apresenta a relação dos tipos de OAM PDU com os correspondentes valores de *OpCode*.

OpCode value	OAM PDU type	OpCode relevance for MEPs/MIPs
OpCodes common with IEEE 802.1		
1	CCM	MEPs
3	LBM	MEPs and MIPs (connectivity verification)
2	LBR	MEPs and MIPs (connectivity verification)
5	LTM	MEPs and MIPs
4	LTR	MEPs and MIPs
0, 6-31, 64-255	Reserved (Note 1)	
OpCodes specific to this Recommendation		
33	AIS	MEPs
35	LCK	MEPs
37	TST	MEPs
39	Linear APS	MEPs
40	Ring APS	MEPs
41	MCC	MEPs
43	LMM	MEPs
42	LMR	MEPs
45	1DM	MEPs
47	DMM	MEPs
46	DMR	MEPs
49	EXM	Outside the scope of this Recommendation
48	EXR	Outside the scope of this Recommendation
51	VSM	Outside the scope of this Recommendation
50	VSR	Outside the scope of this Recommendation
32, 34, 36, 38, 44, 52-63	Reserved (Note 2)	
NOTE 1 – Reserved for definition by IEEE 802.1. NOTE 2 – Reserved for future standardization by ITU-T.		

Figura 8-19 - OAM PDUs com os correspondentes valores de OpCode.

Como se verifica nessa figura, os 5 primeiros valores de *OpCode* são comuns com aqueles do padrão IEEE 802.1Q-2014.

No interior das OAM PDUs podem ser transportados os diferentes tipos de TLV, aplicados em função do tipo de OAM PDU.

Os tipos de TLV definidos para OAM PDUs encontram-se relacionados na Figura 8-20.

Type value	TLV name
Types common with IEEE 802.1	
0	End TLV
3	Data TLV
5	Reply ingress TLV
6	Reply egress TLV
7	LTM egress identifier TLV
8	LTR egress identifier TLV
2, 4, 9-31, 64-255	Reserved (Note 1)
Types specific to this Recommendation	
32	Test TLV
33-63	Reserved (Note 2)
NOTE 1 – Reserved for definition by IEEE 802.1. NOTE 2 – Reserved for future standardization by ITU-T.	

Figura 8-20 - Tipos de TLV de OAM PDUs.

Como se observa nessa figura, o único TLV acrescentado pela recomendação ITU-T G.8013/Y.1731 ao padrão IEEE 802.1Q-2014 é o *Test TLV* (tipo 32).

8.2.4. SOAM CFM: Recomendação ITU-T G.8013/Y1731

A recomendação ITU-T G.8013/Y.1731 considera as seguintes funções de OAM para gerenciamento de falhas, ou seja, para SOAM CFM:

- *Ethernet Continuity Check* (ETH-CC);
- *Ethernet Loopback* (ETH-LB);
- *Ethernet Link Trace* (ETH-LT);
- *Ethernet Alarm Indication Signal* (ETH-AIS);
- *Ethernet Remote Defect Indication*(ETH-RDI);
- *Ethernet Locked Signal* (ETH-LCK);
- *Ethernet Test Signal* (ETH-*Test*);
- *Ethernet Automatic Protection Switching* (ETH-APS);
- *Ethernet Maintemance Communication Channel* (ETH-MCC);
- *Ethernet Experimental OAM* (ETH-EXP);
- *Ethernet Vendor-Specific* OAM (ETH-VSP);
- *Ethernet Client Signal Fail* (ETH-CSF).

As primeiras 7 dessas funções, pela sua similaridade com as correspondentes funções definidas no padrão IEEE 802.1Q-2014, serão abordadas parcialmente neste item.

A função ETH-APS, que diz respeito à comutação de proteção, será abordada adiante em outro item deste parágrafo.

8.2.4.1. Funções ETH-CC, ETH-LB, ETH-LT e ETH-RDI

As funções ETH-CC, ETH-LB, ETH-LT e ETH-RDI foram apresentadas quando da abordagem do padrão IEEE 802.1Q-2014.

As LBM PDUs, LBR PDUs, LTM PDUs e LTR PDUs possuem os mesmos formatos que as correspondentes PDUs no padrão IEEE 802.1Q-2014, e que, portanto, não serão apresentados neste item.

Entretanto, será apresentada, a seguir, a formatação das CFM CCM PDUs, como ilustração e por apresentar algumas particularidades próprias da recomendação ITU-T G.8013/Y.1731.

As CCM PDUs (CFM CCM PDUs) são utilizadas como suporte à função ETH-CC, à função ETH-RDI e à função *Dual-Ended* ETH-LM.

Na Figura 8-21 está representado o formato das CCM PDUs conforme a recomendação ITU-T G.8013/Y.1731.

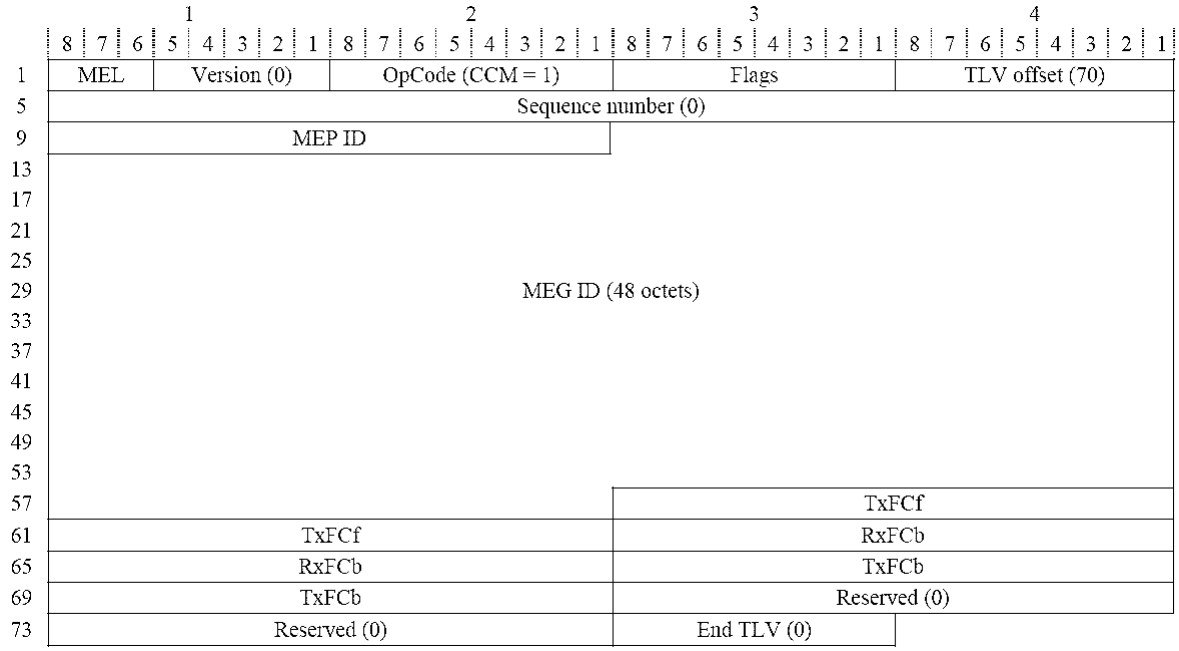


Figura 8-21 - Formato das CCM PDUs.

Nesse formato ressaltam-se os seguintes pontos:

— O campo *Flags* indica, pelo primeiro *MSB bit* (bit 8), o uso do RDI, e, pelos bits 3 a 1, o valor do período de transmissão adotado;
— O campo TLV *Offset* é setado para o valor 70.

Os valores de períodos de retransmissão, com os respectivos códigos binários, encontram-se representados na Figura 8-22.

Flags[3:1]	Period value	Comments
000	Invalid value	Invalid value for CCM PDUs
001	3.33 ms	300 frames per second
010	10 ms	100 frames per second
011	100 ms	10 frames per second
100	1 s	1 frame per second
101	10 s	6 frames per minute
110	1 min	1 frame per minute
111	10 min	6 frames per hour

Figura 8-22 - Períodos de transmissão das CCM PDUs.

Ao contrário das CCM PDUs no padrão IEEE 802.1Q-2014, as CCM PDUs na recomendação ITU-T G.8013/Y.1731 não utilizam TLVs à exceção do *End*TLV (0).

8.2.4.2. Funções ETH-AIS e ETH-LKC

A função ETH-AIS (*Ethernet Alarm Indication Signal*) é utilizada para suprimir alarmes subseqüentes à detecção de condição de defeito na camada servidora. Tais alarmes são acionados em consequência da detecção de perda de continuidade de sinal.

Um MEP, ao detectar uma condição de defeito, pode transmitir quadros AIS no sentido oposto, para o seu MEP par. A periodicidade de transmissão recomendada é igual a 1 segundo, e essa transmissão deve se repetir até a remoção da condição de defeito.

A função ETH-LCK (*Ethernet Locked Signal*) é utilizada para comunicar o bloqueio administrativo de um MEP e a consequente interrupção de tráfego de dados no sentido do MEP que receberia esse tráfego. Essa função permite aos MEPs diferenciar entre uma condição de defeito e um bloqueio deliberado.

Um exemplo de aplicação que poderia requerer o bloqueio administrativo de um MEP é a realização de um teste (função *ETH-Test*) em um momento desejado.

O formato de uma AIS PDU e de uma LCK consiste exclusivamente do cabeçalho comum das SOAM PDUs. Cada uma dessas funções é identificada pelo valor de *OpCode* desse cabeçalho comum, que é igual a 33 para as AIS PDUs e igual a 35 para as LCK PDUs.

Para as funções ETH-AIS os três últimos LSB *bits* do campo *Flags* indica o período de transmissão de quadros AIS em condição de defeito. O valor 100 indica o período de 1 segundo, enquanto o valor 110 indica o período de 1 minuto.

8.2.4.3. Função ETH-Test

A função *ETH-Test* é utilizada para a realização de testes unidirecionais (*one-way*) em demanda, tanto em serviço (*in-service*) quanto fora de serviço (*out-of-service*). Esses testes objetivam a verificação de taxas de vazão (*throughput*), perda de quadros, erros de bits, etc.

Quando configurado com esse propósito, um MEP insere quadros com informações *ETH-Test*, especificando taxas de vazão de tráfego com tamanhos de quadros e padrões de transmissão.

Quando realizada em serviço, a função *ETH-Test* não requer a interrupção do tráfego de dados, utilizando-se uma parte limitada da banda para o serviço em curso.

Na hipótese de *ETH-Test* fora de serviço, o tráfego de dados é previamente interrompido mediante a transmissão de quadros LCK.

O formato de uma TST (*test*) PDU encontra-se na Figura 8-23.

1								2								3								4							
8	7	6	5	4	3	2	1	8	7	6	5	4	3	2	1	8	7	6	5	4	3	2	1	8	7	6	5	4	3	2	1
MEL			Version (0)					OpCode (TST=37)								Flags (0)								TLV Offset (4)							
Sequence Number																															
[Test TLV]																															
																								End TLV (0)							

Figura 8-23 – Formato de uma TST PDU.

Como se observa nessa figura, o valor de *OpCode* para a TST PDU é igual a 37. O *TLV Offset* é fixado no valor 4 (octetos).

8.2.4.4. Funções ETH-APS, ETH-MCC, ETH-EXP, ETH-VSP e ETH-CSF

A função ETH-APS, especificada na recomendação ITU-T G.8031/Y.1342, será abordada posteriormente neste capítulo, no item relativo a essa recomendação.

A função ETH-MCC (*Ethernet Maintenance Communication Channel*) provê um canal de comunicação entre um par de MEPs. Essa função pode ser utilizada para realizar gerenciamento remoto.

Um MEP pode enviar um quadro (ou mensagem) MCC para o seu MEP par para fins de manutenção remota, contendo uma solicitação, uma resposta ou uma notificação. O MEP receptor repassa a informação ETH-MCC para o agente de gerenciamento, que pode ou não enviar uma resposta. O *OpCode* das mensagens MCC utiliza o valor 41.

Os MIPs são transparentes à função ETH-MCC, e não requerem, portanto, qualquer informação de configuração quanto a essa função.

A função ETH-EXP (*Ethernet Experimental*) pode ser utilizada em um domínio administrativo de forma temporária. Para isso, foram definidas as EXM PDUs (*Experimental OAM Message PDUs*), com o *OpCode* igual a 49, e as EXR PDUs (*Experimental OAM Reply PDUs*), com o *OpCode* igual a 48.

A função ETH-VSP (*Ethernet Vendor – Specific*) é também especificada pela recomendação ITU-T G.8013/Y.1731. As VSM PDUs (*Vendor Specific Message PDUs*) utilizam o *OpCode* igual a 51, enquanto as VSR PDUs (*Vendor Specific Reply PDUs*) utilizam o *OpCode* igual a 50.

A função ETH-CSF (*Ethernet Client Signal Fail*) é utilizada por um MEP com a finalidade propagar para um MEP par a detecção de um evento de falha ou de defeito em um sinal Ethernet de cliente, quando o próprio cliente não suporta mecanismos apropriados de detecção e de propagação de defeitos ou falhas. Em outras palavras, quando o cliente não dispões de mecanismos tais como ETH-CC ou ETH-AIS.

A função ETH-CSF é aplicável apenas em *links* Ethernet de transporte ponto a ponto.

8.2.5. **SOAM PM: Recomendação ITU-T G.8013/Y.1731**

As funções de OAM para PM (*Performance Monitoring*), definidas na recomendação ITU-T G.8013/Y.1731, possibilitam a medição de diferentes parâmetros de desempenho. Foram definidos funções e métodos de medição para conexões ponto a ponto e para conectividade multiponto.

Foram consideradas as seguintes funções para monitoramento de desempenho (PM):

- Medições de Perdas de Quadros (*Frame Loss Measurement*), referidas como ETH-LM;
- Medições de Retardos de Quadros *(Frame Delay Measurement*), referidas como ETH-DM;
- Medições de Vazões de Tráfego *(Throughput Measurement*);
- Medições de Perdas Sintéticas (*Synthetic Loss Measurement*), referidas como ETH-SLM.

Para a realização dessas funções, foram especificados os seguintes parâmetros a serem medidos:

- Razão de Perda de Quadros (*Frame Loss Ratio*), referida como FLR;
- Retardo de Quadros (*Frame Delay),* referido como FD;
- Variação de Retardo de Quadros (*Frame Delay Variation),* referido como FDV.
- Disponibilidade (*Availability*);
- Taxa de Vazão de Tráfego (*Throughput*), com base na RFC 2544.

A definição de disponibilidade de redes Ethernet encontra-se na recomendação ITU-T Y.1563 (*Ethernet Frame Transfer and Availability Performance*). Embora os mecanismos definidos nessa recomendação possam contribuir para a implementação de métodos para a medição de disponibilidade, a recomendação ITU-T G.8013/Y.1731 postergou para estudos futuros o detalhamento desses métodos.

8.2.5.1. Medição de Perda de Quadros

A Razão de Perda de Quadros (*Frame Loss Ratio*) é a razão entre o número de quadros de dados não entregues e o número total de quadros enviados, medida em um intervalo de tempo T e expressa sob a forma de porcentagem.

A Razão de Perda de Quadros pode ser medida usando quadros de serviço ou quadros sintéticos, pertencentes a uma mesma CoS.

O uso de quadros de dados para medição de perda de quadros, referida como ETH-LM, restringe-se a conectividade ETH ponto a ponto.

A medição de perda de quadros sintéticos é referida como ETH-SLM (ETH *Synthetic Loss Measurement)*. O uso de quadros sintéticos pode aplicar-se também no caso de conetividade ETH multiponto.

A medição do envio e da entrega de quadros ocorrem, respectivamente, no ponto de fluxo ETH (MEP) de ingresso e no ponto de fluxo ETH (MEP) de egresso.

8.2.5.2. Função ETH-LM

A função ETH-LM (*Ethernet Loss Measurement*) é utilizada para coletar valores de medição de envio e de entrega de quadros de dados, em contadores respectivamente situados no MEP de ingresso e no MEP de egresso.

Um MEP mantém os seguintes dois tipos de contadores locais para cada MEP par e para cada classe de prioridade:

— TxFC1: contador para o número de quadros de dados transmitidos para o MEP par;
— RxFC1: contador para o número quadros de dados recebidos do MEP par.

Devem ser especificamente configuradas em um MEP as seguintes informações para suporte à função ETH-LM:

— Nível de MEG;
— Período de transmissão ETH-LM;
— Prioridade de transmissão de quadros com informações ETH-LM (configuração por operação);
— Elegibilidade para descarte de quadros contendo informações ETH-LM (sempre inelegíveis para descarte, o que torna não obrigatória essa configuração).

Os MIPs são transparentes à função ETH-LM.

A função ETH-LM pode se realizar de dois modos:

— Modo *Dual-Ended* ETH-LM;
— Modo *Single-Ended* ETH-LM.

A função ETH-LM utiliza os elementos de informação TxFCf, RxFCb e TxFCb. A semântica desses elementos de informação depende do modo de realização da função ETH-LM, sendo os seus valores utilizados para a medição das razões de perda de quadros de dados.

As medições de perdas de quadros podem se realizar separadamente para cada classe de serviço (CoS) correspondente aos quadros utilizados nas medições.

8.2.5.3. Dual-Ended ETH-LM

O modo *Dual-Ended* ETH-LM é utilizado como um mecanismo de OAM proativo destinado à medição de perda de quadros, sendo também utilizado para gerenciamento de falhas.

Nesse modo, cada MEP envia periodicamente CCM PDUs. O envio é feito para o MEP par de um MEP em uma ME ponto a ponto.

Quando configurado para ETH-LM proativo, um MEP envia periodicamente CCM PDUs contendo os elementos de informação TxFCf, RxFCb e TxFCb.

Sendo as CCM PDUs transmitidas em ambos os sentidos, os dois MEPs (supostamente configurados para ETH-LM proativo) realizam medições de perdas de quadros *near-end* e *far-end* utilizando, para isso, os valores TxFCf(*tc*), RxFCb(*tc*) e TxFCb(*tc*) de uma CCM PDU recebida e o valor do contador local RxFC1 no momento dessa recepção (momento *tc*).

Esse MEP receptor utiliza também os valores TxFCf(*tp*), RxFCb(*tp*) e TxFCb(*tp*) da CCM PDU anterior e o valor do contador local RxFC1(*tp*) no momento de recepção da CCM PDU anterior (momento *tp*).

As perdas de quadros serão obtidas, então, pelas seguintes fórmulas:

— Perda de quadros na extremidade distante (*far-end*): [TxFCb(*tc*)-TxFCb(*tp*)] – [RxFCb(*tc*) – RxFCb(*tp*)];
— Perda de quadros na extremidade próxima (*near-end*): [TxFCf(*tc*) – TxFCf(*tp*)] – [RxFC1(*tc*) – RxFC1(*tp*)].

8.2.5.4. Single-Ended ETH-LM

No modo *Single-Moded*, a função ETH-LM utiliza dois tipos de LM PDU:

— LMM (*Loss Measurement Message*) PDUs;
— LMR (*Loss Measurement Reply*) PDUs.

No modo *Single-Ended* ETH-LM são realizadas medições nos dois sentidos, sendo que as medições ocorrem normalmente no MEP que envia as LMM PDUs.

O modo *Single-Ended* ETH-LM é utilizado para OAM proativo ou para OAM em demanda, por meio de LMM PDUs e de LMR PDUs. Um MEP configurado para operar nesse modo envia LMM PDUs periodicamente.

A distinção entre OAM proativo e OAM em demanda realiza-se com base no valor de um campo *flag* existente nas LM PDUs.

- **LMM PDUs**

As LMM (*Loss Measurement Message*) PDUs são utilizadas para suportar requisições de medição de razões de perdas de quadros com uma única terminação (*Single-Ended*), de forma proativa ou em demanda.

Uma LMM PDU transporta o elemento de informação TxFCf, com quatro octetos, que contém o valor do contador TxFC1 no momento da transmissão da LMM PDU pelo MEP de origem.

O formato de uma LMM PDU encontra-se na figura 8-24.

1								2								3								4							
8	7	6	5	4	3	2	1	8	7	6	5	4	3	2	1	8	7	6	5	4	3	2	1	8	7	6	5	4	3	2	1
MEL			Version (1)					OpCode (LMM=43)								Flags								TLV Offset (12)							
TxFCf																															
Reserved for RxFCf in LMR																															
Reserved for TxFCb in LMR																															
End TLV (0)																															

Figura 8-24 – Formato de uma LMM PDU.

Como se verifica nessa figura, o valor de *OpCode* utilizado é igual a 43, enquanto o campo TLV tem o valor fixo de 12 (octetos).

O campo *Flags*, em seu *LSB bit*, representa um *Type bit*, utilizado para indicar se a operação é proativa (*Type bit* setado para 1) ou se é em demanda (*Type bit* setado para zero).

- **LMR PDUs**

As LMR PDUs são as respostas às solicitações de medição de razões de perdas de quadros feitas por meio de LMM PDUs

A Figura 8-25 representa o formato de uma LMR PDU.

1								2								3								4							
8	7	6	5	4	3	2	1	8	7	6	5	4	3	2	1	8	7	6	5	4	3	2	1	8	7	6	5	4	3	2	1
MEL			Version					OpCode (LMR=42)								Flags								TLV Offset							
TxFCf																															
RxFCf																															
TxFCb																															
End TLV (0)																															

Figura 8-25 – Formato de uma LMR PDU.

Como se observa nessa figura, o valor de *OpCode* utilizado é igual a 42, enquanto o campo *TLV Offset* tem o valor copiado da LMM PDU recebida.

O valor do campo *Flags* é também copiado da última LMM PDU recebida.

Os elementos de informação transportados em uma LMR PDU têm os seguintes significados:

— TxFCf: transporta o valor do TxFCf copiado da última LMM PDU recebida pelo MEP emisssor da LMR PDU;
— RxFCf: transporta o valor do contador RxFC1 do MEP emissor da LMR PDU no momento do recebimento da última LMM PDU;
— TxFCb: transporta o valor do contador TxFC1 do MEP transmissor da LMR PDU no momento de sua transmissão.

- **Exemplo de Uso da Função ETH-LM no modo Single-Ended**

A Figura 8-26 apresenta um exemplo de uso da função ETH-LM no modo *single-ended*.

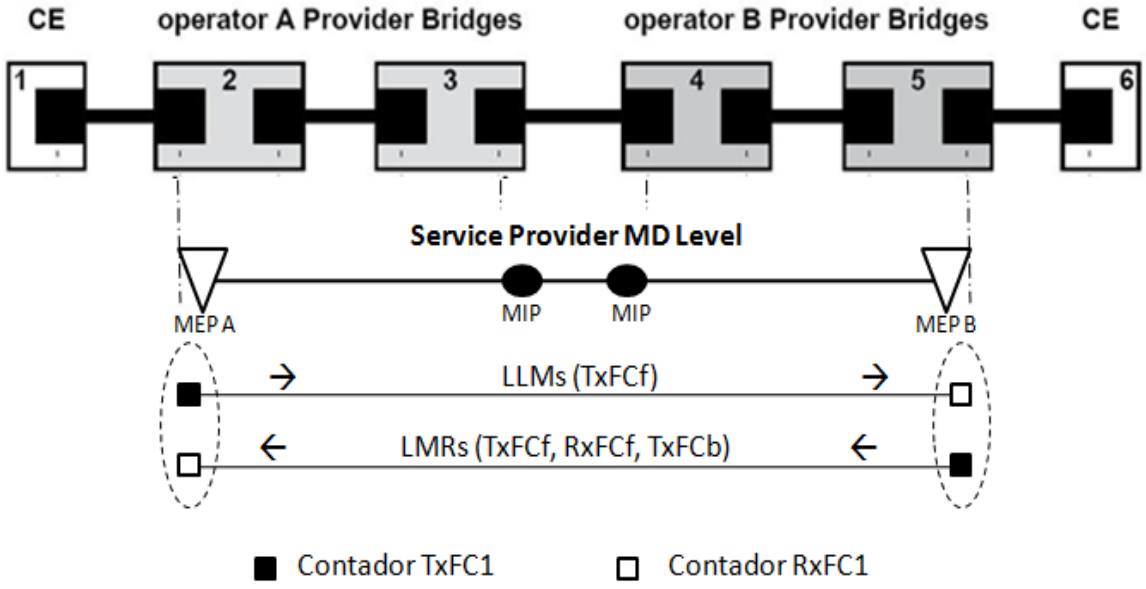


Figura 8-26 –Função ETH-LM no modo single - ended.

Nessa figura, o MEP que se encontra à esquerda representa a extremidade próxima (*near-end*), que é aquela que transmite as LLM PDUs. O seu MEP par, que se encontra à direita, representa a extremidade distante (*far-end*), que é aquela que responde com as LMR PDUs.

Considerando-se o momento de início de uma medição como *tp* e o momento de conclusão da medição como *tc*, e considerando-se também que os contadores já possuíam determinados valores no momento *tp*, as perdas de

quadros de dados são obtidas utilizando-se as seguintes fórmulas (que representam as diferenças entre quadros transmitidos e quadros recebidos):

— Na extremidade distante (*far-end*):
[TxFCf(*tc*) – TxFCf(*tp*)] – [RxFCf(*tc*) – RxFCf(*tp*)];
— Na extremidade próxima (*near end*):
[TxFCb(*tc*) – TxFCb(*tp*)] – [RxFC1(*tc*) – RxFC1(*tp*)].

8.2.5.5. Função ETH-SLM

A função ETH-SLM (*Ethernet Synthetic Loss Measurement*) é um mecanismo para medir perdas de quadros em redes Ethernet, acrescentado à recomendação ITU-T G.8013/Y.1731 em 2011, que utiliza quadros sintéticos para as medições.

A função ETH-SLM é utilizada para a realização de teste no modo proativo ou no modo em demanda.

A função ETH-SLM utiliza o modo *Dual-Ended* ou o modo *Single-Ended*.

- **Modo Dual-Ended ETH-SLM**

A função ETH-SLM no modo *Dual-Ended ETH-SLM* realiza-se pela transmissão de 1SL (*One-Way Synthetic Loss Measurement*) PDUs. Da mesma forma que no modo *Single-Ended ETH SLM*, o modo *Dual-Ended ETH-SLM* suporta OAM proativo ou OAM em demanda.

AS 1SL PDUs possuem o mesmo formato que as SLM PDUs, variando apenas os valores do campo *OpCode* (igual a 53 para as 1SL PDUs) e do campo *TLV Offset* (igual a 16 para as 1SL PDUs).

- **Modo Single-Ended ETH-SLM**

O modo *Single-Ended ETH-SLM* é utilizado para OAM proativo ou para OAM em demanda, tanto para conexão Ethernet ponto-a-ponto quanto para conectividade Ethernet multiponto. Esse modo possibilita a um MEP iniciar e reportar medições de perda na extremidade distante (*far-end*) e na extremidade próxima (*near-end*), de forma associada a um MEP par ou associada a um conjunto de MEPs pares que pertencem a um mesmo MEG.

As PDUs utilizadas no modo *Single-Ended* ETH-SLM são as SLM (*Synthetic Loss Message*) PDUs, com o valor de *OpCode* igual a 55, e as SLR (*Synthetic Loss Reply*) PDUs com o valor de *OpCode* igual 54.

As SLM PDUs são enviadas periodicamente durante um período de testes. O período de envio pode ser, por exemplo, de um segundo, de 10 segundos ou de um minuto.

Os contadores (TxFC1 e RXFC1) são os mesmos que os contadores da função ETH-LM.

As SLM PDUs, assim como as LMM PDUs, transportam o elemento de informação TxFCf.

As SLR PDUs, no entanto, diferentemente do que ocorre na função ETH-LM, transporta apenas o elemento de informação TxFCf (copiado das correspondentes SLM PDUs) e o elemento de informação TxFCb (que indica o valor do contador local RxFC1 no momento da transmissão da SLR PDU).

Observa-se que como uma SLR PDU é sempre recebida quando uma SLM PDU é gerada, o valor do contador RxFC1 no MEP respondedor é igual ao número de SLM PDUs recebidas e também igual ao número de SLR PDUs enviadas. Em outras palavras, no MEP respondedor RxFC1=TxFC1.

8.2.5.6. Medição de Retardos de Quadros (ETH-DM)

A função ETH-DM engloba a medição de retardos de quadros e a medição da variação desses retardos (*jitter* de quadros).

Essa função pode ocorrer no modo proativo ou em demanda, mediante o envio periódico, por um MEP, de PDUs com informações ETH-DM para o seu MEP par e mediante o retorno de PDUs, também com informações ETH-DM, por parte desse MEP par. Essa troca de PDUs ocorre durante uma sessão de medição proativa e/ou durante um intervalo de diagnóstico no caso de OAM em demanda.

As informações configuradas específicas requeridas por um MEP para suportar ETH-DM são as seguintes:

— Nível de MEG do MEP;
— Endereço MAC *unicast* do MEP par;
— Aplicação DM (*Delay Measurement*), que pode corresponder a proativo ou em demanda;
— *Data*, que representa um elemento de informações cujo comprimento é configurável no MEP;
— Prioridade, que é configurável por operação, sendo que um MEP pode ativar monitoramento múltiplo em diferentes tipos de CoS simultaneamente;
— Elegibilidade para descarte, que corresponde sempre a inelegível para descarte, que não é necessariamente configurada;
— Test ID, que pode ser utilizado opcionalmente se medições múltiplas são ativadas simultaneamente.

Os MIPs são transparentes à função ETH-DM.

Da mesma forma que na função ETH-LM, a função ETH-DM pode se realizar para a totalidade das DM PDUs ou para cada CoS separadamente nessas PDUs. As DM PDUs são sempre marcadas como inelegíveis para descarte.

Os quadros específicos para ETH-DM transportam o parâmetro *TxTimeStampf*, que contém o instante (*timestamp*) em que ocorreu a sua transmissão.

A variação de retardo (*jitter*) entre quadros é medida pela diferença entre os retardos desses quadros.

A função ETH-DM pode ocorrer de dois modos:

— Modo *Dual-Ended ETH-DM*;
— Modo *Single-Ended ETH-DM*.

Em versões prévias da recomendação ITU-T G.8013/Y.1731, esses dois modos eram conhecidos como *One-Way* ETH-DM e *Two-Way* ETH-DM.

- **Modo Dual-Ended ETH-DM**

No modo *Dual-Ended ETH-DM*, cada MEP envia quadros com informações específicas para seu MEP par, para efetivar a medição dos retardos de quadros em um sentido ou a medição da variação desses retardos nesse sentido.

O quadro ETH-DM utilizado é o quadro 1DM (*One-Way Delay Measurement*), sendo a correspondente PDU referida como 1DM PDU.

As 1DM PDUs utilizam o mesmo formato que as DMM PDUs, ressalvada a eliminação dos campos reservados para as DMR PDUs, utilizadas apenas no modo *Single-Ended* ETH-DM.

Um MEP receptor pode comparar o *TxTimeStampf* recebido com o instante de recepção (referido como *RxTimef*), obtendo, por diferença, o retardo do quadro. Ou seja, o retardo é igual a [*RxTimef-TxTimeStampf*].

- **Modo Single-Ended ETH-DM**

Para a efetivação de medições no modo *Single-Ended* ETH-DM, um MEP envia DMM PDUs para o seu MEP par, e recebe DMR PDUs como retorno. Dessa forma são medidos retardos de quadros e as respectivas variações de retardos nos dois sentidos de transmissão.

As PDUS correspondentes são as DMM (*Delay Measurement Message*) PDUS e as DMR (*Delay Measurement Reply*) PDUS.

O MEP receptor de uma DMR PDU compara o *TxTimeStampf* com o *RxTimeb* (que é o momento da recepção da DMR PDU) e calcula o *two-way delay* pela fórmula [*RxTimeb – TxTimeStampf*].

O formato de uma DMM PDU encontra-se na Figura 8-27.

1								2								3								4							
8	7	6	5	4	3	2	1	8	7	6	5	4	3	2	1	8	7	6	5	4	3	2	1	8	7	6	5	4	3	2	1
MEL			Version (1)					OpCode (DMM=47)								Flags								TLV Offset (32)							
TxTimeStampf																															
Reserved for DMM receiving equipment (0) *(for RxTimeStampf)*																															
Reserved for DMR (0) *(for TxTimeStampb)*																															
Reserved for DMR receiving equipment (0)																															
[optional TLV starts here; otherwise End TLV]																															
																								End TLV (0)							

Figura 8-27 – Formato de uma DMM PDU.

No campo *Flags*, o *LSB bit* indica o tipo 1 (operação proativa) ou o tipo zero (operação em demanda).

Uma DMR PDU transporta, além do *TxTimeStampf* copiado da correspondente DMM PDU, o campo *RxTimeStampf* (*timestamp* do momento da recepção da correspondente DMM PDU) e o campo *TxTtimeStampb* (*timestamp* do momento da transmissão da DMR PDU).

O campo *Flags* é copiado do campo *Flags* da correspondente DMM PDU.

8.2.5.7. Medição de Throughput

Throughput é a taxa média de vazão de tráfego escoada com sucesso em um canal de comunicação. O *throughput* é medido tipicamente sob condições de teste, isto é, com o canal de comunicação fora de serviço (*out-of-service*), quando não existe tráfego de serviço no canal de comunicação.

A recomendação ITU-T Y.1564 (*Ethernet Service Activation Test Metodology*) define uma metodologia para testar um serviço baseado em Ethernet fora de serviço, durante a fase de sua ativação. Essa recomendação descreve testes de configuração de serviços Ethernet para verificar perfis de vazão de tráfego (*bandwidth*) e outros atributos de serviço.

A RFC 2544 especifica a medição de *throughput* (taxa de vazão de tráfego) pelo envio de quadros em taxas crescentes (até o limite teórico máximo), computando a porcentagem de quadros recebidos e reportando a taxa na qual os quadros começam a ser descartados. De um modo geral, os valores de *throughput* dependem do tamanho dos quadros utilizados para a medição.

A medição de *throughput* pode se efetivar mediante a utilização de quadros *unicast* ETH-LB (quadros LBM e LBR com *payload* de dados), ou de quadros ETH-*Test* com *payload* de dados.

As medições podem ser no modo unidirecional (*One-Way*) ou no modo bidirecional (*Two-Way*).

8.2.5.8. Endereços de Quadros de OAM

Os quadros de OAM são identificados por um único código *EtherType*, cujo valor é igual a 0x8902.

O processamento e a filtragem de quadros de OAM em um MEP são baseados nesse valor de código *EtherType* e nos respectivos níveis de MEG.

Os endereços MAC de destino de quadros de OAM podem ser *unicast* ou *multicast*, dependendo da função de OAM específica. Os endereços MAC de origem são sempre *unicast*.

A Figura 8-28 apresenta um sumário dos tipos de DA aplicáveis para diferentes tipos de OAM.

OAM type	DAs for frames with OAM PDU
CCM	Multicast class 1 DA or Unicast DA
LBM	Unicast DA or Multicast class 1 DA
LBR	Unicast DA
LTM	Multicast class 2 DA
LTR	Unicast DA
AIS	Multicast class 1 DA or Unicast DA
LCK	Multicast class 1 DA or Unicast DA
TST	Unicast DA or Multicast class 1 DA
Linear APS	Multicast class 1 DA or Unicast DA
Ring APS	Multicast class 1 DA or Unicast DA
MCC	Unicast DA or Multicast class 1 DA
LMM	Unicast DA or Multicast class 1 DA
LMR	Unicast DA
1DM	Unicast DA or Multicast class 1 DA
DMM	Unicast DA or Multicast class 1 DA
DMR	Unicast DA
EXM, EXR, VSM, VSR	Outside the scope of this Recommendation

Figura 8-28 - Sumário dos tipos de DA para quadros OAM.

Como se observa nessa figura, existem dois tipos de *multicast* DA para quadros de OAM:

— *Multicast* DA de classe 1: para quadros de OAM endereçados apenas à totalidade de MEPs de um MEG (por exemplo, CC, *multicast* LBM, AIS etc..);

— *Multicast* DA de classe 2: para quadros de OAM endereçados à totalidade de MEPs e de MIPs de um MEG (por exemplo, LTM).

Normalmente um único valor de *multicast* DA de classe 1 e um único valor de *multicast* DA de classe 2 seriam suficientes. No entanto, para aplicação da recomendação ITU-T G.8013/Y.1731 em equipamentos *Ethernet* atuais a curto prazo, os endereços *multicast* de destino podem conduzir implicitamente o nível de MEG.

Nesse caso, são necessários 8 *multicast* DAs distintos para a classe 1, e 8 *multicast* DAs distintos para a classe 2. Valores específicos para esses *multicast* DAs são os seguintes:

— *Multicast* DA 01-80-C2-00-00-3*x* para a classe 1: *x* representa níveis de MEG no intervalo 0-7;
— *Multicast* DA 01-80-C2-00-00-3*y* para a classe 2: *y* representa níveis de MEG no intervalo 8-F.

8.2.5.9. Condições de Defeito

O apêndice I da recomendação G.8013/Y.1731 apresenta os seguintes tipos de condição de defeito:

— Condição *Loss of Continuity* (LOC);
— Condição *Mismerge* (*Mismerge*);
— Condição *Unexpected MEP* (*UnexpectedMEP*);
— Condição *Unexpected MEG Level* (*UnexpectedMEGLevel*);
— Condição *Unexpected Period* (*UnexpectedPeriod*);
— Condição *Signal Fail* (*SignalFail*);
— Condição *Alarm Indication Signal* (AIS);
— Condição *Remote Defect Indication*(RDI);
— Condição *Ethernet Locked Signal* (LCK).

8.2.6. OAM de Link: Cláusula 57 do Padrão IEEE 802.3

As edições do padrão IEEE 802.3 a partir de 2008 incluem as cláusulas referentes a *Ethernet* para redes de acesso de usuários (*Ethernet Subscriber Access Networks*), que obsoletaram e incorporaram os termos do padrão IEEE 802.3ah-2004 (*Ethernet in the First Mile* – EFM).

A cláusula 57 (*Operations, Administration, and Maintenance*) da Seção 5 da última versão do padrão IEEE 802.3 (de 2012) incorporou funções de OAM, com o propósito de prover mecanismos para a monitoração da operação de *links*, com propósitos tais como indicação de falhas e *loopback* remoto.

De maneira geral, as funções de OAM possibilitam aos operadores de redes monitorar a saúde da rede e determinar rapidamente a localização de enlaces com defeitos ou pontos de falha. As funções de OAM descritas na cláusula 57 do padrão

IEEE 802.3 fornecem mecanismos que complementam aplicações eventualmente localizadas nas camadas mais altas.

As informações de OAM de *Link* são contidas em quadros designados como OAMPDUs. As OAMPDUs contêm as informações apropriadas de controle e de *status* usadas para monitorar, testar e diagnosticar falhas em *links* que suportam OAM. As OAMPDUs trafegam por um único enlace, entre entidades pares de OAM.

As OAMPDUs são do tipo *Slow Protocol*. O conceito de *Slow Protocols* encontra-se no subitem 2.7.7 psc do capítulo 2 deste livro. Esse subitem tem como base o anexo 57A do padrão IEEE 802.3-2012 (Cláusula 5), que define os requisitos para o suporte desse tipo de protocolo.

A subcamada OAM definida no padrão IEEE 802.3 não inclui funções tais como gerência de estações, alocação de bandas passantes e provisonamento. Ela tem como propósito atender as funções de OAM em *links* ponto a ponto ou em *links* ponto a ponto emulados (como nas EPONs).

Os objetivos da subcamada OAM são os seguintes:

— Indicação remota de falhas;
— *Loopback* remoto;
— Monitoramento de *links*;
— Outros (*miscellaneous*).

8.2.6.1. OAM na Arquitetura EFM

A subcamada EFM OAM posiciona-se na arquitetura EFM conforme a Figura 8-29.

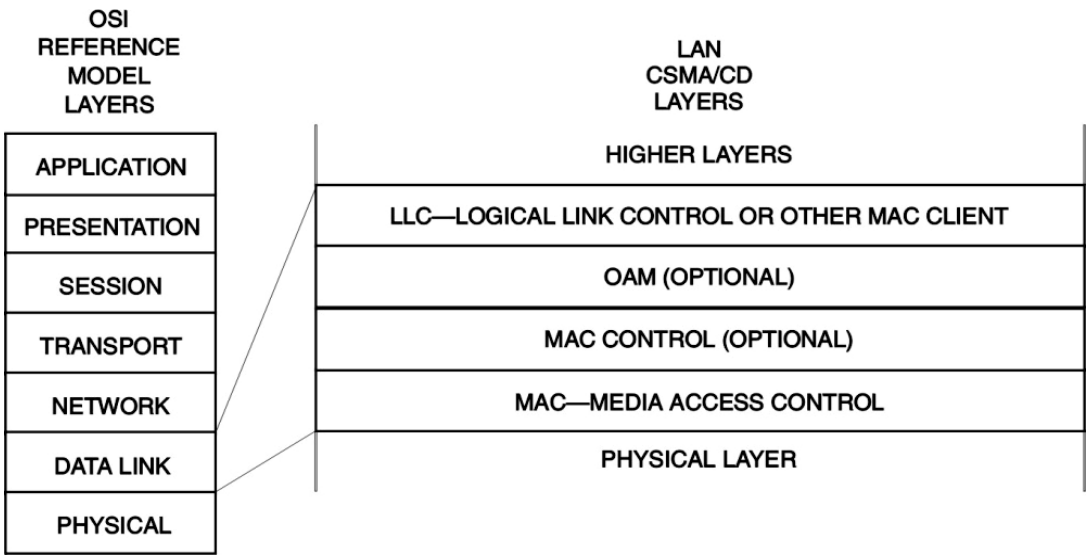


Figura 8-29 - OAM na arquitetura EFM.

Embora não esteja claramente especificado nessa figura, a subcamada de OAM presta serviços à subcamada cliente de OAM (*OAM Client*) ou à subcamada

cliente de MAC (*Mac Client*). Tais serviços são prestados mediante a troca de primitivas de serviço, como mostra a Figura 8-30.

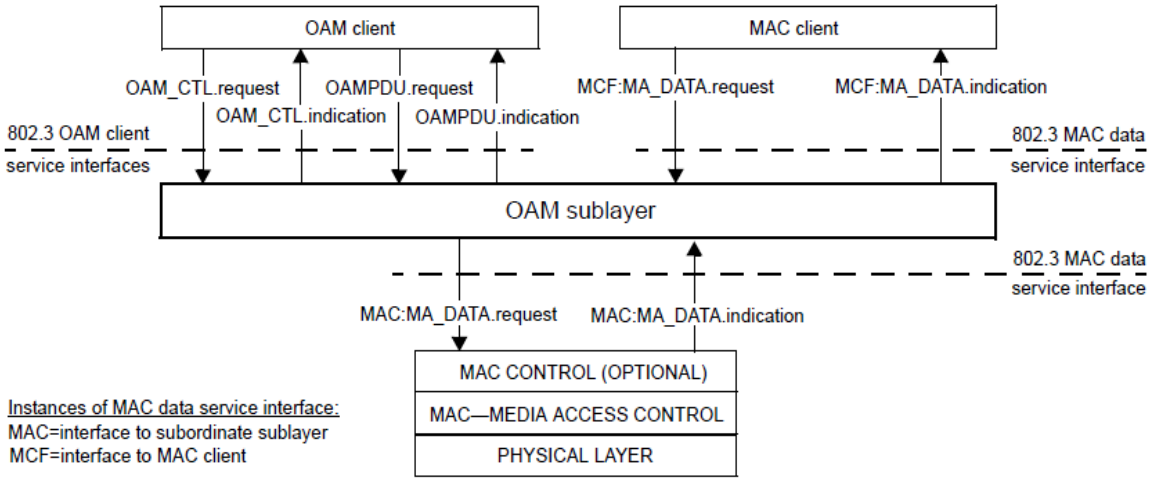


Figura 8-30 – Primitivas de serviço envolvendo a subcamada de OAM.

Observa-se nessa figura também as primitivas de serviço intercambiadas com a camada inferior ao OAM, que é a camada MAC ou, opcionalmente, a subcamada *MAC Control*.

Quando da transmissão de OAM PDUs, a subcamada de OAM utiliza um módulo de multiplexação para atendimento das duas subcamadas cliente. Na recepção de OAM PDUs, a subcamada de OAM, além da função de demultiplexação, utiliza também um módulo analisador (*parser*).

8.2.6.2. Estrutura Comum de OAMPDUs

A estrutura comum de OAMPDUs encontra-se na Figura 8-31.

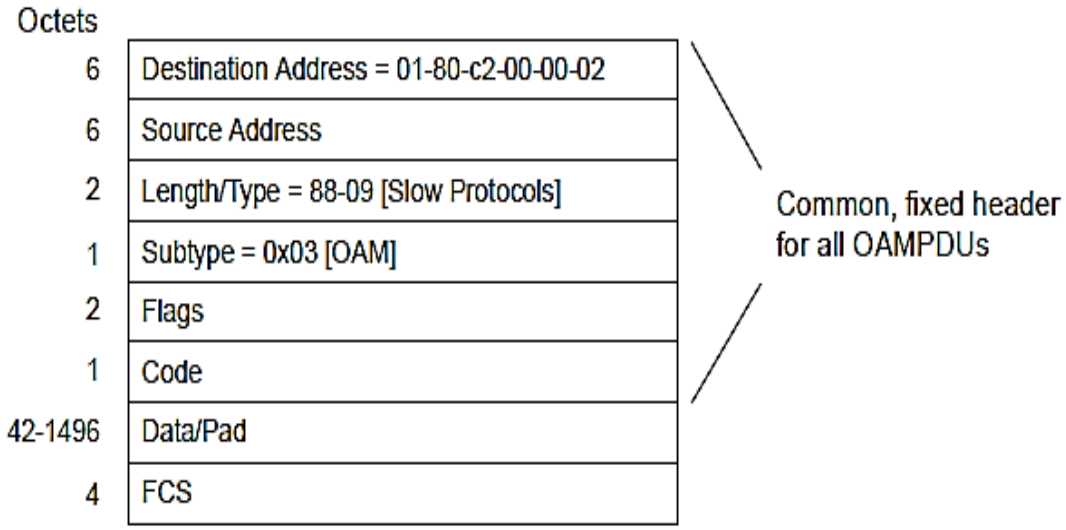


Figura 8-31 – Estrutura comum das OAMPDUs.

Os campos de uma OAMPDU, estruturada conforme essa figura, têm os seguintes significados:

— *Destination Address* (DA): é o *Slow Protocols_Multicast Address*, especificado no anexo 57A do padrão IEEE 802.3-2012 (cláusula 5);
— *Source Address* (SA): é o endereço MAC individual associado à porta pela qual a OAMPDU é transmitida;
— *Length/Type*: as OAMPDUs são sempre codificadas como tipo, e utilizam o valor OX8809 (88-09) indicando o uso de *Slow Protocols*;
— *Sub-Type*: esse campo identifica o *Slow Protocol* específico. O seu valor para OAMPDUs é igual a 0X03;
— *Flags*: esse campo contém valores de *flags* especificados a seguir;
— *Code*: esse campo identifica o tipo específico de OAMPDU, como será visto adiante neste item;
— *Data/Pad*: esse campo contém o *payload* da OAMPDU e eventuais *padding bits*;
— FCS: esse campo representa a FCS da subcamada MAC.

Os valores utilizados para o campo *Flags* encontram-se na Figura 8-32.

Bit(s)	Name	Description
15:7	*Reserved*	Reserved bits shall be set to zero when sending an OAMPDU, and should be ignored on reception for compatibility with future use of reserved bits.
6	Remote Stable	When remote_state_valid is set to TRUE, the Remote Stable and Remote Evaluating values shall be a copy of the last received Local Stable and Local Evaluating values from the remote OAM peer. Otherwise, the Remote Stable and Remote Evaluating bits shall be set to 0.
5	Remote Evaluating	
Bit(s)	**Name**	**Description**
4	Local Stable	Local Stable and Local Evaluating form a two-bit encoding shown below: 4:3 0x0 = Local DTE Unsatisfied, Discovery can not complete 0x1 = Local DTE Discovery process has not completed 0x2 = Local DTE Discovery process has completed 0x3 = Reserved. This value shall not be sent. If the value 0x3 is received, it should be ignored and not change the last received value.
3	Local Evaluating	
2	Critical Event	1 = A critical event has occurred. 0 = A critical event has not occurred.
1	Dying Gasp	1 = An unrecoverable local failure condition has occurred. 0 = An unrecoverable local failure condition has not occurred.
0	Link Fault	The PHY has detected a fault has occurred in the receive direction of the local DTE (e.g., link, Physical layer). 1 = Local device's receive path has detected a fault. 0 = Local device's receive path has not detected a fault.

Figura 8.32 – Valores do campo Flags.

Os três últimos *Flags* dessa figura representam os eventos de *link* classificados como críticos. A cláusula 57 do padrão IEEE 802.3, contudo, não define as falhas envolvendo esses eventos críticos, que são o *Link Fault*, o *Dying Gasp* e o *Critical Event*, considerando essa definição uma questão de implementação.

Além de se basear na indicação de ocorrência de eventos críticos no *link* pelo campo *Flags* de OAMPDUs recebidas, um cliente OAM local detecta a ocorrência de outros eventos de falha pela recepção de *Event Notification OAMPDUs*.

Os valores do campo *Code*, que representam os diversos tipos de OAMPDUs, encontram-se na figura 8-33.

Code	OAMPDU	Comment	Source
00	Information	Communicates local and remote OAM information.	OAM client / OAM sublayer
01	Event Notification	Alerts remote DTE of link event(s).	OAM client
02	Variable Request	Requests one or more specific MIB variables.	OAM client
03	Variable Response	Returns one or more specific MIB variables.	OAM client
04	Loopback Control	Enables/disables OAM remote loopback.	OAM client
05-FD	*Reserved*	*Reserved*	OAM client
FE	Organization Specific	Reserved for Organization Specific Extensions, distinguished by Organizationally Unique Identifier	OAM client
FF	*Reserved*	*Reserved*	OAM client

Figura 8-33 – Valores do campo Code.

8.2.6.3. Tipos de OAMPDU

Conforme o subitem anterior , os diferentes tipos de OAMPDU são identificados pelos códigos indicados nos respectivos campos *code*.

Serão apresentadas a seguir as OAMPDUs de informação e os demais tipos de OAMPDU.

- **OAMPDUs de Informação**

As OAMPDUs de informação (*Information OAMPDUs*), identificadas pelo código 0x00, são utilizadas para o envio de informações de estado de OAM para o DTE remoto. A Figura 8-34 apresenta a estrutura desse tipo de OAMPDU.

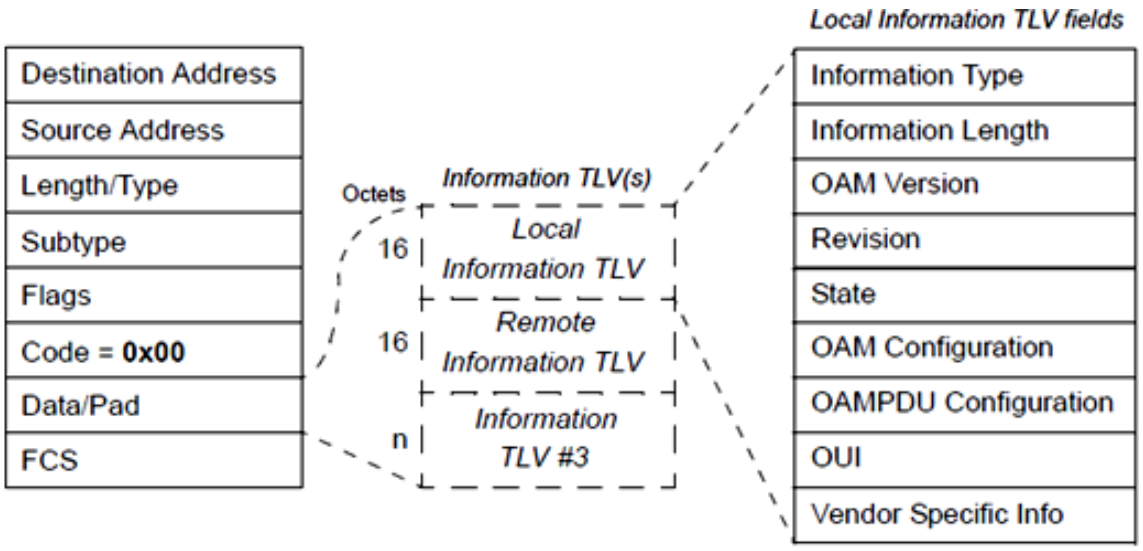


Figura 8-34 – Estrutura de Information OAMPDUs.

A existência do *Remote Information TLV* nessa figura é opcional, e, caso exista,constitui-se em cópia do último *Local Information TLV* recebido do par OAM remoto, com exceção do valor do campo *Information TLV Type*, que passa a ser igual a 0x02 (esse valor é igual a 0x01 para o *Local Information TLV*).

Os *Organization Specific Information TLVs*, do tipo 0xFE, utilizam o campo OUI (*Organization Unique Identifier*), complementado pelo campo *Organization Specific Value*, definido a critério da organização.

- **Demais Tipos de OAMPDU**

A *Event Notification OAM PDU*, identificada pelo código igual a 0x01, é utilizada para alertar o DTE remoto quanto a alguns eventos ocorridos no *link*, tais como determinados tipos de erro.

A *Variable Request OAMPDU,* identificada pelo código igual a 0x02, é utilizada para requisitar um ou mais variáveis da MIB do DTE remoto. Cada uma dessas variáveis é contida em um campo *Variable Descriptor,* identificado por um número de sequência a partir de #1.

A *Variable Response OAMPDU,* identificada pelo código igual a 0x03, é utilizada para retornar uma ou mais variáveis de MIB solicitadas ao DTE remoto. O campo *Data* de uma *Variable Response* OAMPDU pode conter um ou mais *Variable Containers,* onde são transportados as variáveis de MIB solicitadas.

A *Organization Specific OAMPDU*, identificada pelo código igual a 0xFE, é utilizada para extensões específicas da organização responsável pela rede. Os três primeiros octetos do campo *Data* desse tipo de OAMPDU representam o valor de OUI atribuido à organização. Os formatos e as funções do restante desse campo *Data* dependem do valor de OUI e não foram especificados no padrão IEEE 802.3.

A *Loopback Control* OAMPDU, identificada pelo código 0x04, é utilizada para controlar o estado de *loopback* do DTE remoto. O campo *Data* desse tipo de OAMPDU consiste exclusivamente no parâmetro *OAM Remote Loopback Command*.

Foram especificados os seguintes valores para esse comando, sendo que os demais possíveis valores encontram-se reservados:

— 0x01: *Enable OAM Remote Loopback*;
— 0x02: *Disable OAM Remote Loopback.*

8.2.6.4. – OAM Remote Loopback

A Cláusula 57 do padrão IEEE 802.3 suporta, opcionalmente, o modo *loopback*, que opera remotamente. Esse modo de operação pode ser utilizado para a localização de falhas e para teste de desempenho do *link*.

A Figura 8-35 apresenta o caminho percorrido pelos quadros através da arquitetura , desde o DTE local até o DTE remoto.

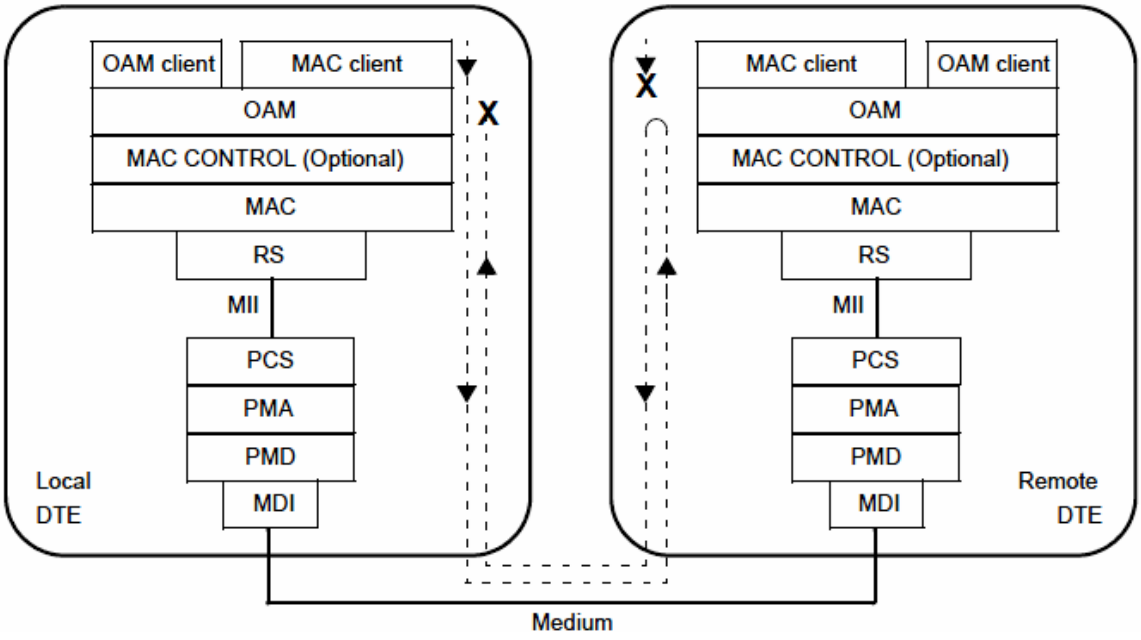


Figura 8-35 – Loopback remoto em OAM de Link.

Para iniciar a operação em *loopback*, o cliente de OAM local envia uma *Loopback Control OAMPDU* com o comando *Enable OAM Remote Loopback*. Após a recepção desse comando, o DTE remoto realiza algumas ações específicas, e envia então uma *Information OAMPDU* para o DTE local, indicando o seu estado atual quanto a OAM. O cliente de OAM local seta então o respectivo *local_max_action* para FWD (*forwarding*).

Estando o modo *loopback* habilitado, ocorrem as seguintes ações:

— O DTE local transmite quadros vindos do cliente de MAC local ou OAM PDUs vindas do cliente de OAM local ou da subcamada OAM local;

— Dentro da entidade da subcamada OAM remota, qualquer quadro que não seja uma OAMPDU, mesmo que outros quadros de *Slow Protocols*, são retornados pelo *loop* sem qualquer alteração;
— As OAMPDUs recebidas pelo DTE remoto são repassadas para o cliente OAM remoto;
— Ambos os DTEs devem enviar OAMPDUs para o DTE par para manter o processo *Discovery*.

Para encerrar o teste de OAM *loopback* remoto, o cliente de OAM local seta o respectivo *local_mux_action* para *DISCARD*, e envia uma *Loopback Control OAMPDU* com o comando *Disable OAM Remote Loopback*.

Ao receber esse comando, o cliente de OAM remoto reseta os parâmetros ativados e envia uma *Information OAMPDU* para o cliente de OAM local. O *parser* remoto reassume a função de aceitar quadros vindos do cliente de MAC local.

8.2.6.5. Operação OAM Unidirecional

O padrão IEEE 802.3 define uma forma para que o OAM de *link* suporte um mecanismo baseado em OAMPDUs, para notificar o DTE remoto quando um sentido do *link* encontra-se fora de operação. Para isso, o *link* passa a operar de modo unidirecional.

8.2.6.6. Modos de Operação de OAM

Os DTEs que incorporam a subcamada de OAM podem operar no modo ativo ou no modo passivo. Quando o OAM é habilitado, um DTE capacitado para operar em ambos os modos deve selecionar um desses modos.

No modo ativo, um DTE pode realizar as seguintes ações, que não podem ser realizadas pelos DTEs no modo passivo:

— Iniciar o processo OAM *discovery*;
— Enviar *Variable Request* OAMPDUs e *Loopback Control* OAMPDUs.

A Figura 8-36 sintetiza o comportamento de DTEs operando no modo ativo e no modo passivo.

Capability	Active DTE	Passive DTE
Initiates OAM Discovery process	Yes	No
Reacts to OAM Discovery process initiation	Yes	Yes
Required to send Information OAMPDUs	Yes	Yes
Permitted to send Event Notification OAMPDUs	Yes	Yes
Permitted to send Variable Request OAMPDUs	Yes	No
Permitted to send Variable Response OAMPDUs	Yes [a]	Yes
Permitted to send Loopback Control OAMPDUs	Yes	No
Reacts to Loopback Control OAMPDUs	Yes [a]	Yes
Permitted to send Organization Specific OAMPDUs	Yes	Yes

[a]Requires the peer DTE to be in Active mode.

Figura 8-36 – Modos de operação de DTEs.

8.3. MECANISMOS DE PROTEÇÃO EM ETHERNET

Serão abordados neste item os seguintes padrões:

— Recomendação ITU-T G.8031/Y.1342 (*Ethernet Protection Switching*);
— Recomendação ITU-T G.8032/Y.1344 (*Ethernet Ring Protection Switching*), que define a tecnologia ERP;
— RFC 3619 (*Extreme Networks' Ethernet Automatic Protection Switching*), que define a tecnologia *EAPS*;
— Padrão IEC 62439.

8.3.1. Recomendação ITU-T G.8031/Y.1342 (Ethernet Protection Switching)

A recomendação ITU-T G.8031/Y.1342 (*Ethernet Protection Switching*) define o protocolo APS (*Automatic Protection Switching*), doravante referido como protocolo *Linear* APS, assim como mecanismos para comutação de proteção linear, para conexões *Ethernet* VLAN-*based* ponto a ponto em redes *Ethernet* com topologias lineares. Outros esquemas de proteção, incluindo aqueles para conexões ponto a multiponto e conexões multiponto a multiponto, ficam para posterior estudo.

A conectividade ponto a ponto de que trata essa recomendação é aquela aplicável entre dois pontos de fluxo em um domínio de fluxo ETH.

A comutação de proteção associada ao protocolo APS deve ser mais rápida que aquela obtida por outros mecanismos de sobrevivência, tais como o RSTP, com um tempo de transferência (*Tt*) inferior a 50 ms.

Arquiteturas lineares para comutação de proteção 1+1 e 1:1 encontram-se definidas na última versão da recomendação ITU-T G.8031/Y.1342, de 06/2011.

Como será visto adiante, a recomendação ITU-T G.8032/Y.1344 define o processo de comutação de proteção para redes *Ethernet* com arquiteturas em anel (*ring*).

8.3.1.1. Tipos de Comutação de Proteção Linear

A comutação de proteção provida pelo protocolo *Linear* APS é um mecanismo de sobrevivência alocado de forma total. Isso significa que a rota e as larguras de banda da entidade de transporte de proteção ficam reservadas para uma determinada entidade de transporte de serviço.

- **Comutação de Proteção Unidirecional e Bidirecional**

No caso de transmissão bidirecional é possível ocorrer comutação de proteção unidirecional ou bidirecional. Na comutação de proteção unidirecional, os seletores em cada terminação são independentes entre si.

Na comutação de proteção bidirecional existe a necessidade de coordenação entre as terminações, para que ambas adotem o mesmo arranjo de seletores, mesmo que a falha seja unidirecional. Essa coordenação ocorre pelo intercâmbio de informações pelo protocolo *Linear* APS.

- **Comutação de Proteção Reversiva e Não Reversiva**

Na comutação de proteção reversiva, o tráfego *Ethernet* normal retorna para a entidade de transporte principal após a eliminação da condição de defeito.

No caso de um comando (*Forced Switch*, por exemplo), a reversão é imediata. Ocorrendo uma condição de defeito, a reversão realiza-se normalmente após a expiração de um temporizador, referido como "*Wait-to-Restore*" *timer*, utilizado para evitar instabilidade em caso de defeitos intermitentes.

Em comutação de proteção não reversiva, o tráfego normal permanece escoando pela entidade de transporte de proteção mesmo após a eliminação da condição de defeito.

8.3.1.2. Arquiteturas de Comutação de Proteção Linear

O protocolo *Linear* APS opera com duas arquiteturas básicas:

— Arquitetura de comutação de proteção 1:1 linear;
— Arquitetura de comutação de proteção 1+1 linear.

- **Arquitetura 1:1**

Na arquitetura linear de comutação de proteção 1:1, a entidade de transporte de proteção é dedicada à entidade de transporte de serviço. No entanto, o tráfego normal é transportado exclusivamente em uma das entidades de transporte,

utilizando para isso uma *bridge* seletora nas duas terminações do domínio protegido.

Na arquitetura de comutação linear 1:1, como há necessidade de coordenação para garantir que a *bridge* seletora na origem e o seletor no destino (*sink*) selecionem a mesma entidade de transporte, torna-se indispensável o uso do protocolo *Linear* APS.

- **Arquitetura 1+1**

Na arquitetura linear de comutação de proteção 1+1, a entidade de transporte de proteção é também dedicada à proteção da entidade de transporte de serviço, mas em uma das terminações, referida como *head end*, o tráfego protegido é permanentemente comutado tanto para entidade de transporte de serviço quanto para a de proteção.

O tráfego normal é transmitido, nos dois sentidos, simultaneamente em ambas as entidades de transporte. No entanto, em um dado instante, apenas o tráfego transmitido em uma dessas entidades de transporte é recebido pelo destino.

A comutação de proteção 1+1 pode ocorrer de duas formas:

— Comutação de proteção 1+1 unidirecional;
— Comutação de proteção 1+1 bidirecional.

Na comutação de proteção 1+1 unidirecional, apenas a direção afetada pelo defeito é comutada, e o tráfego na direção não afetada continua a escoar-se pela entidade de transporte de serviço. Como na comutação de proteção 1+1 unidirecional cada direção é comutada independentemente, a coordenação pelo protocolo *Linear* APS não é necessária.

Na comutação de proteção 1+1 bidirecional, ambas direções são comutadas, mesmo no caso de uma falha unidirectional. Consequentemente, a coordenação pelo protocolo *Linear APS* é necessária.

8.3.1.3. Protocolo Linear APS

As PDUs do protocolo *Linear* APS são estruturadas em conformidade com a recomendação ITU-T G.8013/Y.1731. Assim, tais PDUs utilizam o formato comum representado na Figura 8-37.

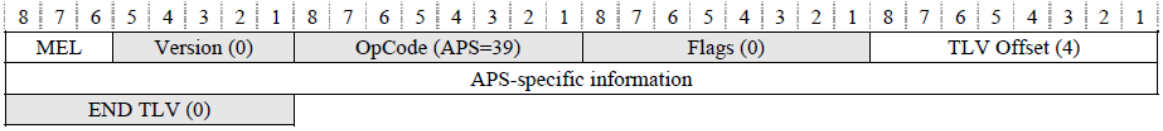


Figura 8-37 – Formato da Linear APS PDU.

O nível de MEG no qual uma *Linear APS PDU* é inserido e indicado no campo MEL (MEG *Level*).

O campo *Version* assume o valor 0x00, o valor de OpCode é igual a 39 (0x27), e os campos *Flags* e END TLV têm o valor 0x00. O campo *TLV Offset* possui o valor 4.

O campo *APS-Specific Information* é formatado de acordo com o item 11.1 da recomendação G.8031/Y.1342 do ITU-T conforme representação na Figura 8-38.

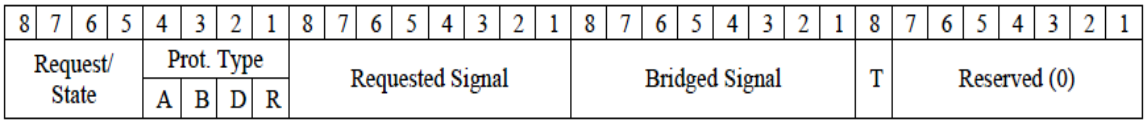


Figura 8-38 – Formato do campo APS-Specific Information.

Os valores de código para o campo *APS-Specific Information* encontram-se na tabela 2 da recomendação ITU-T G.8031/Y.1342.

Alguns valores de tipo de proteção são inválidos por não existirem as combinações dos quatro parâmetro A, B, D e R. Os valores válidos, com as correspondências com os valores de tipo de *bridge*, encontram-se na figura 8-39.

Protection type	Bridge type	Valid configuration
000x	N/A	1+1 unidirectional, no APS communication
100x	N/A	1+1 unidirectional w/APS communication
101x	N/A	1+1 bidirectional w/APS communication
111x	0	1:1 bidirectional w/APS communication using a selector bridge
111x	1	1:1 bidirectional w/APS communication using a broadcast bridge

Figura 8-39 – Valores válidos de tipo de proteção e valores de tipo de bridge.

8.3.1.4. Interação entre Linear APS e STP

Para evitar interação indesejada entre a comutação de proteção linear automática em *Ethernet* pelo protocolo *Linear APS* e o protocolo *Spanning Tree* (STP), um domínio STP não pode incluir portas de uma *bridge* localizadas em um domínio protegido.

Uma maneira de alcançar esse objetivo seria desabilitar o STP no domínio protegido. Os domínios fora desse domínio protegido podem permanecer com o STP habilitado.

Outra maneira seria fazer com que as entidades de transporte principal e de proteção pertencessem a domínios STP distintos.

A Figura 8-40 ilustra o primeiro desses cenários, no qual os domínios protegido e STP são separados verticalmente e não há intercessão entre os mesmos. As *bridges* A e B encontram-se nos limites entre os domínios STP e protegido, e interconectam os dois domínios STP sem que haja perigo de formação de *loops*.

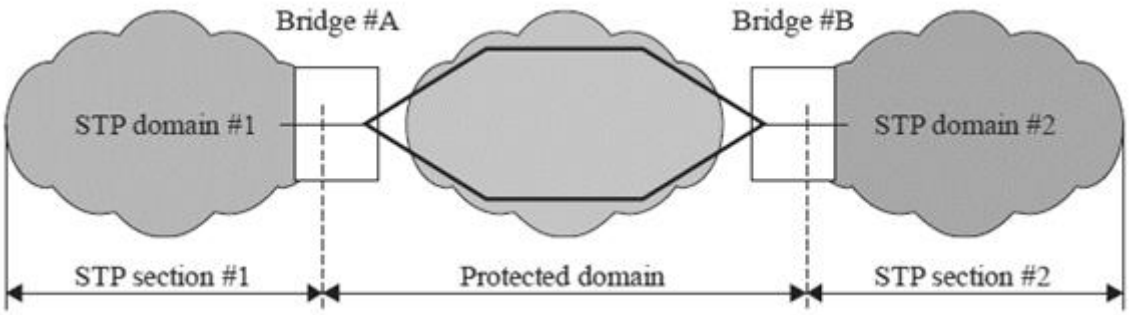


Figura 8-40 – Separação entre domínio Linear APS e domínios STP.

O segundo cenário é ilustrado na Figura 8-41. Configuram-se dois domínios STP, sendo o domínio protegido segmentado horizontalmente para que uma das entidades de transporte (por exemplo, a principal) fique em um domínio STP e a outra (por exemplo, a de proteção) fique no outro domínio STP.

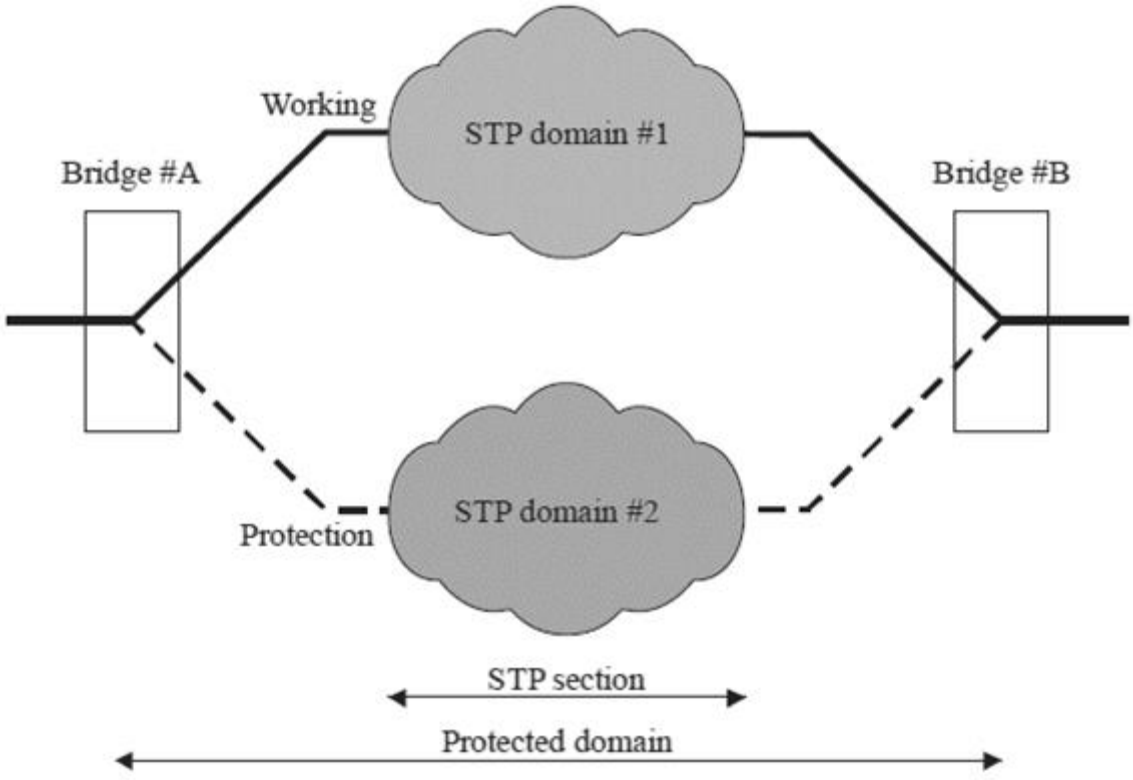


Figura 8-41 – Superposição entre domínio Linear APS e domínios STP.

A Figura 8-42 mostra como, nesse segundo cenário, as entidades de transporte principal e de proteção são provisionadas de forma separada para VLANs distintas. Essa configuração permite que cada VLAN e o recurso de rede sejam usados de maneira eficaz.

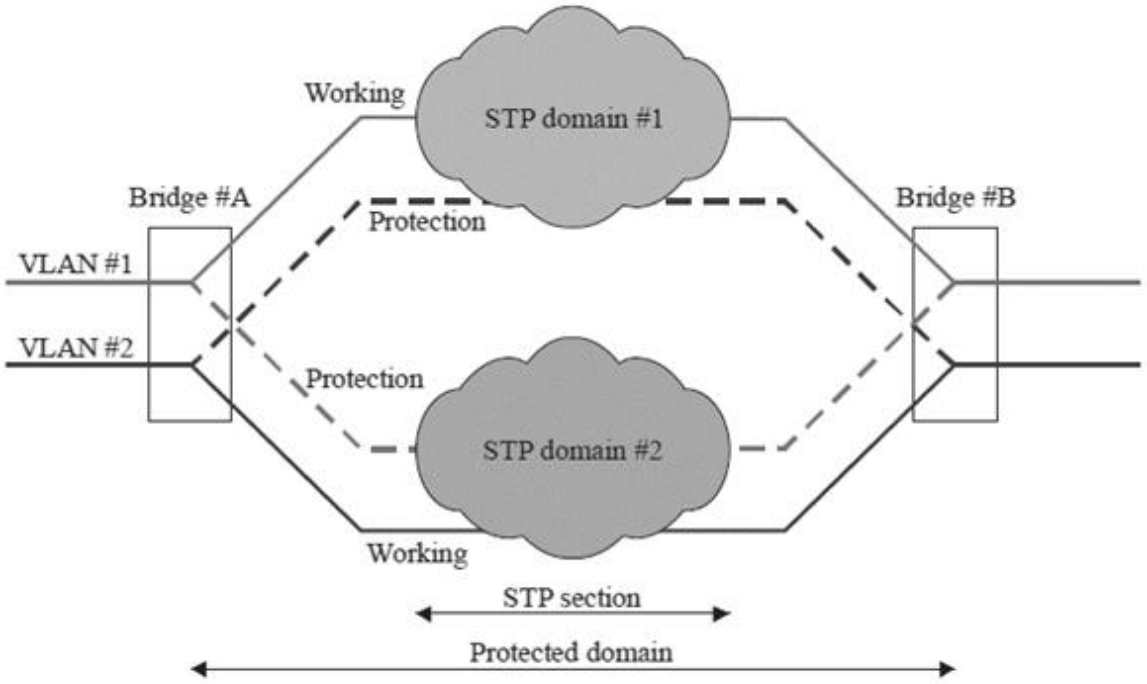


Figura 8-42 - Superposição entre domínio Linear APS e domínios STP por VLAN.

Nessa figura, caso uma falha no domínio STP#1 (domínio em serviço) afete a VLAN#1 mas não afete a VLAN#2, somente o tráfego da VLAN#1 será transferido para o domínio STP#2 (domínio de proteção), continuando o tráfego da VLAN#2 a cursar pelo domínio STP#1.

8.3.2. **Recomendação ITU-T G.8032/Y1344 (Ethernet Ring Protection Switching)**

Essa recomendação, intitulada *Ethernet Ring Protection Switching*, referente à tecnologia ERP, define o protocolo *Ring* APS (R-APS), assim como mecanismos para comutação de proteção, para conexões na ETH *layer* em redes *Ethernet* com topologias em anel (*ring*).

O protocolo de comutação de proteção definido nessa recomendação possibilita conectividade ponto a ponto, ponto a multiponto e multiponto a multiponto dentro de um anel ou dentro de um conjunto de anéis interconectados. Esses conjuntos de anéis interconectados são denominados redes *multi-ring/ladder*.

Um anel da ETH *layer* é mapeado na estrutura de um anel na camada física. Os mecanismos e protocolo são definidos de forma a evitar a formação de *loops*.

Diferentemente da comutação de proteção linear, o ERP não separa as entidades de transporte de serviço e de proteção, mas reconfigura a entidade de transporte durante a comutação de proteção

A arquitetura de proteção no ERP baseia-se na existência do protocolo *Ring* APS (R-APS) para coordenar as ações de proteção no anel *Ethernet*.

Assim como na comutação de proteção linear, a comutação de proteção no ERP pode ser reversiva ou não reversiva.

8.3.2.1. Funcionamento Básico

Os princípios fundamentais de funcionamento da arquitetura de proteção no ERP são os seguintes:

— Prevenção de *loops*;
— Utilização dos mecanismos de aprendizagem, de transmissão e de FDB (*Forwarding Database*) definidos na função ETH-FF (ETH *Flow Forwarding*).

Os protocolos *spanning tree* não são utilizados no ERP.

Para o funcionamento do anel *Ethernet* de forma a evitar-se a ocorrência de *loops* e de possibilitar a comutação de proteção, um dos *links* do anel encontra-se bloqueado ao tráfego na fase de funcionamento normal do anel.

Esse *link* bloqueado é referido como *Ring Protection Link* (RPL). Um dos nós do anel *Ethernet* que delimita o RPL é denominado nó RPL *owner*, que é o responsável pelo bloqueio e desbloqueio do RPL.

O outro nó do anel *Ethernet* que delimita o RPL, referido como nó RPL *neighbour*, pode também ser o responsável pelo bloqueio e desbloqueio do RPL em sua respectiva terminação.

Em caso de falha no anel *Ethernet*, cabe ao protocolo R-APS a coordenação das ações de proteção do anel. Em cada nó da rede as ações de comutação de proteção são controladas pela função *Flow Forwarding* (ETH-FF).

Além da topologia consistindo de um único anel *Ethernet*, o mecanismo ERP pode suportar redes *multi-ring/ladder*, que consistem de conjuntos integrados de anéis *Ethernet* que se interconectam por um ou mais pontos. Os anéis interconectados podem desempenhar a função *major ring* ou a função *sub-ring*.

8.3.2.2. Modelos ERP em Anel Isolado

A Figura 8-43 apresenta um exemplo de modelo ERP em anel isolado, na fase de funcionamento normal (sem falhas na rede).

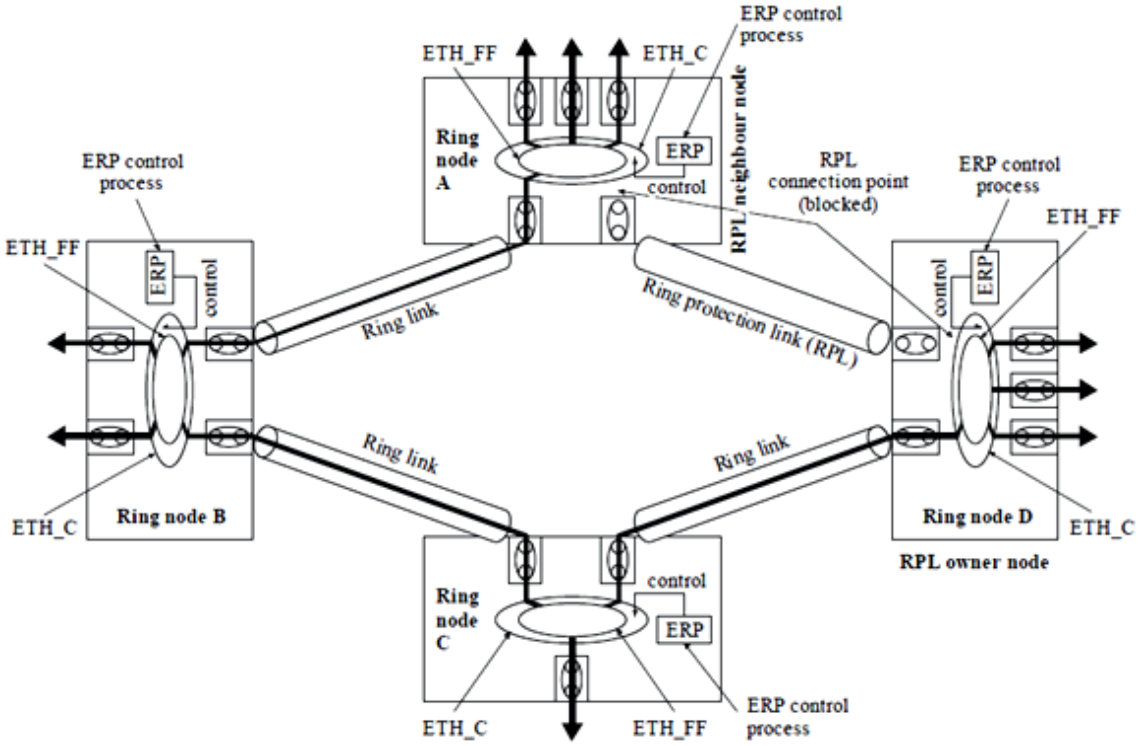


Figura 8-43 – ERP em anel isolado em condições normais de tráfego.

Como se observa nessa figura, o nó D representa o nó RPL *owner*, enquanto o nó A representa o nó *neighbour*. Esses dois nós são os responsáveis por bloquear a transmissão e a recepção de tráfego sobre o RPL quando não existe *requests* no anel Ethernet.

Um quadro MAC com endereço individual desconhecido, um quadro *multicast* ou um quadro *broadcast*, que entra no anel *Ethernet* por um determinado nó, é transmitido nos dois sentidos. Ressalva-se, contudo, a possível filtragem que pode ocorrer em decorrência da inexistência de registros do endereço MAC de grupo e do VID da VLAN, indicados no quadro MAC.

Nos nós sucessivos os quadros são enviados pelas portas que não aquelas pelas quais o quadro ingressou. Esse envio sucessivo encerra-se quando o quadro atinge em um sentido, ou nos dois sentidos, os nós de borda do RPL.

A Figura 8-44 representa a figura anterior após a ocorrência de uma falha de sinal (*SF condition*).

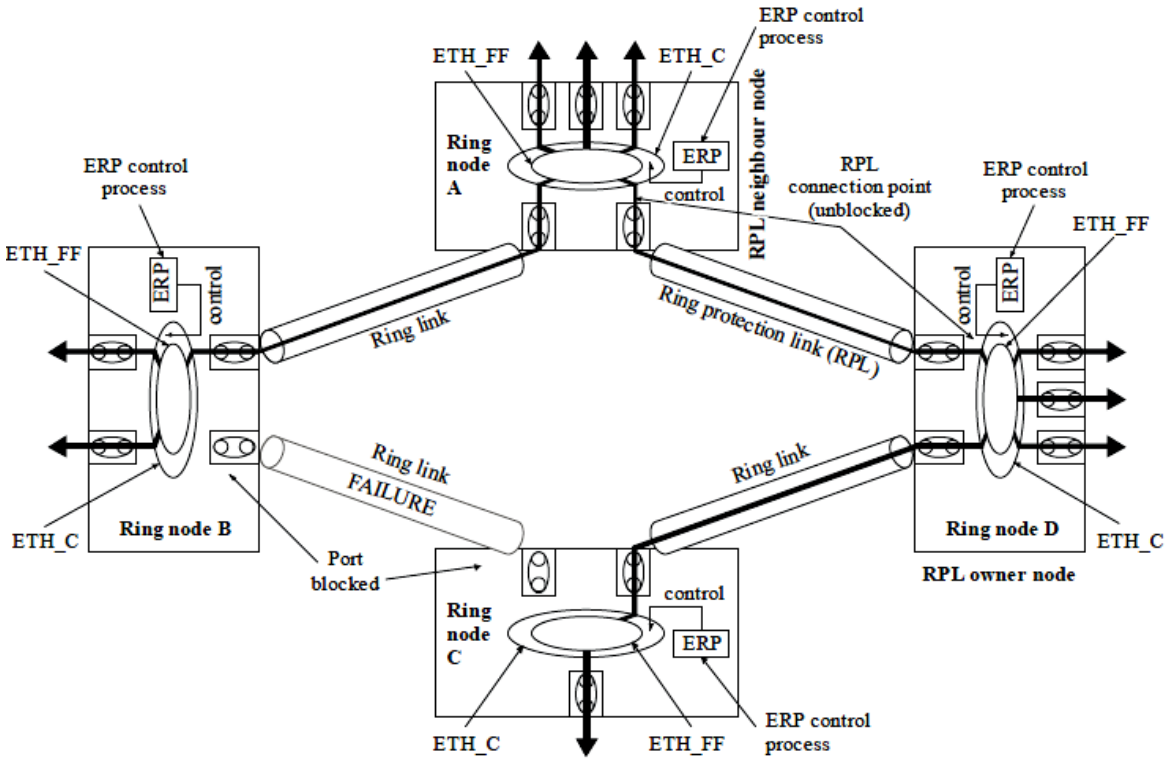


Figura 8-44 – Figura anterior após ocorrência de falha.

Como se observa nessa figura, o RPL foi desbloqueado bidirecionalmente após a detecção da falha.

Para fins de detecção de condições de falha no anel, podem ser configurados MEPs nos diferentes nós do anel, ocorrendo então a troca de CCM PDUs nos *links* do anel.

8.3.2.3. Protocolo R-APS

Da mesma forma que no caso do protocolo Linear APS, as PDUs do protocolo R-APS são estruturadas em conformidade com a recomendação ITU-T G.8013/Y.1731. Assim, tais PDUs utilizam o formato comum representado na Figura 8-45.

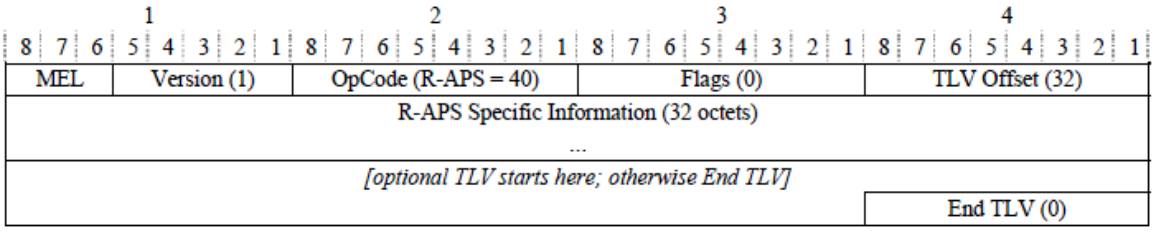


Figura 8-45 – Formato das R-APS PDUs.

O campo *Version* assume o valor 0x01 (versão atual da recomendação G.8032/Y.1344), o valor de *OpCode* é igual a 40 (0x28), os campos *Flags* e END TLV têm o valor 0x00 e o campo *TLV Offset* o valor 32 (0x20).

O campo *R-APS-Specific Information* é formatado de acordo com a Cláusula 10 da recomendação G.8032/Y.1344 do ITU-T, conforme a Figura 8-46.

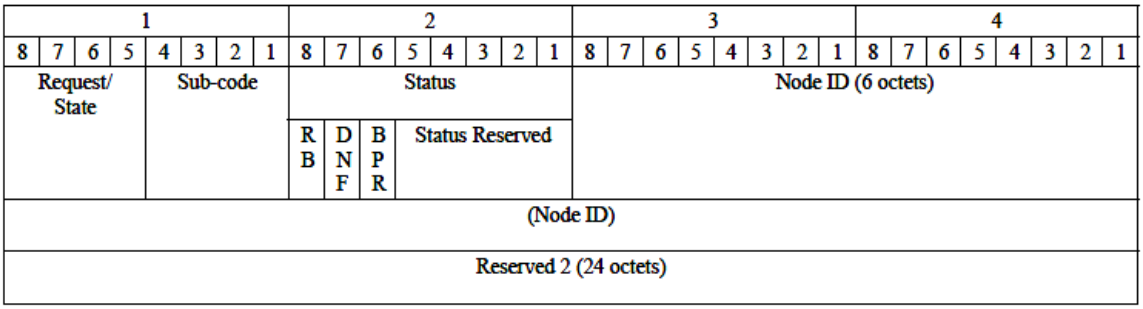


Figura 8-46 – Formato do campo R-APS-Specific Information.

O campo *Request/State* (com 4 bits) pode assumir os valores da Figura 8-47.

Field	Value	Description
Request/state	1101	Forced switch
	1110	Event
	1011	Signal fail (SF)
	0111	Manual switch (MS)
	0000	No request (NR)
	Other	Reserved for future international standardization

Figura 8-47 – Valores do campo Request/State.

O campo *Sub-Code* (com 4 bits) apresenta valores associados com alguns valores de *request/state*. Esses sub-códigos indicam *Flush Request* (0000), estando os demais valores reservados.

O campo *Status* (com 8 bits), indica, em seus três primeiros bits, os seguintes valores (os demais bits encontram-se resevados):

— *RB – RPL Blocked*, onde RPL significa *Ring Protection Link*, sendo que o valor 1 indica RPL bloqueado;
— DNF (*Do Not Flush*), sendo que o valor 1 indica que um RDB (*Filtering Database*) *flush* deveria não ser acionado pela recepção dessa PDU;
— *BRP - Blocked Port Reference*, cujo valor indica qual anel (anel 0 ou anel 1) encontra-se bloqueado.

8.3.3. **RFC 3619 (EAPS)**

A RFC 3619 (*Extreme Networks´ Ethernet Automatic Protection Switching – Version 1*) descreve a tecnologia EAPS, desenvolvida pela *Extreme Networks* para detecção de falhas e comutação de proteção em anéis *Ethernet* (*Ethernet rings*). A resiliência é comparável à dos anéis SONET, mas com custo mais reduzido e com menos restrições (como tamanho do anel, por exemplo).

8.3.3.1. Tempo de Convergência

O EAPS apresenta um tempo de convergência menor que 1 segundo, podendo ser, frequentemente, menor que 50 milisegundos.Não existe limite para o número de nós no anel, e o tempo de convergência independe desse número.

8.3.3.2. Operação do EAPS

Um domínio EAPS opera em um único anel *Ethernet*. Qualquer VLAN a ser protegida é configurada em todas as portas no anel em um dado Domínio EAPS. Cada Domínio EAPS possui um único nó *master* (*master node*). Todos os demais nós no anel são referidos como nós de trânsito.

Uma das portas do nó *master* é designada como porta primária, enquanto a outra é designada como porta secundária. Em operação normal o nó *master* bloqueia a porta secundária para o tráfego de dados, evitando assim a ocorrência de *loops*.

Dessa forma os mecanismos STP tornam-se desnecessários. Para o tráfego de controle, contudo, a porta secundária permanece desbloqueada, sendo para isso utilizada uma VLAN de Controle (*Control VLAN*).

Se o nó *máster* detectar uma falha no anel, ele desbloqueia a sua porta secundária e passa a permitir a passagem de quadros de dados através dessa porta.

Os processos de comutação *Ethernet* e de aprendizagem de endereços MAC normais continuam sendo utilizados.

O EAPS utiliza um mecanismo de alarme e um mecanismo de *polling*.

- **Mecanismo de Alarme**

Quando um nó de trânsito detecta uma falha (*link down*) em uma de suas portas, esse nó envia um quadro *Link Down* de controle para o nó *master*, utilizando a VLAN de Controle.

Ao receber o quadro *Link Down*, o nó *master* muda do estado *normal* para o estado *ring-fault* e desbloqueia a porta secundária para o tráfego de dados. Todas as entradas nas tabelas de aprendizagem são deletadas em todos os nós do anel, sendo os nós de trânsito para isso instruídos pelo nó *master* mediante o envio de um quadro de controle.

Após deletar a sua tabela de *bridging*, cada nó do anel inicia o processo de aprendizagem da nova topologia.

- **Mecanismo de Polling**

Independentemente do estado do anel, o nó *master* envia periodicamente quadros *Health-Check* pela porta primária, que transitam ao longo de todo o anel.

Quando a rede, estando no estado *normal*, apresenta falha, obviamente esse quadro passa a não retornar para o nó *master*. Se o período sem retorno de quadros ultrapassar o *fail-period timer*, o nó master passa do estado *normal* em que se encontrava para o estado *ring-fault*.

Quando a rede é restabelecida, e um primeiro quadro *Health-Check* é retornado ao nó *master* pela porta secundária, o nó *master* comuta de volta para o estado *normal*. O nó *master* deve deletar a sua tabela de *bridging* e enviar um quadro de controle para que os demais nós do anel procedam da mesma forma e iniciem então os seus processos de aprendizagem da nova topologia.

Para evitar *loops* transientes, ocorre um estado *pre-forwarding*, em que os nós de trânsito mantêm as VLANs protegidas bloqueadas até o recebimento de instruções para deletar os registros de suas tabelas de aprendizagem. Após esse recebimento, tais registros são deletados e os nós retornam ao estado normal.

8.3.3.3. Outros Aspectos do EAPS

Um *switch Ethernet* habilitado como nó EAPS pode participar de mais de um anel, ou seja, pode pertencer simultaneamente a mais de um domínio EAPS. Nesse caso, aplica-se no *switch* um processo independente para cada domínio ao qual pertence.

Por outro lado, em um anel *Ethernet* podem ser configurados múltiplos domínios EAPS, cada um com o seu nó *master*. Dessa forma, facilita-se o reuso espacial da largura de banda do anel.

A RFC 3619 recomenda a utilização de criptografia no EAPS, com base em mecanismos definidos pelo IEEE. Alternativamente, recomenda também o uso de mecanismos de segurança em camadas superiores quando apropriado.

No encapsulamento dos quadros de controle do EAPS em um quadro Ethernet, os campos do MAC/LLC têm a seguinte semântica:

— O MAC-DA *multicast* é sempre igual a 0x00E02B000004;
— O valor *EtherType* é igual a 0x8100;
— O campo LSAP é igual a 0xAAAA (indicando o uso do SNAP);
— O campo OUI é igual a 0x00E02B;
— O protocolo EAPS é indentificado pelos códigos 0x00BB, 0x999 e 0x0B.

No interior do *payload* dos quadros de controle do EAPS destacam-se os seguintes campos:

— *EAPS Version*: valor atual é igual a 0x0001;
— *EAPSTYPE* (*EAPS Type*): indicador de cada tipo de quadro de controle do EAPS;
— *CTRL_VLAN_ID*: VLAN ID da VLAN de Controle;
— *SYSTEM_MAC_ADDRESS*: Endereço MAC do nó que envia o quadro de controle;
— *HELLO_TIMER:* valor setado pelo nó *master*;
— *FAIL_TIMER*: valor setado pelo nó *master*;
— *STATE:* valor indicativo do estado em que se enconta o anel EAPS.

8.3.4. **Padrão IEC 62439**

A IEC (*International Electrotechnical Comission*) emitiu o padrão IEC 61850 (*Communication Networks and Systems in Substations*), que se constitui na referência para a automação de sub-estações elétricas, possibilitando, de forma padronizada pela primeira vez, a engenharia de proteção, de medição e de controle. São utilizados com esse propósito os equipamentos referidos como *Intelligent Eletronic Devices* (IEDs). Dessa forma, equipamentos de controle de diferentes fabricantes passaram a interoperar utilizando uma rede Ethernet, um barramento (*bus*) de estação ou um barramento de processo.

No entanto, quando se requer um alto nível de disponibilidade, o padrão IEC 61850 não satisfaz plenamente no que tange à redundância de rede. De início, essa lacuna foi preenchida por soluções proprietárias, o que obviamente não satisfazia plenamente as necessidades das operadoras de sistemas elétricos.

Foi publicado então pelo comitê SC65WG15 da IEC, o padrão IEC 62439 (*Industrial Communication Networks – High Availability Automation Networks*), cuja última versão foi emitida em 2012. Esse especifica vários métodos de redundância, como mostra a Figura 8-48.

Protocolos	**Padrão IEC**	**Tempo de Recuperação**

Media Redundancy Protocol (MRP)	62439-2	10 a 500 ms
Parallel Redundancy Protocol (PRP)	62439-3	0
High-Availability Seamless Redundancy (HSR)	62439-3	0
Cross-Network Redundancy Protocol (CRP)	62439-4	400 ms a 1 s
Beacon Redundancy Protocol (BRP)	62439-5	10 ms
Distributed Redundancy Protocol (DRP)	62439-6	100 ms

Figura 8-48 – Protocolos de redundância do padrão IEC 62439.

Nessa figura encontram-se os sub-padrões que especificam cada tipo de protocolo, assim como valores aproximados dos tempos necessários para a recuperação da rede após a ocorrência de falha.

A IEC optou, para subestações elétricas, pelo protocolo PRP (*Parallel Rendundancy Protocol*) para a proteção de redes paralelas. Para redes em anel, o IEC adotou uma variação do PRP, que é o protocolo HSR (*High-Availability Seamless Redundancy*). Esses protocolos operam sem interrupção da rede em caso de falha (*seamless rendundancy*), com o tempo de recuperação portanto igual a zero.

Os demais protocolos do padrão IEC 62439, assim como os próprios PRP e HSR, podem ser aplicados em ambientes industriais , com utilização de redes Ethernet industriais. Dentre esses protocolos merece destaque o MRP (*Media Redundancy Protocol*), que será abordado adiante neste item.

8.3.4.1. PRP (Paralell Redundancy Protocol)

O PRP é uma forma de constituição de redes Ethernet industriais em paralelo, padronizada na Cláusula 4 do padrão IEC 62439-3, que possibilita a continuidade do tráfego de dados sem interrupção (*seamless*) em caso de ocorrência de falhas em uma das redes paralelas.

O PRP é aplicável à maioria das aplicações de Ethernet Industrial devido à sua independência de protocolos. Dentre tais aplicações encontra-se a automação de subestações elétricas.

No PRP, cada estação possui duas portas Ethernet conectadas a duas diferentes *bridged Ethernet LANs* de topologias similares. Essas *bridged Ethernet LANs* devem ser totalmente independentes no que se refere a falhas.

A Figura 8-49 apresenta um exemplo de operação do PRP em redes Ethernet paralelas duplicadas.

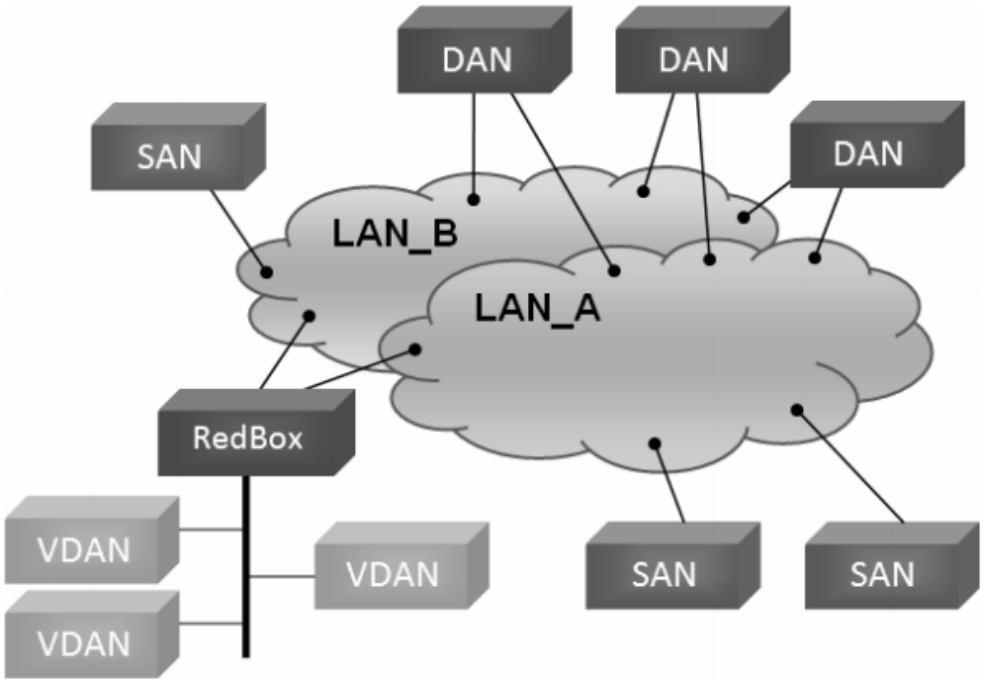


Figura 8-49 – Exemplo de PRP em redes Ethernet paralelas duplicadas.

Essa figura foi extraída do tutorial "Tutorial on Parallel Redundancy Protocol (PRP)" do professor Hans Weibel.

Nessa figura, as estações com acesso às duas redes Ethernet paralelas são referidas como DANs (*Doubly Attached Nodes*). Cada DAN possui duas portas, cada uma delas conectadas a uma das *bridged Ethernet LANs*.

Os DANs são também referidos como DANPs (*Doubly Attached Nodes Using PRP*).

Um DAN possui um gerenciador de redundância, que distribui as informações para as duas portas de acesso às redes paralelas.

O tráfego gerado pelos DANs é enviado simultaneamente pelas duas portas, nas duas redes paralelas. Da mesma forma, as estações DANs recebem duas cópias dos quadros enviados, também pelas suas duas portas.

Em condições normais de operação, a primeira cópia recebida por uma estação é processada, enquanto a segunda cópia é descartada.

Em caso de falha em uma das redes que afete a recepção de um quadro, a cópia que continua sendo recebida é processada normalmente.

Podem ser também utilizadas estações que se conectam apenas a uma das redes paralelas. Essas estações são referidas como SANs (*Single Attached Nodes*). É óbvio que o tráfego dessas estações não usufruem da redundância e da consequente proteção oferecidas pelo PRP.

Um ou mais SANs, contudo, mediante a sua adaptação para se tornar um VDAN (*Virtual Doubly Attached Node*), podem se conectar de forma dupla às redes

paralelas através de um dispositivo referido como *Redundancy Box* (*Red Box*). A figura anterior ilustra o uso de um *Red Box*.

É possível a utilização de barramentos de estação duplicados em paralelo com o PRP. Esses barramentos podem se constituir em topologias abertas ou em topologias em anel (o que aumenta o nível de rendundância da rede).

8.3.4.2. HSR (High-Availability Seamless Redundancy)

O HSR, padronizado na Cláusula 5 do padrão IEC 62439-3, adota os mesmos princípios que a PRP, mas diferentemente, utiliza tipicamente uma topologia em um único anel que possibilita operação sem *switches* dedicados, o que torna a rede mais econômica.

No HSR, as estações, que operam com duas portas, possuem um elemento de comutação (*switch element*), que ordena a transmissão de quadros para o anel.

A Figura 8-50 apresenta um exemplo de operação do HSR.

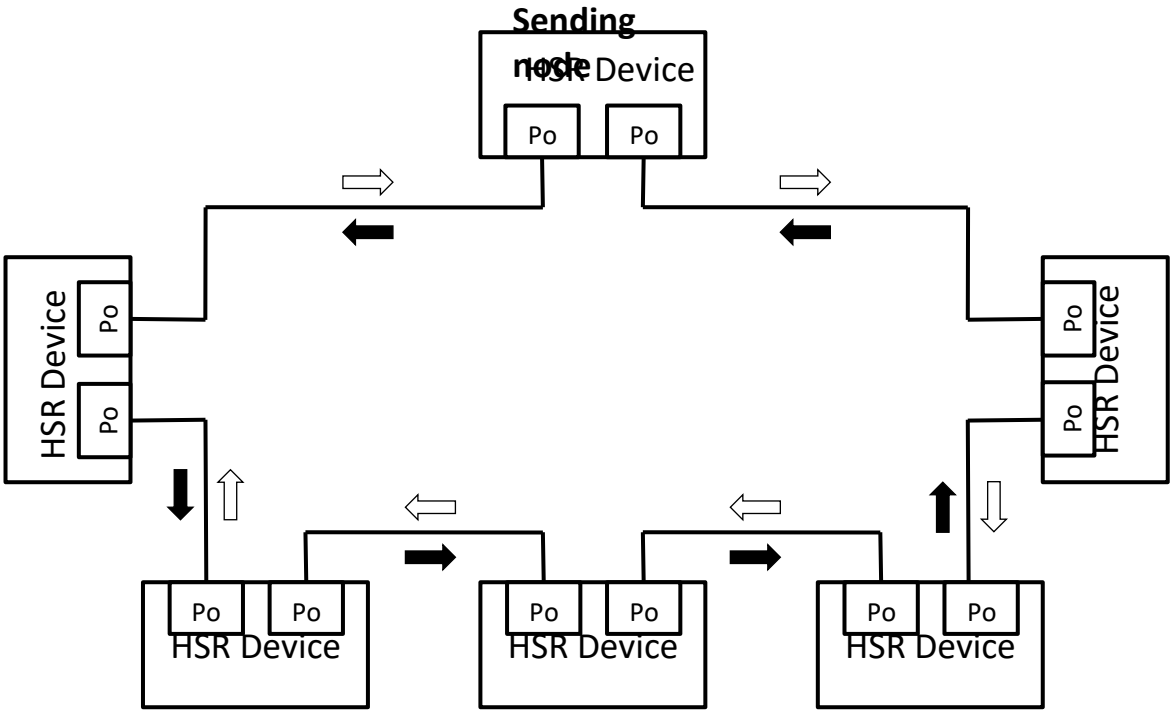


Figura 8-50 – Exemplo de operação do HSR.

Como no caso do PRP, duas cópias de um quadro de dados transmitido no HSR atingem o nó receptor, sendo a primeira a chegar utilizada e a segunda descartada. Caso um quadro de dados retorne ao nó que o enviou, esse quadro de dados é descartado, mitigando-se, dessa forma, a enventual ocorrência de *loops*.

Em caso de falha em um dos *links* do anel, uma das cópias continuará chegando ao receptor. No caso de duas ou mais falhas, o anel fica impedido de atender a totalidade das estações.

8.3.4.3. MRP (Media Redundancy Protocol)

O MRP é um protocolo especificado no padrão IEC 62439-2, utilizado por diversos *Ethernet switches* industriais disponibilizados no mercado.

O MRP utiliza uma topologia em anel que, na realidade, opera sempre como uma topologia linear.

Em um "anel", um dos *switches* desempenha a função MRM (*Media Redundancy Manager*), enquanto os demais *switches* constituem-se em MRCs (*Media Redundancy Clients*). As portas do MRM e dos MRCs suportam três estados:

- *Disabled* (desabilitado): a porta descarta todos os quadros recebidos;
- *Blocked* (bloqueado): a porta descarta todos os quadros recebidos exceto quadros de controle do MRP;
- *Forwarding* (transmitindo): a porta aceita ou retransmite todos os quadros recebidos.

A rede, por sua vez, pode se encontrar no estado *Ring-Closed* ou no estado *Ring-Open.*

A Figura 8-51 apresenta um exemplo de rede operando com o MRP no estado *Ring Closed*.

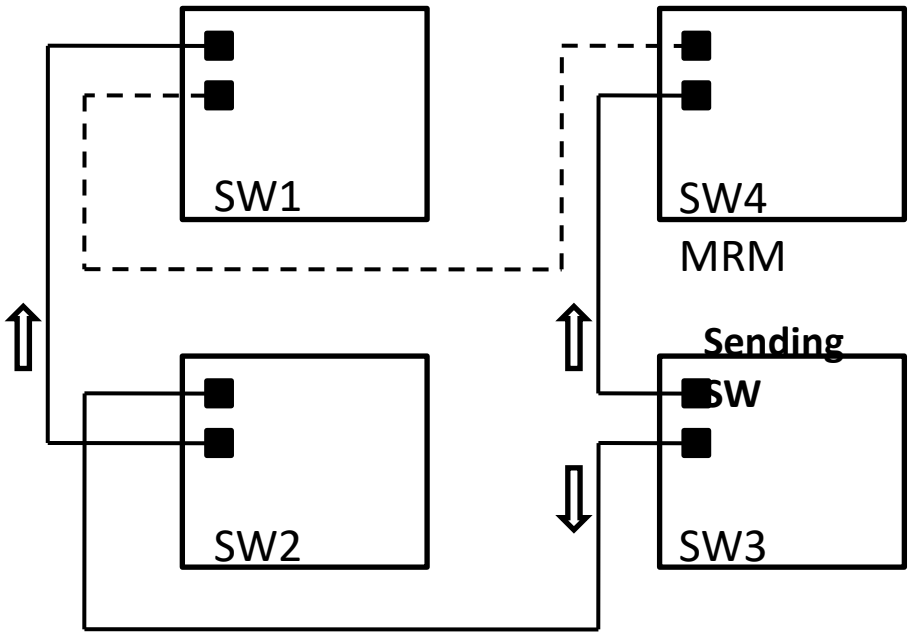


Figura 8-51 – Rede operando com MRP no estado Ring-Closed.

Como se observa nessa figura, uma das portas do MRM, assim como a porta correspondente em um dos MRCs vizinhos do MRM, encontram-se no estado *blocked*. Dessa forma, todos os *switches* da rede podem intercambiar quadros Ethernet sem a ocorrência de *loops*.

Caso um dos *links* falhe, a rede entra no estado *Ring-Open,* como mostra a Figura 8-52.

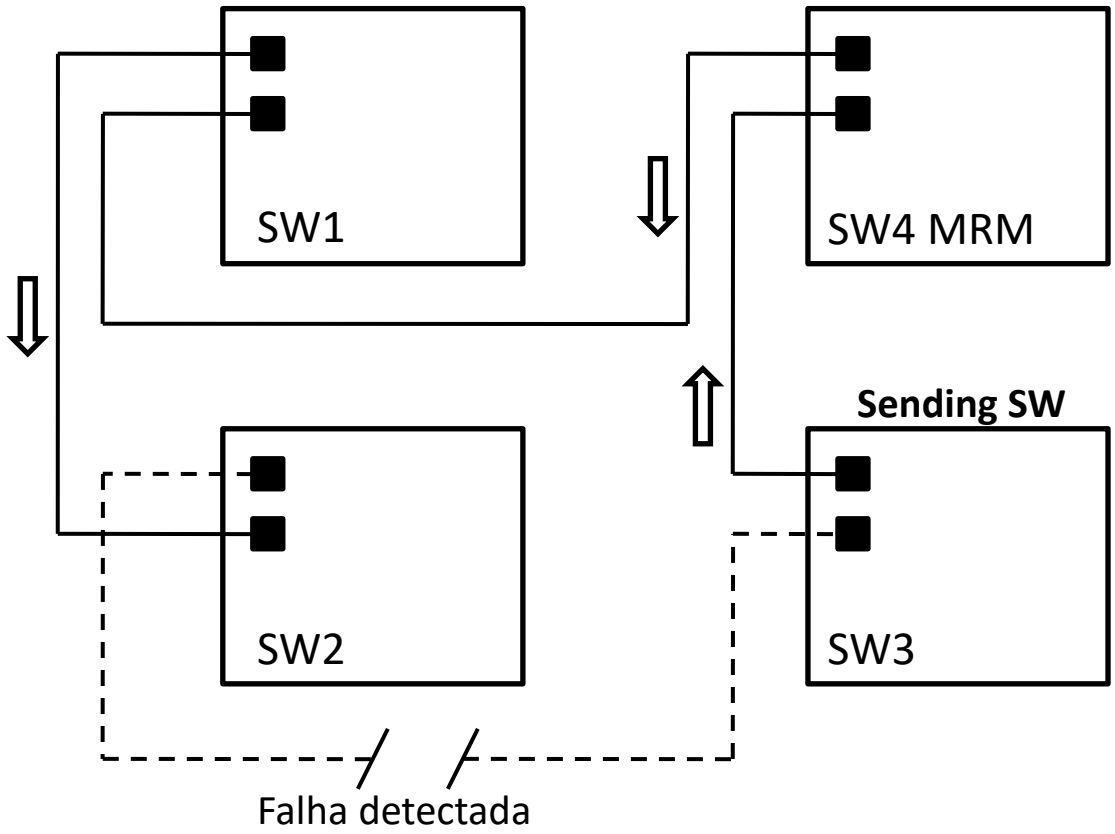


Figura 8-52 – Rede da figura anterior no estado Ring-Open.

Como se observa nessa figura, ocorreram na rede, sob o comando do MRM, as seguintes ações:

— A porta do MRM e a porta do MRC vizinho (SW1) que se encontravam no estado *blocked*, passaram para o estado *forwarding*;
— As portas do MRC SW2 e do MRC SW3 que terminavam o *link* que falhou passam do estado *forwarding* para o estado *blocked*.

Dessa forma, a rede continua plenamente operacional, sem a possibilidade de *loops*, tendo ocorrido apenas a alternância de *links* em operação.

8.4. PADRÃO IEEE 802.1AB-2009

O padrão IEEE 802.1AB (*Station and Media Access Control Connectivity Discovery*), cuja última edição ocorreu em 2009 (IEEE 802.1AB-2009), especifica o Protocolo Descobrimento da Camada de Enlace de Dados (*Link Layer Discovery Protocol*), referido como LLDP.

Para a operação harmoniosa de uma *bridged* LAN, é imprescindível a correta configuração dos diferentes parâmetros operacionais em seus sistemas (*bridges*, particularmente). Essa questão vem se agravando com a expansão dessas redes e pelo aumento do número de parâmetros operacionais a serem configurados.

Muitos desses parâmetros operacionais requerem a sua configuração de forma associada em diferentes sistemas da rede. Qualquer discrepância na configuração associada torna-se um entrave para a operação da rede.

Tome-se como exemplo os valores de VID (inclusive possivelmente um PVID) configurados em uma porta de saída de uma *bridge*. Esses valores devem ser igualmente configurados na porta de entrada da *bridge* vizinha conectada ao *link* comum entre essas *bridges*.

É óbvio que o armazenamento dos valores de parâmetros configurados para cada sistema de forma descentralizada dificulta a consistência entre esses valores, caso os demais sistemas que também necessitem configurar esses parâmetros não tenham conhecimento dos valores configurados.

Além de dificultar a configuração harmoniosa de parâmetros em diferentes sistemas , a descentralização do processo de configuração torna mais dificultosa e lenta a identificação de erros de configuração e de sua retificação.

O LLDP permite que um sistema de uma *bridged* LAN informe, aos demais sistemas, as configurações por ele realizadas e os correspondentes valores configurados. Em consequência do LLDP, cada sistema da *bridged* LAN passa a registrar, além das configurações por ele realizadas, as configurações realizadass pelos sistemas remotos pares, nos diferentes processos . Isso naturalmente possibilita a harmonia entre as configurações e a verificação de erros de configuração para a tomada de medidas corretivas.

As informações distribuídas por meio do LLDP são armazenadas pelos receptores em uma MIB padronizada, de modo a permitir que essas informações possam ser acessadas mais facilmente por um NMS (*Network Management System*) utilizando um protocolo de gerenciamento tal como o SNMP (*Simple Network Management Protocol*).

8.4.1.1. Princípios de Operação

O protocolo LLDP foi concebido para divulgar informações úteis para o descobrimento (*Discovery*) de informações significativas a respeito de uma porta remota e para alimentar MIBs de topologia, assim como para prover um meio para repassar as informações descobertas para entidades de gerenciamento de camadas superiores. As informações distribuídas pelas LLDPDUs são armazenadas em uma ou mais MIBs.

O LLDP é um protocolo de camada de controle de enlace que possibilita a uma estação de uma LAN IEEE 802 divulgar os recursos e os estados atuais do sistema associado com um MSAP (*Mac Service Access Point)*. O MSAP provê o serviço MAC para uma entidade LLC (*Logical Link Control*), e essa entidade LLC

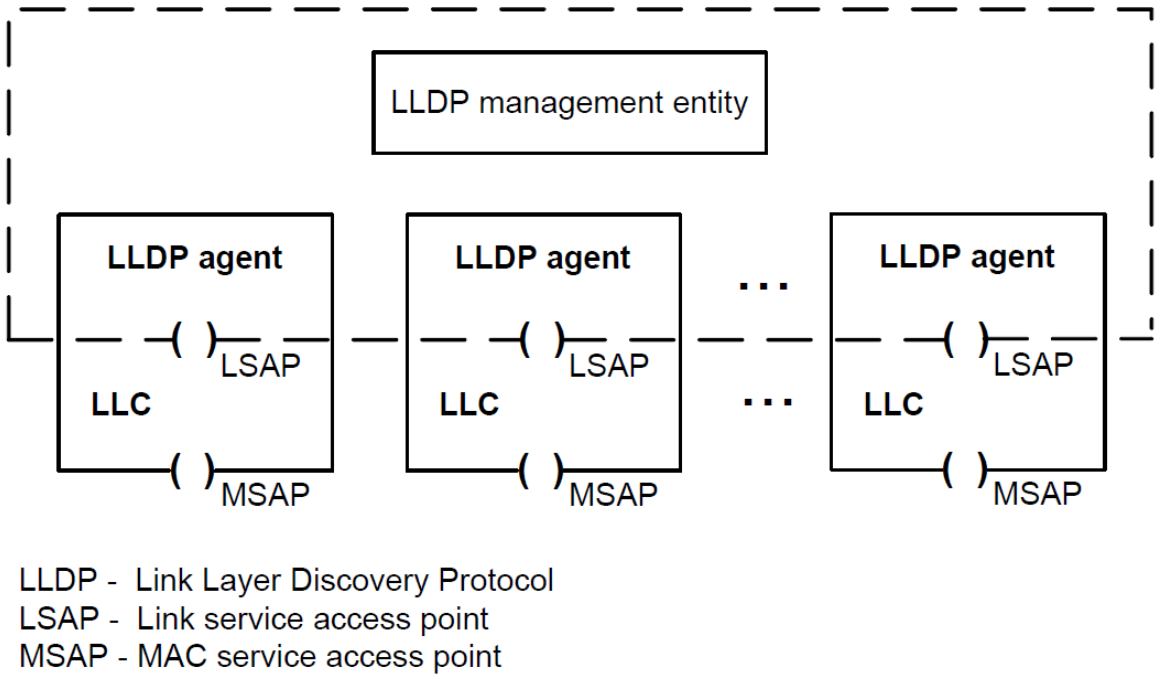
provê um LSAP (*Link Service Access Point)* para um Agente LLDP que transmite e recebe informações para e de Agentes LLDP de outras estações conectadas à mesma LAN IEEE 802.

Um sistema que implementa o LLDP suporta um ou mais Agentes LLDP.

Observa-se que a cada Agente LLDP corresponde um único LSAP, e a este LSAP corresponde um único MSAP, de modo que um Agente LLDP pode ser identificado pelo respectivo LSAP ou pelo respectivo MSAP.

Um dado sistema pode requerer o uso de múltiplos Agentes LLDP porque esse sistema abrange naturalmente múltiplos MSAPs (como no exemplo de uma *bridge* que possua no mínimo duas portas) ou porque ele necessita ser capaz de transmitir diferentes informações (ou seja, diferentes conjuntos de TLVs) em suas LLDPDUs para diferentes conjuntos de Agentes LLDP pares.

A Figura 8-53 ilustra um sistema com múltiplos Agentes LLDP e a relação entre os Agentes LLDP, as Entidades LLC e os MSAPs, e a Entidade de Gerenciamento LLDP (*LLDP Management Entity*), desse sistema.



LLDP - Link Layer Discovery Protocol
LSAP - Link service access point
MSAP - MAC service access point

Figura 8-53 – Sistema com múltiplos Agentes LLDP.

Para a sua plena utilização, o LLDP opera segundo o diagrama de relacionamentos da Figura 8-54.

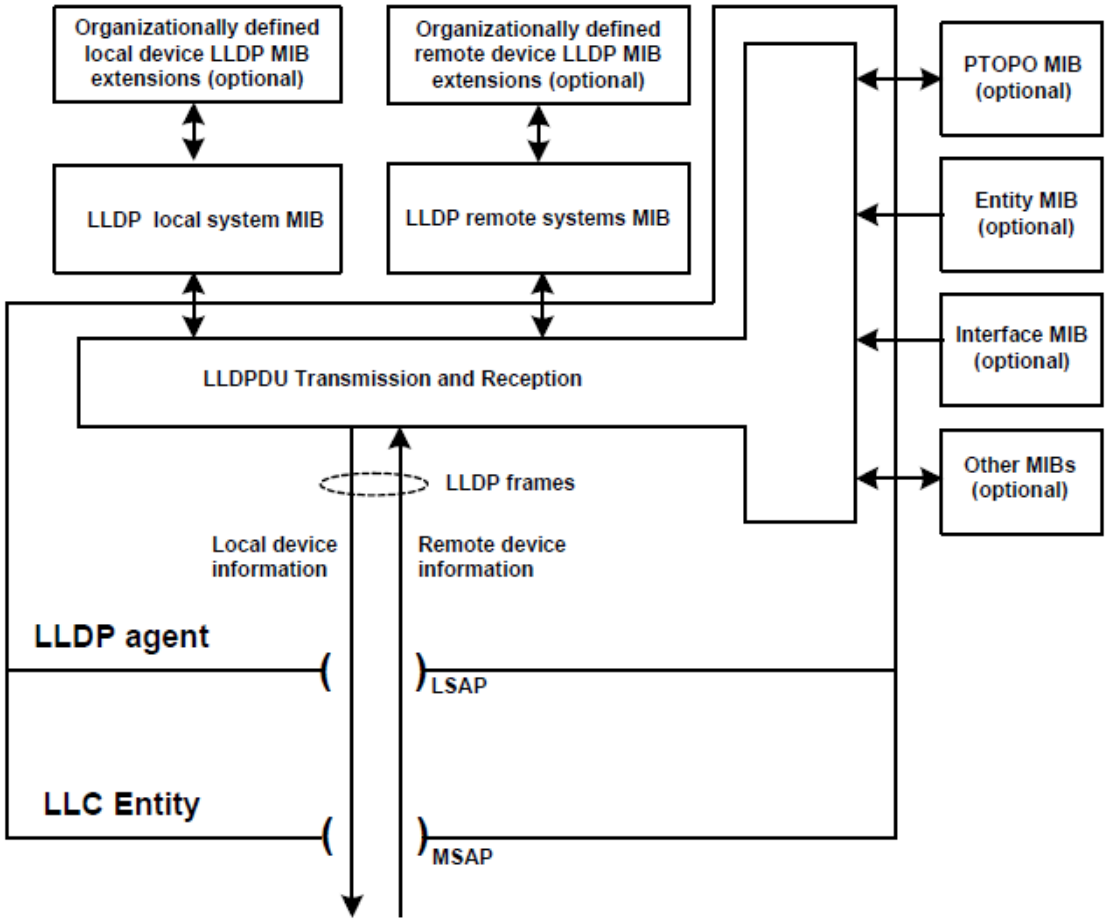


Figura 8-54 – Diagrama de relacionamentos do LLDP.

Conforme essa figura, um sistema possui os dois módulos abaixo relacionados, que são constituidos por LLDP MIBs, e representam a Entidade de Gerenciamento de LLDP desse sistema:

— Módulo LLDP de Transmissão: esse módulo é constituído pela LLDP *Local System MIB* e, opcionalmente, por uma *Organizationally Defined Local Device LLDP MIB Extension*;

— Módulo LLDP de Recepção: esse módulo é constituído pela LLDP *Remote Systems MIB* e, opcionalmente, por uma *Organizationally Defined Remote Device LLDP MIB Extension*.

As configurações realizadas em um sistema são armazenadas na LLDP *Local System MIB* desse sistema. Para a plena alimentação dessa MIB, o sistema utiliza informações derivadas de outras MIBs padronizadas.

Dentre tais MIBs padronizadas encontram-se a *Physical Topology* (PTOPO) *MIB* (RFC 2922), a *Entity MIB* (RFC 2737) e a *Interfaces MIB* (RFC 2863), além de outras MIBs opcionalmente utilizadas para alimentar a *LLDP Local System MIB*.

A transmissão de uma LLDPDU pode ser iniciada pela expiração de um contador de tempo (*countdown*) de transmissão ou por uma mudança de valor de

um ou mais elementos de informação (objetos gerenciados) associados com o sistema local.

Quando um ciclo de transmissão é iniciado, a Entidade de Gerenciamento de LLDP do sistema local extrai os objetos gerenciados da *LLDP Local System MIB*, e formata essas informações como TLVs (*Type, Length, Values*). Tais TLVs são inseridos em uma LLDPDU, que é passada para o respectivo Agente LLDP, que adiciona parâmetros de endereçamento e a transmite pela devida porta, utilizando as camadas Ethernet inferiores (LLC, MAC e Camada Física).

Do lado da recepção, o Agente LDP que recebe uma LLDPDU, após processar e remover os parâmetros de endereçamento acrescidos pelo Agente LLDP transmissor, repassa as informações contidas nessa LLDPDU para a Entidade de Gerenciamento de LLDP do sistema remoto receptor.

Essa entidade valida os TLVs contidos na LLDPDU recebida e insere as informações contidas nesses TLVs na *LLDP Remote MIB*.

Se a comparação entre informações contidas na *LLDP Local System MIB* de um sistema e aquelas contidas na *LLDP Remote System MIB* desse sistema apresentar inconsistências, tais inconsistências podem ser retificadas pela utilização de processos adequados. Tais processos, contudo, não são especificados pelo padrão IEEE 802.1AB-2009, sendo definidos proprietariamente de forma específica para cada implementação.

8.4.1.2. Endereçamento de LLDPDUs

O LLDP utiliza o serviço provido pelo LLDP/LSAP e pelo LLC para transmitir e receber LLDPDUs. Cada LLDPDU é transmitida como uma única solicitação (*request*) ao serviço MAC por uma entidade LLC que utiliza uma única instância do serviço MAC provido em um MSAP. Cada LLDPDU é recebida no MSAP pela entidade LLC como uma indicação (*indication*) do serviço MAC.

O valor de *EtherType* (igual 88-CC) que identifica o LLDP pode ser utilizado de duas diferentes formas, em função do método de controle de acesso ao meio da rede Ethernet utilizada.

Quando a entidade LLC utiliza um MSAP que é suportado por um método de controle de acesso ao meio que suporta diretamente a codificação *EtherType* (*direct-encoded LLDP*), como é o caso de IEEE 802.3, a entidade LLC codifica o *EtherType* (88-CC) correspondente a LLDP no cabeçalho de dois octetos da MSDU (*Mac Service Data Unit*) da respectiva solicitação de serviço MAC (*MAC service request*). Esse valor *EtherType* passa a constar do cabeçalho da subcamada MAC.

Quando a entidade LLC utiliza um MSAP que é suportado por um método de controle de acesso ao meio que não suporta diretamente a codificação *EtherType*, como no exemplo de redes IEEE 802.11, o LLC é formatado via SNAP (*SNAP-encoded LLDP*), com a codificação AA-AA-03-00-00-00 (indicativa da codificação *EtherType*), sucedida pelo valor igual a 88-CC que indica o uso do LLDP.

O conjunto de endereços MAC de grupo que podem ser utilizados pelo LLDP estão especificados na Figura 8-55.

Name	Value	Purpose
Nearest bridge	01-80-C2-00-00-0E	Propagation constrained to a single physical link; stopped by all types of bridge
Nearest non-TPMR bridge	01-80-C2-00-00-03	Propagation constrained by all bridges other than TPMRs; intended for use within provider bridged networks
Nearest Customer Bridge	01-80-C2-00-00-00	Propagation constrained by customer bridges; this gives the same coverage as a customer-customer MACSec connection

Figura 8-55 - Endereços MAC de grupo utilizados pelo LLDP.

A escolha do endereço MAC de grupo utilizado realiza-se em função do alcance da propagação de LLDPDUs dentro de uma bridged LAN, como a seguir:

— O endereço MAC de grupo *nearest bridge* é um endereço que nenhum tipo de *bridge* pode retransmitir. Consequentemente, a comunicação de LLDPDUs restringe-se às estações conectadas a um segmento de LAN;
— O endereço MAC de grupo *nearest non-TPMR* (n*on two-port MAC relay*) *bridge* é um endereço que nenhum tipo de *bridge*, à exceção das *TPMR Bridges*, pode retransmitir;
— O endereço MAC de grupo *nearest customer bridge* é um endereço que apenas as *S-VLAN bridges* e as *TPMR bridges* podem retransmitir. As *C-VLAN bridges* e as *provider edge bridges*, que juntas constituem as denominadas *costumer bridges*, são impossibilitadas de retransmitir esse tipo de endereço MAC de grupo.

8.4.1.3. Formatação de LLDPDUs e Tipos de TLV

A LLDPDU deve conter uma sequência ordenada de três TLVS obrigatórios seguida de zero ou mais TLVs opcionais, além de mais um último TLV obrigatório referido como End Of LLDPDU TLV, conforme a Figura 8-56.

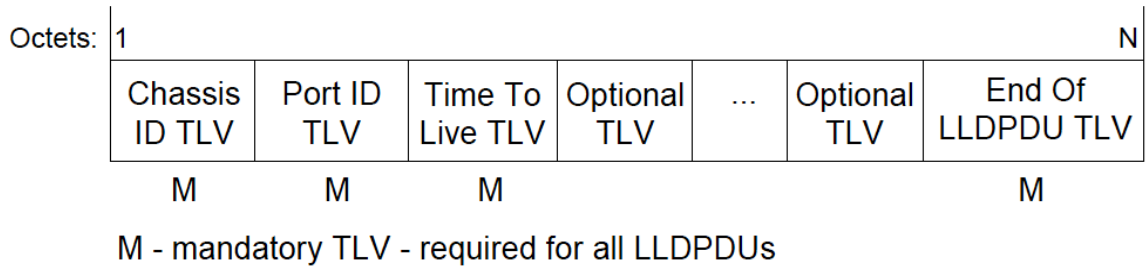


Figura 8-56 - Formato da LLDPDU.

Um subconjunto de TLVs é considerado básico para o gerenciamento das estações da rede e constitui um recurso requerido para todas as implementações de LLDP. Outros subconjuntos representam TLVs específicos para organizações, que são definidos por grupos de padronização tais como o IEEE 802.1 e o IEEE 802.3, e se destinam a operar com protocolos e/ou meios particulares.

Na Figura 8-57 encontra-se a relação dos tipos de TLV utilizados no LLDP com os respectivos valores de identificação.

TLV type [a]	TLV name	Usage in LLDPDU	Reference
0	End Of LLDPDU	Mandatory	8.5.1
1	Chassis ID	Mandatory	8.5.2
2	Port ID	Mandatory	8.5.3
3	Time To Live	Mandatory	8.5.4
4	Port Description	Optional	8.5.5
5	System Name	Optional	8.5.6
6	System Description	Optional	8.5.7
7	System Capabilities	Optional	8.5.8
8	Management Address	Optional	8.5.9
9–126	Reserved for future standardization	—	—
127	Organizationally Specific TLVs	Optional	Annex Annex

[a]TLVs with type values 0–8 are members of the basic management set.

Figura 8-57 - Relação dos tipos de TLV.

As referências indicadas nessa figura dizem respeito aos respectivos itens do padrão IEEE 802.1AB-2009.

Os quatro TLVs obrigatórios têm os seguintes significados:

— *Chassis ID TLV*: identifica o *chassis* do dispositivo da *bridge* LAN 802. Esse TLV representa o componente de *chassis* dos identificadores MSAP das portas dos dispositivos;

— *Port ID TLV*: identifica a porta pela qual a LLDPDU é transmitida. Esse TLV representa o componente de porta do identificador MSAP da porta do dispositivo;
— *TTL (Time to Live*): indica o período de tempo (em segundos) durante o qual as informações contidas na LLDPDU devem ser consideradas válidas;
— *End Of LLDPDU TLV*: esse TLV encerra obrigatoriamente as sequências de TLVs de todas as LLDPDUs.

Como visto acima, a cada Agente de LLDP corresponde um valor de identificador de MSAP. O valor de identificador de MSAP de uma porta de dispositivo é formado por um componente de *chassis* desse dispositivo (representado pelo *chassis ID*) e por um componente de porta do dispositivo (representado pelo *port ID*).

No caso de recebimento de um TTL TLV com valor zero em uma LLDPDU, o sistema receptor dessa LLDPDU deve remover de sua *LLDP Remote System MIB*, todas as informações relativas à *bridge* transmissora da LLDPDU.

Em qualquer dos TLVs específicos para organizações, cujo tipo de TLV é sempre igual a 127, cada organização é identificada pelo campo OUI (*Organizationally Unique Identifier*), com três octetos, no primeiro subcampo do campo *TLV Information String* (que possui de 4 a 511 octetos).

Os TLVs não obrigatórios relacionados nessa figura são distribuídos por três diferentes conjuntos:

— Um conjunto formado pelos *Basic Management TLVs*: são os TLVs cujos valores de tipo se encontram na faixa de 4 a 8 na Figura 8-57;
— Conjunto de TLVs IEEE 802.1 *Organizationally Specific*;
— Conjunto de TLVs IEEE 802.3 *Organizationally Specific.*

Os TLVs do conjunto IEEE 802.1 *Organizationally Specific* encontram-se na Figura 8-58:

IEEE 802.1 subtype	TLV name
01	Port VLAN ID
02	Port And Protocol VLAN ID
03	VLAN Name
04	Protocol Identity
05	VID Usage Digest
06	Management VID
07	Link Aggregation
08–FF	Reserved

Figura 8-58 – IEEE 802.1 Organizationally Specific TLVs.

O *Port VLAN ID*, por exemplo, é um TLV opcional de tamanho fixo que permite a uma porta de uma VLAN *bridge* informar o valor configurado de PVID (*Port`s VLAN Identifier*), PVID esse que representa o VLAN ID a ser atribuído a quadros *untagged* ou a quadros *priority-tagged* que ingressam por uma porta de *bridge.*

Os TLVs do conjunto IEEE 802.3 *Organizationally Specific* encontram-se na Figura 8-59:

IEEE 802.3 subtype	TLV name
1	MAC/PHY Configuration/Status
2	Power Via Medium Dependent Interface (MDI)
3	Link Aggregation (deprecated)
4	Maximum Frame Size
5–255	Reserved

Figura 8-59 – IEEE 802.3 Organizationally Specific TLVs.

As LLDPDUs podem ser do tipo normal ou do tipo *shutdown*.LLDPDUs normais provêem informações de gerenciamento a respeito de uma estação para as estações vizinhas. Esse tipo de LLDPDU contém os quatro TLVs obrigatórios, e podem conter TLVs opcionais.

As *Shutdown LLDPDUs* são utilizados em um procedimento especial que existe para o caso em que um Agente LLDP toma conhecimento de que uma porta associada encontra-se na iminência de tornar-se não operacional. Nesse caso, o Agente LLDP tenta enviar uma *shutdown LLDPDU* final com o intuito de comunicar

a sua desabilitação. As *shutdown LLDPDUs* contêm exclusivamente os quatros TLVs obrigatórios.

8.4.2. **LLDP-Media Endpoint Discovery (LLDP-MED)**

O protocol LLDP-MED é uma extensão do LLDP, definida no padrão ANSI/TIA-1057 em 2006, que objetiva o tratamento de algumas questões que surgiram com relação à comunicação entre dispositivos de *media* (*media devides*) em *bridged LANs*. Destacam-se, dentre esses dispositivos de *media*, aparelhos de voz utilizados em VoIP.

O LLDP-MED utiliza alguns TLVs do LDP com extensões. Um exemplo desses TLVs do LLDP extendidos para o LLDP-MED é o *Power via Media Dependent Interface TLV* (MDI TLV), pela importância que a potência de transmissão assume no LLDP-MED.

Além do uso de TLVs do LLDP extendidos, o padrão ANSI/TIA-1057 define também TLVs opcionais específicos para o LLDP-MED.

8.5. **PADRÃO IEEE 802.1AX-2008 (LINK AGGREGATION)**

O padrão IEEE 802.1AX-2008 (*Link Aggregation*), que obsoletou e substituiu o padrão IEEE 802.3ad, tem como propósito a definição do mecanismo Agregação de *Links* (*Link Aggregation*), também referido como *Link Agg*.

8.5.1. **Considerações Gerais**

O mecanismo de agregação de *links* possibilita o estabelecimento de agregados de *links*, de forma tal que tais agregados, referidos como *Link Aggregation Groups* (LAGs), possam ser tratados como um único *link* com maior vazão de tráfego.

Com esse propósito, o padrão IEEE 802.1AX-2008 especifica o estabelecimento de *links* lógicos que consitem de *n* instâncias paralelas de *links full-duplex* ponto a ponto operando à mesma velocidade. Define também a capacidade de agregação de *links* independentemente da camada MAC, além de informações gerais relevantes para tipos de MAC que suportam o mecanismo Agregação de *Links*.

A ocorrência de agregação de *links* tem as seguintes condicionantes:

— Os *links* devem ser ponto a ponto com o MAC operando no modo *full-duplex;*
— Os *links* devem utilizar exclusivamente o IEEE 802.3 MAC (ou seja, *Ethernet);*
— Todos os *links* agregados devem operar na mesma velocidade de transmissão.

Os LAGs são identificados pelos respectivos LAG IDs.

As funções de agregação de *links* efetivam-se na subcamada opcional Agregação de *Links*, que se posiciona na arquitetura IEEE 802.3 conforme a Figura 8-60.

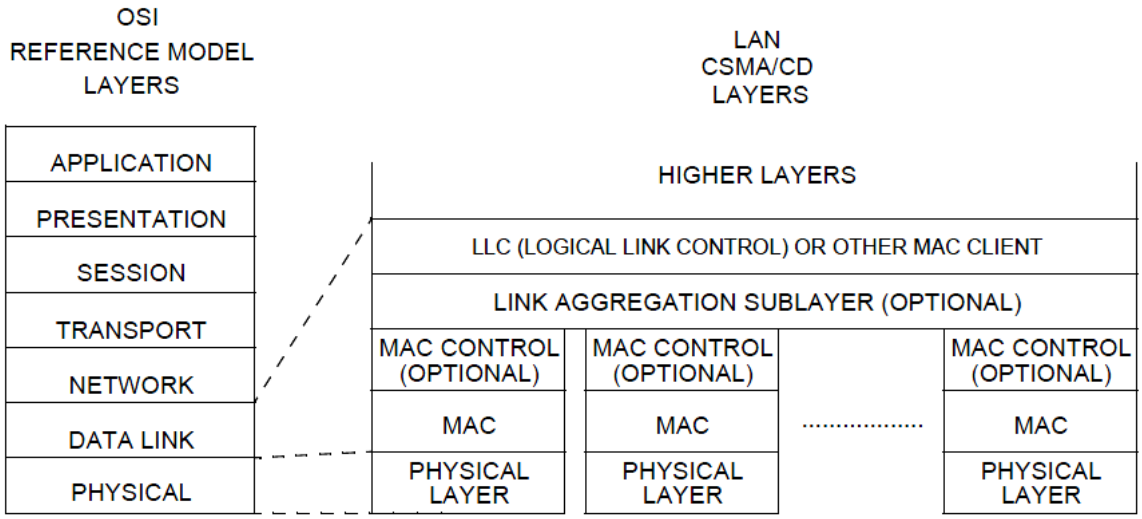


Figura 8-60 – Posicionamento da subcamada opcional Agregação de Links.

É possível o uso opcional da subcamada Agregação de *Links* para algumas portas de um sistema, sem utilizá-la, contudo, para as demais portas.

Não é permitida a duplicação de quadros de dados em *links* agregrados. A ordem dos quadros de dados transmitidos deve ser preservada.

A especificação do serviço MAC no padrão ISO/IEC 15802-1 estabelece que a ordem deve ser preservada para quadros de dados com uma *tuple* [MAC-SA, MAC-DA, *priority*]. A rigor, porém, podem existir múltiplas conversações independentes entre um par MAC-SA/MAC-DA em uma mesma prioridade. O requisito fundamental é o de preservar a ordem de quadros de dados em uma conversação.

O padrão IEEE 802.1AX-2008 define o *Link Aggregation Control Protocol* (LACP) e o *Marker Protocol*.

8.5.2. LACP (Link Aggregation Control Protocol)

O LACP estabelece formas padronizadas para a troca de informações entre as terminações de um *link*, de forma a possibilitar às suas instâncias *Link Aggregation Control* as seguintes ações:

- Obter um acordo quanto à identificação do LAG ao qual o *link* vai ser incorporado;
- Incluir o *link* nesse LAG;
- Possibilitar as funções de transmissão e de recepção de quadros nesse *link*.

O LACP, assim como o protocolo *Marker*, são exemplos do grupo de protocolos referidos como *Slow Protocols,* abordado nos subitem 2.7.7 do capítulo 2 deste livro.

O LACP depende da transmissão de informações e de estados, e não de comandos. As LACPDUs são enviadas pela primeira parte (*Actor*), e são recebidas pela segunda parte (*Actor's Protocol Partner*). As informações enviadas são de conhecimento do *Actor*, tanto no que diz respeito a seu próprio estado quanto ao estado do *Partner*.

Uma LACPDU é encapsulada em um quadro MAC, cujos campos têm os seguintes significados:

- Endereço MAC-DA: é o endereço *Slow_Protocol_Multicast* (01-80-C2-00-00-02);
- Endereço MAC-SA: é o endereço individual associado à porta da *bridge* pela qual a LACPDU é transmitida;
- *Lenght/Type*: é sempre codificado como *Type* e transporta o parâmetro *Slow_Protocol_Type*, que corresponde ao valor *EtherType* 88-09.

No interior do campo de informações de cliente MAC do LACP destacam-se os seguintes campos:

- *Sub-type*: esse campo identifica o *Slow Protocol* sendo encapsulado. O LACP utiliza o valor de subtipo igual a 0x01;
- *Version Number*: a versão atual do LACP é identificada pelo valor 0x01;
- *TLV_Type = Actor Information*: Esse campo indica a natureza da informação do *Actor* transportada na LACPDU, sendo utilizado o valor 0x01;
- *TLV_Type = Partner Information*: Esse campo indica a natureza da informação do *Partner* transportada na LACPDU, sendo utilizado o valor 0x02.

8.5.3. **Marker Protocol**

Como no caso do LACP, o protocolo *Marker*, também opcional, utiliza a função *Actor* e a função *Actor's Protocol Partner* (ou simplesmente *Partner*).

A subcamada Agregação de *Links* do *Actor* utiliza a função *Frame Distribution* para o envio de solicitações (*Marker PDUs*). Tais PDUs são recebidas pela função *Frame Collection* do *Partner*.

Uma *Marker* PDU é respondida pelo *Partner* mediante o envio de uma *Marker Response* PDU pelo mesmo *link* pelo qual ocorreu a transmissão da *Marker* PDU. A *Marker Response* PDU é recebida pela função *Frame Distribution* do *Actor*.

O funcionamento do protocolo *Marker* encontra-se ilustrado na Figura 8-61.

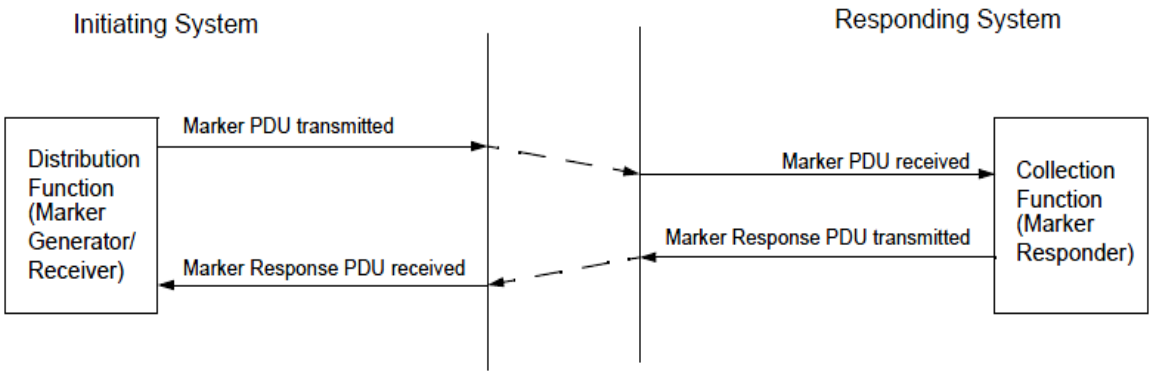


Figura 8-61 – Funcionamento do protocolo Marker.

As PDUs do protocolo *Marker* são envelopadas por um quadro Ethernet da mesma forma que as LACPDUs. O endereço MAC-DA é o endereço *Slow_Protocols_Multicast* (01-80-C2-00-00-02). O valor de *EtherType* utilizado é o *Slow_Protocols_Type* (88-09).
Na Figura 8-62 encontra-se representada a estrutura da *Marker PDU* e da *Marker Response PDU*.

Marker PDU	Octets	Marker Response PDU
Destination Address	6	Destination Address
Source Address	6	Source Address
Length/Type	2	Length/Type
Subtype = Marker	1	Subtype = Marker
Version Number	1	Version Number
TLV_type = Marker Information	1	TLV_type = Marker Response Information
Marker_Information_Length= 16	1	Marker_Response_Information_Length = 16
Requester_Port	2	Requester_Port
Requester_System	6	Requester_System
Requester_Transaction_ID	4	Requester_Transaction_ID
Pad = 0	2	Pad = 0
TLV_type = Terminator	1	TLV_type = Terminator
Terminator_Length = 0	1	Terminator_Length = 0
Reserved	90	Reserved
FCS	4	FCS

Figura 8-62 – Estrutura da Marker PDU e da Marker Response PDU.

O campo *Subtype* apresenta o valor 0x02, que identifica o protocolo *Marker* dentre os *slow protocols*. A versão atual é representada pelo valor 0x01.

Os demais campos encontram-se especificados no subitem 5.5.3 do padrão IEEE 802.1AX-2008.

8.5.4. **Função Frame Collection e Função Frame Distribution**

O anexo A do padrão IEEE 802.1AX-2008 descreve, em caráter informativo, o funcionamento das funções *Frame Collection* e *Frame Distribution*, que se encontram, respectivamente, no *Partner* e no *Actor*.

A função *Frame Collection* contém um *Frame Collector* e um *Marker Responder*.

O *Frame Collector* é responsável por repassar os quadros de dados recebidos nas diversas portas agregadas para o Cliente de MAC (*MAC Client*).

O *Marker Responder* é utilizado pelo protocolo *Marker*, recebendo as *Marker PDUs* e enviando, em resposta, as *Marker Responses PDUs* nas mesmas portas em que as *Marker PDUs* foram recebidas.

A função *Frame Distribution* contém um *Frame Distributor* e um *Marker Generator/Receiver* opcional.

O *Frame Distributor* é responsável por acolher os quadros de dados transmitidos pelo Cliente MAC, e repassá-los para transmissão na porta apropriada, com base no algoritmo de distribuição de quadros empregado.

O padrão IEEE 802.1AX-2008 não especifica qualquer algoritmo de distribuição de quadros, mas requer que o algoritmo adotado não venha a causar a perda de ordem de quadros de dados pertinentes a uma conversação, e que não possa causar a duplicação de quadros.

O *Marker Generator/Receiver* é opcional uma vez que o próprio protocolo *Marker* é opcional. Alguns algoritmos de distribuição de tráfego pelos *links* agregados podem não requerer o uso desse protocolo. Entretanto, a habilitação do *Partner* para receber e responder *Marker PDUs* é obrigatória.

Caso sejam utilizados, o *Marker Generator* é responsável pelo envio de *Marker PDUs*, enquanto o *Marker Receiver* é responsável pelo recebimento de *Marker Response PDUs*, como visto na Figura 8-61 anterior deste item.

Para a escolha do tipo da função *Distribution* a ser utilizada, são considerados diferentes cenários topológicos, como mostra a Figura 8-63.

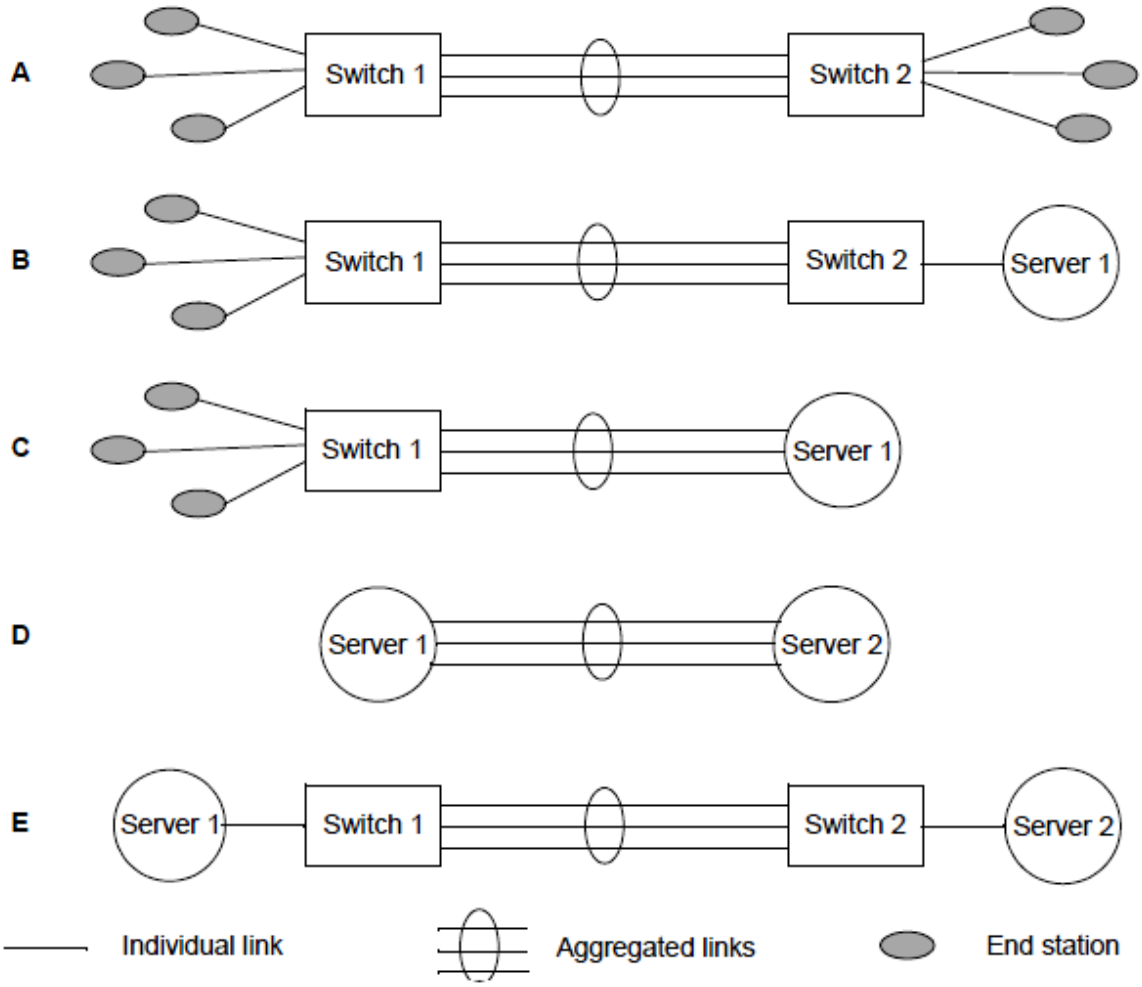


Figura 8-63 – Cenários topológicos para a escolha da função Distribution.

A escolha de um cenário topológico é obviamente importante para a definição, de forma proprietária, do algoritmo de distribuição de quadros de dados a ser utilizado.

No cenário A é possível, para cada *switch*, alocar conversações para *links* simplesmente com base no MAC-SA e no MAC-DA.

Nos cenários B e C, múltiplas estações finais comunicam-se com um único servidor pelo agregado de *links*. Nesses dois casos, a distribuição no sentido do servidor para as estações finais ocorre com base nos endereços MAC-DA (das estações finais). No sentido oposto, a distribuição ocorre também com base nos endereços das estações finais, agora, contudo, como endereços MAC-SA.

Nos cenários D e E, onde o tráfego começa e termina em servidores, a opção apropriada para o algoritmo de distribuição de conversações entre os *links* agregados é baseada em informações de camada superior (identificadores de *sockets* da camada de transporte, no caso).

No cenário E, alternativamente, é possível a alocação de conversações com base em endereços MAC, se os servidores forem capazes de utilizar os endereços MAC e de realizar a alocação.

O uso do protocolo *Marker* pode possibilitar à função *Distribution* um meio de determinar o ponto no qual um conjunto de conversações pode ser realocado com segurança de um *link* agregado para outro, sem o risco de ocorrência de perda do ordenamento dos quadros de dados desse conjunto de conversações.

O anexo A do padrão IEEE 802.1AX-2008 relaciona os procedimentos a serem adotados para que o protocolo *Marker* possa assegurar a não ocorrência de perda de ordem de quadros.

O protocolo *Marker* provê uma otimização que pode resultar em maior rapidez na realocação de conversações. Essa otimização é obtida pelo fato de que o protocolo *Marker* permite a realocação sem a necessidade de esgotamento do período de *timeout* utilizado.

8.6. GERÊNCIA DE BRIDGES

A Gerência de *bridges* encontra-se especificada na Cláusula 12 (*Bridge Management*) e na Cláusula 17 (*Management Information Base* – MIB), ambas do padrão IEEE 802.1Q-2014.

8.6.1. Bridge Management

A Cláusula 12 do padrão IEEE 802.1Q-2014 define o conjunto de objetos gerenciados e as suas funcionalidades, que permitem a configuração administrativa de VLANs.

Foram relacionadas e definidas as funções básicas de gerenciamento sob a ótica do gerenciamento de *bridge*:

— Gerência de Configuração (*Configuration Management);*
— Gerência de Falhas (*Fault Management*);
— Gerência de Desempenho (*Performance Management*);
— Gerência de Segurança (*Security Management*);
— Gerência de Contabilidade (*Accounting Management*).

Foram também relacionados os 20 tipos de informação (*data types*) utilizados na gerência de *bridges*, conforme o padrão IEEE 802.1Q-2014.

Foram então relacionados os recursos gerenciados de uma VLAN *bridge* ou de uma *bridge* com base nos processos e entidades apresentados na Seção 8.3 do próprio padrão IEEE 802.1Q-2014, como a seguir:

— *The Bridge Management Entity (12.4 and 8.12);*
— *The individual MAC Entities associated with each Bridge Port (12.5, 8.2, and 8.5);*
— *The Forwarding Process of the MAC Relay Entity (12.6, 8.2, and 8.6);*

— *The FDB of the MAC Relay Entity (12.7 and 8.8);*
— *The Bridge Protocol Entity (12.8 and 8.10);*
— *MRP participants (Clause 10);*
— *MVRP participants (12.10, Clause 11);*
— *MMRP participants (12.11, Clause 10);*
— *The MST Configuration Table (12.12);*
— *Additional objects to support Provider Bridge management (12.13);*
— *The CFM Entities (12.14);*
— *The DDCFM Entities (12.17);*
— *The PBB-TE protection switching Entities (12.18);*
— *The TPMR Entities (12.19);*
— *The management entities for FQTSS (12.20);*
— *The congestion notification entities (12.21);*
— *The Shortest Path Bridging managed objects;*
— *The EVB entities (12.26).*

Observa-se que os itens mencionados entre parênteses na relação acima dizem respeito ao padrão IEEE 802.1Q-2014, que descreve a gerência de cada um desses recursos em termos dos objetos gerenciados e das correspondentes operações.

A título de exemplo são apresentados a seguir os seguintes recursos:

— Entidade de Gerenciamento de *Bridge* (*The Bridge Management Entity*);
— FDB da Entidade *MAC Relay* (*The FDB of the MAC Relay Entity*).

8.6.1.1. Entidade de Gerência de Bridge

Esse recurso utiliza os seguintes objetos gerenciados:

— *The Bridge Configuration;*
— *The Port Configuration for each Port.*

As operações de gerenciamento envolvendo o objeto gerenciado *The Bridge Configuration* são os seguintes:

— *Discover Bridge;*
— *Read Bridge;*
— *Set Bridge Name;*
— *Reset Bridge;*
— *Read component table entry;*
— *Update component table entry.*

As operações de gerenciamento envolvendo o objeto gerenciado *The Port Configuration for each Port* são as seguintes:

— *Read port table entry;*
— *Update port table entry.*

8.6.1.2. FDB da Entidade MAC Relay

Esse recurso utiliza os seguintes objetos gerenciados:

— *The Filtering Database;*
— *The Static Filtering Entries;*
— *The Dynamic Filtering Entries;*
— *The MAC Address Registration Entries;*
— *The Static VLAN Registration Entries;*
— *The Dynamic VLAN Registration Entries;*
— *The Permanent Database.*

Também a título de exemplo, encontram-se relacionados, a seguir, as operações de gerenciamento envolvendo o objeto gerenciado *The Filtering Database*:

— *Read Filtering Database*;
— *Set Filtering Database Ageing Time*;
— *Read Permanent Database*;
— *Create Filtering Entry*;
— *Delete Filtering Entry*;
— *Read Filtering Entry*;
— *Read Filtering Entry Range*.

8.6.2. **Management Information Base – MIB**

A Cláusula 17 do padrão IEEE 802.1Q-2014 contém a completa descrição da IEEE 802.1Q MIB, incluindo todos os seus módulos de MIB.

As versões prévias do padrão IEEE 802.1Q-2014 baseavam-se em módulos de MIB especificados pelo *IETF Bridge Working Group*. A partir da emissão da RFC 4663 (*Transferring MIB Work From IETF Bridge MIB WG to IEEE 802.1 WG),* em setembro de 2006, a responsabilidade por essa atribuição foi transferida para o IEEE 802.1 *Working Group*.

8.6.2.1. Conceitos e Funcionamento de MIBs

Os objetos gerenciados são acessados através de uma base de dados de informações virtuais, denominada *Management Information Base*, ou simplesmente MIB. Os objetos gerenciados de uma MIB são armazenados de forma distribuída em módulos de MIB.

Os objetos gerenciados dos módulos de MIB são definidos pela utilização de um subconjunto adaptado da *Abstract Syntax Notation One* (ASN-1), da ISO, subconjunto esse denominado SMIv2 (*Structure of Management Information version 2*).

O acesso aos objetos de uma MIB ocorre, geralmente, pela utilização do SNMP (*Simple Network Management Protocol*).

A SMIv2 foi especificada na RFC 2580 (*Conformance Statement for SMIv2)* e na RFC 2579 (*Textual Conventions for SMIv2*).

Textual Conventions for SMIv2 (*TCs for SMIv2*) representam a definição, com base na SMIv2, de um novo tipo de objeto gerenciado, com um novo nome e com uma sintaxe similar, porém com uma semântica mais precisa. A RFC 2579 tem como propósito definir o conjunto inicial de TCs para SMIv2 disponíveis para todos os módulos de MIB da IEEE 802.1Q MIB.

8.6.2.2. IEEE 802.1Q MIB

A IEEE 802.1Q MIB é dividida em 22 módulos de MIB, de modo correspondente aos projetos que criaram as respectivas funcionalidades.

A relação de módulos de MIB da IEEE 802.1Q MIB encontra-se na tabela 17.1 do padrão IEEE 802.1Q-2014.

O primeiro desses módulos, que é o módulo IEEE 8021-TC-MIB, tem como propósito definir as TCs utilizadas em todos os módulos da própria IEEE 802.1Q MIB. Foram definidas 25 TCs, que se encontram relacionadas na tabela 17-2 do padrão IEEE 802.1Q-2014.

Referências Bibliográficas

Livros
Padrões do IEEE
Recomendações do ITU-T
Outros Padrões

Padrões do IEEE

IEEE 802-2001: Overview and Management
IEEE 802.1AB-2009: Station and Media Access Control Connectivity Discovery
IEEE 802.1AX-2008: Link Aggregation
IEEE 802.1D-2004: Media Access Control (MAC) Bridges
IEEE 802.1Q-2014: Bridges and Bridged Networks
IEEE 802.2-1998: Logical Link Control - LLC
IEEE 802.3 -2012: CSMA/CD Access Method and Physical Layer Specifications

Padrões do IETF (RFCs)

RFC 1195 *Use of OSI IS-IS for Routing in TCP/IP and Dual Environments*
RFC 2205 R*SVP Version 1 Functional Specification*
RFC 2328 *Open Shortest Path First (OSPF) Version 2*
RFC 3107 *Carrying Label Information in BGP-4*
RFC 3209 *RSVP-TE: Extensions to RSVP for LSP Tunnels*
RFC 3471 GMPL*S Signaling Functional Description*
RFC 3473 *GMPLS Signaling RSVP-TE Extensions*
RFC 3619 *Extreme Networks´ Ethernet Automatic Protection Switching – Version 1*
RFC 3630 *Traffic Engineering Extensions to OSPF Version 2*
RFC 3916 *Requirements for Pseudo-Wire Emulation Edge-to-Edge (PWE3)*
RFC 3945 *Generalized Multiprotocol Label Switching (GMPLS) Architecture*
RFC 3985 *Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture*
RFC 4202 R*outing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)*
RFC 4203 *OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)*
RFC 4204 *Link Management Protocol – LMP*)
RFC 4271 A *Border Gateway Protocol 4 (BGP-4)*
RFC 4364 *BGP/MPLS IP Virtual Private Network*s
RFC 4447 *Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)*
RFC 4448 *Encapsulation Methods for Transport of Ethernet over MPLS Networks*
RFC 4461 *Signaling Requirements for Point-to-Multipoint TE MPLS LSPs*
RFC 4618: *Encapsulation Methods for Transport of PPP/HDLC over MPLS*

RFC 4619 *Encapsulation Methods for Transport of Frame Relay over MPLS*

RFC 4655 *A Path Computation Element (PCE)-Based Architecture*

RFC 4659 *BGP/MPLS IP Virtual Private Network Extension for IPv6 VPN*

RFC 4664 *Framework for L2VPNs*

RFC 4665 *Service Requirements for Layer 2 Provider-Provisioned Virtual Private Networks*

RFC 4717 *Encapsulation Methods for Transport of ATM over MPLS*

RFC 4760 *Multiprotocol Extensions for BGP-4*

RFC 4762 *VPLS Using LDP Signaling*

RFC 4873 *GMPLS Segment Recovery*

RFC 4875 *Extensions to RSVP-TE for Point-to-Multipoint TE Label Switching Paths (LSPs)*

RFC 4974 *GMPLS RSVP-TE Signaling Extensions in Support of Calls*

RFC 5120 *M-ISIS: Multi-Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)*

RFC 5250 *The OSPF Opaque LSA Option*

RFC 5305 *IS-IS Extensions for Traffic Engineering*

RFC 5307 *IS-IS Extensions in Support of Generalized Multiprotocol Label Switching (GMPLS)*

RFC 5342 IANA *Considerations and IETF Protocol Usage for IEEE 802 Parameters*

RFC *5420* *Encoding Attributes for MPLS LSP Establishment Using RSVP-TE*

RFC 5556 *Transparent Interconnection of Lots of Links* (TRILL): *Problem and Applicability Statement*

RFC 5586 MPLS *Generic Associated Channel*

RFC 5828 *Generalized Multiprotocol Label Switching (GMPLS) Ethernet Label Switching Architecture and Framework*

RFC 5880 *Bidirectional Forwarding Detection*

RFC 6060 *Generalized Multiprotocol Label Switching (GMPLS) Control of Ethernet Provider Backbone Traffic Engineering (PBB-TE)*

RFC 6165 *Extensions to IS-IS for Layer-2 Systems*

RFC 6325 *Routing Bridges* (*RBridges*): *Base Protocol Specification*

RFC 6329 *IS-IS Extensions Supporting IEEE 802.1aq Shortest Path Bridging*

RFC 6387 *GMPLS Asymmetric Bandwidth Bidirectional LSPs*

RFC 6361 *PPP TRILL Protocol Control Protocol*

RFC 6439 *RBridges*: *Appointed Forwarders*

RFC 6850 *Definitions of Managed Objects for RBridges*

RFC 6905 *Requirements for OAM in TRILL*

RFC 7172 TRILL: *Fine-Grained Label*

RFC 7173 *TRILL Transport Using PWs*

RFC 7174 *TRILL OAM Framework*

RCF 7175 *TRILL: Bidirectional Forwarding Dectection (BFD)*

RFC 7176 *TRILL Use of IS-IS*

RFC 7177 *TRILL: Adjacency*

RFC 7178 TRILL: *RBridge Channel Support*

RFC 7179 *TRILL: Header Extension*

RFC 7180 T*RILL: Clarifications, Corrections, and Updates*

RFC 7356 *IS-IS Flooding Scope Link State PDUs*

RFC 7357 *End Station Address Distribution Information (ESADI) Protocol*

RFC 7379 *Problem Statement and Goal for Active-Active Connection at the TRILL Edge*

RFC 7455 *TRILL: Fault Management*;

RFC 7456 *Loss and Delay Measurement in TRILL*

Recomendações do ITU-T

G. 7041/Y. 1303: *Generic Framing Procedure*

G.7042/Y.1305: *Link Capacity Adjustment Scheme* (*LCAS*) *for Virtual Concatenated Signals*

G.7043/Y.1343: *Virtual Concatenation of Plesiochronous Digital Hierarchy – PDH*

G.707/Y.1322: *Network Node Interface for the Synchronous Digital Hierarchy – SDH*

G.709/Y.1331: *Interfaces for the Optical Transport Networks*

G.8010/Y.1306: *Architecture of Ethernet Layer Networks*

G.8013/Y.1731: *OAM Functions and Mechanisms for Ethernet Networks*

G.8031/Y.1342: *Ethernet Protection Switching*

G.8032/Y1344: *Ethernet Ring Protection Switching*

G.783: *Characteristcs of Synchronous Digital Hierarchy* (SDH) *Equipment Functional Blocks*

G.798: *Characteristics of Optical Network Hierarchy Equipment Functional Blocks*

G.872: *Architecture of Optical Transport Networks*

G.984.2*: Gigabit-Capable Passive Optical Networks (GPONs): Physical Media Dependent*

(PMD) Layer Specification

G.991.1: High bit rate digital subscriber line (HDSL) transceivers
G.991.2: *Single-Pair High-Speed Digital Subscriber Line (SHDSL) Transceivers*
G.992.1: *Asymmetric digital subscriber line (ADSL) transceivers*
G.992.2: *Splitterless asymmetric digital subscriber line (ADSL) transceivers*
G.992.3: *Asymmetric digital subscriber line transceivers 2 – ADSL2*
G.992.4: *Splitterless asymmetric digital subscriber line transceivers 2 - splitterless ADSL2*
G. 992.5: *Asymmetric digital subscriber line 2 transceivers (ADSL2) - Extended bandwidth ADSL2 - ADSL2plus*
G.993.1: *Very high speed digital subscriber line transceivers – VDSL*
G.993.2: *Very high speed digital subscriber line transceivers 2 – VDSL2*
G.sup43: *Transport of IEEE 10GBASE-R in Optical Transport Networks (OTN)*
X.85/Y.1321: *IP over SDH using LAPS*
X.86//1323: *Ethernet over LAPS*
Y.1563: *Ethernet Frame Transfer and Availability Performance*
Y.1564: *Ethernet Service Activation Test Metodology*
Y.1730*: Requirements for OAM Functions in Ethernet-Based Networks and Ethernet Services*

Outros Padrões

ANSI T1.424*: Interface between Networks and Customer Installations: Very-high bit-rate Digital Subscriber Line (VDSL) Metallic Interface (DMT Based)*
ANSI/TIA 1057*: Link Layer Discovery Protocol for Media Endpoint Devices(LLDP-MED)*
ANSI –X.3.230-1994: *Fibre Channel Physical and Signaling Interface (FH-PH)*
IEC 61850*: Communication Networks and Systems in Substations*
IEC 62439*: Industrial Communication Networks – High Availability Automation Networks*
ISO/IEC 8802-2*: Logical Link Control (LLC)*

Relação de Siglas

Relação de Siglas

ADSL	*Asymetric* DSL
ADSL2	*Asymetric DSL 2*
ADSL2+	*Extended Bandwidth ADSL2-ADSL2plus*
ANSI	*American National Standard Institute*
AP	*Appointed Forwarder*
APON	*ATM PON*
APS	*Automatic Protection Switching*
ARP	*Address Resolution Protocol*
BASE	*Baseband*
BEB	
BCB	
BER	*Bit Error Rate*
BGP-4	*Border Gateway Protocol version 4*
BPDU	*Bridge Protocol Data Unit*
BPON	*Broadband PON*
BROAD	*Broadband*
B-tag	*Backbone tag*
B-VID	*Backbone VID*
CAM	*Content Addressable Memory*
CE	*Customer Edge Equipment*
CFI	*Canonical Format Indicator*
CIST	*Common and Internal Spanning Tree*
CLNP	*Connectionless Network Protocol*
GMPLS	*Generalized Multiprotocol Label Switching*
CO	*Connection-Oriented*
CRC	*Cyclic Redudancy Check*
CSMA-CD	*Carrier Sensing Multiple Access - Collision Detection*
CST	*Common Spanning Tree*
C-tag	*Customer tag*
C-VID	*Customer VID*
C-VLAN	*Customer VLAN*
CWDM	*Coarse WDM*
DDCFM	*Data-Driven and Data-Dependent CFM*
DIX	*Digital, Intel, Xerox*
DMT	*Discrete Multi Tone*
DQDB	*Distributed Queue Dual Bus*
DSL	*Digital Subscriber Line*
DWDM	*Dense WDM*
EAPS	*Extreme Networks' Ethernet Automatic Protection Switching*
EC	*Executive Commitee*

EFM	*Ethernet in the First Mile*
EGP	*Exterior Gateway Protocol*
EISS	*Enhanced Internal Sublayer Service*
E-LAN	*Ethernet LAN*
E-Line	*Ethernet Line*
EoM	*Ethernet over* MPLS
EoMPLS	*Ethernet over MPLS*
EoOTN	*Ethernet over OTN*
EoSDH	*Ethernet over SDH*
EPON	*Ethernet PON*
ERP	*Ethernet Ring Protection*
EtherType	*Ethernet Type*
ETHoETH	*Ethernet over Ethernet*
E-Tree	*Ethernet Tree*
EVC	*Ethernet Virtual Connection*
FCS	*Frame Check Sequence*
FDB	*Forwarding Database*
FDDI	*Fiber Distributed Data Interface*
FEC	*Forward Error Correction*
FIB	*Forwarding Information Base*
FID	*Filtering Identifier*
FTTC	*Fiber-to-the-Curb*
FTTH	*Fiber-to-the-Home*
FTTP	*Fiber-to-the-Premises*
GARP	*Generic Attribute Registration Protocol*
GELS	GMPLS *Ethernet Label Switching*
GFP	*Generic Frame Procedure*
GPON	*Gigabit-Capable PON*
HDSL	*High Bit Rate DSL ou High Speed DSL*
HSR	*High Availability Seamless Redundancy*
IEC	*International Eletrotecnical Commission*
IEEE	*Institute of Electrical and Electronics Engineering*
IETF	*Internet Engineering Task Force*
IGMP	*Internet Group Management Protocol*
IGP	*Interior Gateway Protocol*
IPG	*Inter-Packet Gap*
IPX	*Internetwork Packet Exchange*
IS	*Intermediate System*
I-SID	*Instance Service Identifier*
IS-IS	*Intermediate System-to-Intermediate System*
ISO	*International Organization for Standardtization*
ISP	*Internet Service Provider*

ISS	*Internal Sublayer Service*
IST	*Internal Spanning Tree*
I-tag	*Instance tag*
ITU	*International Telecommunication Union*
ITU-T	*ITU Telecommunication Standardization Sector*
IVL	*Independente VLAN Learning*
L2VPN	*Layer 2 Virtual Private Networks*
L3VPN	*Layer 3 Virtual Private Network*
LACP	*Link Aggregation Control Protocol*
LAN	*Local Area Network*
LAP	*Link Access Procedure*
LAP-B	*LAP- Balanced Mode*
LAPS	*Link Access Procedure-SDH*
LCAS	*Link Capacity Adjustment Scheme*
LLC	*Logical Link Control*
LLDP	*Link Layer Discovery Protocol*
LLID	*Logical Link Identifier*
LSB	*Least Significant Bit*
LSP	*Label Switching Path*
LSP	*Link State Packet ou Link State Protocol Data Unit*
MAC	*Media Access Control*
MAC-DA	*MAC Destination Address*
MAC-SA	*MAC Source Address*
MAD	*MRP Attribute Declaration*
MAN	*Metropolitan Area Network*
MAP	*MRP Attribute Propagation*
MDI	*Medium Dependent Interface*
MEF	*Metro Ethernet Forum*
MIB	*Management Information Base*
MIRP	*Multiple I-SID Registration Protocol*
MMRP	*Multiple MAC Registration Protocol*
MP2MP	*Multipoint-to-Multipoint*
MP-BGP	*Multiprotocol BGP*
MPCP	*Multi-Point Control Protocol*
MPLS-TP	MPLS *Transport Profile*
MPMC	*Multi-Point MAC Control*
MRP	*Media Redundancy Protocol*
MRP	*Multiple Registration Protocol*
MSAP	*Mac Service Access Point*
MSB	*Most Significant Bit*
MSP	*Multiple Stream Protocol*
MSRP	*Multiple Stream Registration Prococol*

MST	*Multiple Spanning Tree*
MSTI	*Multiple Spanning Tree Instance*
MSTID	*MSTI Identifier*
MSTP	*Multiple Spanning Tree Protocol*
MSTP	*Multiple Spanning Tree Protocol*
MVRP	*Multiple VLAN Registration Protocol*
NG-SDH	*Next Generation SDH*
OAM	*Operation, Administration and Maintenance*
ODN	*Optical Distribution Network*
OLT	*Optical Line Terminal*
ONU	*Optical Network Unit*
OSI	*Open Systems Interworking*
OTN	*Optical Transport Network*
OUI	*Organizationally Unique Identifier*
P	*Provider Equipment*
P2MP	*Point-to-Multipoint*
P2P	*Point-to-Point*
PASS	*Passband*
PB	*Provider Bridging (Bridges)*
PBB	*Provider Backbone Bridging (Bridges)*
PBBN	*Provider Backbone Bridging Network*
PBB-TE	*Provider Backbone Bridging-Traffic Engineerig*
PBN	*Provider Bridging Network*
PBT	*Provider Backbone Transport*
PCP	*Priority Code Point*
PCS	*Physical Coding Sublayer*
PDH	*Plesiochronous Digital Hierarchy*
PDU	*Protocol Data Unit*
PE	*Provider Edge Equipment*
PFC	*Priority-based Flow Control*
PhSAP	*Physical SAP*
PHY	*Physical*
PID	*Protocol Identifier*
PLS	*Physical Layer Signaling*
PM	*Perfomance Monitoring*
PMA	*Physical Coding Sublayer*
PMD	*Physical Mediun Dependent*
PON	*Passive Optical Network*
POP	*Point of Presence*
POTS	*Plain Old Tephone Service*
PPP	*Poin-to-Point Protocol*
PRP	*Parallel Redundancy Protocol*

PTN	*Packet Transport Network*
PVID	*Port VLAN ID*
PW	*Pseudowire*
PWE3	*Pseudo Wire Emulation Edge-to-Edge*
QAM	*Quadrature Amplitude Modulation*
QoS	*Quality of Service*
R-APS	*Ring Automatic Protection Switching*
RFC	*Request for Comment*
RSTP	*Rapid Spanning Tree Protocol*
RSVP	*Resource Reservation Protocol*
SAP	*Service Access Point*
SDH	*Synchronous Digital Hierarchy*
SDU	*Service Data Unit*
SFD	*Start Frame Delimiter*
SG	*Study Group*
SHDSL	*Single-pair High-speed DSL*
SOAM	*Service OAM*
SOAM CFM	*SOAM Connectivity Fault Management*
SOAM PM	*SOAM Performance Monitoring*
SONET	*Synchronous Optical Network*
SPB	*Shortest Path Bridging*
SRP	*Stream Registration Protocol*
SST	*Single Spanning Tree*
S-tag	*Service tag*
STP	*Spanning Tree Protocol*
S-VID	*Service VID*
SVL	*Shared VLAN Learning*
S-VLAN	*Service VLAN*
TCA	*Topology Change Acknowledgement*
TCI	*Tag Control Information*
TCN	*Topology Change Notification*
TC-PAM	*Trellis Coded Pulse Amplitude Modulation*
TDMA	*Time-Division Multiple Access*
TE	*Traffic Engineering*
TPID	*Tag Protocol Identifier*
TPMR	*Two-Port Media Access Control Relay*
TRILL	*Transparent Interconnection of Lots of Links*
U/L bit	*Universal/Local Bit*
UTP	*Unshielded Twisted Pair*
VCAT	*Virtual Concatenation*
VDSL	*Very High Speed DSL*
VDSL2	*Very high Speed DSL 2*

VID	*VLAN Identifier*
VLAN	*Virtual* LAN
VPLS	*Virtual Private* LAN *Service*
VPWS	*Virtual Private Wire Service*
VRF	*Virtual Routing Function*
VSI	*Virtual Switching Instance*
WAN	*Wide Area Network*
WDM	*Wavelength Division Multiplexing*
WMAN	*Wireless Metropolitan Area Networks*
WPAN	*Wireless Personal Area Network*
IIH	*IS-IS Hello* PDUs
CSNPs	*Complete Sequence Number PDUs*
PSNPs	*Partial Sequence Number PDUs*
ECMP	*Equal-Cost Multi-Path*
SPBV	*Shortest Path Bridging – VID*
SPBM	*Shortest Path Bridging –MAC*
PLSB	*Provider Link State Bridging*
ISIS-SPB	*IS-IS Shortes Path Bridging*
NLPID	*Network Layer Protocol Identifier*
SPT	*Shortest Path Tree*
MDT	*Multicast Distribution Tree*
NNI	*Network-Node Interface ou Network-to-Network Interface*
ECT	*Equal Cost Trees*
MT ID	*Multi-Topology Identifier*
SYSID	*System ID*
PATHID	*Path Identifier*
SPVID	*Shortest Path VID*
MT	*Multi Topology*
VL TRILL	*VLAN-Labelled TRILL*
FGL TRILL	*Fine-Grained Label TRILL*
ESADI	*End Station Address Distribution Information*
SNPA	*Sub-Network Point of Attachment*
RPF	*Reverse Path Forwarding*
TNCP	*TRILL Network Control Protocol*
TNP	*TRILL Network Protocol*
TLSP	*TRILL Link State Protocol*
LCP	*Link Control Protocol*
BCP	*Bridging Control Protocol*
EAP	*Extensible Authentication Protocol*
DRB	*Designated RBridge*
LAALP	*Local Active-Active Link Protocol*
MC-LAG	*Multi-Chassis Link Aggregation*

AAE	*Active-Active Edge RBridge Group*
RBv	*RBridge Virtual*
RBridge	*Routing Bridge*
BFD	*Bidirectional Forwarding Detection*
MEP	*Maintance End Point*
MIP	*Maintance Intermediate Point*
MD-L	*Maintenance Domain Level*
MTVM	*Multi-Destination Tree Verification Message*
ME	*Maintenance Entity*
MA	*Maintenance Association*
MD	*Maintenance Domain*
CCM	*Continuity Check Message*
1DM	*One-way Delay Measurement*
2DM	*Two-way Delay Measurement*
SLA	*Service Level Agreement*
1SL	*One-way Synthetic Loss Measurement*
IPPM	*Inter-Packet Performance Monitoring*
SLM	*Synthetic Loss Measuremente Message*
SLR	*Synthetic Loss Measurement Reply*
DMM	*Delay Measurement Message*
DMR	*Delay Measurement Reply*
SMI	*Structure of Managemente Information*
SNMP	*Simple Network Management Protocol*