Multiprotocol Label Switching

O que é, como funciona e como utilizar em redes de telecomunicações e data centers

Lucas Rayan Guerra

Introdução

As redes de telecomunicações modernas carregam volumes de tráfego impensáveis há duas décadas. Vídeo em tempo real, serviços de nuvem geograficamente distribuídos, VoIP de alta densidade e aplicações industriais com latência crítica colocam sob pressão constante a infraestrutura de transporte. O roteamento IP tradicional, baseado na análise do cabeçalho de cada pacote em cada dispositivo, mostrou-se insuficiente para atender aos requisitos de desempenho, engenharia de tráfego e qualidade de serviço dessas aplicações.

O MPLS (Multiprotocol Label Switching) surgiu como resposta a esse desafio. Definido inicialmente pela IETF nos anos 2000 e continuamente evoluído desde então, combina a flexibilidade do roteamento IP com a velocidade e previsibilidade da comutação por rótulos, formando a espinha dorsal das maiores redes de operadoras de telecomunicações, data centers e backbones corporativos do planeta [1].

Esta cartilha técnica apresenta o MPLS de forma aprofundada, desde seus conceitos fundamentais e arquitetura de funcionamento, passando pelos protocolos de sinalização LDP, RSVP-TE e BGP, até os casos de uso práticos como VPNs L2/L3, Engenharia de Tráfego, Fast Reroute e a evolução para o ambiente de redes programáveis com Segment Routing. O material é dirigido a profissionais de redes, engenheiros de NOC e estudantes que desejam compreender em profundidade como o tráfego é transportado nas grandes redes de hoje.

O Problema do Roteamento IP Tradicional

Para entender por que o MPLS foi criado, é necessário primeiro compreender as limitações do modelo de encaminhamento IP clássico. Em uma rede IP pura, cada roteador, ao receber um pacote, realiza as seguintes operações de forma independente:

- **Extração do IP de destino:** O roteador lê o campo de endereço IP de destino no cabeçalho do pacote.

O MPLS foi projetado para resolver todos esses problemas de forma elegante, inserindo uma camada de encaminhamento baseada em rótulos entre as camadas de rede (L3) e enlace (L2) do modelo OSI, razão pela qual é frequentemente chamado de tecnologia de **"camada 2,5"**.

O Que é MPLS e Como Funciona

O MPLS (Multiprotocol Label Switching) é uma tecnologia de encaminhamento de pacotes que utiliza rótulos (labels) curtos e de comprimento fixo para tomar decisões de encaminhamento, em vez de analisar o cabeçalho IP completo a cada salto. Isso transforma roteadores em comutadores de rótulos de alta velocidade, simplificando e acelerando radicalmente o processo de encaminhamento [3].

O Rótulo MPLS (Label)

O rótulo MPLS é um campo de **32 bits** inserido entre o cabeçalho de enlace (Ethernet, por exemplo) e o cabeçalho IP de um pacote. Sua estrutura é definida pelo RFC 3032 [4]:

Campo	Bits	Descrição
Label Value	20	Valor do rótulo (0–1.048.575). Identifica o LSP.
TC (Traffic Class)	3	Anteriormente chamado EXP. Usado para QoS (DiffServ). Até 8 classes.
S (Bottom of Stack)	1	Indica se este é o último rótulo da pilha (S=1 → último rótulo).
TTL (Time to Live)	8	Previne loops de encaminhamento, equivalente ao TTL do cabeçalho IP.

O campo **Label Value** de 20 bits é o coração do MPLS. Ele identifica univocamente um LSP (Label Switched Path), o caminho predeterminado que os pacotes seguirão pela rede.

Importante: o valor do rótulo tem significado apenas local entre dois dispositivos adjacentes; o mesmo LSP pode ter rótulos diferentes em cada segmento do caminho.

Pilha de Rótulos (Label Stack)

Uma das características mais poderosas do MPLS é a possibilidade de empilhar múltiplos rótulos sobre um mesmo pacote, formando uma **pilha de rótulos (label stack)**. O pacote é sempre encaminhado com base no rótulo do topo da pilha. Isso viabiliza cenários complexos como VPNs aninhadas, Engenharia de Tráfego sobre VPNs e Fast Reroute.

- **Rótulo de topo (outer/transport label):** Determina o caminho físico pelo backbone MPLS. É trocado (swap) a cada salto pelos LSRs do core.

- **Rótulo inferior (inner/VPN label):** Identifica o cliente VPN ou serviço específico no destino. Transportado intacto pelo backbone e analisado apenas no PE de saída.

Os Atores do MPLS: LER e LSR

A rede MPLS é composta por dois tipos de dispositivos com funções bem definidas:

Dispositivo	Nome Completo	Descrição
LER de Ingresso (PE entrada)	Label Edge Router	Classifica o pacote IP recebido, determina o LSP adequado e empilha os rótulos MPLS. É a porta de entrada do backbone.
LSR (P Router)	Label Switching Router	Recebe pacote MPLS, troca o rótulo do topo (SWAP) e encaminha. Opera apenas com rótulos, sem inspecionar o IP interno.
LER de Egresso (PE saída)	Label Edge Router	Remove os rótulos MPLS (POP) e entrega o pacote IP original ao destino ou à VRF do cliente.

Nas redes de operadoras, os LERs são chamados de **PE (Provider Edge)** e os LSRs de **P (Provider core)**. Os roteadores do cliente são denominados **CE (Customer Edge)** e não precisam ter conhecimento de MPLS, comunicam-se com o PE via roteamento IP normal.

As Três Operações Fundamentais do MPLS

O encaminhamento MPLS é governado por apenas três operações sobre a pilha de rótulos. A simplicidade dessas operações é justamente o que confere ao MPLS sua excepcional velocidade de processamento:

Operação	Nome Completo	Onde ocorre
PUSH	Empilha um ou mais rótulos sobre o pacote. Converte um pacote IP em MPLS, ou adiciona rótulo de transporte sobre VPN label já existente.	LER de Ingresso (PE entrada)
SWAP	Troca o rótulo do topo da pilha pelo novo valor mapeado na LFIB. É a operação padrão nos roteadores core (LSRs).	LSR (P Router) intermediário
POP	Remove o rótulo do topo da pilha. Pode ser feito no penúltimo salto (PHP) ou no PE de saída, expondo o próximo rótulo ou o IP original.	LSR penúltimo (PHP) ou LER de Egresso

Penultimate Hop Popping (PHP)

O PHP é uma otimização importante: o rótulo de transporte externo é removido pelo **penúltimo LSR** antes de encaminhar o pacote ao PE de saída. Isso evita que o PE precise realizar duas operações em sequência (pop do outer label + lookup do inner label), reduzindo a carga no dispositivo de borda. Quando PHP está ativo, o LSR penúltimo recebe o rótulo especial **3 (Implicit NULL)** via LDP, sinalizando que deve remover o rótulo antes de encaminhar.

FIB e LFIB: As Tabelas de Encaminhamento MPLS

Enquanto o roteamento IP clássico usa a FIB (Forwarding Information Base) para decisões baseadas em prefixos IP, o MPLS adiciona uma tabela específica para encaminhamento por rótulos, a LFIB:

Tabela	Chave de Busca	Ação	Usado por
FIB	Prefixo IP destino (Longest Prefix Match)	Encaminha pacote IP ou realiza PUSH de rótulo(s) MPLS.	LER ingresso, pacotes IP sem rótulo
LFIB	Rótulo de entrada (Exact Match)	Realiza SWAP (troca rótulo) ou POP (remove rótulo) e encaminha.	LSR e LER egresso, pacotes MPLS

A LFIB é a razão da alta velocidade do MPLS: em vez de realizar uma operação complexa de Longest Prefix Match sobre centenas de milhares de prefixos, o LSR realiza uma simples **busca por chave exata (exact match)** sobre o valor do rótulo, uma operação muito mais rápida, implementável diretamente em hardware ASIC.

Label Switched Path (LSP) e Protocolos de Sinalização

Um LSP é o caminho unidirecional, predefinido, que os pacotes com um determinado rótulo seguem da origem ao destino. Diferente do roteamento IP hop-by-hop, em que cada roteador decide independentemente, no MPLS o caminho completo é estabelecido antes que o primeiro pacote seja enviado.

LDP — Label Distribution Protocol

O LDP (RFC 5036) é o protocolo mais simples e amplamente utilizado para distribuição de rótulos. Seu princípio é direto: para cada prefixo IP que um roteador conhece via IGP (OSPF ou IS-IS), ele gera um rótulo local e anuncia o par (prefixo, rótulo) para seus vizinhos LDP.

- **Descoberta de vizinhos:** Via mensagens Hello UDP na porta 646, enviadas para o multicast 224.0.0.2. Suporta também descoberta direcionada (targeted LDP) para sessões entre nós não adjacentes.

- **Sessão TCP confiável:** Após descoberta, uma sessão TCP é estabelecida (porta 646) para troca de mensagens de mapeamento de rótulos com garantia de entrega.

- **Distribuição de rótulos:** Cada roteador anuncia rótulos para todos os prefixos de sua FIB. O vizinho usa essas informações para montar sua LFIB (Label Forwarding Information Base).

- **Convergência automática:** Quando o IGP converge após uma falha, o LDP reconverge automaticamente e redistribui rótulos para os novos caminhos.

O LDP cria LSPs que seguem exatamente o melhor caminho calculado pelo IGP, mas não oferece controle explícito sobre o caminho. Para isso, é necessário o RSVP-TE.

RSVP-TE — Resource Reservation Protocol - Traffic Engineering

O RSVP-TE (RFC 3209) é o protocolo de sinalização usado para criar LSPs com Engenharia de Tráfego: caminhos explicitamente definidos pelo operador ou calculados por algoritmos de otimização, independentemente do caminho de menor custo calculado pelo IGP [5].

- **PATH Message:** Enviada da origem ao destino, carregando o caminho explícito (ERO, Explicit Route Object) e os requisitos de recursos (banda mínima, prioridade de setup e hold).

- **RESV Message:** Enviada de volta do destino à origem, confirmando a reserva de recursos e distribuindo os rótulos salto a salto na direção reversa ao PATH.

- **Soft state com refresh periódico:** RSVP mantém o LSP ativo por meio de mensagens de refresh periódicas. Se os refreshes pararem, o LSP é removido automaticamente por um mecanismo de autorecuperação.

- **Preempção por prioridade:** LSPs de maior prioridade podem preemptar LSPs de menor prioridade quando recursos são escassos, usando os campos Setup Priority (0–7) e Hold Priority.

Uma rede com RSVP-TE requer que os roteadores conheçam a topologia completa e os recursos disponíveis em cada enlace. Para isso, extensões TE são adicionadas ao OSPF (RFC 3630) ou IS-IS (RFC 5305), que passam a anunciar atributos como banda disponível, cor de enlace (admin groups) e métrica TE.

Os Três Pilares de Serviços MPLS

A implementação eficaz do MPLS se organiza em três grandes pilares de serviços, que frequentemente coexistem sobre o mesmo backbone físico:

Pilar	Descrição	Protocolo Principal
VPN Layer 3 (L3VPN)	Isolamento de roteamento por cliente com VRFs separadas. Múltiplos clientes compartilham o backbone físico com total independência de tabelas de roteamento.	MP-BGP + LDP/RSVP-TE
VPN Layer 2 (L2VPN)	Extensão de conectividade Ethernet sobre MPLS. Pseudowires ponto a ponto e VPLS multiponto simulam circuitos de Layer 2 sobre o backbone.	LDP + Signaling (LDP-based PW)
Engenharia de Tráfego (TE)	Controle explícito de caminhos, reserva de banda e proteção FRR. Permite utilizar a capacidade da rede de forma otimizada e previsível.	RSVP-TE + IGP-TE Extensions

MPLS VPN Layer 3 (L3VPN) — RFC 4364

As L3VPNs são o caso de uso mais popular do MPLS em operadoras de telecomunicações. Elas permitem que múltiplos clientes compartilhem o mesmo backbone físico com total isolamento de roteamento, como se cada cliente tivesse sua própria rede privada dedicada.

- **VRF (Virtual Routing and Forwarding):** Cada cliente recebe uma tabela de roteamento separada nos PEs. Prefixos de clientes diferentes nunca se misturam, mesmo que usem os mesmos blocos de endereços IP privados (RFC 1918).

- **Route Distinguisher (RD):** Campo de 8 bytes adicionado a cada prefixo IPv4 antes de ser anunciado via MP-BGP, tornando-o globalmente único. Formato típico: AS:número (ex: 65000:100).

- **Route Target (RT):** Atributo BGP Extended Community que controla quais prefixos são exportados de uma VRF e importados em outra. Viabiliza topologias complexas como hub-and-spoke e extranet entre clientes.

- **MP-BGP (Multiprotocol BGP):** Extensão do BGP para transportar prefixos VPNv4 (RD + prefixo IPv4) entre PEs. Opera sobre o backbone MPLS, encapsulando prefixos de clientes em atualizações BGP seguras.

O fluxo de encaminhamento em uma L3VPN usa **dois rótulos em pilha**: o rótulo externo (transporte LDP/RSVP-TE) que guia o pacote pelo backbone até o PE de saída, e o rótulo interno (VPN label) que identifica a VRF do cliente no PE de destino.

MPLS VPN Layer 2 — Pseudowires e VPLS

Enquanto a L3VPN opera na camada de rede (roteamento), as L2VPNs estendem a camada de enlace sobre o backbone MPLS, permitindo que sites remotos pareçam estar conectados por um cabo de Layer 2 virtual:

- **Pseudowire (PW):** Emula um circuito ponto a ponto de Layer 2 (Frame Relay, ATM, Ethernet, TDM) sobre MPLS. Definido pelo RFC 3985 (PWE3). Ideal para conectar dois sites como se fossem adjacentes no mesmo switch.

- **VPLS (Virtual Private LAN Service):** Emula uma LAN Ethernet multiponto sobre MPLS. Todos os sites do cliente parecem estar no mesmo domínio de broadcast. O PE aprende endereços MAC dos clientes e encaminha frames via MPLS.

- **EVPN (Ethernet VPN):** Evolução moderna do VPLS que usa MP-BGP para distribuir informações de MAC/IP, eliminando flooding excessivo e melhorando escalabilidade. Suporta multi-homing ativo-ativo.

Engenharia de Tráfego (MPLS-TE)

A Engenharia de Tráfego com MPLS permite definir explicitamente quais caminhos os fluxos de tráfego devem seguir, independentemente do caminho de menor custo calculado pelo IGP. Isso permite ao operador utilizar a capacidade da rede de forma balanceada e eficiente:

- **Balanceamento de carga por fluxo:** Diferentes fluxos podem ser direcionados para caminhos fisicamente distintos, evitando que o IGP concentre todo o tráfego no enlace de menor custo.

- **Reserva de banda (CAC):** O RSVP-TE pode reservar largura de banda garantida para fluxos críticos como VoIP ou videoconferência, assegurando latência e jitter dentro dos limites acordados em SLAs.

- **Restrições de caminho por cor (affinity):** Administradores podem marcar enlaces com atributos administrativos e especificar que certos LSPs só utilizem enlaces com determinadas características.

- **DiffServ sobre MPLS-TE:** O campo TC (3 bits) do rótulo MPLS é usado para marcar filas de QoS, permitindo tratamento diferenciado para diferentes classes de serviço no backbone.

Fast Reroute (FRR) — Proteção em Milissegundos

O Fast Reroute é o mecanismo de proteção contra falhas mais sofisticado do MPLS-TE. Diferente da convergência IGP, que pode levar de 1 a 30 segundos, o FRR permite redirecionar o tráfego para um caminho de proteção em **menos de 50 milissegundos** após a detecção de uma falha [6].

O FRR exige que LSPs de proteção sejam pré-provisionados antes das falhas ocorrerem. Existem dois modelos:

- **Link Protection (Bypass LSP):** Um LSP de bypass é pré-estabelecido contornando apenas o enlace falho. No momento da falha, o tráfego é imediatamente desviado para o bypass pelo PLR (Point of Local Repair), antes mesmo de o IGP convergir.

- **Node Protection (Bypass de nó):** O LSP de bypass contorna tanto o enlace quanto o nó subsequente. Protege contra falhas de roteador completo, não apenas de enlace. Mais robusto, requer mais recursos de signaling.

O FRR é viabilizado pelo BFD (Bidirectional Forwarding Detection), que detecta falhas de enlace em milissegundos através de mensagens de controle de baixo overhead enviadas continuamente entre vizinhos. Quando o BFD detecta uma falha, o PLR ativa imediatamente o LSP de bypass, sem aguardar sinalização IGP/RSVP.

Comparativo de Tempos de Recuperação

Mecanismo	Tempo Típico de Recuperação	Método de Detecção
Convergência IGP pura	1–30 segundos	Dead timer OSPF/IS-IS (hello × dead = 40s padrão)
IGP com BFD	300ms–3 segundos	BFD detecta falha e notifica IGP para reconvergência
MPLS FRR (Link Protection)	< 50 milissegundos	BFD + ativação imediata do bypass LSP pré-estabelecido
MPLS FRR (Node Protection)	< 50 milissegundos	BFD + bypass contornando o nó completo

QoS sobre MPLS: Garantindo Qualidade de Serviço

Um dos principais atrativos do MPLS para operadoras é a capacidade de implementar QoS de ponta a ponta com granularidade controlada. O campo TC de 3 bits no cabeçalho MPLS é o mecanismo central, viabilizando até 8 classes de serviço distintas.

Modelos de QoS em MPLS

- **Pipe Model:** O campo TC é definido na entrada (PE) e transportado inalterado pelo backbone. O tratamento de QoS no core é baseado no TC do rótulo, não no DSCP original do IP. Mais simples e mais comum.

- **Uniform Model:** O DSCP do pacote IP é copiado para o campo TC do rótulo MPLS e vice-versa em cada salto. Garante consistência entre o tratamento dentro e fora do backbone, mas é mais complexo de operar.

- **Short Pipe Model:** Variante do Pipe Model onde o PE de egresso usa o DSCP original (não o TC) para tratamento na saída. Oferece flexibilidade adicional para políticas de egresso.

Classes de Serviço Típicas

TC (bits)	Nome	Serviços Típicos	Tratamento
111 (7)	Network Control	BGP, OSPF, LDP, RSVP	Priority Queue
110 (6)	Voice (EF)	VoIP, Videoconferência	Strict Priority
100 (4)	Video (AF4x)	Streaming, IPTV	LLQ / CBWFQ
010 (2)	Business (AF2x)	Aplicações críticas	CBWFQ garantido
000 (0)	Best Effort (BE)	HTTP, email, downloads	FIFO / WFQ

Segment Routing: A Evolução do MPLS

O Segment Routing (SR) é a evolução arquitetural mais significativa do MPLS desde sua criação. Padronizado pela IETF (RFC 8402), o SR simplifica radicalmente a rede ao eliminar LDP e RSVP-TE como protocolos independentes, substituindo-os por uma abordagem de **source routing**: o nó de origem instrui o caminho completo através de uma sequência de segmentos no cabeçalho do pacote [7].

SR-MPLS: Segment Routing com Plano de Dados MPLS

No SR-MPLS, segmentos são representados como rótulos MPLS comuns. Cada nó recebe um **Node-SID** (Segment Identifier) anunciado globalmente via IGP (OSPF ou IS-IS com extensões SR). Para direcionar tráfego por um caminho específico, o PE de ingresso empilha a sequência de SIDs correspondente, sem sinalização hop-by-hop via RSVP-TE:

- **Node-SID:** Identifica um nó específico na rede. Rótulo globalmente significativo, onde qualquer roteador sabe que o Node-SID 16001 corresponde ao roteador A.

- **Adjacency-SID:** Identifica uma adjacência específica (enlace) entre dois nós. Tem significado apenas local e é usado para forçar o tráfego por um enlace específico.

- **Prefix-SID:** Identifica um prefixo IP específico. Derivado do Node-SID mais um offset configurável pelo operador.

SR vs. MPLS Tradicional: Comparativo

Aspecto	MPLS Tradicional	Segment Routing (SR-MPLS)
Sinalização	LDP + RSVP-TE (dois protocolos independentes)	Apenas IGP com extensões SR (OSPF-SR ou IS-IS-SR)
Estado nos LSRs	Per-LSP state em cada LSR, consumo crescente de memória com escala	Apenas per-node state (Node-SID). Sem estado por fluxo nos LSRs.
TE e FRR	RSVP-TE com setup e manutenção complexos de LSPs	TI-LFA (Topology-Independent LFA) para FRR automático sem sinalização adicional
Programabilidade	Difícil: caminhos definidos estaticamente ou por CSPF local	Alta: controladores SR-PCE calculam e instalam caminhos via BGP-LS + PCEP
Complexidade operacional	Alta: dois protocolos, tabelas de estado separadas, adjacências RSVP	Baixa: protocolo único, configuração simplificada, operação mais intuitiva

SRv6: Segment Routing sobre IPv6

O SRv6 é a implementação do Segment Routing usando IPv6 como plano de dados, em vez de rótulos MPLS. Segmentos são endereços IPv6 de 128 bits inseridos em um cabeçalho de extensão IPv6 chamado **SRH (Segment Routing Header)**. O SRv6 elimina completamente o cabeçalho MPLS e viabiliza programabilidade ainda maior, pois permite codificar funções de rede diretamente no endereço IPv6 (Network Programming / SRv6 uSID) [8].

Configuração Prática: Exemplos Comentados

A seguir são apresentados exemplos de configuração para os principais cenários MPLS, usando sintaxe do Cisco IOS-XE como referência. As configurações são simplificadas para fins educacionais.

SRv6: Segment Routing sobre IPv6

```
! Habilitar LDP globalmente
mpls ip
mpls label protocol ldp
!
! Habilitar MPLS na interface do backbone
interface GigabitEthernet0/0/0
 description Link para P1
 ip address 10.0.0.1 255.255.255.252
 mpls ip
!
! Verificacao - vizinhos LDP e bindings
show mpls ldp neighbor
show mpls ldp bindings
show mpls forwarding-table
```

Segment Routing com IS-IS (SR-MPLS)

```
! IS-IS com extensoes Segment Routing
router isis CORE
 net 49.0001.0000.0000.0001.00
 is-type level-2-only
 metric-style wide
 segment-routing mpls
!
! Node-SID no loopback
interface Loopback0
 ip address 10.255.255.1 255.255.255.255
 ip router isis CORE
 isis prefix-sid absolute 16001
!
! Verificar SIDs distribuidos
show isis segment-routing prefix-sid-map
show mpls forwarding-table labels 16001-16010
```

Configurando L3VPN MPLS (PE Router)

```
! Criar VRF para o cliente
vrf definition ClienteA
 rd 65000:100
 address-family ipv4
  route-target export 65000:100
  route-target import 65000:100
 exit-address-family
!
! Interface CE-PE dentro da VRF
interface GigabitEthernet0/1/0
 description Link para CE ClienteA
 vrf forwarding ClienteA
 ip address 192.168.1.1 255.255.255.252
!
! MP-BGP para distribuir rotas VPN entre PEs
router bgp 65000
 neighbor 10.255.255.2 remote-as 65000
 neighbor 10.255.255.2 update-source Loopback0
 address-family vpnv4
  neighbor 10.255.255.2 activate
  neighbor 10.255.255.2 send-community extended
 exit-address-family
 address-family ipv4 vrf ClienteA
  redistribute connected
 exit-address-family
```

Comandos de Verificação e Diagnóstico

```
! Tabela de encaminhamento MPLS (LFIB)
show mpls forwarding-table

! Rotas VPNv4 no BGP
show bgp vpnv4 unicast all summary

! VRF específica
show ip route vrf ClienteA

! Trace do caminho MPLS
traceroute mpls ipv4 10.0.10.0/24

! Verificar LSPs RSVP-TE
show mpls traffic-eng tunnels brief

! Estado do BFD
show bfd neighbors detail
```

Cenários Práticos e Exemplos Reais

Criando a Segmentação com VLANs

Uma operadora de telecomunicações atende milhares de clientes corporativos com diferentes necessidades. Com MPLS, oferece múltiplos serviços sobre a mesma infraestrutura física:

- **L3VPN para empresas com múltiplas filiais:** Cada empresa recebe uma VRF isolada. Os CEs trocam rotas com os PEs via OSPF ou BGP, e a operadora garante isolamento total, mesmo que clientes usem os mesmos blocos IP privados.

- **Pseudowires para substituição de circuitos TDM:** Clientes com sistemas legados que dependem de Frame Relay ou ATM têm esses serviços emulados sobre MPLS via Pseudowires, sem necessidade de substituir equipamentos de ponta.

- **QoS diferenciada por SLA:** Clientes premium têm tráfego marcado com TC=6 (Expedited Forwarding) com latência máxima garantida ponta a ponta; clientes básicos usam TC=0 (Best Effort).

- **MPLS-TE para SLAs de banda garantida:** LSPs RSVP-TE com banda reservada são provisionados para clientes que necessitam de throughput garantido, como links de backup de data center.

Data Center Interconnect (DCI)

A interconexão de data centers é um dos casos de uso mais críticos da infraestrutura moderna. Com MPLS e EVPN, múltiplos DCs podem ser interligados transparentemente na camada 2, viabilizando migração de VMs entre sites e clustering de aplicações geograficamente distribuídas:

- **EVPN-MPLS para L2 Stretch:** Segmentos Ethernet (VLANs) são estendidos entre data centers via MPLS com EVPN. Servidores em sites diferentes aparecem na mesma sub-rede, sem os problemas de escalabilidade do VPLS tradicional.

- **SR-TE para DCI com latência determinística:** LSPs SR-TE estabelecidos com banda garantida entre DCs evitam que tráfego de migração de VMs afete outras aplicações. O controlador SR-PCE recalcula caminhos automaticamente em falhas.

- **Fast Reroute com TI-LFA:** Com SR e TI-LFA, falhas de enlace são contornadas em menos de 50ms, então aplicações críticas não percebem interrupções durante eventos de manutenção ou falhas físicas.

Rede de Ensino e Pesquisa (RNP / Internet2)

Redes de ensino e pesquisa como a RNP (Rede Nacional de Ensino e Pesquisa) utilizam MPLS em seu backbone para conectar instituições de todo o país com QoS diferenciada para diferentes tipos de tráfego acadêmico:

- **Pseudowires para laboratórios remotos:** Experimentos científicos que requerem conectividade direta entre laboratórios em diferentes universidades usam Pseudowires Ethernet para obter conectividade L2 transparente com latência determinística.

- **Multicast sobre MPLS (mVPN):** Transmissões ao vivo de conferências, aulas e defesas de tese usam IP Multicast encapsulado em MPLS VPN (mVPN, RFC 6514), replicando o vídeo apenas para os sites que solicitaram, sem desperdício de banda.

- **TE para grandes transferências científicas:** LSPs com banda reservada são provisionados para transferências massivas de dados científicos (como dados do LHC do CERN), garantindo que o volume não impacte o tráfego regular.

Boas Práticas para Redes MPLS em Produção

Segurança no Plano de Controle

- **Autenticação LDP:** Habilite autenticação MD5 em sessões LDP entre todos os LSRs para prevenir injeção de rótulos por dispositivos não autorizados na rede.

- **BGP com senhas e TTL Security:** Proteja sessões MP-BGP com senhas MD5 e habilite GTSM (RFC 5082) para bloquear pacotes BGP com TTL menor que o esperado, mitigando ataques de spoofing.

- **Controle de rótulos especiais:** Rótulos 0–15 são reservados (Implicit NULL, Explicit NULL, OAM). Nunca os distribua para prefixos de clientes, configure filtros LDP explícitos.

- **Isolamento do plano de gerência:** O acesso SSH e NETCONF a dispositivos MPLS deve ser feito exclusivamente via VRF de gerência separada, nunca via interfaces do backbone.

Monitoramento e Observabilidade

- **MPLS OAM com LSP-Ping e LSP-Traceroute:** Ferramentas definidas no RFC 8029 permitem verificar a integridade de LSPs de ponta a ponta, identificando exatamente onde pacotes são descartados ou mal encaminhados.

- **BFD agressivo em todas as interfaces:** Configure BFD com timers agressivos (300ms hello, 3× multiplier = 900ms detecção) para acionar FRR antes que os protocolos de roteamento detectem a falha.

- **Telemetria por streaming (gRPC/gNMI):** Substitua polling SNMP por telemetria streaming para visualização em tempo real do uso de banda, contadores de erros e estado de LSPs com resolução de segundos.

- **NetFlow/IPFIX nos PEs:** Habilite exportação de NetFlow v9 nos PEs para análise de padrões por cliente/VPN, capacitando engenharia de capacidade e detecção de anomalias de tráfego.

Planejamento de Capacidade e Resiliência

- **Regra dos 50%:** Dimensione enlaces do backbone para utilização normal abaixo de 50% da capacidade. Em falha de um enlace com FRR ativo, o tráfego redirecionado não deve causar congestionamento no bypass.

- **Diversidade de caminhos física:** LSPs de trabalho e de proteção devem seguir rotas fisicamente distintas, diferentes dutos, condutos e centros de distribuição. Diversidade lógica sem diversidade física não protege contra falhas de infraestrutura civil.

- **Documentação centralizada:** Mantenha registro de rótulos LDP, Node-SIDs, Route Distinguishers e Route Targets em ferramenta de inventário (como NetBox). Conflitos de RD/RT causam problemas difíceis de diagnosticar.

- **Testes de failover regulares:** Execute testes planejados de FRR e convergência IGP periodicamente, desativando interfaces em janelas de manutenção e medindo o tempo real de reconvergência, não apenas o tempo teórico.

Conclusão

O MPLS representa um dos pilares mais sólidos das redes de telecomunicações modernas. Sua capacidade de abstrair o encaminhamento para uma camada independente do protocolo de rede, combinada com mecanismos sofisticados de Engenharia de Tráfego, QoS e proteção contra falhas em milissegundos, fez dele a escolha natural para backbones de operadoras, redes corporativas de grande escala e interconexão de data centers no mundo inteiro.

A evolução para o Segment Routing simplifica operacionalmente o MPLS eliminando LDP e RSVP-TE como protocolos independentes, mas mantém o plano de dados MPLS familiar e amplamente suportado. O SRv6 aponta para o futuro com programabilidade nativa sobre IPv6, mas ainda enfrenta desafios de adoção e suporte por parte dos fabricantes.

Para profissionais de redes, dominar o MPLS é fundamental para trabalhar em qualquer ambiente de backbone de médio a grande porte. A progressão natural de aprendizado é:

1. **Comece pelo LDP:** Entenda a distribuição de rótulos e a construção da LFIB.

2. **Implemente L3VPNs:** Configure VRFs, RDs, RTs e MP-BGP para isolar clientes.

3. **Adicione MPLS-TE e FRR:** Habilite RSVP-TE para engenharia de tráfego e proteção em milissegundos.

4. **Migre para SR-MPLS:** Simplifique a operação com Segment Routing e TI-LFA.

Referências

[1] IETF RFC 3031. (2001). Multiprotocol Label Switching Architecture. Rosen, E. et al.

[2] IETF RFC 4271. (2006). A Border Gateway Protocol 4 (BGP-4). Rekhter, Y. et al.

[3] IETF RFC 3031. (2001). MPLS Architecture — visão geral do plano de dados.

[4] IETF RFC 3032. (2001). MPLS Label Stack Encoding. Rosen, E. et al.

[5] IETF RFC 3209. (2001). RSVP-TE: Extensions to RSVP for LSP Tunnels. Awduche, D. et al.

[6] IETF RFC 4090. (2005). Fast Reroute Extensions to RSVP-TE for LSP Tunnels. Pan, P. et al.

[7] IETF RFC 8402. (2018). Segment Routing Architecture. Filsfils, C. et al.

[8] IETF RFC 8754. (2020). IPv6 Segment Routing Header (SRH). Filsfils, C. et al.

[9] IETF RFC 4364. (2006). BGP/MPLS IP Virtual Private Networks (L3VPN). Rosen, E.; Rekhter, Y.

[10] IETF RFC 7432. (2015). BGP MPLS-Based Ethernet VPN (EVPN). Sajassi, A. et al.

[11] IETF RFC 5036. (2007). LDP Specification. Andersson, L. et al.

[12] IETF RFC 8029. (2017). Detecting MPLS Data-Plane Failures (LSP-Ping). Kompella, K. et al.

Apêndice A - Checklist de Verificação MPLS

Plano de Controle LDP

□ MPLS habilitado em todas as interfaces do backbone (mpls ip)
□ Sessões LDP estabelecidas entre todos os vizinhos (show mpls ldp neighbor)
□ Rótulos distribuídos para todos os loopbacks dos PEs (show mpls ldp bindings)
□ Autenticação MD5 habilitada em todas as sessões LDP
□ Filtros LDP aplicados para bloquear rótulos reservados (0–15)

L3VPN

□ VRFs criadas com Route Distinguishers únicos em todos os PEs
□ Route Targets de exportação/importação corretos para a topologia
□ Sessões MP-BGP estabelecidas entre todos os PEs (full-mesh ou via Route Reflector)
□ Prefixos VPNv4 propagados corretamente (show bgp vpnv4 unicast all)
□ Conectividade ponta a ponta testada entre CEs de todos os sites

Resiliência e Monitoramento

□ BFD habilitado em todas as interfaces do backbone com timers adequados
□ FRR configurado e testado, com bypass LSPs pré-estabelecidos e verificados
□ Marcação de QoS (TC bits) validada para cada classe de serviço
□ Nenhuma interface do backbone com utilização acima de 50% em estado normal
□ Telemetria ou NetFlow habilitados nos PEs para visibilidade de tráfego
□ Logs centralizados com retenção mínima de 90 dias
□ Documentação atualizada: RDs, RTs, Node-SIDs e mapa de LSPs
□ Testes de failover agendados e realizados ao menos trimestralmente