

**LIGA DE ENSINO DO RIO GRANDE DO NORTE
FACULDADE NATALENSE PARA O DESENVOLVIMENTO DO RN
ESPECIALIZAÇÃO EM REDES DE COMPUTADORES**

A TECNOLOGIA MPLS E SEUS SERVIÇOS

JOSSENILSON DE MELO BEZERRA

**NATAL/RN
2010**

JOSSENILSON DE MELO BEZERRA

A TECNOLOGIA MPLS E SEUS SERVIÇOS

Monografia apresentada ao curso de Especialização em Redes de Computadores da Faculdade Natalense para o Desenvolvimento do Rio Grande do Norte como requisito parcial para a obtenção do título de Especialista em Redes de Computadores.

**Orientador:
Aluizio Ferreira da Rocha Neto (M.Sc.)**

**NATAL/RN
2010**

AGRADECIMENTOS

Agradeço a Deus, autor do conhecimento universal e de todas as luzes por ter inspirado em mim a busca do conhecimento e iluminado os meus caminhos nos momentos de escuridão.

Agradeço aos meus pais que me apoiaram durante toda vida para alcançar meus objetivos profissionais e pessoais.

Agradeço ao meu orientador, Aluizio Ferreira da Rocha Neto, por ser muito gentil e competente na orientação desta etapa final do curso.

Agradeço ao meu Gerente Técnico da Embratel/RN, Alfredo Gama de Carvalho Junior e toda sua equipe (meus amigos e amigas de trabalho) pela contribuição no meu crescimento profissional, agregando conhecimento e experiência, as quais serviram de embasamento empírico para elaboração deste trabalho.

JOSSENILSON DE MELO BEZERRA

A TECNOLOGIA MPLS E SEUS SERVIÇOS

Monografia apresentada e aprovada em ____ de ____ de _____, pela banca examinadora
composta pelos seguintes membros

Aluizio Ferreira da Rocha Neto., M.Sc.
Orientador

Gleydson A. F. Lima, M.Sc.

Natal/RN
2010

RESUMO

A presente Monografia tem por objetivo apresentar um estudo realizado sobre a tecnologia MPLS com enfoque nos seus principais serviços. O MPLS (*Multi-Protocol Label Switching*) é uma tecnologia emergente que está sendo largamente utilizada por muitos Provedores de Serviços de Internet e vem substituindo com eficiência outras tecnologias, como por exemplo, o Frame-Relay. O MPLS utiliza os núcleos das redes dos ISP (*Interne Service Provider*) para promover uma integração entre os paradigmas de comutação e de roteamento, desta forma esta Tecnologia aumenta o desempenho das redes atualmente instaladas, além de facilitar as implementações de vários serviços, como por exemplo: Qualidade de Serviço, Redes Privadas Virtuais e Engenharia de Tráfego.

Palavras Chaves: MPLS. Qualidade de Serviços. Rede Privada Virtual. Engenharia de Tráfego.

ABSTRACT

This monograph aims to present a study on the MPLS technology with a focus on its core services. The MPLS (Multi-Protocol Label Switching) is an emerging technology that is being widely used by many Internet Service Providers and is replacing efficiency with other technologies such as Frame-Relay. MPLS uses the cores of the networks of ISPs (Internet Service Provider) to promote integration between the paradigms of switching and routing, so this technology increases the performance of currently installed networks, and facilitate the implementations of various services, such as Example: Quality of Service, Virtual Private Networks and Traffic Engineering.

Keywords: MPLS. Quality of Service. Virtual Private Network. Traffic Engineering.

SUMÁRIO

1	INTRODUÇÃO	11
1.1	OBJETIVOS	12
1.1.1	Objetivos Específicos	12
1.2	ORGANIZAÇÃO DO TRABALHO	12
2	REFERENCIAL TEÓRICO	13
2.1	PRINCÍPIOS BÁSICOS.....	13
2.1.1	Roteamento.....	13
2.1.2	Comutação	14
2.2	COMPARAÇÃO ENTRE REDES ROTEADAS E COMUTADAS	15
2.3	MPLS (MULTI PROTOCOLO LABEL SWITCHING)	16
2.3.1	Label (Rótulo)	17
2.3.2	Espaço de Rótulos	18
2.3.3	LER (Label Edge Router).....	19
2.3.4	LSR (Label Switch Router)	19
2.3.5	LSP (Label Switch Path)	20
2.3.6	FEC (Forwarding Equivalence Class)	21
2.3.7	NHLFE (Next Hop Label Forward Entry)	21
2.3.8	LFIB (Label Forwarding Information Base)	22
2.3.9	Empilhamento de Rótulos	23
2.3.10	LDP (Label Distribution Protocol).....	24
2.3.11	Funcionamento da rede MPLS	26
2.3.12	Funcionamento do LDP	28
2.4	PROTOCOLOS DE ROTEAMENTOS	31
2.4.1	OSPF (Open Short Path First)	32
2.4.1.1	Características.....	32
2.4.1.2	Definições do OSPF	33
2.4.1.3	Funcionamento do protocolo OSPF	34
2.4.2	BGP-4	39
2.4.2.1	Características.....	39
2.4.2.2	Funcionamento do BGP	39
2.4.2.3	Atributos do BGP	41
2.4.2.4	Processo Decisório	42
2.5	QUALIDADE DE SERVIÇO.....	43
2.5.1	Serviços Diferenciados (DiffServ)	45
2.5.1.1	Arquitetura DiffServ	46
2.5.2	Serviços Integrados (IntServ)	49
2.5.2.1	Arquitetura IntServ	50
3	SERVIÇOS.....	55
3.1	VIRTUAL PRIVATE NETWORK (VPN)	55
3.1.1	Funcionamento da VPN MPLS	56
3.1.1.1	VRF (Virtual Routing and Forwarding)	57
3.1.1.2	Divulgação de Rotas VPN IPv4	58
3.1.1.3	Transporte do Tráfego do Cliente.....	59
3.2	QUALIDADE DE SERVIÇO (QoS)	60
3.3	ENGENHARIA DE TRÁFEGO (TE)	62

3.3.1 Funcionamento	63
4 CONSIDERAÇÕES FINAIS.....	66
REFERÊNCIAS	67

LISTA DE FIGURAS

Figura 1: Rede Roteada	14
Figura 2: Rede Comutada	15
Figura 3: Rótulo MPLS	17
Figura 4: Tabela do Nó com Espaço de Rótulos GLOBAL	18
Figura 5: Tabela do Nó com Espaço de Rótulos LOCAL	18
Figura 6: Label Edge Router – LER	19
Figura 7: Label Switch Router – LSR	20
Figura 8: Label Switch Path – LSP	20
Figura 9: Empilhamento de Rótulos	24
Figura 10: PDU	26
Figura 11: Exemplo de funcionamento do MPLS	26
Figura 12: Mensagens do Label Distribution Protocol	28
Figura 13: Relação entre Sistemas Autônomos, EGP e IGP	32
Figura 14: Conexões em redes com e sem DR/BDR	35
Figura 15: Exemplo de Topologias de Redes	36
Figura 16: Troca de pacotes para preencher as bases de dados	37
Figura 17: Bases de Dados nos Roteadores	37
Figura 18: Caminho Menor criado pelo OSPF baseado no Custo	38
Figura 19: Funcionamento do BGP	40
Figura 20: Divulgação de rotas com atributo Path	42
Figura 21: Exemplo comparativo de Qualidade de Serviço.	44
Figura 22: Exemplo da arquitetura DiffServ	46
Figura 23: Localização do DSCP no pacote IP	47
Figura 24: Exemplo de Rede com DiffServ.	48
Figura 25: Exemplo de Balde de Fichas	52
Figura 26: Exemplo de Objetos transportados nas Mensagens do RSVP	53
Figura 27: Sinalização do protocolo RSVP	53
Figura 28: Rede emulada em laboratório.	56
Figura 29: VRF (<i>virtual routing and forwarding</i>)	57
Figura 30: Divulgação das rotas VPN IPv4	58
Figura 31: Transferência dos Dados das VPNs	59
Figura 32: No E-LSP cada PHB é identificado pelo EXP de cada pacote	60
Figura 33: No L-LSP cada PHB é identificado pelo LSP correspondente	60
Figura 34: Modo Uniform	61
Figura 35: Modo Short-Pipe e Pipe	62
Figura 36: Rede com MPLS-TE	63

LISTA DE QUADROS

Quadro 1: Diferenças entre redes comutadas e roteadas.....	16
Quadro 2: FEC (Forwarding Equivalence Class)	21
Quadro 3: NHLFE (Next Hop Label Forward Entry).	22
Quadro 4: FEC-To-NHLFE (FTN).....	22
Quadro 5: Incoming Label Map – ILM.....	23
Quadro 6: Relação entre tráfego e parâmetros de rede.....	43
Quadro 7: Relação entre IP Precedence do TOS e DSCP	47

INTRODUÇÃO

Visando aumentar o desempenho dos dispositivos de comutação de redes, entre outras soluções, foi desenvolvida uma nova tecnologia de comutação de pacote¹. Esta nova tecnologia de comutação de pacotes tinha por princípio, a leitura de uma pequena quantidade de informação, conhecida como rótulo, que é acrescentada ao cabeçalho para direcionar o pacote, em oposição à leitura completa cabeçalho do mesmo, usada na tecnologia TCP/IP². Em princípio com menos informação, há um tempo menor para leitura do pacote e mais rápido o mesmo é repassado, conseqüentemente aumenta a eficiência de comutação das redes. Esta foi a idéia inicial da nova tecnologia de comutação de pacotes, mas com a evolução do hardware esta nova tecnologia logo ficou obsoleta, pois passaram a ser construídos os “wirespeed”, equipamentos que são capazes de comutar todos os pacotes que lhe são entregues, na mesma velocidade da interface. Atualmente foram desenvolvidas novas funcionalidades para obsoleta tecnologia de comutação de pacotes, que ficou conhecida como MPLS (*Multiprotocol Label Switching*). O MPLS fornecer serviços como VPN³, Qos⁴ e TE⁵, que o mercado necessita, devido a novas aplicações. Com esta nova visão o IETF (*Internet Engeneering Task Force*) padronizou o MPLS, com sua definição básica descrita na RFC (*Request For Comments*) 3031.

Na tecnologia MPLS, o MP (*Multi-Protocol*) significa que se pode transportar vários tipos de protocolos encapsulados através comutação de rótulos, expresso por LS (*Label Switching*), ou seja, MPLS tem por objetivo funcionar sobre várias tecnologias de enlace, bem como, transportar vários tipos de pacotes.

¹ Pacote - Unidade de informação que compõe um fluxo de dados
² TCP/IP conjunto de protocolos utilizados na Internet
³ Vitual Private Network – Rede Privada Virtual
⁴ Quality Of Service – Qualidade de Serviço
⁵ TE (Traffic Engeneering) – Engenharia de Tráfego

1.1 OBJETIVOS

Este projeto tem como objetivo estudo detalhado da Tecnologia MPLS, expondo os principais conceitos envolvidos na tecnologia necessário a sua implementação, bem como, também destacar os principais serviços do MPLS utilizados atualmente.

1.1.1 Objetivos Específicos

- Realizar pesquisa dos principais protocolos usados em uma rede MPLS.
- Analisar o uso de Protocolos de Roteamento.
- Realizar pesquisa dos principais dos principais serviços oferecidos pelo MPLS.
- Realizar pesquisa dos modelos de qualidade de serviço.
- Implementações, em laboratórios emulados, dos serviços.

1.2 ORGANIZAÇÃO DO TRABALHO

No capítulo dois é apresentado os princípios básicos da usados na tecnologia MPLS, como também, toda a terminologia usada. Desta forma, pode-se ter um melhor embasamento nos capítulos subseqüentes.

Depois de familiarizado com os termos e princípios básicos, ainda no capítulo dois será descrito o funcionamento de outros protocolos usados em conjuntos com Tecnologia MPLS com enfoque nos mais utilizadas no mercado atual, a saber, OSPF, BGP e RSVP.

No Capítulo três será descrito as características e funcionamento dos principais serviços que a Tecnologia MPLS dispõe, como também, as implementações destes pela tecnologia.

No Capítulo quatro há algumas considerações finais sobre o assunto, descrevendo os principais benefícios que a tecnologia MPLS dispõe.

2 REFERENCIAL TEÓRICO

2.1 PRINCÍPIOS BÁSICOS

De acordo com PAQUET e TEARE (2003, p.3):

O mecanismo de aprendizado e manutenção do conhecimento da topologia de rede é considerado como a função de **Roteamento**. O movimento real do tráfego transiente por meio do roteador, da interface de entrada para uma interface de saída é uma função separada considerada como a função de **Comutação**.

Segue abaixo um exemplo de uma rede roteada e uma rede comutada e logo após seguem as principais diferenças entre ambas.

2.1.1 Roteamento

Em uma rede de roteada, conforme exemplificado na figura 1, os pacotes são transmitidos de um ponto a outro da rede através de vários nós⁶ interligados, cada nó possui uma **Tabela de Roteamento** que foi preenchida dinamicamente, pelos protocolos de roteamento, ou manualmente, pelo administrador da rede. A tabela de roteamento possui o mapeamento *Destino do Pacote* para *Interface de Saída*, para cada pacote que entra na interface, o nó lê o endereço de destino do pacote e consulta a tabela de roteamento, se houver uma igualdade entre o Endereço de destino do pacote e endereço de destino da tabela de roteamento o pacote é enviado pela interface de saída, se não o pacote é enviado pela interface de saída padrão, se estiver configurada, conhecida como **Rota Default**.

⁶ Dispositivo que permite a comutação ou o roteamento de pacotes que pode ser um switch, um roteador ou ambos

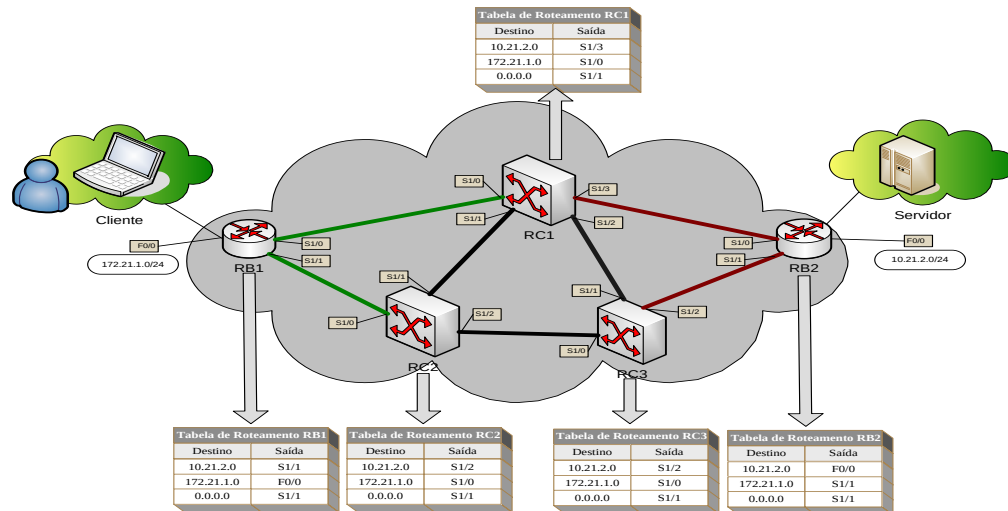


Fig. 1: Rede Roteada
Fonte: Elaborado pelo Autor (2009)

Este processo é repetido a cada nó que o pacote passa até chegar ao destino, desta forma existe a vantagem de não precisar conhecer o caminho prévio até o destino, pois o caminho é feito no momento da transmissão do pacote, outra vantagem é de poder haver vários caminhos, aumentando a disponibilidade da rede no caso de falha de algum deles. No caso de haver inconsistência no preenchimento das tabelas de roteamento o pacote pode ficar circulando na rede por um período máximo definido no pacote, esta é uma desvantagem e tal circunstância é conhecida como **Loop**.

2.1.2 Comutação

No processo de comutação, diferentemente do roteamento, o caminho deve ser previamente estabelecido e somente há um caminho na transmissão da informação, se houver falha neste caminho, deverá ser refeito outro caminho de forma manual. Na comutação, semelhantemente a Tabela de roteamento há uma **Tabela de Encaminhamento**, conforme exemplificado na figura 2.

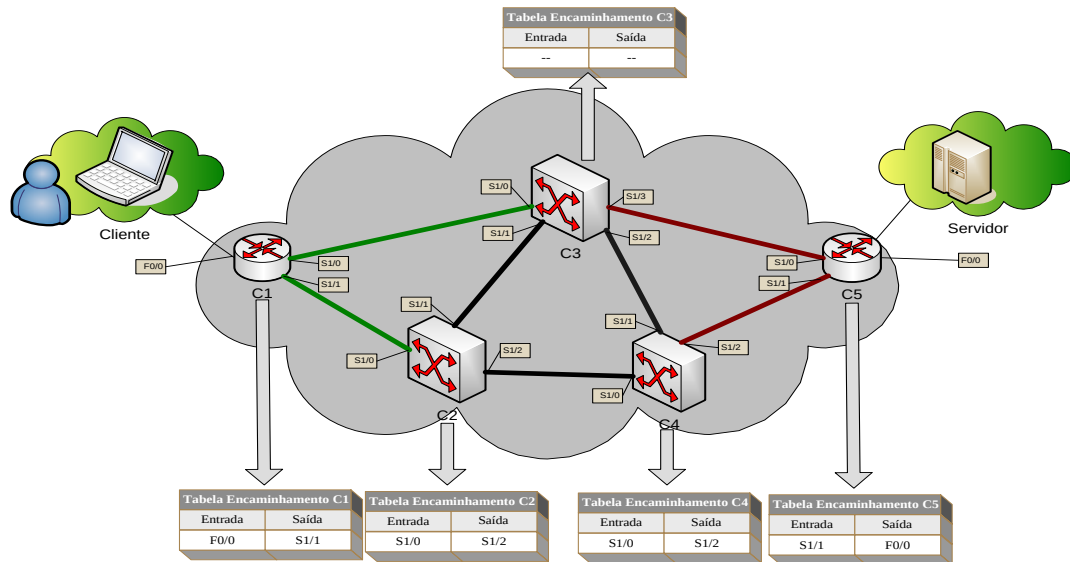


Fig. 2: Rede Comutada
Fonte: Elaborado pelo Autor (2009)

Mas no caso da rede comutada as informações de encaminhamento são estáticas e o comutador simplesmente comuta o pacote que chega por uma interface de entrada para uma interface de saída, ao contrário da tabela de roteamento, na qual as informações são dinâmicas, pois se não houver uma rota para destino existe uma rota default, que no caso do comutador **não existe**, pois que o caminho a ser seguido deve ser o especificado previamente, assim se houver uma falha no link a comunicação permanece inoperante até que seja reconfigurado outro caminho.

2.2 COMPARAÇÃO ENTRE REDES ROTEADAS E COMUTADAS

Nas seção anterior foi explicado o funcionamento conceitual das redes comutadas e roteadas. No Quadro 1 seguem as principais diferenças encontradas em ambas as redes, cada modelo de rede possui vantagens e desvantagens, de acordo com a necessidade do serviço pode-se escolher o modelo de rede que mais se adapte ao projeto.

	Roteamento	Comutação
Configuração de Circuitos	Desnecessária	Obrigatória
Endereçamento	Cada pacote contém os endereços de origem e destino completos	Cada pacote contém um único número que identifica o circuito
Informações sobre o estado	Os roteadores não armazenam informações sobre o estado das conexões	Cada circuito requer espaço na tabela do nó, para manter as informações dos circuitos.
Roteamento	Cada pacote é roteado independente	A rota é escolhida quando o circuito estabelecido e todos os pacotes seguem a mesma rota
Efeitos de falhas no nó	Nenhum com exceção dos pacotes perdidos durante a falha	Todos os circuitos que tiverem passando pelo nó defeituoso serão encerrados
Qualidade de Serviço	Difícil	Fácil
Controle de Congestionamento	Difícil	Fácil

Quadro 1: Diferenças entre redes comutadas e roteadas.

Fonte: Adaptado de Tanenbaum (2003, p.37)

2.3 MPLS (*Multi Protocolo Label Switching*)

O paradigma MPLS é uma estratificação do que há de melhor no paradigma de roteamento com o que há de melhor no paradigma de comutação, ou seja, há uma junção da vantagem da dinâmica do roteamento com as vantagens de controle de encaminhamento e rapidez da comutação.

O MPLS utiliza o roteamento baseado no IP para prover a dinâmica inerente às redes roteadas e a tecnologia de rótulos para prover controle no encaminhamento, que é uma característica inerente a redes comutadas.

Nas seções seguintes serão apresentados os termos e conceitos usados na arquitetura de uma rede MPLS.

2.3.1 Label (Rótulo)

O rótulo (Label) é um identificador que é usado no pacote para encaminhá-lo pela rede, baseado em um caminho previamente criado. É através do rótulo que o pacote pode alcançar o seu destino na rede MPLS. Há dois modos de se identificar um rótulo:

- a) **Frame** – Na figura 3 temos um exemplo deste modo amplamente difundido, o rótulo de quatro bytes é inserido entre o cabeçalho de nível dois e o cabeçalho de nível três do pacote, desta forma há uma alteração no tamanho do pacote original. O rótulo possui um campo identificador de rótulo 20 bits, entre outros campos que influenciam na forma como o pacote é comutado na rede, totalizando 32bits (04 bytes).

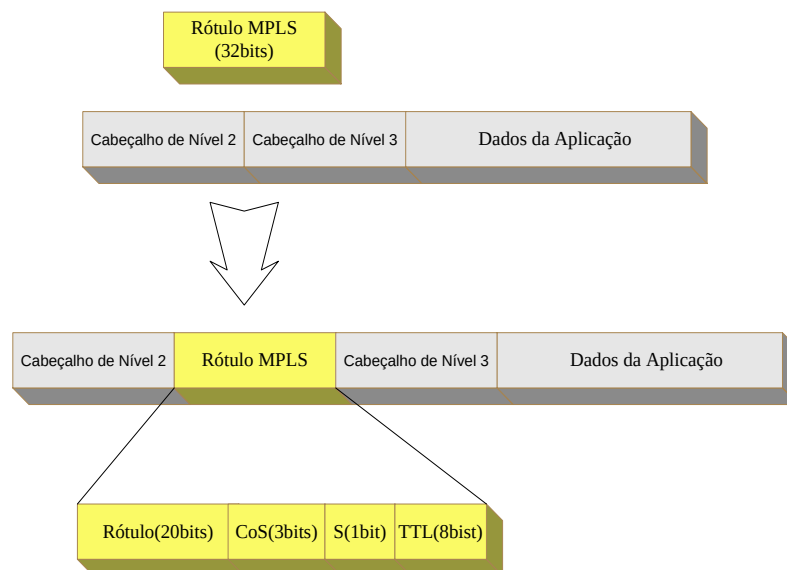


Fig. 3: Rótulo MPLS
Fonte: Elaborado pelo Autor (2009)

- **Rótulo ID (20bits)** – É um identificador numérico no intervalo de 2^{20} . Os valores de 0 a 15 são reservados, mas os valores de 16 até 1.048.575 são válidos.
- **COS (03bits)** – Também conhecido como EXP (experimental bits), inicialmente foram criados para uso experimental, mas geralmente são usados para classificar o pacote, baseado na qualidade de serviço contratado na rede, através de três bits é disponibilizado um total de 08 classes.
- **S (01bit)** – Usado para indicar que o pacote é o último de uma pilha, empilhamento de rótulos será explicado posteriormente.
- **TTL (08bits)** – Geralmente é uma cópia do campo TTL do cabeçalho IP e possui a

mesma função, evitar que o pacote circule indefinidamente pela rede.

- b) **Célula** - o rótulo é identificado por outros campos, como por exemplo, no caso das redes ATM que o rótulo é identificado pelo par VCI/VPI, é geralmente utilizado no plano de controle do ATM em substituição da sinalização para troca dos pares VCI/VPI, é um modo pouco utilizado, sendo citado apenas a título de conhecimento.

2.3.2 Espaço de Rótulos

O espaço de rótulos é a forma como este endereçamento será usado para construir os rótulos e atribuí-los as interfaces de uma plataforma (Nó). Há dois modos de espaço de rótulos:

- Por Plataforma (Global) – O espaço de rótulos é único para toda a plataforma, ou seja, em cada interface há um único rótulo, que não pode ser duplicado para ser usado em outra interface do mesmo nó, conforme exemplo da figura 4.

Global	
Rótulo	Interface
40	S1/0
32	S2/0
50	F0/0
12	F0/0

Fig. 4: Tabela do Nó com Espaço de Rótulos GLOBAL
Fonte: Elaborado pelo Autor (2009)

- Por Interface (Local)– Na figura 5, o espaço de rótulos é único por interface, ou seja, não pode haver rótulos iguais na mesma interface, mas pode haver rótulos iguais na mesma plataforma, entretanto em interfaces diferentes.

Local	
Rótulo	Interface
40	S1/0
40	S2/0
50	S3/0
50	F0/0

Fig. 5: Tabela do Nó com Espaço de Rótulos LOCAL
Fonte: Elaborado pelo Autor (2009)

2.3.3 LER (Label Edge Router)

É o nó de ingresso na borda da rede, ou seja, é o nó pelo qual todos os pacotes normais, sem rótulos inseridos, entram na rede. O LER tem por objetivo inserir o rótulo no pacote, baseado em um conjunto de características definidas anteriormente, para então enviar o novo pacote para o próximo nó, geralmente um LSR. A figura 6 é um exemplo de funcionamento do LER.

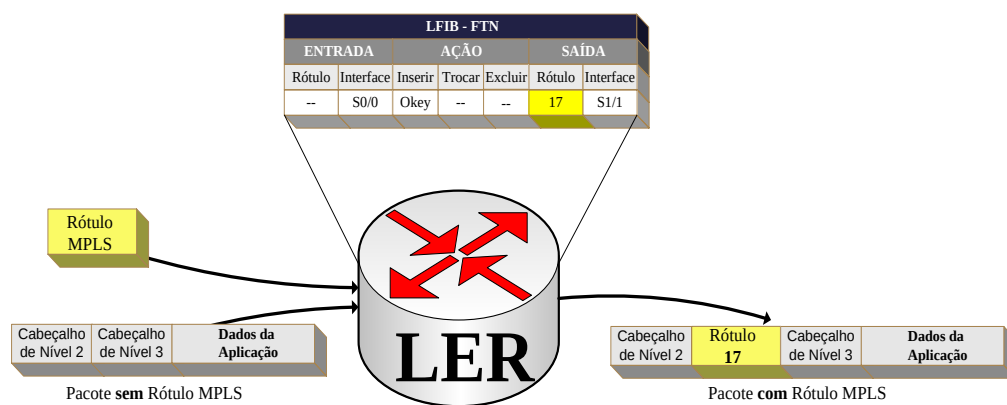


Fig. 6: Label Edge Router – LER
Fonte: Elaborado pelo Autor (2009)

2.3.4 LSR (Label Switch Router)

É o nó que recebe o pacote com o rótulo inserido, oriundo do LER, e o repassa para o próximo LSR alterando apenas o rótulo de acordo com tabela de encaminhamento anteriormente preenchida. A tabela de encaminhamento pode ser preenchida estática ou dinamicamente. O LSR normalmente altera apenas no valor do rótulo e o repassa, esta é sua função principal, mas pode opcionalmente inserir ou retirar rótulos se for necessário empilhá-los, o empilhamento será descrito mais adiante. A figura 7 demonstra um exemplo do LSR.

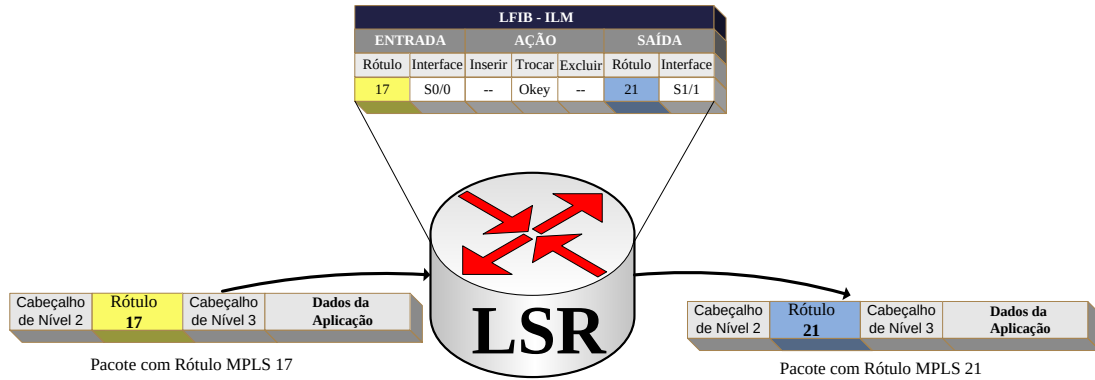


Fig. 7: Label Switch Router – LSR
Fonte: Elaborado pelo Autor (2009)

2.3.5 LSP (Label Switch Path)

É o caminho percorrido pelo pacote, dentro da rede MPLS, ou seja, a sequência de associações definidas em cada uma das Tabelas de Encaminhamento dos LER e LSR, que associam um rótulo de entrada por uma interface a um rótulo de saída por outra, desta forma criam, através desta lista ordenada, um caminho único pelo quais todos os pacotes deverão passar, conforme exemplificado na figura 8.

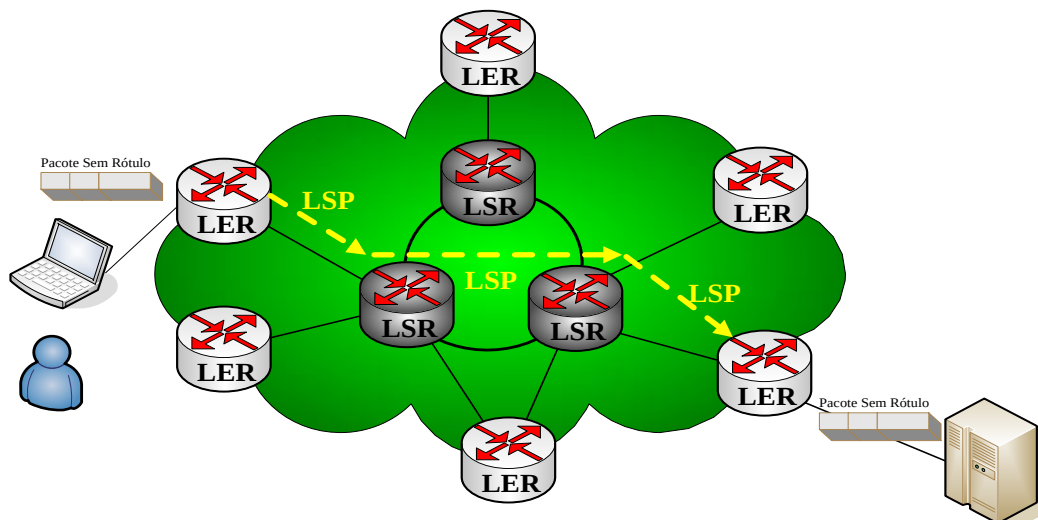


Fig. 8: Label Switch Path – LSP
Fonte: Elaborado pelo Autor (2009)

2.3.6 FEC (Forwarding Equivalence Class)

A FEC (Classe de Equivalência de Encaminhamento) é um conjunto de parâmetros que definem um fluxo de dados⁷, que são mapeados para um rótulo específico, ou seja, um grupo de pacotes que possuem as mesmas características e recebem o mesmo rótulo. A FEC dispõe de enormes variedades de parâmetros que podem ser selecionados, como por exemplo, o endereço IP, porta, DSCP⁸ entre outros.

O fluxo de uma FEC é unidirecional, sendo necessários dois fluxos, em sentido contrário para se estabelecer uma conexão full-duplex⁹.

No exemplo do Quadro 2 foram definidas três FEC, baseado em alguns parâmetros, por exemplo, podemos definir como FEC 1 todos os pacotes com destino a porta 80, podemos definir como FEC 2 todos os pacotes com destino a 10.21.2.0 na porta 21 e como FEC 3 todos os pacotes com origem 172.21.1.0, destinados a 10.21.2.0 na porta 22 e com o campo DSCP configurado para EF. Toda esta classificação é feita no LER, o qual tem a função de classificar o pacote e atribuí-lo ao rótulo.

FEC(Forwarding Equivalence Class)					
FEC	ORIGEM		DESTINO		DSCP
	Endereço	Porta	Endereço	Porta	
1	--	--	--	80	--
2	--	--	10.21.2.0	21	--
3	172.21.1.0	--	10.21.2.0	22	EF

Quadro 2: FEC (Forwarding Equivalence Class)
Fonte: Elaborado pelo Autor (2009)

2.3.7 NHLFE (Next Hop Label Forward Entry)

O NHLFE (Registro de Encaminhamento por Rótulos para o Próximo Passo) é um registro em uma tabela que mapeia uma interface para um rótulo de saída e qual ação deve ser executada (inserção, troca ou exclusão do rótulo). O Quadro 3 possui vários exemplos de NHLFE.

⁷ Sequência de um conjunto de pacotes produzidos por uma determinada aplicação.

⁸ Campo no cabeçalho do pacote IP usado para qualificar o tipo de serviço oferecido pela rede ao pacote.

⁹ Transmissão de dados em sentidos contrários e concomitantemente.

NHLFE (Next Hop Label Forward Entry)				
Interface	Ação			Rótulo
	Inserir	Trocar	Excluir	
S1/0	Okey	--	--	17
S1/1	--	Okey	--	21
S1/0	--	--	Okey	35
S1/2	Okey	--	--	70

Quadro 3: NHLFE (Next Hop Label Forward Entry).
Fonte: Elaborado pelo Autor (2009)

2.3.8 LFIB (Label Forwarding Information Base)

É uma tabela que mapeia pacotes ou rótulos que entram no nó para suas respectivas interfaces de saídas, existem dois tipos de LFIB:

- FTN (FEC-to-NHLFE) – É uma LFIB que mapeia FEC para um ou mais registro NHLFE, isto é, para cada pacote, com características definidas, que entra é associado através de FEC a um ou mais NHLFE, que possui o rótulo de saída e ação que deve ser executado, que neste caso deve ser sempre a inserção ou retirada de rótulo, conforme exemplificado no Quadro 4. A FTN é utilizada no geral no LER.

FTN(FEC-to-NHLFE)										
Interface	Ação			Rótulo	FEC	ORIGEM		DESTINO		DSCP
	Inserir	Trocar	Excluir			Endereço	Porta	Endereço	Porta	
S1/0	Okey	--	--	17	1	--	--	--	80	--
F0/0	--	--	Okey	21	2	--	--	10.21.2.0	21	--
S1/0	Okey	--	--	35	3	172.21.1.0	--	10.21.2.0	22	EF
S1/2	Okey	--	--	70	3	172.21.1.0	--	10.21.2.0	22	EF

Quadro 4: FEC-To-NHLFE (FTN)
Fonte: Elaborado pelo Autor (2009)

- ILM (Incoming Label Map) – É uma LFIB que mapeia rótulos de entradas em rótulos de saída (NHLFE), o Quadro 5 é um exemplo. Para cada rótulo entrante é associado (mapeado) a um rótulo de saída, além de associar a ação a ser executada, que neste caso pode ser a troca, inserção de um novo rótulo formando uma pilha ou a retirada do rótulo da pilha ou do pacote. Se for necessário fazer o balanceamento de carga deve-se associar o rótulo de entrada a vários NHLFE.

ILM(Incomming Label Map)						
Entrada		Ação			Saída	
Interface	Rótulo	Inserir	Trocar	Excluir	Rótulo	Interface
F0/0	17	--	Okey	--	21	S1/1
S1/0	21	Okey	--	--	27	S1/3
S1/1	35	--	Okey	--	43	S1/0
S1/2	70	--	--	Okey	40	S1/2

Quadro 5: Incomming Label Map – ILM
Fonte: Elaborado pelo Autor (2009)

2.3.9 Empilhamento de Rótulos

Durante o trajeto do pacote pela rede, pode ser necessário empilhá-lo, seja para adicionar novas características (VPN e TE) ou para passar por diversos túneis no core¹⁰ da rede.

O empilhamento ocorre quando um novo rótulo é acrescentado a um pacote rotulado. O roteador somente enxerga o rótulo mais externo, isto é, todo o processamento do pacote é baseado no último rótulo inserido, para o roteador é como se não existissem outros rótulos internos ao pacote.

Conforme for necessário o pacote será desempilhado, retirando-se sempre o rótulo mais externo, se na retirada do rótulo mais externo houver ainda rótulo interno, o roteador passa encaminhar o pacote baseado neste rótulo interno que será, a partir deste ponto, considerando rótulo mais externo.

¹⁰ Core é núcleo da rede, onde estão situados os nós de maior capacidade de processamento.

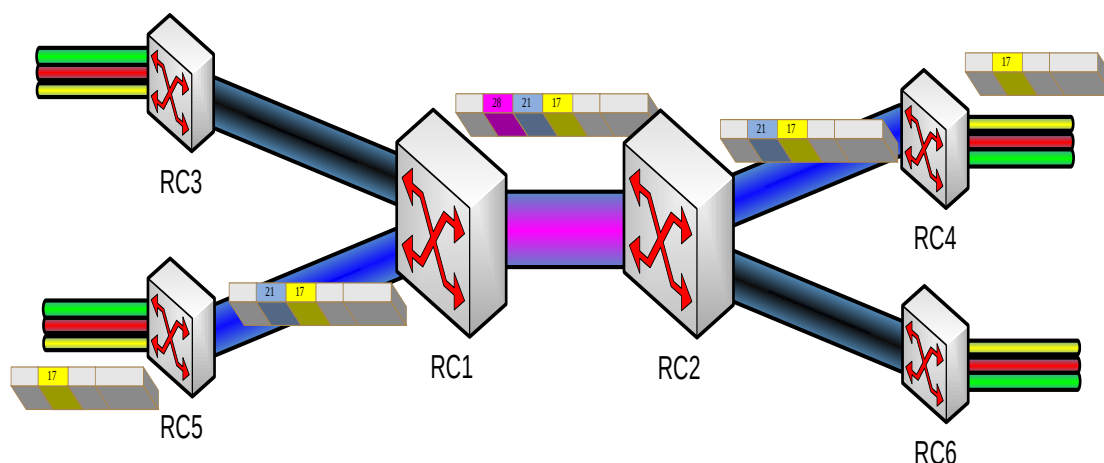


Fig. 9: Empilhamento de Rótulos
Fonte: Elaborado pelo Autor (2009)

Conforme a figura 9, o RC5 recebe o pacote com rótulo 17, este nó contém a informação que deve acrescentar o rótulo 21, para pacotes com rotulo 17, e enviar para RC1. O pacote com rótulo 21 ao chegar no RC1 que também possui a informação que deve acrescenta o rótulo 28 aos pacotes com rótulo 21. O RC1 não enxerga o rótulo 17, pois não possui nenhuma informação do rótulo 17, é como ele não existisse, apesar de estar no pacote, mas nó somente enxerga o rótulo mais externo. O pacote ao chegar no RC2, este somente tem conhecimento que deve retirar rótulo 28, ao retirar o rótulo 28, passa então a enxergar o rótulo 21, para então enviá-lo para o RC4, o RC4 por sua vez sabe que deve retirar o rótulo 21 e passa então a enxergar o rótulo 17, para então enviá-lo a interface correta.

2.3.10 LDP (Label Distribution Protocol)

É o protocolo de distribuição de rótulos, é através do LDP que as tabelas FTN e ILM são preenchidas automaticamente tornando a rede tolerante a falhas e escalável, pois se a tabelas forem preenchidas manualmente através de telnet¹¹ ou SNMP¹² não haveria tolerância a falhas, nem escalabilidade. Há duas formas de distribuição de rótulos pelo LDP:

¹¹ Protocolo para acesso remoto a um terminal ou nó.

¹² Simple Network Manager Protocol - Protocolo utilizado no gerenciamento de rede.

- ❖ DOWNSTREAM – DOU (Não Solicitado) – Neste caso o nó de destino fluxo de dados (egresso) envia para o nó de origem (ingresso) um rótulo para uma determinada FEC sem que seja feita uma solicitação, portanto haverá inúmeros rótulos não solicitados dependendo da tabela de roteamento, pois para cada rota no destino é enviado um rótulo para a origem.
- ❖ UPSTREAM – DOD (Sob Demanda) – Diferentemente do caso anterior, é necessário que o nó de origem (ingresso) solicite ao nó de destino um rótulo, que terá o valor escolhido pelo nó de destino (egresso), portanto somente haverá rótulos solicitados que serão de escolhidos de acordo com o tipo fluxo definido pela FEC.

Há também dois modos como os rótulos, recebidos de outros nós e não usados pelo receptor, são tratados:

- ✓ CONSERVADOR – Os rótulos são liberados desocupando recursos do nó, desta forma pode-se evitar que rótulos para próximos nós inválidos ocupem recursos, uma desvantagem é que se for necessário um rótulo para o próximo nó que havia sido recebido, deve ser feita uma nova solicitação, pois o anterior foi descartado.
- ✓ LIBERAL – Os rótulos são armazenados no nó para posterior utilização, neste modo aumenta a utilização de recursos do nó, entretanto aumenta a rapidez na utilização de um novo rótulo, não sendo necessário fazer uma nova solicitação, pois o rótulo fica armazenado no receptor.

Os pacotes LDP são formados por conjunto lógico de segmentos encapsulados, chamado de PDU (Protocol Data Unit), a **unidade de dados do protocolo (PDU)** é a informação que deve ser trocada entre os nós que possuem LDP, cada PDU é composta por basicamente por um cabeçalho, que são os mesmos em todas as unidades de informação do LDP, e mais várias mensagens.

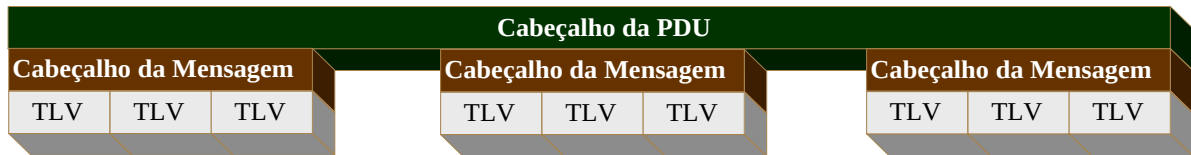


Fig. 10: PDU
Fonte: Elaborado pelo Autor (2009)

Cada mensagem é composta por várias unidades, no formato TLV (Type-Length-Value)¹³, estas últimas unidades possuem um tamanho arbitrário que torna o protocolo altamente extensível. A figura 10 é um exemplo de PDU e suas respectivas partes.

2.3.11 Funcionamento da rede MPLS

Na descrição do funcionamento do MPLS terá como base a Figura 11, levando-se em consideração que todas as tabelas de encaminhamento (LFIB), foram preenchidas, manualmente ou dinamicamente, assim como, todas as tabelas de roteamento.

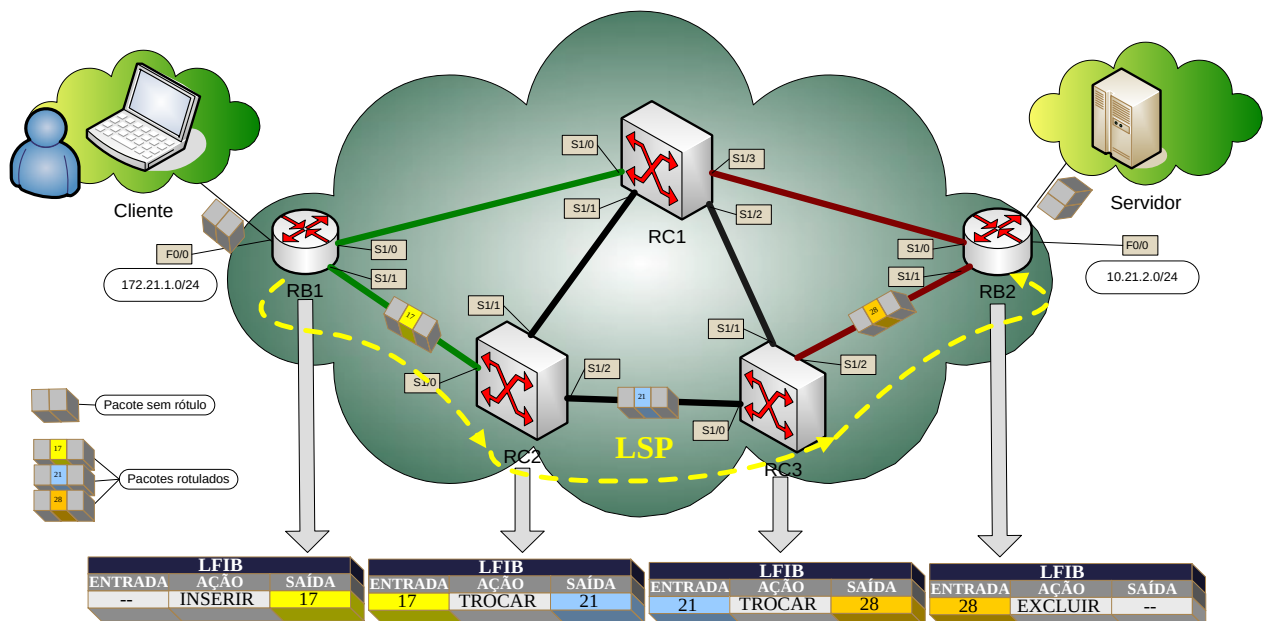


Fig. 11: Exemplo de funcionamento do MPLS
Fonte: Elaborado pelo Autor (2009)

¹³ Tipo-Tamanho-Valor da informação é o formato padrão de transmissão de parâmetros e seus respectivos valores, sendo utilizado por vários outros protocolos

No exemplo explicativo ilustra uma rede WAN¹⁴ MPLS, composta por três LSR que são os roteadores de core RC1, RC2 e RC3 e dois LER que são roteadores de borda RB1 e RB2 que conectam duas redes LAN¹⁵ em locais distintos, o fluxo de uma rede para outra é formado por uma seqüência de pacotes, representado no exemplo pelos pacotes cinza (sem rótulo), cinza e amarelo (rótulo 17), cinza e azul (rótulo 21), cinza e laranja (rótulo 28) e cinza (sem rótulo), estes pacotes possuem informações diferentes no cabeçalho, porém a mesma informação a ser transportadas no payload¹⁶. A seqüência de roteadores pelos quais passa o fluxo de pacotes MPLS é conhecida como LSP, que na figura 11 está representado pela linha amarela tracejada.

Inicialmente os pacotes originados nas estações de trabalho, que foram direcionados para outras estações de trabalho na mesma rede LAN não são direcionados para o roteador de borda RB1, somente são direcionados para RB1 os destinados a outras redes, como por exemplo, a rede do servidor.

Os pacotes oriundos das estações de trabalho não possuem cabeçalhos MPLS, apenas cabeçalho IP. O pacote, ao entrar no roteador de borda RB1 representado na figura 11 por pacote cinza (sem rótulo), originado nas estações de trabalho com destino a outra rede LAN, é classificado de acordo com a FEC, a FEC é então mapeada para NHLFE que possui o rótulo associado a interface de saída do roteador. Neste exemplo um conjunto de parâmetros que classifica o fluxo (FEC), são todos os pacotes da LAN com destino a qualquer outra rede na porta 80, que receberão o rótulo 17(NHLFE), que é inserido no pacote e sairão pela interface 3. Nos próximos roteadores não há inserção de rótulo, mas sim a troca dos rótulos, conforme a seqüência, 17 para 21, 21 para 28. No último roteador do LSP, ou seja, o roteador de borda RB2 não há troca ou inserção de rótulos no pacote, mas somente retira o rótulo MPLS do pacote, pois na sua LFIB há uma indicação, na coluna ação, que informa ao roteador para retirar o rótulo, desta forma o pacote é entregue na LAN, sem nenhuma alteração, pois a rede MPLS é totalmente transparente para o cliente.

Há um LSP para o sentido de cada fluxo, os pacotes que seguem o sentido contrário, como por exemplo, as respostas as requisições HTTP, devem seguir por outro LSP, outra seqüência de rótulos MPLS.

¹⁴ Rede de longa distância

¹⁵ Rede de baixa distância

¹⁶ Parte do pacote que contém a informação do usuário/aplicação.

2.3.12 Funcionamento do LDP

Na explicação do funcionamento do MPLS, partiu-se do princípio que as LFIB já estavam preenchidas, segue-se a explicação de como a LFIB é preenchida dinamicamente pelo protocolo LDP, podemos ter outros protocolos para tal objetivo, mas o LDP é o foco desse estudo.

O LDP é um protocolo definido basicamente pela RFC3036 do IETF, semelhantemente a outros protocolos no LDP existem quatro fases de funcionamento (Descoberta de vizinhos, Estabelecimento da Sessão, Anúncio de Rótulos e Finalização da Sessão) conforme ilustrado na figura 12 e que serão descritas abaixo:

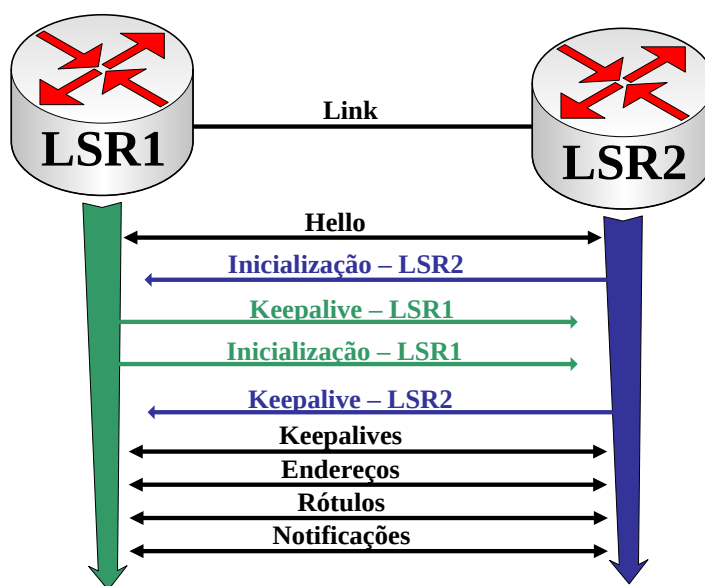


Fig. 12: Mensagens do Label Distribution Protocol
Fonte: Elaborado pelo Autor (2009)

- **DESCOBERTA DE VIZINHOS** – antes de qualquer troca de rótulo é necessário que os haja conexão entre os nós e que os mesmos conheçam o protocolo LDP, cada nó que conhece o LDP é considerado um vizinho, desde que haja conexão, esta conexão é feita através da descoberta desses roteadores vizinhos por meio do envio de PDU LDP e as mensagem com o campo tipo igual HELLO. Há dois modos de envio das mensagens de HELLO, estes modos dependem do tipo roteadores vizinhos.

- ✓ **Vizinhos Diretos** – São conectados a roteadores pelo mesmo enlace, camada nível 2 da ISO¹⁷, com este tipo de vizinho as mensagens HELLO são enviadas no endereço de multicast 224.0.0.2 (todos os roteadores no enlace) com UDP na porta 646.
- ✓ **Vizinhos Indiretos** - São conectados roteadores remotos, em enlaces diferentes, com este tipo de vizinho as mensagens HELLO são enviadas no endereço de unicast do vizinho, este endereço deve ser previamente configurado, no UDP na porta 646.
- **ESTABELECIMENTO DE SESSÃO** – Após identificar os vizinhos diretos ou indiretos o LDP passa para segunda fase, que é a de estabelecer a uma sessão TCP para troca de rótulos, mas antes de estabelecer a sessão TCP é necessário configurar os parâmetros da sessão TCP, como em toda sessão TCP deve haver um passivo, que recebe solicitações, e um ativo que envia solicitações de conexões, esta eleição é feita através do RouterID, que é um número enviado no cabeçalho da PDU LDP que identifica unicamente cada roteador, o roteador de MAIOR RouterID é eleito o ATIVO. Após definido o ativo e a sessão TCP estabelecida, são enviadas mensagens de INICIALIZAÇÃO pelo ativo, nestas mensagens há a configuração que o ativo deseja usar, por exemplo, modo de distribuição, loopback, entre outros, se o passivo não aceitar rejeita a conexão e fecha a sessão TCP, se aceitar, o passivo envia uma mensagem de INICIALIZAÇÃO e um KEEPALIVE, a mensagem de inicialização possui a configuração de sessão desejada pelo passivo e o keepalive é uma resposta positiva a mensagem de inicialização do ativo. Quando o ativo recebe a inicialização do passivo e o keepalive, ele envia o seu próprio keepalive se aceitar a pedido configuração do passivo, que foi enviado através da mensagem de inicialização, se não fecha a conexão TCP, se qualquer uma das partes, ativo ou passivo, não concordar com as configurações desejadas pela outra parte a conexão é fechada. Após estabelecida e configurada a sessão TCP, esta permanece ativa para troca de rótulos, através de mensagem KEEPALIVE que são enviadas periodicamente, pois os rótulos somente são enviados quando necessário.

¹⁷

Open Systems Interconnection – padrão que facilita a comunicação entre redes de diversos fabricantes

- **ANÚNCIO DE RÓTULOS** – como citado anteriormente os rótulos somente podem ser trocados após o estabelecimento de uma sessão LDP. São necessárias várias mensagens para que haja uma troca de rótulos de forma coerente. Abaixo segue a descrição e função de cada mensagem usada na troca de rótulos.
 - ✓ **Endereço (*Address*)** – Esta mensagem informa ao nó, que a recebe, todos os endereços das interfaces do transmissor, para que código LDP nó receptor possa associar a o rótulo a interface de saída correta.
 - ✓ **Retirada de Endereço (*Withdraw Address*)** – Indica ao receptor para retirar da LFIB o endereço e rótulo, pois o mesmo não está mais em uso, o receptor deve responder com uma mensagem de Liberação de Rótulo.
 - ✓ **Solicitação de Rótulo (*Label Request*)** – Usada no modo DoD (Sob Demanda) o qual é necessário solicitar um rótulo e endereço ou lista de endereços, em resposta a esta mensagem é enviada uma mensagem Endereço e uma Mapeamento de Rótulo, se for possível, se não é enviado uma mensagem de Notificação indicando a falha.
 - ✓ **Mapeamento de Rótulo (*Label Mapping*)** – A principal informação enviada nesta mensagem é a FEC, o Rótulo associado e a Identificação do nó transmissor. Não é enviada nesta mensagem o endereço do nó, o qual é enviado na mensagem *Address*, conforme explicado anteriormente.
 - ✓ **Liberação de Rótulo (*Label Release*)** – Esta mensagem confirma a liberação de rótulos.
 - ✓ **Pedido de Cancelamento de Rótulo (*Abort*)** - Se um nó transmissor decidir que não precisa mais usar um rótulo, mas já enviou uma mensagem de Mapeamento de Rótulo, ele envia uma mensagem de Pedido de Cancelamento de Rótulo, para anular o efeito da mensagem de Mapeamento de Rótulo.
- **FINALIZAÇÃO DE SESSÃO** - é a ultima fase usada para finalizar uma sessão, devidos a erros arbitrários que são informados nas mensagens de notificação. Nem toda mensagem de notificação encerrar a sessão, algumas apenas informam problemas, pois há dois tipos de mensagens de notificação:
 - ✓ Notificação de Erros – Usada quando o LSR encontra um erro fatal, resultando no encerramento da sessão.
 - ✓ Notificações Consultivas – São usadas como advertência ou status. O LSR pode recuperar tais problemas.

2.4 PROTOCOLOS DE ROTEAMENTOS

Devido o aumento substancial da internet, o preenchimento manual das tabelas de roteamento tornou-se inviável, então foram criados os protocolos de roteamento, os quais tem por objetivo preencher as tabelas de roteamento de cada nó na rede de forma automatizada. O plano de controle¹⁸ do MPLS é todo baseado na tecnologia IP, pois como foi descrito acima o MPLS utiliza o roteamento baseado no IP para prover a dinâmica inerente às redes roteadas, portanto é necessário primeiramente o funcionamento da tecnologia de roteamento IP, para posterior envio dos dados, definido pelo plano de encaminhamento¹⁹, o qual é implementado pela comutação de rótulo, da rede MPLS.

Os protocolos de roteamento contribuem para escalabilidade de uma rede IP. Os nós com protocolos de roteamento configurados enviam suas informações de rotas para os outros nós, que também possuem os mesmo protocolos de roteamento configurados. A principal diferença entre os protocolos de roteamento está no modo como estas informações de roteamento são distribuída entre os nós. Existem vários tipos e protocolos de roteamento cada um com suas peculiaridades.

Os protocolos de roteamento não transportam informações do usuário, mas tem por objetivo controle do tráfego de rede, ou seja, preenchendo as tabelas do roteamento IP, bem como, informar o melhor caminho baseado no valor da menor métrica²⁰. Todos os dados das aplicações dos usuários são transportados pelo IP, pelo melhor caminho.

Como citado anteriormente houve enorme aumento da internet e até mesmo com protocolos de roteamento torna-se inviável um único tipo de protocolo de roteamento para toda internet, criou-se então o conceito de Sistema Autônomo (AS - *Autonomous Systems*). De acordo com PAQUET e TEARE (2003, p.289)

Sistema Autônomo é um conjunto de roteadores sob uma única administração técnica usando um Interior Gateway Protocol e a métrica comum para rotear os pacotes dentro do AS, e usando um Exterior Gateway Protocol para rotear os pacotes para os outros [sistemas autônomos].

Com este conceito a internet foi dividida em unidades administrativas e desta forma cada unidade pode utilizar seu próprio protocolo de roteamento para rotear o tráfego dentro do AS, independente do protocolo de roteamento de outros AS, bem como, usar um protocolo de

¹⁸ Informações de roteamento que definem quais caminhos disponíveis.

¹⁹ Informações de cabeçalho dos pacotes para envio dos dados.

²⁰ Valor associado uma rota, quando há varias rotas para o mesmo destino, a melhor rota possui a menor métrica

roteamento comum entre os AS para trocar informações de roteamento, conforme a figura 13. Cada AS enxerga o outro como um único plano de endereçamento, há um acoplamento fraco entre os AS pois algumas mudanças dentro de um AS são irrelevantes para os outros.

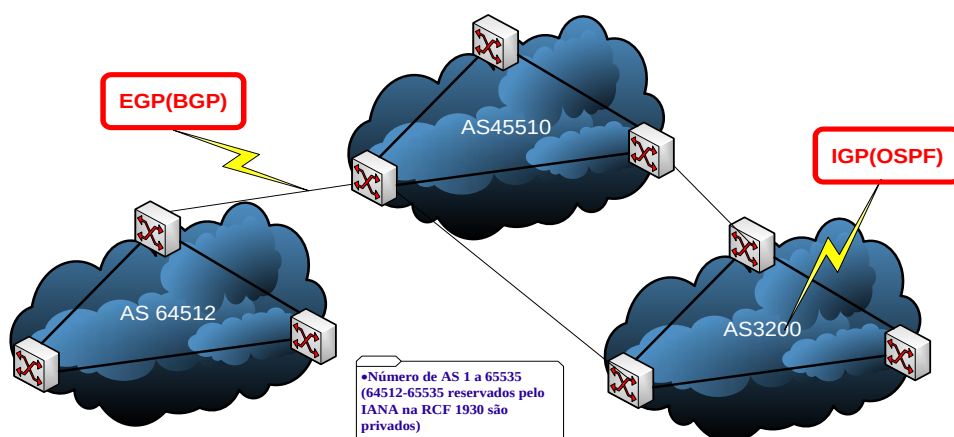


Fig. 13: Relação entre Sistemas Autônomos, EGP e IGP
Fonte: Elaborado pelo Autor (2009)

Neste trabalho serão apresentados, o OSPF (*Open Short Path First*) como protocolo de roteamento interior ao AS e o BGP (*Border Gateway Protocol*) para troca de informações de roteamento entre os AS, pois são atualmente os mais difundidos.

2.4.1 OSPF (Open Short Path First)

2.4.1.1 Características

Sua definição básica está descrita pelo IETF na RFC 2328 e foi desenvolvido para abordar necessidade de grandes redes escaláveis, bem como fornecer suporte a autenticação.

De acordo com a definição, o OSPF é um protocolo de roteamento IGP²¹, pois geralmente não propaga rotas para nós que não fazem parte do seu AS, além de possuir cálculo uniforme da métrica em todo o AS.

O OSPF utiliza o modelo de estado de enlace para obter o caminho da origem até o destino, onde cada nó tem uma visão completa da topologia de rede, pois cada nó cria uma

²¹ Interior Gateway Protocol – protocolo utilizado dentro de um sistema autônomo.

estrutura de árvore²² sendo o próprio nó a raiz, desta forma há um único caminho do nó origem (raiz) até o destino (folhas), dentro do AS.

O consumo do link²³ na inicialização do protocolo é alto devido à troca completa de tabelas de roteamento, entretanto, posteriormente é baixo, pois somente atualizações são enviadas, exceto pelo fato que no intervalo de 30 minutos são re-enviados novamente as tabelas de roteamento para ratificar a consistência da base de dados, mantendo-as sincronizadas.

A forma como as informações de roteamento são trocadas entre os nós dependem da topologia e do tipo de rede definida, que será descrita adiante, entretanto será antes apresentado algumas definições usado no protocolo.

2.4.1.2 Definições do OSPF

Para um melhor entendimento da tecnologia usada se faz necessário o conhecimento de alguns termos e conceitos usados nesta tecnologia.

- LINK-STATE – informação sobre o estado das interfaces dos roteadores, cada interface corresponde a um enlace (rota).
- CUSTO – valor atribuído ao enlace, baseado na largura de banda da interface.
- ÁREA – um conjunto de roteador com uma mesma identificação e que tem a mesma informação da topologia da rede.
- VIZINHO - nó que tem sua interface se comunicando com OSPF em uma mesma rede ou redes diferentes.
- ADJACÊNCIAS – relacionamento de comunicação entre vizinhos.
- HELLO – protocolo usado para criar e manter adjacências ativas.
- BANCO DE DADOS DOS VIZINHOS – uma lista de todos os vizinhos
- BANCO DE DADOS DOS ESTADOS DOS ENLACES – Um banco de dados que possui o estado de todos os enlaces dos roteadores na mesma área.

²² Estrutura organizacional.

²³ Link – enlace dados entre dois nós.

2.4.1.3 Funcionamento do protocolo OSPF

O protocolo OSPF troca mensagens entre os nós para fornecer informações suficientes para execução do algoritmo Dijkstra, que é usado no processo de manutenção da tabelas de roteamento. Após coleta informações dos nós e os valores dos enlaces entre os respectivos nós, o algoritmo constrói o menor caminho entre dois pontos, sendo um ponto o nó que executa o algoritmo e o outro o nó de destino dos dados. Para uma rede com diversos caminhos diferentes para o mesmo destino, somente o menor caminho é utilizado.

Existem vários tipos de mensagens. As primeiras mensagens a serem trocadas são para manter ativas as adjacências, após isso são enviadas mensagens que possuem informação de link-state de cada roteador. As mensagens somente podem ser enviadas a vizinhos adjacentes.

A forma como estas mensagens são trocadas entre os nós depende do tipo de rede a ser usada, as quais serão descritas abaixo:

- a) **MULTIACESSO BROADCAST (Ethernet)** – Neste tipo de rede é selecionado um único nó para enviar e receber atualização de roteamento (link-state), que é chamado de Roteador Designado (DR), e também é selecionado outro nó sobressalente, chamado de Roteador Designado de Backup (BDR), conforme ilustrado na figura 14. Com o conceito de DR e BDR são necessários apenas $2N-3$ adjacências entre os nós, de outra forma seriam necessário $N(N-1)/2$ adjacências neste tipo de rede. A troca de rotas acontece apenas entre os nós com o DR e o BDR, não há troca de rotas entre nós que não sejam DR/BDR, pois estes não possuem adjacências. No caso de queda do DR o BDR assume como DR e é escolhido um novo BDR na rede. As escolhas do DR e BDR são feitas no momento inicial, quando são criadas as adjacências.

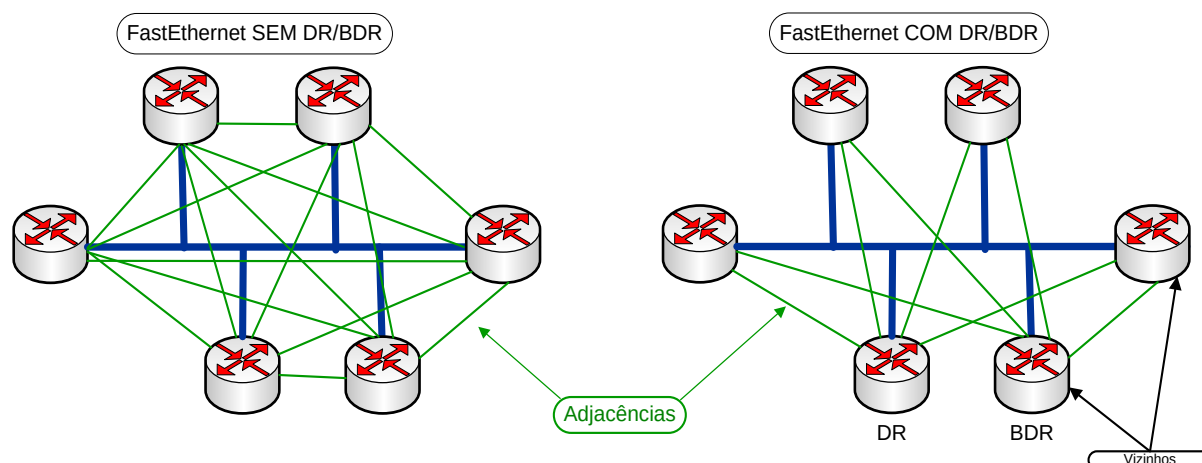


Fig. 14: Conexões em redes com e sem DR/BDR
Fonte: Elaborado pelo Autor (2009)

- b) **PONTO-A-PONTO** – É o modelo mais simples de funcionamento do OSPF, Neste tipo de rede não há o conceito de DR e BDR, o protocolo Hello é apenas enviado no endereço de multicast 224.0.0.5 para estabelecer adjacência e a troca é inicializado por quem possui o maior valor de ROUTERID²⁴.
- c) **NÃO-BROADCAST MULTILANÇO (NBMA)** – Uma rede NBMA é uma rede que tem a capacidade de multi-acesso, mas não dispõe da capacidade de broadcast. O OSPF trabalha neste tipo de rede de acordo com a topologia implementada, que são as seguintes:
- **FULL-MESH** – Topologia a qual todos os nós possuem conexões diretas entre todos os outros nós, conforme ilustrado na figura 15, neste tipo de topologia o OSPF simula uma rede Ethernet com o uso de sub-interfaces²⁵ enviando o broadcast/multicast em todas as elas, elegendo um DR e um BDR.
 - **ESTRELA E PARTIAL MESH** – Na topologia estrela todos os nós conectam os outros nós através de um nó central, conforme ilustrado na figura 15, na topologia estrela e partial-mesh os nós não tem uma conexão direta com todos os outros nós. Nestes tipos de rede os links funcionam como enlaces ponto-a-ponto, não havendo DR e BDR como no caso de NBMA full-mesh.

²⁴

Valor atribuído a um nó para identificá-lo de forma exclusiva.

²⁵

Interface lógica dentro de uma interface física. Pode haver várias lógicas dentro de uma física.

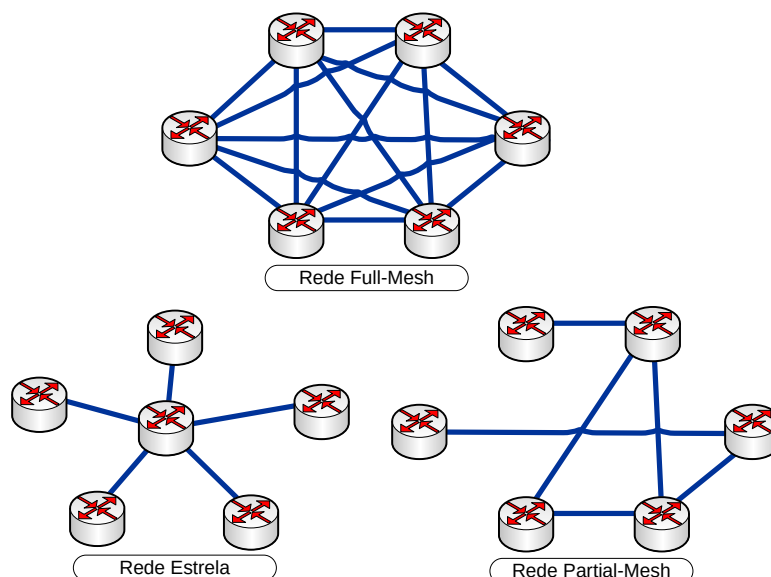


Fig. 15: Exemplo de Topologias de Redes
Fonte: Elaborado pelo Autor (2009)

A partir da escolha do DR/BDR, ou do maior RouterID para rede PPP, é feita a troca de pacotes contendo o resumo dos links-state (LSA – *Link State Advertisement*) entre os roteadores, cada LSA possui um resumo de cada registro²⁶ que estão no banco de dados de estado dos enlaces e seu respectivo número de seqüência. Um roteador ao receber o LSA compara o número de seqüência, para uma provável solicitação do registro completo. Existem vários tipos de pacotes LSA, que dependendo da função pode conter outros tipos de informação.

Caso seja necessário atualizar a base de dados, devido a inconsistências, uma solicitação de registro (LSR – *Link State Request*) é gerada e enviada. O nó receptor de um pacote LSR, gera um pacote contendo os dados completo do registro (LSU – *Link State Update*) para ser enviado ao emissor do LSR, a LSU pode conter várias LSA. O nó gerador do LSR responde com pacote de reconhecimento (LSAck – *LSA Acknowledgment*) em resposta ao LSU.

Após a sincronização da base de dados dos estados do enlaces, somente atualizações são enviadas, devido algum novo evento²⁷ na rede. O protocolo HELLO é usado para manter as adjacências caso não haja nenhum evento e assim mitigar o tráfego do protocolo OSPF na rede.

²⁶ Contém várias informações do enlace entre os nós, como largura de banda, estado UP/DOWN, etc.

²⁷ Qualquer alteração na topologia, estado de interface (UP/DOWN), alteração de métrica, etc...

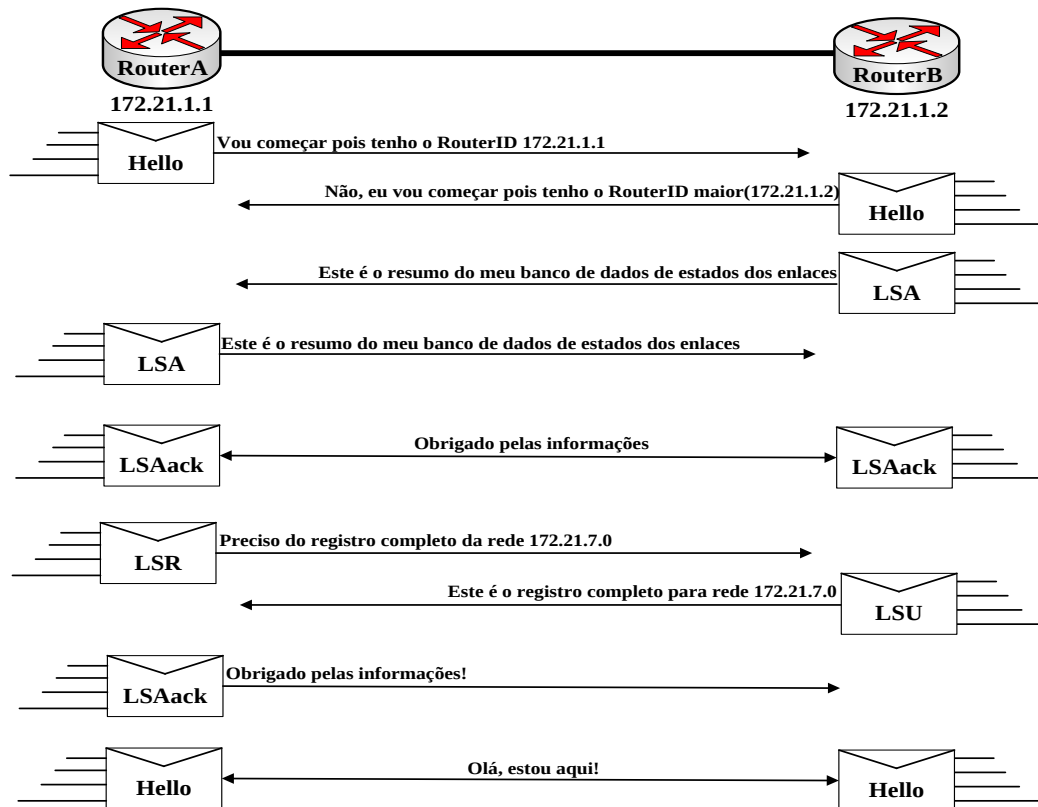


Fig. 16: Troca de pacotes para preencher as bases de dados.
Fonte: Elaborado pelo Autor (2009)

Após este processo, exemplificado na figura 16, todos os roteadores dentro de uma área terão as mesmas bases de dados. Cada roteador OSPF possui basicamente três bases de dados, Banco de Dados dos Vizinhos e Estado de Enlaces, Banco de Dados topológico e a Tabela de roteamento. Conforme exemplificando na figura 17, o OSPF aplica o algoritmo Dijkstra no banco de dados dos vizinhos/estado dos enlaces e gera todos os possíveis caminhos para o destino, criando assim o banco de dados topológico, entretanto somente o melhor caminho será instalado na tabela de roteamento.

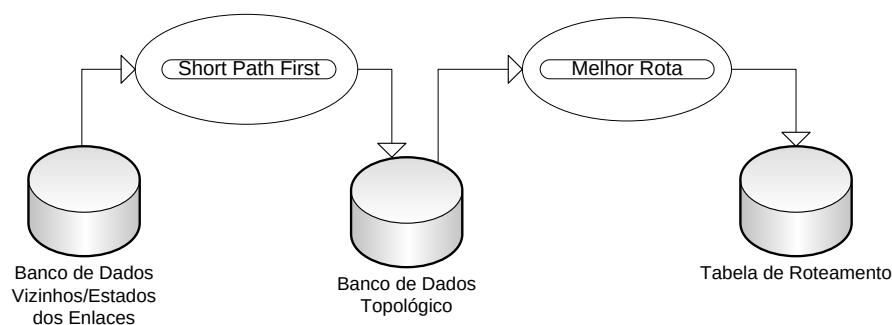


Fig. 17: Bases de Dados nos Roteadores
Fonte: Elaborado pelo Autor (2009)

Toda informação de estado dos enlaces e vizinhos é difundida pela rede, segundo os procedimentos acima descritos para cada tipo de rede.

A fórmula, $\text{Custo} = 10^8 / \text{Bandwidth}^{28}$, é aplicada para o cálculo do custo do enlace.

A rede mostrada pela figura 18 é usada para exemplificar todo este processo, o caminho percorrido pelos pacotes da máquina cliente até o servidor será de R1 para R3, então para R2 e por fim para R4, pois para o OSPF, este é o melhor caminho.

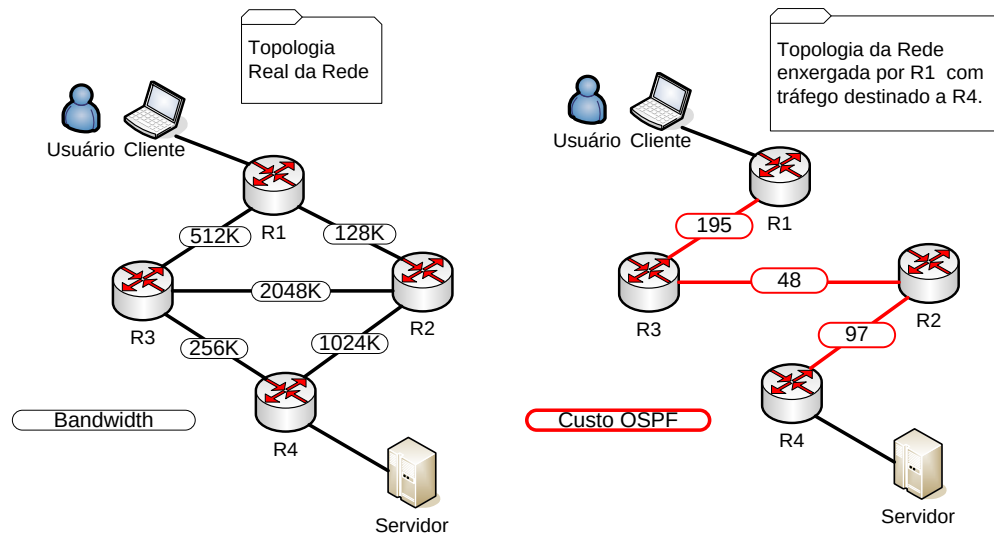


Fig. 18: Caminho Menor criado pelo OSPF baseado no Custo.
Fonte: Elaborado pelo Autor (2009)

²⁸

Largura de banda do enlace em bits por segundos.

2.4.2 BGP-4

2.4.2.1 Características

O protocolo BGP é usado para distribuir rotas entre sistemas autônomos, possuindo a opção de ser usado dentro do AS como IGP. É um protocolo simples, escalável e estável. De acordo com PAQUET e TEARE (2003, p.290) “Seu principal objetivo é fornecer um sistema de roteamento entre domínios que garanta a troca, isenta de loops, das informações de roteamento entre os sistemas autônomos.”.

O BGP-4 usa a porta 179 do protocolo TCP para transportar suas mensagens, bem como, conhece o caminho completo da origem até o destino dos dados a partir das informações recebidas pelos seus vizinhos.

Entre as várias informações recebidas, conhecidas como atributos da(s) rota(s), o número de Sistema Autônomo é usado para evitar loops, outros atributos são usados para definir o melhor caminho baseado em políticas, este modelo é conhecido como PBR – Police Based Routing.

O BGP-4 está definido basicamente na RFC1771, mas existem várias outras RFCs que acrescentam funcionalidades específicas ao protocolo, como por exemplo, a extensão MP (Multi-Protocolo), que acrescenta a capacidade de transportar rotas com endereçamento do tipo IPv6²⁹, além de outros tipos de endereçamento, ao protocolo.

2.4.2.2 Funcionamento do BGP

O BGP utiliza a porta 179 do TCP, portanto necessário a configuração manual das extremidades TCP, estabelecendo assim uma sessão. Cada roteador que usa BGP é chamado de speaker ou parceiro BGP. Um speaker BGP pode ter vários parceiros BGP, mas todos com ligação ponto-a-ponto.

²⁹

Sexta versão do tipo de endereçamento do IP (Internet Protocol)

Conforme exemplificado na figura 19, a troca de todas as mensagens é feita entre os parceiros BGP, depois de estabelecida a sessão TCP, a primeira mensagem ser enviada é a OPEN, a qual possui as informações de configuração de cada sessão BGP, se o receptor da mensagem não concordar com as configurações ele envia uma mensagem de NOTIFICATION, que informa a causa do impasse e fecha a sessão, caso contrário, ele envia sua mensagem OPEN e a partir de então são trocadas mensagens de UPDATE, as quais possuem as informações de rotas com seus respectivos atributos. A mensagem de KEEPALIVE é usada quando não há mensagem de UPDATE a ser enviada, mas precisa-se manter a sessão aberta. Após o recebimento de uma mensagem NOTIFICATION a sessão é encerrada, há também a mensagem de ROUTE-REFRESH que solicita o re-anúncio de rotas, esta última foi acrescentada na RFC 2918.

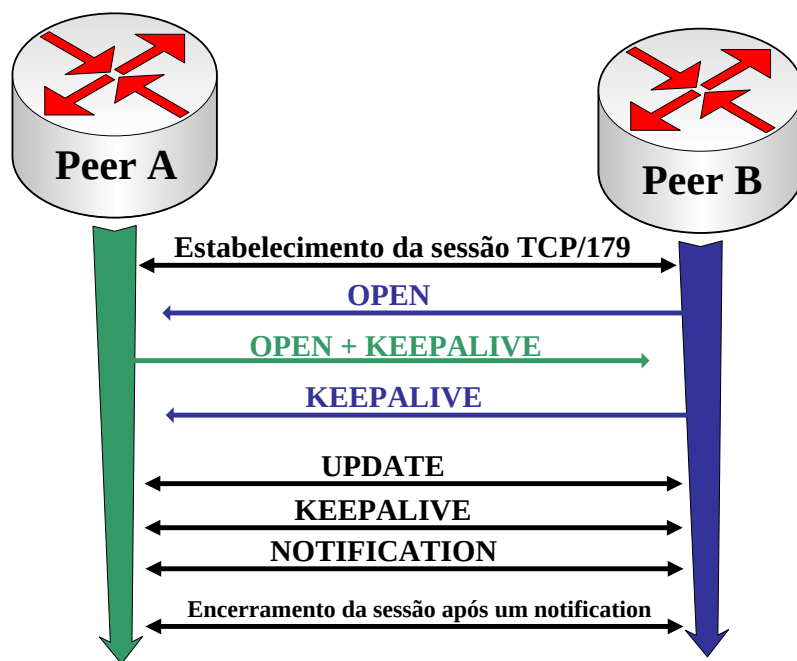


Fig. 19: Funcionamento do BGP
Fonte: Elaborado pelo Autor (2009)

Os nós parceiros possuem tabelas BGP além das tabelas de roteamento. Cada tabela BGP possui todas as rotas aprendidas via o protocolo, poderá haver várias rotas para mesmo destino, porém a tabela de roteamento possui somente a melhor rota para um determinado destino. A decisão de qual rota BGP deve ser instalada na tabela de roteamento é baseada no **processo decisório do BGP**, este analisa os vários atributos de uma rota, baseada nesta análise é escolhida apenas uma única rota (a melhor) a ser instalada na tabela de roteamento, quando houver mais de uma possibilidade. Opcionalmente pode-se instalar mais de uma rota

na tabela de roteamento, dependendo da configuração, com objetivo de fazer balanceamento de carga.

2.4.2.3 Atributos do BGP

São características que atribuídas a uma determinada rota ou a um conjunto de rotas, as quais têm por finalidade servir de subsídio para o processo decisório do BGP. Os principais atributos são classificados em quatro classes:

- **Bem-conhecidos e Mandatórios** - as implementações do BGP devem conhecer e usar estes atributos;
 - AS - Path – É uma lista dos AS pelo qual a rota passou, sempre que uma rota passa de um AS para o outro é anexado esta rota o número do AS, os números de AS estão no intervalo de 1 a 65535, sendo a partir de 64512 privados.
 - Next-Hop – É usado para indicar o próximo roteador, o qual o pacote deve ser enviado para se chegar ao destino.
 - Origin – Indica forma como a rota foi aprendida³⁰, “i” indica que foi aprendida de um IGP, “e” indica que foi aprendida via BGP e “?” indica que não se conhece o método de aprendizagem da rota.
- **Bem-conhecidos e Arbitrários** - as implementações do BGP devem conhecer e podem usar ou não estes atributos;
 - Local Preference – É um atributo trocado dentro do AS, indica para os outros roteadores dentro do AS o caminho preferido de saída, para um AS com várias saídas, o qual é conhecido como Multihoming.
 - Atomic Aggregate – este atributo informa ao AS vizinho que a rota recebida é um resumo, isto é, resultado de uma agregação de várias rotas.
- **Opcionais e Transitivos** - as implementações do BGP não são obrigadas a implementar, mas devem repassar inalterados estes atributos;
 - Aggregator – Identifica o roteador que gerou a rota de resumo e número

³⁰

Neste contexto aprendida significa como foram obtidas as informações de roteamento.

do respectivo AS.

- **Community** – É uma marca (comunidade), que indica para os outros roteadores que compartilham propriedades comuns, assim os roteadores podem tomar decisões nas rotas baseados nesta marcação
- **Opcionais e Não-Transitivos** - as implementações do BGP não são obrigadas a implementar e devem excluir o atributo ao repassar a rota;
 - **Multi-Exit Discriminator (MED)** – É trocado entre AS, mas não é repassada para outros AS, o MED sugere ao AS receptor um caminho preferido, esta é a única forma como o AS pode influenciar o outro. Uma rota com um MED menor indica um melhor caminho.

2.4.2.4 Processo Decisório do BGP

O processo decisório do BGP é muito complexo, mas para exemplificar **didaticamente** será utilizada a figura 20, onde o AS3200 recebe rotas para mesma rede de destino de dois AS diferentes, neste caso as rotas carregam a informação por quais AS devem passar para chegar ao destino.

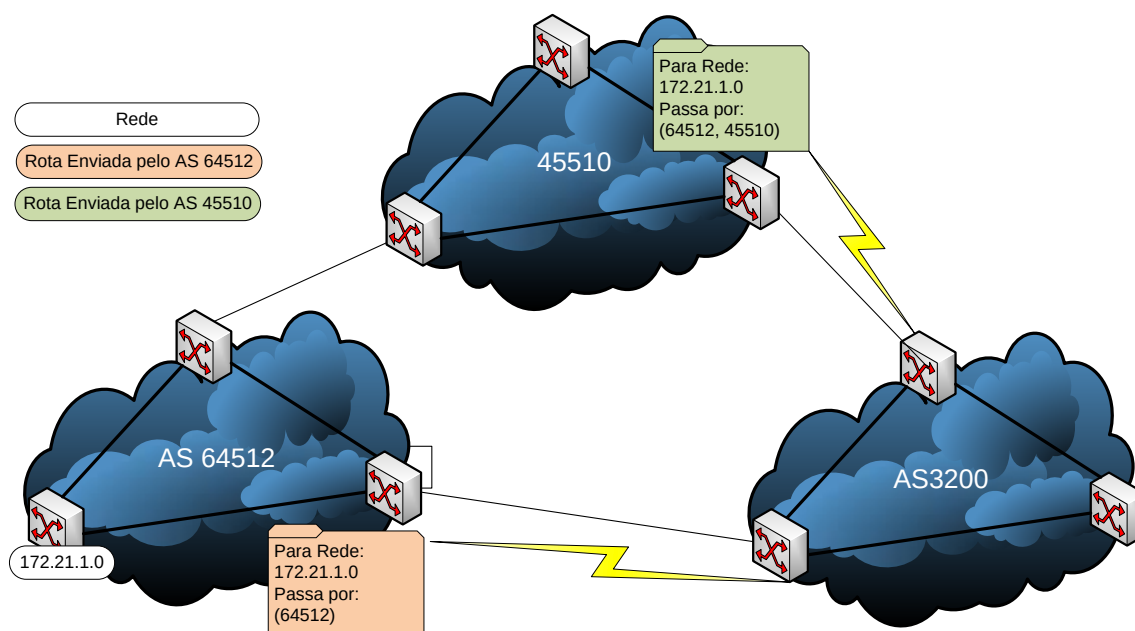


Fig. 20: Divulgação de rotas com atributo Path
Fonte: Elaborado pelo Autor (2009)

O AS3200 ao receber as rotas com destino a rede 172.21.1.0, pode usar qualquer uma das rotas, entretanto deve escolher apenas uma para ser instalada na tabela de roteamento. A escolha da rota está atrelada a política adotada pelo AS3200.

Após definida a política pelo AS3200, a rota será instalada na tabela de roteamento, há vários outros atributos que assim como o Path fornecem informações para decisão do BGP de qual rota é a melhor baseada na política. Melhor rota no caso do BGP não significa necessariamente melhor link, pois características externas, como por exemplo, valor monetário do link pode influenciar nestas decisões.

2.5 QUALIDADE DE SERVIÇO

De acordo com FARREL (2005, p.180):

O modelo de operação tradicional das redes IP foi baseado na remessa de serviço pelo melhor esforço. Nenhuma garantia foi feita sobre a qualidade de serviço fornecida as aplicações ou usuários da rede, e cada pacote foi tratado como um objeto separado e encaminhado dentro da rede sem precedência ou prioridade sobre outros pacotes. Além disso, uma consideração de projeto fundamental do IP e da Internet foi tornar a simplicidade mais importante de que qualquer outra coisa.

O modelo inicial da Internet foi projetado para transferência de dados que são tolerantes a uma alta latência e variação no atraso (jitter), como no caso do tráfego de e-mail e transferência de arquivos, mas com aumento da Internet e surgimento de novas aplicações, conseqüentemente surgiram novos tipos de tráfegos com diferentes requisitos conforme descrito resumidamente no Quadro 6.

	VOZ	DADOS	VIDEO
Largura de Banda	Baixa	Média	Alta
Sensibilidade a Perdas	Baixa	Alta	Baixa
Sensibilidade a Atrasos	Alta	Baixa	Alta
Sensibilidade a Jitter	Alta	Baixa	Alta

Quadro 6: Relação entre tráfego e parâmetros de rede
Fonte: Elaborado pelo Autor (2009)

Surgiu então necessidade do gerenciamento destes parâmetros, o qual atualmente é conhecido como Qualidade de Serviço (QoS – Quality of Service), com o objetivo de condicionar a rede a atender os requisitos das novas aplicações,

Existem várias técnicas usadas para gerenciar alguns destes parâmetros na rede (perda de pacotes, latência, variação no atraso) com o objetivo de manter os níveis destes parâmetros em um determinado patamar, conforme acordado entre o cliente e a operadora da rede, este contrato é conhecido como SLA(Service Level Agreement).

Na Figura 21, segue um exemplo do comportamento de link de rede, onde podem ser observados os três primeiros casos, nos quais a rede não implementa técnicas de QoS, em contraste com o quarto caso que implementa as técnicas de QoS, mantendo a rede sempre em uma situação favorável a atender os requisitos contratados.



Fig. 21: Exemplo comparativo de Qualidade de Serviço.
Fonte: Elaborado pelo Autor (2009)

Visando fornecer QoS para rede foram desenvolvidos duas arquiteturas de gerenciamento de recursos, conhecidos como DiffServ e IntServ. Cada um possui suas peculiaridades, sendo a escolha da arquitetura atrelada ao tipo aplicação. De acordo com FARREL (2005, p.190):

O foco IntServ está no tráfego de dados em tempo real. Ele não é uma exigência para as trocas de dados que não dependem de tempo, e tais fluxos podem ser mais bem tratados pelo DiffServ, onde não há overhead de outro protocolo de sinalização e nenhuma necessidade de reserva de recursos complexas em cada nó. Entretanto se a qualidade de tempo real for exigida, o IntServ oferece um mecanismo formal e simples para descrever os fluxos e os requisitos.

As duas arquiteturas também podem ser usadas em conjuntos na mesma rede.

Face o exposto, o IntServ e o DiffServ serão os dois modelos de arquitetura de QoS que serão estudados neste trabalho.

2.5.1 Serviços Diferenciados (DiffServ)

O DiffServ permite classificar o tráfego em diferentes classes, de acordo com serviço contratado, cada classe define um comportamento diferente do tráfego dentro dos nós da rede, as classe especificam diferentes valores de largura de banda, latência, jitter e descarte de pacotes, entretanto é importante haver um conhecimento coerente em toda rede do nível de serviço (SLA) aplicado aos pacotes dentro de um fluxo, o perímetro deste conhecimento uniforme é conhecido como Domínio DiffServ. Neste modelo não há interação entre a rede e os dispositivos que a ela se conectam, pois cada nó de rede possui uma configuração específica e estática dos recursos a serem reservados, isto é, cada nó pode possui quantidades diferentes de reservas de recursos, como largura de banda, latência, jitter e perda de pacotes, conforme o projeto da rede.

2.5.1.1 Arquitetura DiffServ

A arquitetura DiffServ contém 06 processos básicos para implementação do QoS, conforme mostra a figura 22, todos os processos são implementados em nos nós que faz parte do Domínio DiffServ, normalmente os nós não implementam todos os processos, geralmente são divididos em dois estágios, o primeiro estágio é executada na borda da rede e os outros no núcleo.

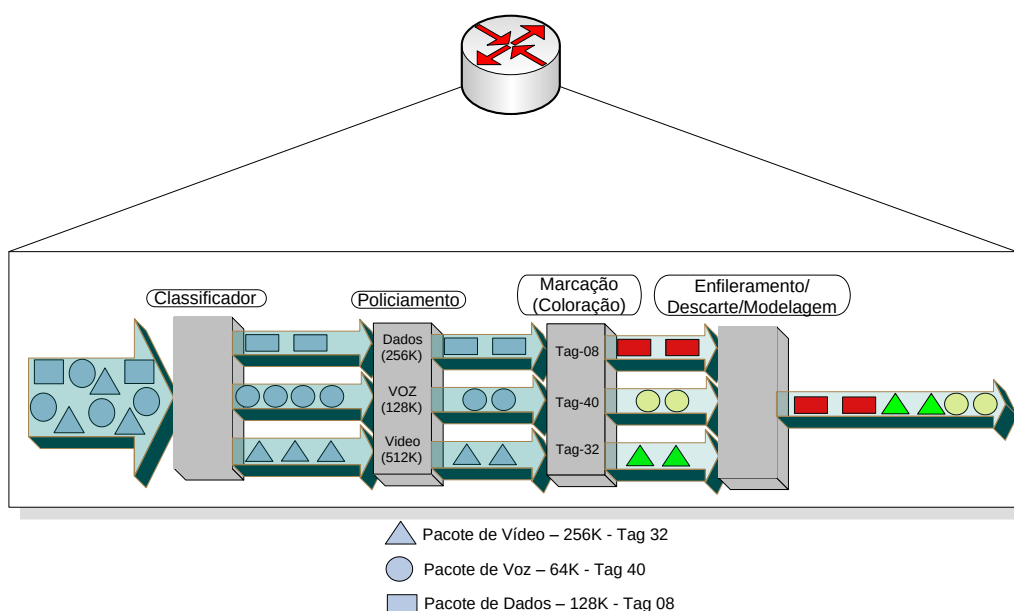


Fig. 22: Exemplo da arquitetura DiffServ
 Fonte: Elaborado pelo Autor (2009)

CONDICIONAMENTO DO TRÁFEGO – Este é o primeiro estágio e conforme foi descrito anteriormente é executado na borda da rede, ou seja, no primeiro nó pelo qual o tráfego entra na rede, este estágio é composto por três processos:

- **Classificação** - A classificação separa os pacotes que entram na rede através de diversas classes de serviço, baseado em alguns parâmetros definidos, como por exemplo, porta, ip de origem, interface, etc.
- **Policiamento (Policing)** – De acordo com o perfil contratado pode-se realizar um policiamento de entrada, desta forma, não permitido mais tráfego do que foi projetado na rede, por exemplo, o tráfego excedente da classe *Real Time*, pode ser descartado e o tráfego excedente das demais é remarcado para *Best Effort*, este processo é usado na entrada do tráfego do nó, sendo o policiamento atrelado ao SLA.
- **Marcação (Coloração)** - A marcação do tráfego é realizada através um byte no

cabeçalho IP definido para a marcação de classes de serviço. Este byte é o campo TOS (*Type of Service*), que ficou obsoleto e atualmente foi redefinido como DSCP (*Differentiated Services Code Point*) na Arquitetura DiffServ, conforme ilustrado na figura 23 e no Quadro 7. O DSCP é o identificador que informa a rede como a qual classe o pacote pertence e conseqüentemente como deve ser tratado. Cada classe especifica um tipo de serviço que é definido pelo byte DSCP. Basicamente há três tipos de serviços:

1. **EF** (*Expedited Forwarding*): Os pacotes são encaminhados com a mais alta prioridade e não são descartados.
2. **AF** (*Assured Forwarding*): Pacotes são encaminhados com alta probabilidade de serem entregues, entretanto pode haver descartes. Existem subclasses onde se pode especificar a precedência de descarte.
3. **BE** (*Best-Effort - Padrão*): Os pacotes são encaminhados sem garantias.

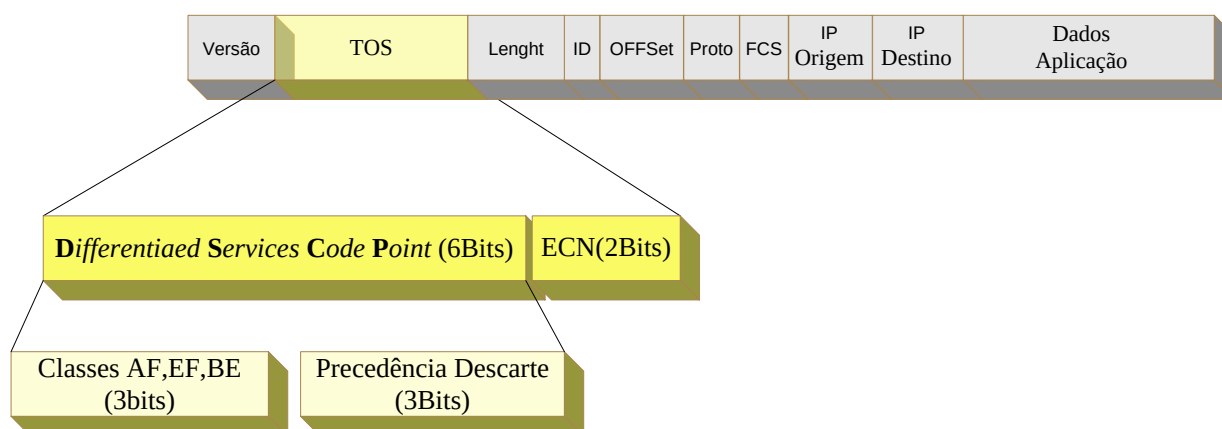


Fig. 23: Localização do DSCP no pacote IP
Fonte: Elaborado pelo Autor (2009)

IP Precedence		DSCP		Classes	CSCP
0	000	0	000000	CS0	Best-Effort
1	001	8	001000	CS1	AF1
2	010	16	010000	CS2	AF2
3	011	24	011000	CS3	AF3
4	100	32	100000	CS4	AF4
5	101	40	101000	CS5	EF
6	110	48	110000	CS6	-
7	111	56	111000	CS7	-

Quadro 7: Relação entre IP Precedence do TOS e DSCP
Fonte: Elaborado pelo Autor (2009)

COMPORTAMENTO POR SALTO (PHB – *Per Hop Behaviour*) – Este é o segundo estágio e conforme foi descrito anteriormente é um comportamento, que é implementado no núcleo da rede, mas pode opcionalmente ser utilizado na borda da rede. Este comportamento está atrelado a classe do pacote, pois cada classe define um comportamento diferente. O PHB É composto basicamente por três processos:

- **Enfileiramento (Queuing)** - Este processo é um mecanismo de filas permite que cada classe de tráfego (EF, AF, Best-Effort) utilize filas diferentes no roteador. Com isso pode se dividir a banda e os buffers³¹ de saída entre as classes
- **Descarte (Dropping)** – É inevitável que ocorram descartes de pacotes, em casos de congestionamento, entretanto foram criados mecanismos de descarte com a finalidade de evitar o congestionamento dos buffers, bem como, gerenciar estes buffers na situação de congestionamento.
- **Modelagem (Shapping)** – É o processo responsável por adequar o tráfego ao perfil contratado. Ele utiliza buffers de saída que armazenam picos ou rajadas de tráfego e transmitem dentro dos limites estabelecidos, visando evitar o descarte de pacotes no nó posterior, este processo é preventivo usado na saída do nó, com objetivo de evitar descartes na entrada do próximo.

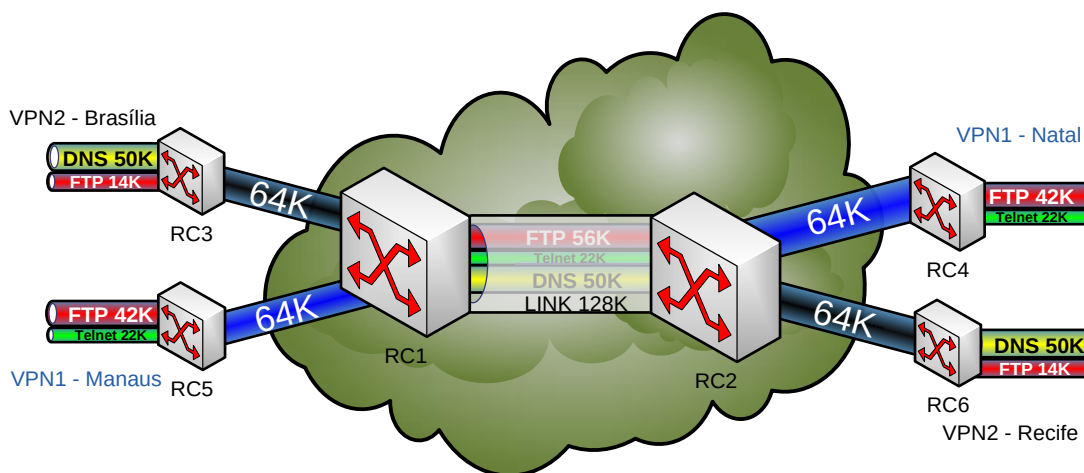


Fig. 24: Exemplo de Rede com DiffServ.
Fonte: Elaborado pelo Autor (2009)

Conforme o exemplo da figura 24, todo tráfego FTP da VPN-1 que entra em RC4, do site localizado em Natal, passa pelo Backbone e chega até RC5, em outro site da mesma VPN-1, localizado em Manaus. A garantia de largura de banda de 42Kbps para o tráfego de FTP é mantida, desde da origem até o destino, isto é, se entrar 42Kbps de FTP em Natal vai

³¹ espaço de memória utilizado para armazenar dados voláteis.

sair 42Kbps de FTP em Manaus. Em uma rede sem QoS, como a internet, por exemplo, não é possível ter esta garantia, pois não há nenhum controle sobre todos os roteadores e por quais caminhos o tráfego vai passar, além de não possui nenhuma política de gerenciamento do tráfego nos roteadores da rede. Neste exemplo os roteadores de borda são RC3, RC4, RC5 e RC6, os quais são implementados os 03 primeiro processos de QoS (classificação, policiamento e marcação) e os roteadores de cores são RC1 e RC2, os quais implementam os outros três processos do QoS (enfileiramento, descarte e modelagem). Em cada roteador possui uma configuração que especifica o comportamento do pacote dentro do nó (PHB - *Per Hop Behaviour*).

2.5.2 Serviços Integrados (IntServ)

Diferentemente do DiffServ, onde cada nó possui uma configuração específica e estática, no modelo do IntServ há uma interação entre a rede e os dispositivos que a ela se conectam, para que seja realizada a configuração dinâmica da reserva de recursos em cada nó. O IntServ também realiza a marcação do campo DSCP, ação que é conhecida como colorir os pacotes.

Os dispositivos, que se conectam a rede, usam a coloração dos pacotes para informar à rede as características do tráfego, através de solicitações de reservas de recursos, estes também recebem o informações dela, através de relatórios, que entre outras informações, relatam se a rede pode ou não fornecer o QoS solicitado. Essa troca de informações é uma característica de configuração dinâmica de reserva de recursos do IntServ, a qual é realizada pelo protocolo RSVP.

O RSVP é um mecanismo de controle e segue o caminho percorrido pelos dados, que são definidos pelo protocolo de roteamento.

O IntServ provê um controle de admissão, o qual aceita ou rejeita a solicitação de reserva de recursos, segundo FARREL(2005, p.190) “ *O IntServ às vezes é descrito como modelo “tudo ou nada”*”.

No DiffServ há três tipos de serviço, EF, AF, e BF, entretanto no IntServ, o RSVP é usado somente para dois tipos de serviços:

- I. Carga Controlada – neste modelo a rede fornece um serviço que simula para o cliente uma rede sem congestionamento, mesmo que a rede esteja saturada por outros tipos de tráfego, sendo equivalente ao AF do DiffServ.
- II. Serviço garantido – neste modelo a rede simula uma linha privada, garantido o limite de tempo e não descarte de pacotes, sendo equivalente ao EF do DiffServ.

Não há necessidade de sinalização para o serviço do tipo BE, pois este é modo padrão que a rede opera, isto é, sem o uso do protocolo RSVP.

2.5.2.1 Arquitetura IntServ

2.5.2.1.1 Balde de Fichas

O IntServ não especifica como os recursos devem ser reservados, mas sim quais os parâmetros e os valores destes parâmetros, ou seja, a implementação da reserva de recursos dentro do roteador é arbitrária, o nó pode implementar qualquer tipo de fila, tamanho de buffer, algoritmo, etc. Entretanto deve fornecer os recursos especificados nas mensagens do protocolo RSVP.

O conceito de Balde de Fichas é usado para definir os parâmetros a serem usados na especificação de reserva de recursos dos nós, é uma maneira padronizada de representar o desempenho desejado, para um determinado fluxo de dados. O tráfego se comporta analogamente a um fluxo de água em um balde furado, isto é, a taxa de saída, vazão e latência, no furo do balde é mantida constante independente da quantidade de entrada. O fluxo deverá transbordar, se houver uma quantidade de entrada maior que a capacidade do balde, que no caso da rede é refletido pela perda de pacotes.

O Balde de Fichas possui os seguintes parâmetros:

- ✓ Taxa de dispersão de dados(r) - mantém estável a taxa média de saída do fluxo do balde.
- ✓ Capacidade de armazenamento(b) - tamanho do balde, especifica a variação máxima do fluxo na entrada para evitar a perda de pacotes.

Os fluxos são caracterizados basicamente por estes dois parâmetros, mas existem outros que são usados para complementar a especificação do fluxo

- ✓ Taxa de pico dos dados (p) – é a taxa de pico máxima que a fonte pode injetar tráfego na rede em determinado intervalo de tempo.
- ✓ Unidade mínima policiada (m) – Para cada unidade que sai é retirado um pacote do balde, a taxa com que as unidades saem regulam o tráfego de saída dos pacotes, pois a unidade é usada para o policiamento do tráfego, onde é comparado o tamanho do pacote com valor de “m”. Para pacotes menores que “m” será contado como sendo igual “m”, para pacotes maiores que “m” será computado seu valor total.
- ✓ Tamanho máximo do pacote (M) – deve ser igual ou menor que MTU³²

As fichas são usadas para manter a taxa de dispersão, pois a saída de cada pacote depende da saída das fichas, se não houver fichas não será possível a saída dos pacotes. O tráfego de saída das fichas é o que regula o tráfego de saída dos pacotes, conforme exemplificado na figura 25.

³² Unidade Máxima de Transferência.

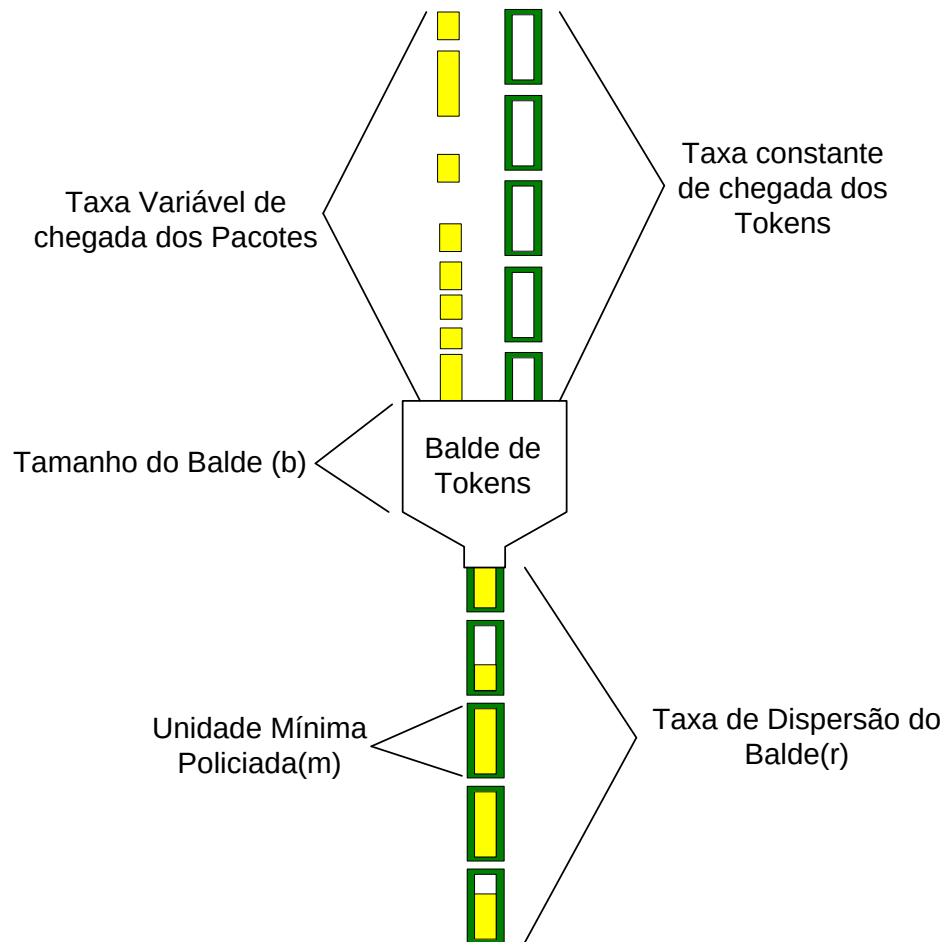


Fig. 25: Exemplo de Balde de Fichas
 Fonte: Elaborado pelo Autor (2009)

2.5.2.1.2 O protocolo RSVP

O funcionamento do protocolo é implementado através da troca de várias mensagens. Cada tipo de mensagem é composta por um conjunto de objetos, onde cada objeto tem uma função específica. O objeto transportado está atrelado ao tipo da mensagem, a qual tem uma determinada finalidade, na figura 26 ilustra três tipos de mensagens PATH, RESV e PATHTEAR.

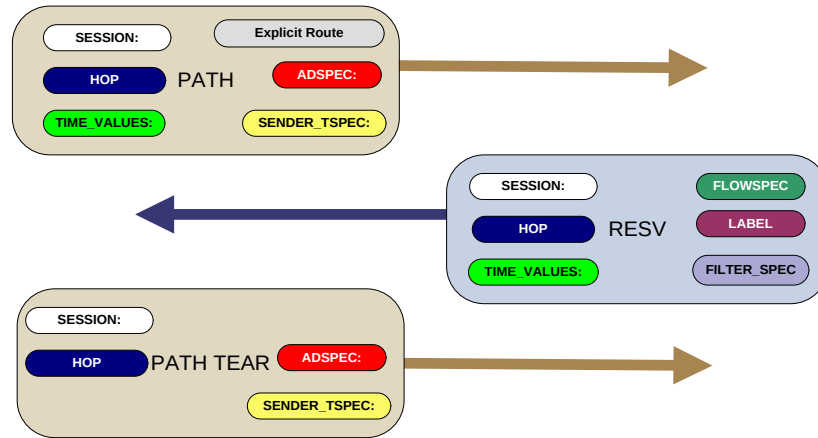


Fig. 26: Exemplo de Objetos transportados nas Mensagens do RSVP
Fonte: Elaborado pelo Autor (2009)

A troca de mensagens do RSVP começa pela mensagem Path, que segue o caminho até o destino dos dados, indicando a cada nó que possivelmente será efetuado uma reserva de recurso. O destino ao receber uma mensagem Path envia as mensagens Resv seguindo o caminho oposto realizado pelas mensagens Path. A mensagem Path também tem a função de indicar o caminho que deve ser seguido pela mensagem Resv, pois aquelas são enviadas pelo caminho que fluem os dados, sendo possível haver tráfego de dados antes da reserva de recursos, por isso a mensagem Resv é usada somente para fazer as reservas de recursos em cada nó no caminho indicado pelas mensagens Path, conforme a figura 27 as mensagens Path seguem o caminho RB1, RC2, RC3 e RB2 e as mensagens Resv seguem o caminho inverso RB2, RC3, RC2, RB1.

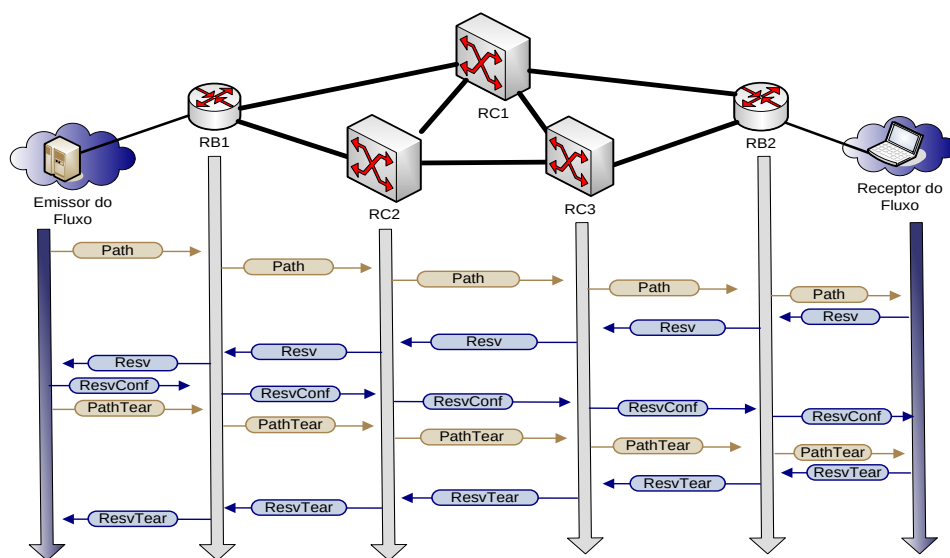


Fig. 27: Sinalização do protocolo RSVP
Fonte: Elaborado pelo Autor (2009)

A mensagem Path transporta a especificação do fluxo, descrito pelo objeto Tspec que é transportado pela mesma, o nó ao receber a mensagem verifica se possui capacidade de receber o fluxo, se for capaz criar outra mensagem Path e envia para o próximo nó, senão envia uma mensagem de erro para o remetente.

Na mensagem Path transporta ainda outro objeto, chamado de Adspec que ao passar por cada nó coleta informações, que são cumulativas, de cada nó ao longo do caminho percorrido por Path.

O nó receptor do tráfego ao receber o Path calcula os recursos que precisam ser reservados e que possam caber dentro dos limites da rede segundo o objeto Adspec, envia de uma mensagem Resv seguindo o caminho de retorno, que é o mesmo percorrido por Path.

Para se manter os reserva de recursos no nó é necessário periodicamente avisá-lo que se deve manter a reserva, isto é feito através do envio de repetidas mensagens Path e Resv periodicamente, senão após um determinado período sem receber nenhuma mensagem o nó libera os recursos para outros fins.

Existe um acoplamento fraco entre as mensagens Path e Resv, pois os períodos entre elas podem ser diferentes, isto é útil para fornece adaptabilidade de mudanças na rede, característica inerente a redes roteadas, além de permitir a mesclagem de fluxos.

A mensagem ResvConf é uma confirmação da chegada de uma mensagem Resv no transmissor, entretanto somente é enviada se em na mensagem Resv houver a indicação que o receptor deseja confirmação.

PathTear e ResvTear são mensagens que segue sempre o caminho definido por Path e indicam que não é mais necessário a reserva de recursos. Existem outros tipos de objetos, além dos mencionados, com características e função mais específicas, que fogem ao escopo deste trabalho.

Independente da reserva de recursos ou não o trafego continua a fluir, pois segue o caminho definido pelo protocolo de roteamento.

3 SERVIÇOS

3.1 VIRTUAL PRIVATE NETWORK (VPN)

VPN(*Virtual Private Network*) é uma rede privada virtual, que pode trafegar informações de forma segura, construída sobre a infra-estrutura de uma rede pública. A infra-estrutura pode contemplar a camada 2 ou 3 do modelo OSI da ISO. Em uma VPN de camada 2 usa-se switch de camada 2, como por exemplo ATM ou Frame-Relay para estabelecer conexão ponto-a-ponto, sendo a troca de informações de roteamento efetuada entre os roteadores do cliente, pois o provedor somente fornece conexão a nível de camada 2. Em uma VPN de camada 3 utiliza-se técnicas de tunelamento, pacotes são transmitidos em um túnel privado que simula uma conexão ponto-a-ponto através da Internet, neste caso a troca de informações de roteamento é efetuada entre os roteadores do cliente e os roteadores do provedor.

As VPNs permitem estender as redes corporativas de uma empresa a pontos distantes. Porém, ao invés de utilizar um grande número de linhas dedicadas para interconexão entre seus diversos pontos, o que onera muito o custo da rede, estas utilizam a infra-estrutura do provedor de serviço para criar uma malha de rede virtual.

A RFC 2547 define como criar uma VPN de camada 3 com funcionalidades MPLS, permitindo o provedor de serviço implementar uma VPN utilizando seu Backbone IP, para prover este serviços aos seus clientes. Os serviços de VPN MPLS utilizam uma extensão do protocolo BGP definida na RFC2283, conhecida com MP-BGP (Multi Protocolo - BGP), para trocar informações de roteamento dos clientes e o MPLS para envio do tráfego dados dos clientes através do Backbone IP. O MP-BGP é uma extensão que define vários espaços de endereço atrelado ao tipo de protocolo de rede, como por exemplo, IPV6, VPN IPv4, entre outros.

3.1.1.1 VRF (Virtual Routing and Forwarding)

Uma VRF é semelhante a um roteador virtual. A figura 29 ilustra um exemplo de VRF, dentro do roteador de borda podemos ter vários roteadores virtuais e cada roteador virtual dispõe de recursos totalmente separados, semelhante a um roteador real, tais como:

- **Tabela de Roteamento IP** – uma tabela de roteamento para cada VRF,
- **Interfaces Associadas** – Interfaces associada exclusivamente a VRF
- **Protocolo de roteamento** – que controla como as informações são inseridas e excluídas da tabela de roteamento da VRF.

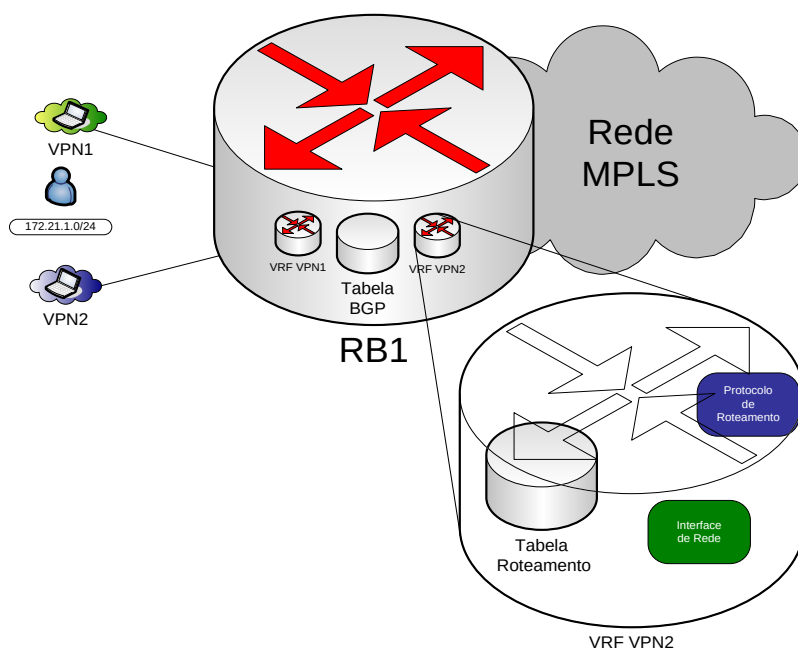


Fig. 29: VRF (virtual routing and forwarding)
Fonte: Elaborado pelo Autor (2009)

Cada VRF define um CE ligado ao PE, somente é possível uma VRF para cada site, mas um site pode estar associado uma ou mais VPN. Cada VRF contém todas as rotas aprendidas das respectivas VPN a qual faz parte, além de possuir um conjunto de Tabela de Roteamento separadas, uma para cada VRF, isto impede que rotas oriundas de outras VPN sejam importadas, exceto em configurações específicas.

3.1.1.2 Divulgação de rotas VPN IPv4

Todas as rotas são exportadas e importadas de cada VRF pelo protocolo BGP, conforme exemplificado na figura 30, o qual é responsável por divulgar as rotas pela rede MPLS. As rotas são trocadas entre os roteadores de borda através do MP-BGP, as principais informações de roteamento enviadas com o anúncio rota são:

- RD (Route Distinguisher) que é um nome associado a VRF, ou seja, é um identificador de VRF.
- Endereço VPN IPv4 que é composto de 12bytes = 8byte do RD + 4bytes da rota.
- Label que identifica a qual VPN o pacote MPLS de dados pertence quando este for desempilhado.
- Atributo Next-Hop como sendo o próprio PE, pois assim informa para onde se devem direcionar os pacotes contendo o tráfego do cliente através de túneis MPLS.

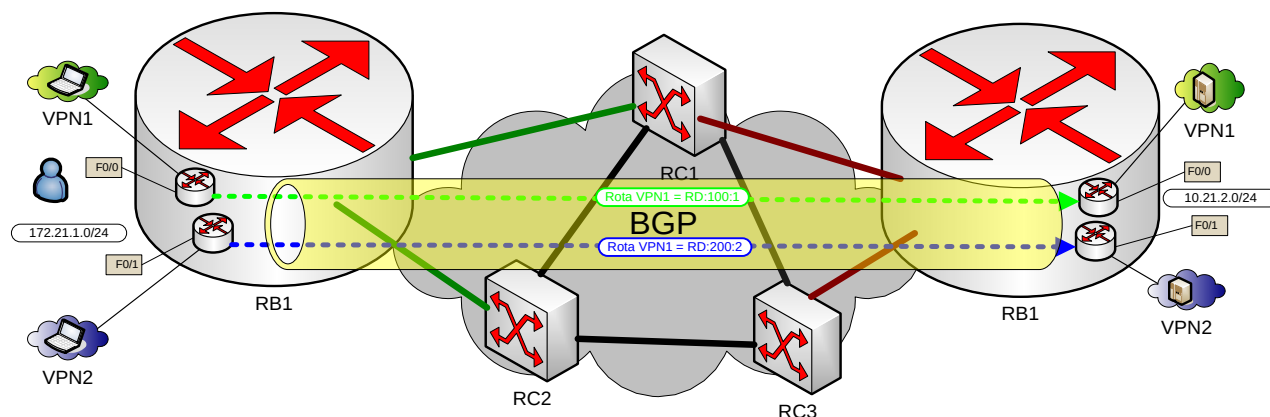


Fig. 30: Divulgação das rotas VPN IPv4
Fonte: Elaborado pelo Autor (2009)

O BGP, através do RD, exportada uma determinada rota somente para sua VRF correspondente, em um determinado roteador de borda. Assim é possível que sites compartilhem espaço de endereços sobreposto (*overlay*), ou seja, um cliente pode usar o mesmo endereçamento IP privado de outro cliente, pois apesar de estarem conectados ao mesmo roteador de borda, possuem VRF distintas. No caso do exemplo da figura 30 ambos os sites de clientes distintos usam do lado esquerdo a rede 172.21.1.0 e do outro lado usam a rede 10.21.1.0, porém eles não interferem entre si.

3.1.1.3 Transporte do Tráfego do Cliente

O core da rede usa o empilhamento de rótulos para transportar o tráfego entre as borda e não se envolvem em com o roteamento das VPN, pois como explicado acima, o plano de controle é implementado pelo protocolo MP-BGP, o qual apenas divulga informação de roteamento, não transporta o tráfego do cliente, entretanto todo tráfego do cliente deve ser conduzido utilizando o protocolo MPLS.

Quando o tráfego do cliente, oriundo da outra extremidade, chega na borda, ele é identificado a qual VPN pertence através do rótulo identificador de VPN que cada pacote transporta. Assim o roteador pode entrega o tráfego a sua VPN correspondente.

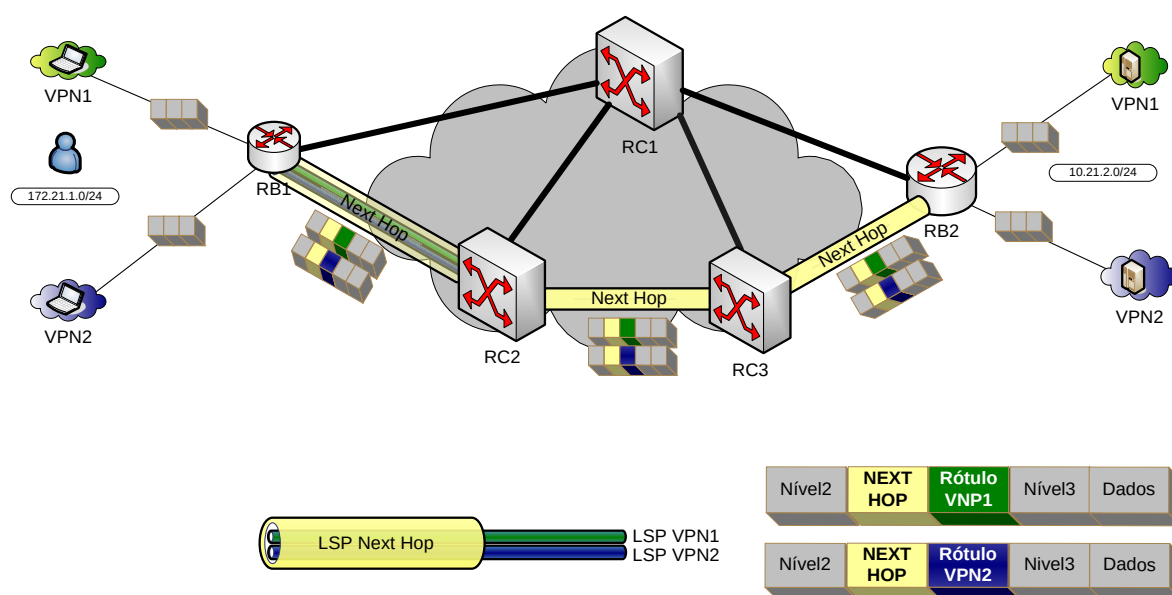


Fig. 31: Transferência dos Dados das VPNs.
Fonte: Elaborado pelo Autor (2009)

Conforme o exemplo didático da figura 31, os pacotes originados na rede 10.21.2.0 com destinos a rede 172.21.1.0, entram em RB2, que por vez inseri em cada pacote o rótulo correspondente a sua VPN, para posteriormente adicionar o rótulo correspondente ao túnel que levará os pacotes até o seu destino, ao chegar no RB1 os pacotes são desempilhados, isto é, retirado o rótulo amarelo, após isso cada pacote é entregue a sua VPN correspondente conforme o rótulo remanescente, que pode ser azul ou verde. Entretanto antes do pacote rotulado ser entregue aos dispositivos dos clientes, os rótulos remanescentes são retirados e os pacotes entregue na rede do cliente sem eles.

3.2 QUALIDADE DE SERVIÇO (QoS)

Na RCF3270 descreve os mecanismo para suporte de MPLS, em redes DiffServ. Em um Domínio DiffServ o pacote, através do DSCP, é atribuído a uma classe, esta por sua vez determina o comportamento do pacote dentro do nó, entretanto toda manipulação dos pacotes MPLS é baseada apenas no rótulo e não no DSCP, então se faz necessário mapear o valor DSCP do pacote para o campo EXP do rótulo MPLS. Apesar de o DSCP ser totalmente independente do EXP ambos tem função identificar a qual classe o pacote pertence, mas podem possuir valores diferentes para um mesmo pacote, pois o DSCP possui 64 valores possíveis e o EXP somente 8. Foram então desenvolvidos duas formas de oferecer o mapeamento DSCP/EXP:

- **E-LSP (*Exp inferred LSP*)** – Para redes que utilizam até 8 PHBs, o LSR utiliza um único LSP, sendo feito o mapeamento direto do DSCP para EXP, neste caso o PHB será atrelado ao pacote, independente do LSP, conforme exemplificado na figura 32.

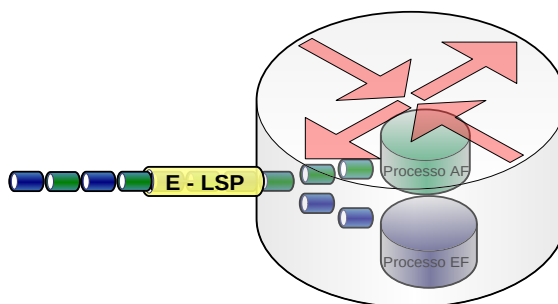


Fig. 32: No E-LSP cada PHB é identificado pelo EXP de cada pacote
Fonte: Elaborado pelo Autor (2009)

- **L-LSP (*Label only inferred LSP*)** - Para redes que utilizam mais de 8 PHB, o PHB é determinado pelo LSP, sendo o EXP o determinante da prioridade de descarte dos pacotes, conforme exemplificado na figura 33.

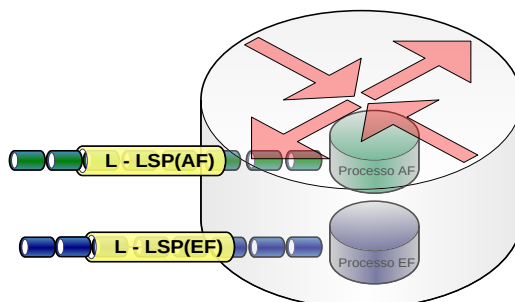


Fig. 33: No L-LSP cada PHB é identificado pelo LSP correspondente
Fonte: Elaborado pelo Autor (2009)

Quando o pacote passa através da rede, o comportamento dele está atrelado a classe ao

qual foi atribuído, entretanto podem haver mudança no valor do EXP durante a sua tramitação. Como o DSCP e EXP são totalmente independentes as mudanças mais externas podem ou não serem refletidas nos rótulos mais internos. A escolha do DSCP ou EXP para identificar a qual classe o pacote pertence está atrelada ao tipo de configuração do nó, pois é através desta que o nó utilizará o PHB e conseqüentemente definirá qual será o comportamento do pacote. Há três tipos de configurações disponíveis:

- ✓ **Uniform** - Neste modo toda mudança no rótulo mais externo é propagada para os demais, inclusive para o DSCP, conforme exemplo da figura 34.

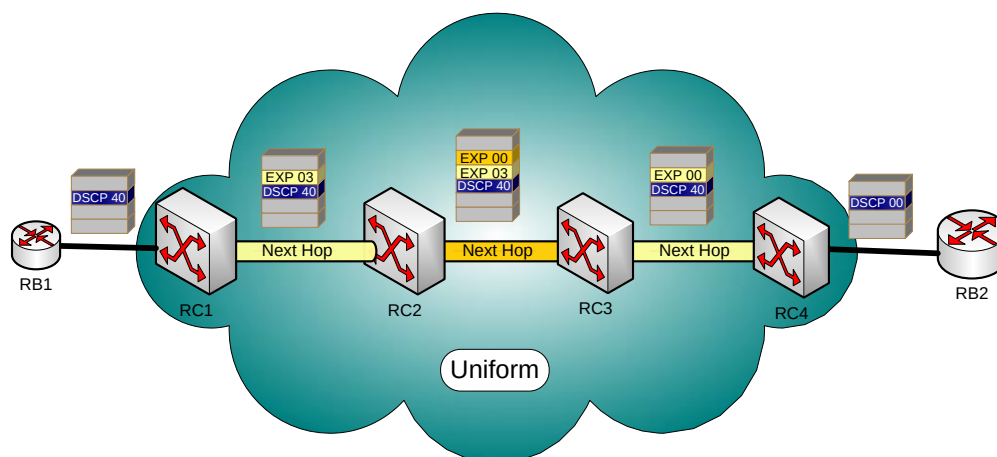


Fig. 34: Modo Uniform
Fonte: Elaborado pelo Autor (2009)

- ✓ **Short-Pipe** – Neste modo o valor do rótulo mais externo é copiado somente para os rótulos mais interno após o desempilhamento, entretanto o valor do DSCP não é alterado, pois este é usado pelo PHB, conforme o exemplo da figura 35.
- ✓ **Pipe** – Neste modo o valor do rótulo mais externo é copiado somente para os rótulos mais interno, após o desempilhamento, entretanto o valor do DSCP não é alterado, mas é o EXP que é usado pelo PHB e não o DSCP, conforme o exemplo da figura 35.

- **Engenharia de Rede** – consiste na manipulação da rede para se ajustar ao seu tráfego, é feita em longa escala de tempo, pois consiste na expansão da rede e sendo necessária aquisição de novos equipamentos, mudança de infra-estrutura, novas ferramentas de gerenciamento, etc.
- **Engenharia de Tráfego** – consiste na manipulação do tráfego para se ajustar a rede, é feita em curto prazo de tempo, através de configuração. Desta forma o tráfego pode-se se ajustar a rede, garantido o bom desempenho das aplicações usadas dentro de limites estabelecidos.

3.3.1 Funcionamento

A figura 36 ilustra um exemplo de rede em que dois tipos de tráfegos diferentes, dados e voz, apesar de terem o mesmo destino possuem características diferentes. Em uma rede IP sem TE todo tráfego segue pelo mesmo caminho, independente da aplicação, que neste exemplo causaria um baixo desempenho para uma das duas aplicações, pois o tráfego de voz e dados ou passaria por um link de satélite, que é ideal para dados, mas não para voz, pois possui alta latência apesar de largura de banda disponível, ou então os dois passariam por um link terrestre, que é ideal pra voz, mas não para dados, pois possui pouca largura de banda apesar de ter baixa latência.

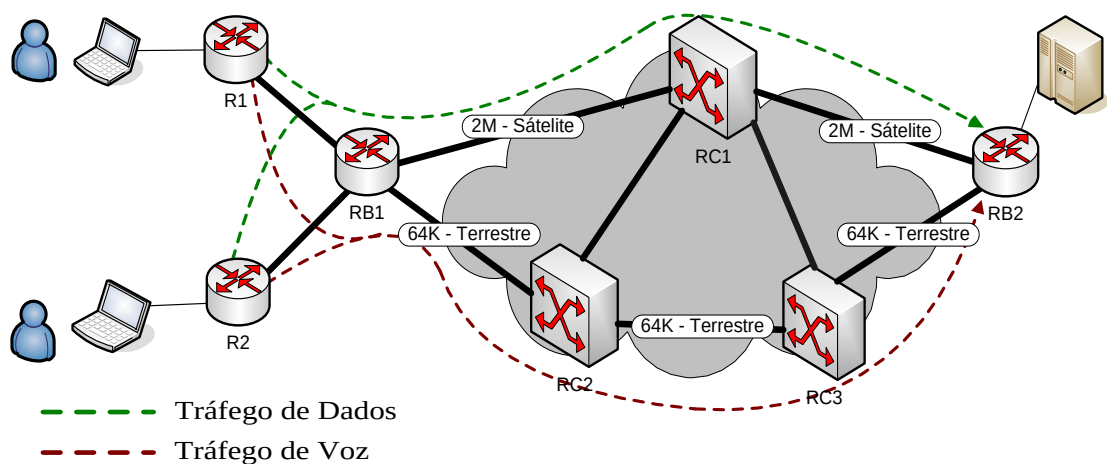


Fig. 36: Rede com MPLS-TE
Fonte: Elaborado pelo Autor (2009)

A situação perfeita é o tráfego de voz pelo enlace de baixa latência e o de dados pelo de maior largura de banda disponível, isto é possível com MPLS-TE. A idéia do MPLS-TE é permitir que sejam criados caminhos usando outras informações que não sejam somente do caminho mais curto usado pelo IP.

Para se obter a divisão do tráfego e os benefícios acima exemplificados conforme ao exemplo da figura 36 é necessário a criação de caminhos alternativos baseados em outras informações, tais como latência largura de banda disponível, jitter, latência, etc. Este caminho são implementados através da criação de túneis virtuais, os quais conduzem o tráfego específico por caminhos que preencham determinados requisitos. Estes túneis virtuais são LSPs do MPLS e cada um é associado a uma interface virtual, pois para o roteador cada interface virtual é apenas mais um enlace disponível.

Entretanto são necessárias três tecnologias (MPLS-TE, RSVP e OSPF-TE) para criação dos túneis. A interação entre estas tecnologias é efetivada por quatro procedimentos básicos:

- **Distribuição da Informação** - uma extensão do protocolo, conhecido como OSPF-TE, acrescenta outras informações a cada interface que participa do processo. As informações adicionais são:
 - ✓ Largura de Banda disponível – cada interface possui uma largura de banda disponível que pode ser reservada. Durante a criação do túnel parte desta largura de banda é reservada e alocada a um determinado túnel, a reserva destes recursos é feita pelo RSVP que dura enquanto o túnel existir.
 - ✓ Prioridade do túnel – podem haver vários túneis em uma interface, sendo alguns prioritários, estes podem expulsar outros do caminho para assumir espaço da banda, caso não haja banda suficiente para criação de um túnel com maior prioridade.
 - ✓ Métrica – é um custo adicionado ao link, geralmente indica o atraso que a interface possui, sendo útil para calcular um caminho baseados na latência.
- **Descoberta do Caminho** - após a informação adicional ser distribuída pela rede o protocolo de roteamento (OSPF-TE), faz a descoberta de um novo caminho baseado nos requisitos estabelecidos para o túnel. Geralmente o caminho é diferente daquele criado pelo processo OSPF padrão, o qual usa menos atributos para o calculo do caminho.

- **Configuração do túnel** - após o caminho ser descoberto, é necessário ser feita a reserva de banda através do caminho que especificado pelo protocolo de roteamento (OSPF-TE), a reservado de banda do túnel é efetuada pelo protocolo RSVP, que possui um objeto ERO (*Explicit Route Object*), o qual especifica todos os nós pelo quais se deve passar para chegar ao destino. O RSVP ainda informa qual o Label deve ser usado para criação do túnel.
- **Encaminhamento dos Dados** - por ultimo deve ser escolhido qual tráfego deve ser encaminhado pelo túnel recém criado. Dependendo da configuração o tráfego pode ser encaminhado manual ou dinamicamente. Todo o túnel é unidirecional, portanto é necessária a criação de outro no sentido contrario para comunicação full-duplex³³.

Após todos estes procedimentos serem implementados os túneis e todo o tráfego, do nosso exemplo, de dados seguirá pelo caminho do satélite e o de voz seguirá pelo caminho terrestre, conforme exemplificado na figura 36. Esta é uma situação ideal, neste exemplo não foi necessário uma Engenharia de Rede para solucionar o problema, mas tão somente uma Engenharia de Tráfego. A decisão de qual tipo de engenharia será usada depende do problema em questão e da estrutura da rede, pois como é notório algumas soluções não se aplicam a determinados tipos de problemas.

³³ Comunicação nos dois sentidos concomitantemente.

4 CONSIDERAÇÕES FINAIS

O MPLS fornece as redes uma série de benefícios tais como:

O domínio MPLS é transparente para o cliente, a rede é vista apenas como uma nuvem encapsulando detalhes irrelevantes aos clientes.

Agilidade no encaminhamento de pacotes, proporcionado pela inspeção do rótulo no núcleo da rede, pois os pacotes IP são analisados na borda da rede, para então serem encaminhados pelos roteadores de núcleo da rede através dos túneis MPLS.

O MPLS permite a implementação de Engenharia de Tráfego, que não é um serviço oferecido aos clientes, mas é uma poderosa ferramenta para gerência da rede.

A criação de túneis MPLS que permitem entre outros benefícios a criação de VPNs em ambientes de grandes escalas, incluindo recursos de qualidade de serviço, para se obter um melhor desempenho da mesmas.

Implementação da tecnologia MetroEthernet, permitindo o transporte quadros ethernet encapsulados em LSPs dentro do Backbone IP, além da integração com muitas outras tecnologias.

Face aos benefícios expostos acima podemos concluir que MPLS possibilita a implementação de redes práticas, econômicas, confiáveis, flexíveis, seguras e inter-operáveis propiciando o uso mais adequado e racional das redes através de um maior poder de gerenciamento e administração.

REFERÊNCIAS

TANENBAUM, Andrew S. **Redes de Computadores 4ª Ed.** Rio de Janeiro: Elsevier, 2003.

FARREL, Adrian. **A Internet e seus Protocolos.** Rio de Janeiro: Elsevier, 2005.

PAQUET, Catherine, TEARE, Diane. **Construindo Redes CISCO Escaláveis.** São Paulo: Pearson Education do Brasil, 2003.

OSBORNE, Eric e SIMHA Ajay. **Engenharia de Tráfego com MPLS.** Rio de Janeiro: Campus, 2002.

ANDERSSON, L. et al. **RFC 3036: LDP Specification**, 2001.
Disponível em: <<http://www.ietf.org/rfc/rfc3036.txt>>.

SWALLOW, George, ANDERSSON, Loa. **MPLS Charter.**
Disponível em: <<http://www.ietf.org/html.charters/mpls-charter.html>>.
Acesso: 12/05/2008