

JULIANA AKEMI NAKAMURA

**EVOLUÇÃO DAS REDES DE
TELECOMUNICAÇÃO E O
MULTIPROTOCOL LABEL SWITCHING
(MPLS)**

Trabalho de Conclusão de Curso
apresentado à Escola de Engenharia de
São Carlos, da Universidade de São Paulo

Curso de Engenharia Elétrica com ênfase
em Eletrônica

ORIENTADORA: Mônica de Lacerda Rocha

São Carlos

2009

FOLHA DE APROVAÇÃO

Autor: _____

Título: _____

**Trabalho de Conclusão de Curso defendido e aprovado
em ____ / ____ / _____,**

com NOTA _____ (____, _____), pela comissão julgadora:

(Assinatura) _____

(Titulação/nome/instituição)

(Assinatura) _____

(Titulação/nome/instituição)

Coordenador da Comissão de Coordenação do
Curso de Engenharia Elétrica (CoC-EE)

Resumo

Este trabalho objetiva a descrição e o estudo da evolução das redes de telecomunicações e suas melhorias, até a tecnologia atualmente mais moderna nas redes dos provedores de serviços, o MPLS. A *internetwork* trata-se da união de redes individuais, conectadas por dispositivos, que funciona como uma única grande rede eficiente e efetiva, mas com o desafio de suportar a comunicação entre diferentes tecnologias. Modelos de referência e padrões foram desenvolvidos e são descritos neste trabalho, até o *Internet Protocol* (IP), base da *Internet*, mas que não provê a qualidade de serviço necessária para gerenciar aplicativos de vídeo, voz sobre IP e outros. Essa exigência é suprida pelas redes ATM, mas existem problemas de escalabilidade. Assim, o MPLS é descrito como uma melhoria dos métodos de encaminhamento de pacotes pela rede utilizando rótulos anexados aos pacotes IP, células ATM ou quadros. Conclui-se que o benefício real do MPLS é que pode ser utilizado sobre qualquer meio e em qualquer camada que possa transmitir pacotes, com suas técnicas implantadas nos *backbones* sem a necessidade de substituição do legado ATM.

Palavras chaves: evolução das redes de computadores, roteamento Camada 3, comutação de rótulos, ATM, IP sobre ATM, MPLS.

Abstract

This work concentrates on describing the evolution of telecommunication networks and its improvements, until the most modern technology nowadays in service provider's backbones, the MPLS. An *internetwork* is a collection of individual networks, connected by intermediate networking devices, that functions as a single large efficient and effective network, but with the challenge of supporting communication among disparate technologies. Reference models and standards were created and are described in this work, until the Internet Protocol (IP), the underlying fabric of the Internet. In order to manage the multitude of applications such as streaming video, Voice over IP (VoIP), and others, a network requires Quality of Service (QoS) provided by networks Layer 2, such as ATM, but there were scalability problems. Thus, MPLS is described as an improved method for forwarding packets through a network using information contained in labels attached to each IP packet, ATM cell, or Layer 2 frame. In conclusion, the real benefit to the use of MPLS is that it can be used on any media at any layer that can pass data packets, with its techniques being implemented in service provider networks without the need to replace the ATM legacy.

Key words: evolution of computer networks, layer 3 routing, label switching, ATM, IP over ATM, MPLS.

Lista de Figuras

FIGURA 1 – DIFERENTES TECNOLOGIAS DE REDE PODEM SER CONECTADAS PARA CRIAR UMA INTERLIGAÇÃO DE REDES [1] ..	19
FIGURA 2 – O MODELO DE REFERÊNCIA OSI [2].....	22
FIGURA 3 - DIVISÃO EM CAMADAS SUPERIORES E INFERIORES DO MODELO OSI [1]	25
FIGURA 4 - ENCAPSULAMENTO NA TRANSMISSÃO DE INFORMAÇÃO [1]	27
FIGURA 5 - O MODELO DE REFERÊNCIA TCP/IP [2]	28
FIGURA 6 - PROTOCOLOS E REDES NO MODELO TCP/IP [2].....	29
FIGURA 7 - DADOS DAS CAMADAS SUPERIORES COMPÕE O QUADRO (<i>FRAME</i>) DA CAMADA DE ENLACE DE DADOS [1]	31
FIGURA 8 - TRÊS COMPONENTES BÁSICOS COMPÕE UM PACOTE (<i>PACKET</i>) DA CAMADA DE REDE [1]	32
FIGURA 9 - DOIS COMPONENTES COMPÕE UMA CÉLULA (<i>CELL</i>) [1].....	32
FIGURA 10 - FORMATO DO QUADRO <i>FRAME RELAY</i> [1].....	38
FIGURA 11 - TRANSMISSÃO DE CÉLULAS ATM VAZIAS PARA COMPLETAR A TAXA DE TRANSMISSÃO [4]	40
FIGURA 12 - VCS CONCATENAM-SE PARA FORMAREM VPS [1]	41
FIGURA 13 - COMUTAÇÃO ATM	41
FIGURA 14 - CONJUNTO DOS PROTOCOLOS DA <i>INTERNET</i> REFERENCIADOS AO MODELO OSI [1]	45
FIGURA 15 - ENDEREÇO IP COMPOSTO DE 32 <i>BITS</i> SEPARADOS EM QUATRO OCTETOS [4]	46
FIGURA 16 - CABEÇALHOS IPv4 e IPv6 [1].....	47
FIGURA 17 - MODELO SOBREPOSTO CONTRA O MODELO INTEGRADO [6]	50
FIGURA 18 - ALGORITMO DE DIJKSTRA [3]	55
FIGURA 19 - EXEMPLO DE CLASSE DE EQUIVALÊNCIA (FEC) [9].....	59
FIGURA 20 - EXEMPLO DE REDE IP BASEADA EM NÚCLEO ATM [9].....	61
FIGURA 21 - REDE QUE SE BENEFICIARIA DA ENGENHARIA DE TRÁFEGO [9].....	63
FIGURA 22 - TOPOLOGIA DE UMA REDE MPLS [9].....	66
FIGURA 23 - PACOTES IP EM UMA REDE MPLS [1].....	73
FIGURA 24 - PROBLEMA DE INTERCALAÇÃO DE CÉLULAS NO DOMÍNIO ATM-LSR [9]	76
FIGURA 25 - RÓTULOS DIFERENTES PARA CÉLULAS COM UM MESMO DESTINO [9].....	76
FIGURA 26 - ARQUITETURA DE UM NÓ MPLS [6].....	77
FIGURA 27 - FORMATOS DE RÓTULO MPLS [6].....	78
FIGURA 28 - <i>LABEL FORWARDING INFORMATION BASE</i> (LFIB) [6]	80
FIGURA 29 - NÍVEIS DE <i>LABEL-SWITCHED PATH</i> [6]	84
FIGURA 30 - LSP COM <i>LOOP</i> [6].....	88
FIGURA 31 - DETECÇÃO DO <i>LOOP</i> COM O DECREMENTO DO CAMPO TTL [6].....	88
FIGURA 32 - MECANISMO DE CONTAGEM DE SALTOS PARA DETECÇÃO DE <i>LOOPS</i> [6]	90

Lista de Siglas

ABR	Available Bit Rate
ARIS	Aggregate Route-Based IP Switching
AS	Autonomous System
ASIC	Application-Specific Integrated Circuit
ATM	Asynchronous Transfer Mode
BGP	Border Gateway Protocol
CBR	Constant Bit Rate
CRC	Cyclic Redundancy Check
CR-LDP	Constraint-based Routed Label Distribution Protocol
CRS	Cell-Switched Router
DLCI	Data-link Connection Identifier
DNS	Domain Name Service
DoS	Denial of Service
EGP	Exterior Gateway Protocol
E-LSR	Edge Label-Switching Router
FCS	Frame Check Sequence
FDDI	Fiber Distributed Data Interface
FEC	Forwarding Equivalency Class
FIB	Forwarding Information Base
FRAD	Frame Relay Access Device
FTP	File Transfer Protocol
GFR	Guaranteed Frame Rate
HTTP	Hypertext Transfer Protocol
IEEE	Institute of Electrical and Eleltrctronics Engineering
IGP	Interior Gateway Protocol
InterNIC	Internet Network Information Center
IP	Internet Protocol

ISDN	Integrated Services Digital Network
ISO	International Organization for Standardization
ISP	Internet Service Provider
ITU-I	International Telecommunication Union Telecommunication Standardization
LAN	Local-area Network
LANE	LAN emulation
LDP	Label Distribution Protocol
LFIB	label forwarding information base
LIB	label information base
LLC	Logical Link Control
LP	Linha privada
LSC	Label Switch Controller
LSP	Label-Switched Path
LSR	Label-Switched Router
MAC	Media Access Control
MBS	Maximum Burst Size
MPLS	Multiprotocol Label Switching
NCP	Network Control Protocol
NNTP	Network News Transfer Protocol
nrt-VBR	Non real time VBR
OSI	Open System Interconnection
OSPF	Open Shortest Path First
PBR	Policy Based Routing
PC	Personal computer
PCR	Peak Cell Rate
PDU	Protocol Data Unit
PDU	protocol data unit
PHP	Penultimate Hop Popping
PNNI	Private Network-to-Network Interface
PVC	Permanent Virtual Connection
QoS	Quality of Service
RIP	Routing Information Protocol
RSVP	Resource Reservation Protocol
RT-VBR	Real time VBR
SCR	Sustainable Cell Rate
SMTP	Simple Mail Transfer Protocol

SVC	Switched Virtual Circuit
TCP	Transmission Control Protocol
TDM	Time Division Multiplexing
TE	Traffic Engineering
TELNET	Teletype Network
ToS	Type of Service
TTL	Time to Live
UBR	Unspecified Bit Rate
UDP	User Datagram Protocol
UNI	User-Network Interface
VBR	Variable Bit Rate
VC	Virtual Channels
VCC	Virtual Channel Connection
VCI	Virtual Channel Identifier
VP	Virtual Path
VPC	Virtual Path Connection
VPI	Virtual Path Identifier
VPN	Virtual Private Network
WAN	Wide-area Network

Sumário

1	Introdução	17
2	Internetwork	19
2.1	Modelo de referência <i>Open System Interconnection</i> (OSI)	20
2.2	O modelo de referência TCP/IP	27
2.3	Serviços orientados a conexão e sem conexão	30
2.4	Formatos da informação	31
3	Redes de Telecomunicações e Tecnologias WANs	33
3.1	Histórico da evolução das redes de telecomunicações	34
3.2	A tecnologia <i>Frame Relay</i>	37
3.2.1	Identificador de conexão de enlace (DLCI)	37
3.2.2	O quadro <i>Frame Relay</i>	38
3.3	<i>Asynchronous Transfer Mode</i> (ATM)	39
3.3.1	Qualidade de serviço ATM (QoS ATM)	42
3.4	A arquitetura da <i>Internet</i>	43
3.4.1	Os protocolos IP e TCP	44
3.4.2	O endereçamento IP	46
3.5	O modelo de sobreposição do IP e ATM	48
3.6	A comutação de rótulos	51
4	O roteamento na Camada de Rede	53
4.1	Componente de encaminhamento	53
4.2	Componente de controle	54
4.2.1	Os algoritmos de roteamento	54
4.2.2	Alguns protocolos de roteamento mais utilizados	56

4.3	Classes de equivalência	58
4.4	As restrições do roteamento tradicional.....	59
4.4.1	O paradigma do roteamento na camada de rede.....	59
4.4.2	Serviços de pacote diferenciados	62
4.4.3	Controle e roteamento independentes	63
4.4.4	Propagação de informações de roteamento externas.....	64
5	Multiprotocol Label Switching (MPLS).....	65
5.1	Benefícios do MPLS	67
5.1.1	MPLS <i>versus</i> IP sobre ATM.....	69
5.1.2	MPLS <i>versus</i> Roteamento em Camada 3.....	70
5.2	Operação MPLS	71
5.2.1	Operação em modo quadro	72
5.2.2	Operação em modo célula.....	74
5.3	Arquitetura do nó MPLS.....	77
5.3.1	Plano de encaminhamento	77
5.3.2	Plano de controle.....	81
5.4	Elementos MPLS.....	82
5.4.1	<i>Label-Switched Router</i> (LSR)	82
5.4.2	<i>Label-Switched Path</i> (LSP).....	83
5.4.3	<i>Label Distribution Protocol</i> (LDP).....	84
5.5	Sobrevivência, detecção, e prevenção de <i>loop</i> no MPLS	86
5.5.1	Sobrevivência de <i>loop</i>	87
5.5.2	Detecção de <i>loop</i>	89
5.5.3	Prevenção de <i>loop</i>	90
5.6	Algumas aplicações do MPLS.....	91
5.6.1	Engenharia de tráfego	91
5.6.2	Qualidade de serviço	92
5.6.3	Redes virtuais privadas.....	93
6	Conclusão	94
	Referências Bibliográficas	97

1 Introdução

Esta monografia apresenta uma revisão bibliográfica sobre a evolução das redes de telecomunicações até as tecnologias mais convergentes, como o *Multiprotocol Label Switching* (MPLS). A motivação para este estudo partiu do estágio realizado na empresa PromonLogicalis, integradora independente de projetos de tecnologia da informação e comunicação (TIC), no qual o conhecimento técnico das tecnologias aqui apresentadas é de grande importância para o desenvolvimento de projetos de redes *backbone* e corporativas.

No início do surgimento das redes de dados, o transporte da informação no formato de pacotes utilizava a rede telefônica como meio de transmissão. Porém, limitações causadas pelas inadequações da rede dimensionadas para o transporte de voz e não de dados, como pequenas taxas de transmissão, motivaram o surgimento de técnicas que aumentassem essa capacidade. A evolução das redes de telecomunicações é baseada no aumento da utilização de aplicativos que cada vez exigem maiores taxas de transmissão

Desde a criação e introdução comercial em 1992, a *Internet* cresceu rapidamente de uma rede de pesquisa para uma rede mundial de comunicação, se tornando um meio conveniente e rentável para a colaboração do usuário, aprendizado, comércio eletrônico e entretenimento. É consenso comum que a *Internet* se tornou um meio no qual voz, vídeo e comunicações de dados se convergem, crescendo em termos de largura de banda, dimensão geográfica e volume de tráfego. Conseqüentemente, as redes de telecomunicações evoluíram para suportar essas demandas que exigiam cada vez mais garantias de qualidade de serviço.

Quando o *Asynchronous Transfer Mode* (ATM) foi lançado, esperava-se que ele dominasse o cenário mundial devido a sua taxa de transmissão consideravelmente maior que a atingida pela *Integrated Services Digital Network* (ISDN) e a agregação de múltiplos tipos de serviços como voz, vídeo e dados em uma pequena célula de tamanho fixo. Porém essa tecnologia

não se integra muito bem com protocolo de rede mais difundido no mundo, o IP, e possui grandes problemas de escalabilidade na utilização IP sobre ATM (IP over ATM). Por isso foi criada a tecnologia de comutação baseada em rótulos, *Label Based Switching* (LBS), possibilitando usufruir a inteligência do roteamento das redes baseadas em pacotes (redes IP) e o desempenho das taxas de transmissão das redes orientadas a conexão (redes ATM).

Muitas soluções comerciais proprietárias do LBS foram criadas, cada qual com suas implementações próprias, dificultando a interoperabilidade entre elas. Para obter uma solução aberta, interoperável e independente de protocolos, o *Internet Engineering Task Force* (IETF) padronizou essas características em um protocolo, aplicável a qualquer protocolo da camada de rede, chamado *Multiprotocol Label Switching* (MPLS), ou multiprotocolo de comutação de rótulos, proporcionando encaminhamento e comutação eficientes de fluxos de dados através da rede.

As motivações do MPLS são a obtenção de maior escalabilidade na rede, mais rapidez no encaminhamento de pacotes, integração IP + ATM, engenharia de tráfego, redes virtuais privadas MPLS e qualidade de serviço. A integração IP + ATM é importante devido ao legado existente nas redes dos provedores de serviços, que acreditavam que o ATM seria a tecnologia dominante nas redes WAN e LAN. Isso não ocorreu nas redes LAN devido ao alto custo dos dispositivos ATM para computadores, o que tornou necessária uma tecnologia que unificasse os benefícios do protocolo de rede IP, amplamente difundido, com os equipamentos ATM existentes.

2 Internetwork

Internetwork é um termo que significa redes interligadas e trata-se de um conjunto de redes individuais, locais, conectadas por intermédio de dispositivos de rede, se comportando com uma única grande rede de comunicação [1]. A Figura 1 ilustra alguns tipos diferentes de tecnologias de rede como a *Ethernet*, *Token Ring* e *Fiber Distributed Data Interface (FDDI)*, que podem ser interconectadas por roteadores e outros dispositivos de rede para criar uma *Internetwork*.

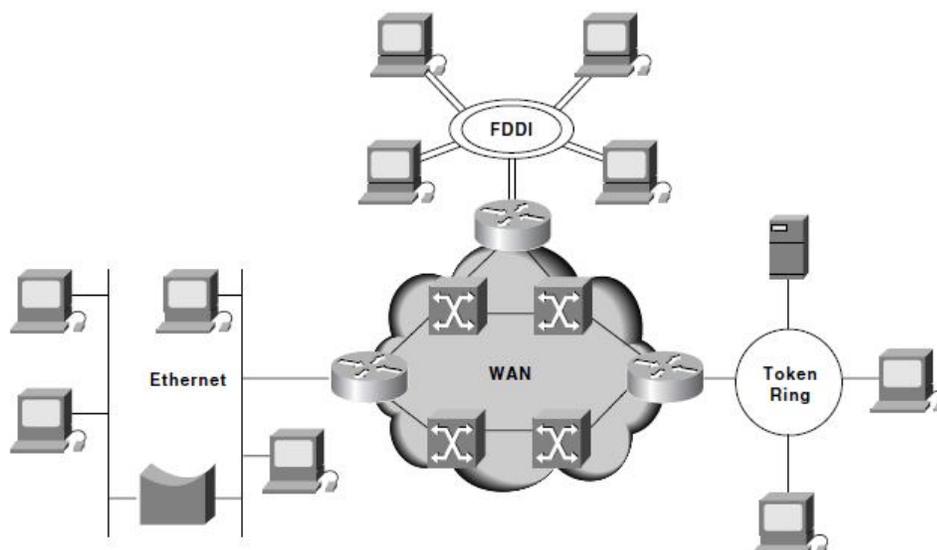


Figura 1 – Diferentes tecnologias de rede podem ser conectadas para criar uma interligação de redes [1]

As redes de áreas locais, *Local-area Networks (LANs)*, evoluíram em torno da revolução do computador pessoal, *personal computer (PC)*, e permitiram que múltiplos usuários, em um espaço geográfico relativamente pequeno, trocassem arquivos e mensagens, assim como acessar recursos compartilhados como servidores de arquivos e impressoras.

As redes de longa distância, *Wide-area Networks* (WANs), interconectam LANs com usuários geograficamente dispersos para criar conectividade. Algumas tecnologias utilizadas incluem ATM, *Frame Relay* e enlaces de rádio. Atualmente, LANs banda larga e redes comutadas são amplamente utilizadas, principalmente porque elas operam com maiores taxas de transmissão, suportando aplicações como multimídia e videoconferência.

As redes evoluíram como solução de três principais problemas: LANs isoladas, duplicação de recursos e falta de gerenciamento de rede. LANs isoladas tornavam a comunicação entre diferentes escritórios ou departamentos impossível, enquanto a duplicação de recursos fazia com que os mesmos *hardware* e *software* fossem fornecidos para cada escritório ou departamento, exclusivamente. Além disso, nenhum método centralizado de gerenciamento e resolução de problemas na rede existia, justificando a falta de gerenciamento de rede.

A implementação de uma rede funcional não é simples. Há muitas dificuldades em relação à conectividade, confiabilidade, gerenciamento de rede e flexibilidade que precisam ser supridas para estabelecer uma rede eficiente.

O desafio para se conectar sistemas variados é suportar a comunicação entre as diferentes tecnologias. Para isto foram criados padrões de referência que possibilitam a comunicação entre equipamentos e tecnologias diferentes, como o modelo de Interconexão dos Sistemas Abertos, ou modelo *Open System Interconnection* (OSI).

Embora os protocolos associados ao modelo OSI raramente sejam usados nos dias atuais, o modelo em si é de fato bastante geral e ainda válido, e as características descritas em cada camada ainda são muito importantes. O modelo *Transmission Control Protocol/Internet Protocol*, ou simplesmente modelo TCP/IP, tem características opostas: o modelo propriamente dito não é muito utilizado, mas os protocolos têm uso geral [2].

2.1 Modelo de referência *Open System Interconnection* (OSI)

O modelo de referência OSI descreve como a informação de uma aplicação de *software* em um computador move-se através da rede para uma aplicação de *software* em outro computador. Trata-se de um modelo conceitual composto por sete camadas, cada uma com funções específicas e particulares, desenvolvido em 1984 pela *International Organization for Standardization* (ISO), ou Organização Internacional para a Padronização, foi considerado o primeiro modelo de arquitetura para comunicação entre computadores.

O modelo OSI divide as tarefas envolvidas com a transmissão da informação entre computadores da rede em sete grupos de tarefas menores e mais gerenciáveis, cada uma correspondente a uma das sete camadas OSI. As tarefas podem ser implementadas de maneira independente em sua camada correspondente, permitindo que as soluções oferecidas por uma camada possam ser atualizadas sem afetar as outras camadas. Em resumo, os princípios aplicados para se alcançar à estruturação em sete camadas foram [2]:

- Uma camada deve ser criada onde houver necessidade de outro grau de abstração;
- Cada camada deve executar uma função bem definida;
- A função de cada camada deve ser escolhida tendo em vista a definição de protocolos padronizados internacionalmente;
- Os limites de camadas devem ser escolhidos para minimizar o fluxo de informações pelas interfaces;
- O número de camadas deve ser suficientemente grande para que funções distintas não precisem ser colocadas desnecessariamente na mesma camada, e pequeno suficiente para que a arquitetura não se torne difícil de controlar.

Essa estruturação modular, com módulos divididos em camadas hierárquicas em que cada camada realiza um conjunto de funções e oferece um serviço à camada superior, possibilita a independência entre as camadas e traz as seguintes vantagens [3]:

- Redução da complexidade de desenvolvimento;
- Desenvolvimentos independentes (modulares);
- Maior flexibilidade e simplicidade de implementação e introdução de alterações numa camada;
- Incorporação de novas tecnologias;
- Adoção de normas (padrões) para garantir que máquinas de diferentes fabricantes se comuniquem entre si (interoperabilidade).

Assim, as sete camadas do modelo de referência OSI, observadas na Figura 2, são:

- Camada 7 – Aplicação (*Layer 7 – Application*);
- Camada 6 – Apresentação (*Layer 6 – Presentation*);
- Camada 5 – Seção (*Layer 5 – Session*);
- Camada 4 – Transporte (*Layer 4 – Transport*);
- Camada 3 – Rede (*Layer 3 – Network*);

- Camada 2 – Enlace (*Layer 2 – Data Link*);
- Camada 1 – Física (*Layer 1 – Physical*).

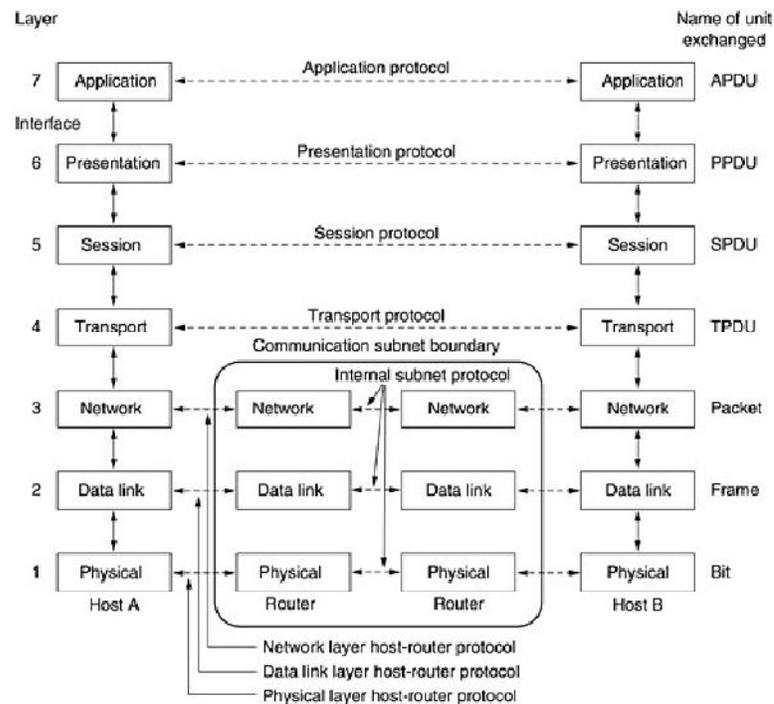


Figura 2 – O modelo de referência OSI [2]

Na Figura 2 observa-se que comunicação acontece entre as camadas, sendo que cada uma possui sua seu tipo de unidade de troca de informações, como por exemplo, o *bit* na camada física, o quadro na camada de enlace e o pacote na camada de rede e a unidade de dados do protocolo, ou *protocol data unit* (PDU), nas outras camadas.

A camada física trata da transmissão de *bits* em um canal de comunicação, das características elétricas e mecânicas do meio, do controle de acesso ao meio, da taxa de transmissão, codificação de linha etc, sendo todas características relacionadas ao meio físico no qual a informação é transportada.

A camada de enlace fornece trânsito seguro dos dados através do enlace físico da rede. Diferentes especificações nessa camada definem redes e características de protocolo diferentes, incluindo endereçamento físico, topologia de rede, notificação de erros, processo de ordenação de quadros e controle de fluxo.

O endereçamento físico (em oposição ao endereçamento de rede) define como os dispositivos estão endereçados na camada de enlace. A topologia de rede consiste das especificações da camada de enlace que frequentemente definem como os dispositivos devem estar fisicamente conectados, tal como uma topologia de barramento ou em anel. As notificações de erro alertam os protocolos das camadas superiores a ocorrência de um erro na transmissão, e o processo de ordenação reordena os quadros transmitidos fora da sequência. O controle de fluxo monitora a transmissão de dados de forma que o dispositivo receptor não seja sobrecarregado com tráfego excessivo.

O Instituto de Engenheiros Elétricos e Eletrônicos (IEEE) subdividiu a camada de enlace em duas subcamadas: controle de enlace lógico, ou *Logical Link Control* (LLC), e controle de acesso ao meio, ou *Media Access Control* (MAC).

A subcamada LLC gerencia comunicações entre dispositivos em um único *link* de uma rede, sendo definida na especificação IEEE 802.2 e suportando serviços orientados a conexão ou não, utilizados por protocolos das camadas superiores. Essa especificação define um número de campos nos quadros da camada de enlace que permitem múltiplos protocolos de camadas superiores compartilharem um único *link* físico de dados. A subcamada MAC gerencia o protocolo de acesso ao meio físico da rede. A especificação IEEE MAC define os endereços MAC que permitem múltiplos dispositivos se identificarem de maneira única na camada de enlace.

A camada de rede define o endereço de rede, que diferentemente do endereço MAC identifica qual o protocolo de rede está sendo utilizado e gera a estrutura de pacote, que são enviados (roteados) de uma máquina de origem para uma de destino por caminhos determinados por algoritmos de roteamento ou tabelas estáticas configuradas manualmente. Algumas implementações da camada de rede, como o *Internet Protocol* (IP), definem endereços de rede de forma que a seleção de rotas possa ser determinada sistematicamente pela comparação do endereço de origem com o de destino e aplicação da máscara de sub-rede. Devido a essa camada definir o plano de rede lógico, roteadores podem utilizá-la para determinar como encaminhar pacotes. Assim, muito do trabalho de projeto e configuração da rede acontece na Camada 3, a camada de rede.

Roteamento é o mecanismo pelo qual se escolhe o caminho (rota) que um pacote deve seguir para atingir seu destino com base nas condições de tráfego da rede e prioridades. Além disso, a camada de rede é responsável pelo controle de congestionamento, quando a

quantidade de pacotes na rede é muito grande e atinge o patamar da capacidade de carga dos equipamentos.

A principal função da camada de transporte é tornar a transmissão pela rede confiável, ou seja, proporcionar um serviço eficiente e de baixo custo aos seus usuários, normalmente usuários da camada de sessão. Outras funções básicas são:

- Receber os dados da camada de sessão, dividi-los em pacotes e enviá-los para a camada de rede;
- Garantir que os pacotes cheguem com a sequência correta à outra extremidade;
- Efetuar a correção de erros;
- Enviar ao transmissor uma informação de recebimento (*acknowledge*), informando que o pacote foi recebido com sucesso e, caso a falta de pacotes seja detectada, solicitar a retransmissão.

A camada de sessão permite que usuários de diferentes máquinas estabeleçam uma sessão de comunicação (*login*), na qual essas aplicações definem como será feita a transmissão de dados e marcações são adicionadas aos dados que estão sendo transmitidos. Em caso de falha na rede, os computadores reiniciam a transmissão dos dados a partir da última marcação recebida pelo computador receptor.

A camada de apresentação trata a sintaxe e a semântica das informações transmitidas, para prover independência nas representações de dados ao traduzi-los do formato do aplicativo para o formato da rede e vice-versa. O objetivo é permitir que computadores que trabalham com diferentes representações se comuniquem apropriadamente.

Por fim, a camada de aplicação provê aplicativos aos usuários, tais como transferência de arquivos, correio eletrônico, acesso a terminais de computadores remotos e funções de gerência.

Podem-se dividir essas camadas em duas categorias: camadas superiores e camadas inferiores, como na observado na Figura 3. As camadas superiores trabalham com questões de aplicação e geralmente são implementadas por *software*, sendo a camada de aplicação a mais próxima do usuário final. As camadas inferiores trabalham com questões de transporte de dados. A camada física e a camada de enlace são implementadas em *hardware* e *software*, sendo a camada física a mais baixa de todas e a mais próxima do meio físico da rede (cabos de rede, por exemplo).

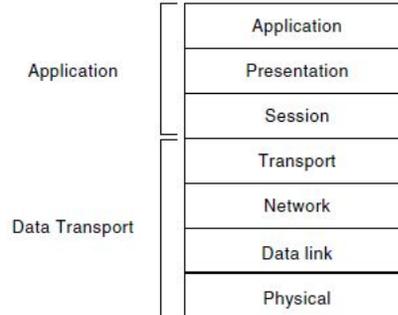


Figura 3 - Divisão em camadas superiores e inferiores do modelo OSI [1]

O modelo OSI fornece uma estrutura conceitual de comunicação entre computadores, mas o modelo em si não é um método de comunicação. A comunicação só é possível por meio de protocolos de comunicação. No contexto das redes de dados, um protocolo é um conjunto de regras e convenções que determinam como os computadores trocarão informações na rede, sendo que um protocolo implementa funções de uma ou mais camadas do modelo OSI.

É necessário, portanto, haver distinção entre os seguintes conceitos no modelo de referência OSI:

- Protocolos consistem nas regras (sintáticas, semânticas e temporais) envolvidas na comunicação entre entidades pares (*peers*) em sistemas diferentes, sendo estruturados em camadas superiores encapsulando as inferiores;
- Interfaces regulam a comunicação entre entidades em camadas adjacentes num mesmo sistema, sendo que com a utilização de serviços fornecidos pela camada inferior, entidades pares cooperam, por meio de um protocolo, com a realização de certas funções e oferecem, por uma interface, um serviço de valor acrescentado à camada superior;
- Serviços correspondem às funções que cada camada possui e são oferecidos por uma camada inferior à uma camada superior.

Uma grande variedade de protocolos existe, incluindo protocolos LAN, WAN, protocolos de rede e de roteamento. Protocolos LAN operam nas camadas física e de enlace, definindo a comunicação por diversos meios LAN, enquanto protocolos WAN operam nas três mais baixas camadas do modelo OSI. Os protocolos de roteamento são protocolos da camada de rede responsáveis pela troca de informação entre roteadores de forma que estes possam

escolher o caminho apropriado para o tráfego da rede, enquanto os protocolos de rede são os diversos protocolos das camadas superiores.

Uma camada OSI se comunica com outra para utilizar os serviços fornecidos pela camada adjacente. Estes serviços ajudam uma dada camada OSI a se comunicar com sua camada par (*peer layer*) em outro sistema.

As sete camadas OSI utilizam diversas formas de informação de controle para se comunicarem com suas camadas pares em outros sistemas de computadores. Esta informação de controle consiste de solicitações e instruções específicas trocadas entre as camadas pares e são tipicamente de dois tipos: cabeçalhos (*headers*) e sufixos (*trailers*) [1].

Os cabeçalhos precedem dados que tenham sido transmitidos de camadas superiores e sufixos procedem aos dados. Porém cabeçalhos, sufixos e dados são conceitos relativos à camada que está analisando a unidade de informação. Na camada de rede, por exemplo, uma unidade de informação consiste do cabeçalho da Camada 3 e dos dados transportados. Na camada de enlace, entretanto, toda a informação transmitida pela camada de rede é tratada como dados. Sendo assim, a quantidade de dados para uma determinada unidade de informação pode conter cabeçalhos, sufixos e dados de todas as camadas mais altas. Esse método é conhecido como encapsulamento, e pode ser observado na Figura 4.

A técnica de encapsulação ilustrada é utilizada para identificar a informação a ser enviada com um cabeçalho que determina, por exemplo, o destino da informação. No destino o cabeçalho deixa de ser útil, pois a informação nele contido já não é mais necessária, e, portanto, é retirado. Essa técnica é utilizada recursivamente em cada elemento de rede adjacente, uma vez que depois de retirado o cabeçalho recebe nova informação sobre o próximo destino dos dados. Cada camada de rede encapsula a informação no formato requerido pela tecnologia de rede empregada naquela camada e, uma vez encapsulada (informação + cabeçalho), passa ser considerada somente informação para a camada seguinte.

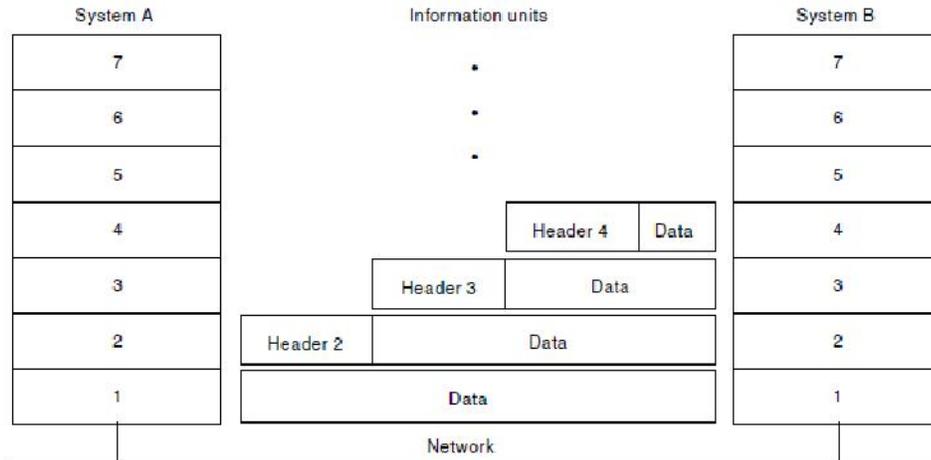


Figura 4 - Encapsulamento na transmissão de informação [1]

O encapsulamento é o fundamento básico do MPLS, pois os quadros são encapsulados por um rótulo (*label*) que contém o endereço para o próximo salto na rede. Uma vez que esse rótulo fica entre a camada de enlace e a camada de rede (*layer 2* e *layer 3*), não há necessidade de análise do conteúdo do pacote para a verificação do seu destino.

2.2 O modelo de referência TCP/IP

Para solucionar o problema das arquiteturas de rede de dados proprietárias com problemas de interoperabilidade, criou-se a arquitetura de comunicações aberta denominada modelo OSI, vista anteriormente, amplamente utilizado como arquitetura de referências, mas com protocolos pouco utilizados. Uma alternativa para a arquitetura aberta, mais simples, foi o modelo TCP/IP, que hoje é um dos padrões mais utilizados mundialmente [4].

Para compreender o objetivo da criação desse modelo de referência é preciso entender a motivação para tal. A ARPANET, primeira rede de comutação de pacotes, era uma rede de pesquisa patrocinada pelo Departamento de Defesa dos Estados Unidos. Aos poucos, centenas de universidades e repartições públicas foram conectadas, utilizando linhas telefônicas dedicadas. Quando foram criadas as redes de rádio e satélite, começaram a surgir problemas com os protocolos existentes, forçando a criação de uma nova arquitetura de referência. Desse modo, a habilidade para conectar várias redes de maneira uniforme foi um dos principais objetivos de projeto, desde o início. Mais tarde, essa arquitetura ficou conhecida como modelo de referência TCP/IP, devido aos seus dois principais protocolos.

Devido à estratégia dessa rede para o Departamento de Defesa dos Estados Unidos, era necessário que as conexões permanecessem intactas enquanto as máquinas de origem e de destino estivessem funcionando, mesmo que algumas máquinas ou linhas de transmissão intermediárias deixassem de operar repentinamente. Além disso, era necessária uma arquitetura flexível, capaz de se adaptar a aplicações com requisitos divergentes, como a transferência de arquivos e a transmissão de dados de voz em tempo real.

Essas necessidades levaram à escolha de uma rede de comutação de pacotes baseada em uma camada de interligação não orientada a conexão (*connectionless*), que será apresentada adiante. Essa camada, chamada camada *Internet*, integra toda a arquitetura com a tarefa de permitir a injeção de pacotes em qualquer rede e garantir que estes trafegarão independentemente até o destino (talvez em uma rede diferente). Os pacotes podem chegar até mesmo em uma ordem diferente daquela em que foram enviados, obrigando as camadas superiores a reorganizá-los.

A camada *Internet* define o formato de pacote e o protocolo IP. A tarefa da camada *Internet* é entregar os pacotes IP onde eles são necessários. O roteamento de pacotes é uma questão de grande importância nessa camada, assim como a necessidade de evitar o congestionamento. Por esses motivos, é razoável dizer que a função da camada *Internet* do modelo TCP/IP é muito parecida com a da camada de rede do modelo OSI. A Figura 5 mostra essa correspondência.

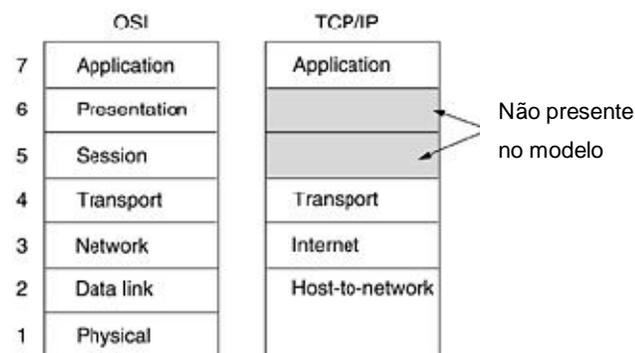


Figura 5 - O modelo de referência TCP/IP [2]

O modelo TCP/IP não tem as camadas de sessão e de apresentação, como ilustra a Figura 5, devido a ausência de necessidades que envolvem essa camada [2]. Isso porque no TCP/IP as funcionalidades dessas camadas estão reunidas na camada de aplicação, uma

vez que essas camadas do modelo OSI normalmente não existem como processos e protocolos separados nas aplicações da *Internet*.

A camada localizada acima da camada *Internet* é a camada de transporte. A finalidade dessa camada é permitir que as entidades pares dos *hosts* de origem e de destino mantenham uma conversação, exatamente como acontece na camada de transporte OSI.

Foi definido o protocolo fim a fim denominado protocolo de controle de transmissão, ou *Transmission Control Protocol* (TCP), que se refere a um protocolo orientado a conexões confiável que permite a entrega sem erros de um fluxo de *bytes* originário de uma determinada máquina em qualquer computador da *Internet* [5]. Esse protocolo fragmenta o fluxo de *bytes* de entrada em mensagens discretas e passa cada uma delas para a camada *Internet*. No destino, o processo TCP receptor volta a montar as mensagens recebidas no fluxo de saída. O TCP também cuida do controle de fluxo, impedindo que um transmissor rápido sobrecarregue um receptor lento com um volume de mensagens maior do que ele pode manipular.

Acima da camada de transporte, está a camada de aplicação. Ela contém todos os protocolos de nível mais alto. Dentre eles estão o protocolo de terminal virtual *Teletype Network* (TELNET), o protocolo de transferência de arquivos *File Transfer Protocol* (FTP) e o protocolo de correio eletrônico *Simple Mail Transfer Protocol* (SMTP), como mostra a Figura 6. O protocolo de terminal virtual permite que um usuário de um computador se conecte a uma máquina distante e trabalhe nela. O protocolo de transferência de arquivos permite mover dados com eficiência de uma máquina para outra. Originalmente, o correio eletrônico era um tipo de transferência de arquivos. No entanto, foi desenvolvido mais tarde um protocolo especializado para essa função, o SMTP.

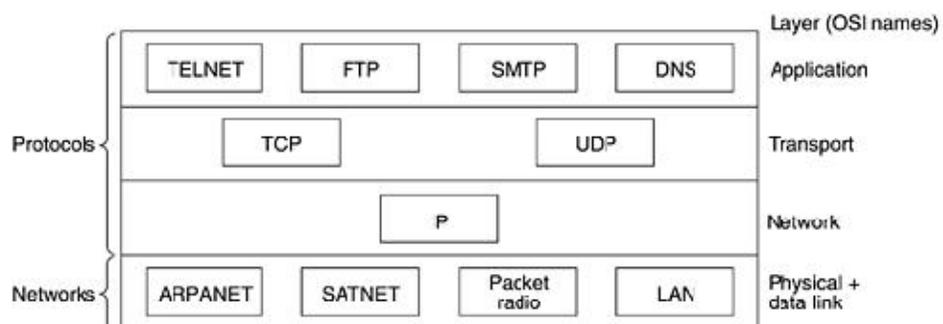


Figura 6 - Protocolos e redes no modelo TCP/IP [2]

Muitos outros protocolos foram incluídos com o decorrer dos anos, como o *Domain Name Service* (DNS), que mapeia os nomes de *hosts* para seus respectivos endereços de rede, o *Network News Transfer Protocol* (NNTP) utilizado para mover novos artigos de notícias da Usenet, e o *Hypertext Transfer Protocol* (HTTP) utilizado para buscar páginas na *World Wide Web*, entre muitos outros.

Abaixo da camada *Internet*, o modelo de referência TCP/IP não é bem especificado. Sabe-se apenas que o *host* tem que se conectar à rede utilizando algum protocolo, para que seja possível enviar pacotes IP. Esse protocolo não é definido e pode variar com o *host* e com a rede. Os livros e a documentação que tratam do modelo TCP/IP raramente descrevem esse protocolo.

2.3 Serviços orientados a conexão e sem conexão

Em geral, protocolos de transporte podem ser caracterizados como orientados a conexão (*connection-oriented services*) ou não orientados a conexão (*connectionless*). Serviços orientados a conexão primeiramente estabelecem uma conexão com o serviço desejado antes de transmitir qualquer informação. Um serviço não orientado a conexão pode enviar dados sem a necessidade de estabelecer uma conexão primeiramente. Em geral, os serviços orientados a conexão oferecem certo nível de garantia de entrega, enquanto o serviço não orientado a conexão, não. O serviço orientado a conexão envolve três fases: estabelecimento da conexão, transferência de dados e término da conexão.

Durante o estabelecimento da conexão, os nós finais devem reservar recursos para a conexão, negociarem e estabelecerem certos critérios para a transferência. Essa reserva de recursos é um dos requisitos explorados em alguns ataques de negação de serviço, os chamados *Denial of Service* (DoS). Um sistema que está sofrendo esse tipo de ataque envia muitos pedidos de estabelecimento de conexão, mas nunca completa a conexão. O computador atacado fica então com recursos alocados para muitas conexões nunca completadas. Então, quando um nó final tenta completar uma conexão real, não há recursos suficientes para a conexão válida.

A fase de transferência de dados ocorre quando a informação real é transmitida em uma conexão. Durante a transferência de dados, a maioria dos serviços orientados a conexão monitoram a perda de pacotes para reenviá-los. O protocolo é geralmente também responsável por colocar os pacotes na sequência correta antes de passar os dados para a

pilha de protocolos. Quando a transferência está completa, o nó final derruba a conexão e libera os recursos antes reservados.

Redes com serviços orientados a conexão têm mais “custos” que os não orientados a conexão, pois esse tipo de serviço deve negociar a conexão, transferir os dados e derrubar a conexão, enquanto o outro apenas transfere os dados sem os custos de criar e derrubar a conexão. Porém, devido às exigências de algumas aplicações, é fundamental a implantação de redes orientadas a conexões que fornecem QoS e boas taxas de transmissão, como é o caso da utilização do MPLS nas redes *backbones* dos provedores de serviços.

2.4 Formatos da informação

As informações de dados e de controle transmitidos através das redes carregam uma variedade de formas. Os termos utilizados para se referirem a estes formatos das informações nem sempre são consistentes, mas algumas vezes utilizados indistintamente. Formatos de informação comuns incluem quadros (*frames*), pacotes (*packets*), datagramas (*datagrams*), segmentos, mensagens, células (*cells*) e unidades de dados (*data units*).

Um quadro (*frame*) é uma unidade de informação cuja origem e destino são entidades da camada de enlace de dados. É composto do cabeçalho da camada de enlace (e possivelmente do sufixo) e dos dados das camadas superiores. O cabeçalho e o sufixo contêm informações de controle destinadas a camada de enlace no sistema de destino. A Figura 7 ilustra os componentes básicos de um quadro da camada de enlace.

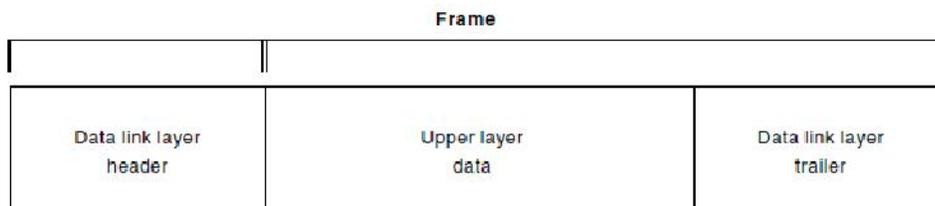


Figura 7 - Dados das camadas superiores compõem o quadro (*frame*) da camada de enlace de dados [1]

Um pacote (*packet*) é uma unidade de informação cuja origem e destino são entidades da camada de rede. É composto do cabeçalho da camada de rede (e possivelmente um sufixo) e dados das camadas superiores. A Figura 8 ilustra os componentes básicos de um pacote da camada de rede.

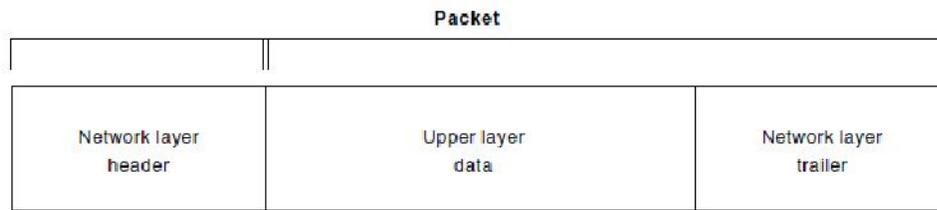


Figura 8 - Três componentes básicos compõe um pacote (*packet*) da camada de rede [1]

O termo datagrama (*datagram*) normalmente refere-se a uma unidade de informação cuja origem e destino são entidades da camada de rede e que utilizam serviços não orientados a conexão, enquanto o termo segmento (*segment*) normalmente se refere a uma unidade de informação cuja origem e destino são entidades da camada de transporte. Já a mensagem é uma informação cuja origem e destinos existem acima da camada de rede (geralmente na camada de aplicação).

Uma célula (*cell*) é uma unidade de informação de tamanho fixo cuja origem e destinos são entidades da camada de enlace. Células são utilizadas em ambientes comutados, assim como nas redes ATM e *Multimegabit Data Service* (MDS) e são compostas de cabeçalho (*header*) e campo de informação (*payload*). O cabeçalho contém a informação de controle destinada a camada de enlace de dados e tem tipicamente de 5 *bytes*. O campo de informação contém informação das camadas superiores que é encapsulada no cabeçalho da célula e possui tipicamente 48 *bytes*. O comprimento do cabeçalho e do campo de informação é fixo em cada célula, como ilustra a Figura 9.

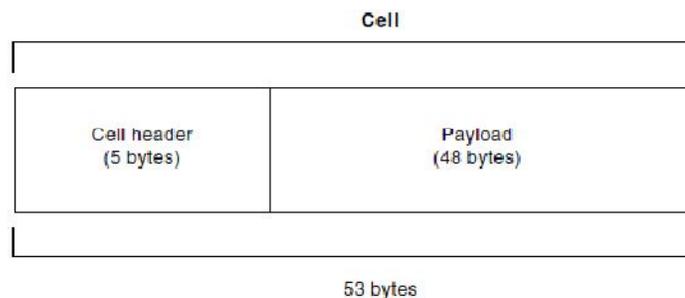


Figura 9 - Dois componentes compõe uma célula (*cell*) [1]

3 Redes de Telecomunicações e Tecnologias WANs

As redes de telecomunicações aqui referidas são aquelas construídas sobre a comunicação de voz, mas que sofreram transformações nas últimas décadas. A convergência de voz e dados iniciou-se com tentativas de transmitir dados sobre as redes de voz, desde 1970. Porém, essas tentativas fizeram com que se lidasse com o tráfego de dados da mesma maneira que o de voz, sendo que as redes de voz eram ruidosas e propensas a erros.

Com o aumento significativo do tráfego de dados, surgiu a necessidade de criação de uma infraestrutura de rede de núcleo (*backbone*), com tecnologias WAN, que possibilitasse a implantação de uma rede com as seguintes características [3]:

- Rede única para comutação de todos os tipos de serviço: voz, dados e vídeo;
- Tecnologia de pacotes, ou seja, as informações a serem encaminhadas são enviadas em blocos denominados pacotes, com campos para inserção da informação a ser transmitida e das informações para encaminhamento destes;
- Flexibilidade para possibilitar uma rápida ampliação da rede devido ao crescimento do tráfego (escalabilidade);
- Qualidade de serviço, ou *Quality of Service* (QoS), equivalente a da rede telefônica para o transporte de voz;
- Baixo custo operacional e economicamente viável para a oferta de serviços de baixo custo.

O transporte de dados começou a ser realizado em redes projetadas para o serviço de voz, sendo apenas posteriormente implantadas redes de dados independentes e paralelas às redes de voz, com sistemas de gerência próprios e alto custo operacional devido à falta de integração com as tecnologias já instaladas e a utilização do legado. Neste cenário surgiu nas redes de telecomunicações uma variedade de tecnologias de comutação de dados baseadas em quadros (*frames*), pacotes (*packets*) e células (*cells*). Em termos cronológicos,

as mais importantes foram o *Frame Relay*, ATM e o MPLS, com o objetivo principal de transportar todos os tipos de aplicações numa plataforma única e seguindo padrões internacionais.

Além disso, com a explosão das redes *Internet*, duas correntes tecnológicas foram crescendo em paralelo: a primeira oriunda da área de telecomunicações, baseada em padrões e com fortes restrições de qualidade de serviço, e a segunda oriunda do mundo da *Internet*, cujo objetivo era a operação em qualquer plataforma e em qualquer parte do mundo. No mundo da *Internet* foi padronizado o *Internet Protocol* (IP), aumentando as aplicações como voz, dados e vídeo, bem como o desenvolvimento de equipamentos voltados para esse protocolo.

Esse contexto foi favorável para o surgimento da plataforma MPLS que possibilitou a união das redes de telecomunicações tradicionais, orientadas a conexão, e das redes *Internet*, pois é baseada em IP e provê QoS estritas e adequadas às aplicações, além de ser padronizada e permitir a interoperabilidade dos equipamentos.

3.1 Histórico da evolução das redes de telecomunicações

As redes de telecomunicações iniciaram-se para a transmissão de telefonia da origem até destino através de centrais de comutação telefônica intermediárias, que tinham a finalidade de fazer o encaminhamento do tráfego com a rota mais adequada até seu destino. Com o surgimento da telefonia digital, passou-se a utilizar o procedimento denominado comutação por circuitos, nos quais os circuitos são os canais telefônicos com taxa de 64 *kbits/s*.

A comunicação via comutação de circuitos envolve a fase de estabelecimento da chamada, transferência da informação e liberação da chamada, conforme explicado no item 2.3 nos serviços orientados a conexão. Na fase de estabelecimento da chamada, o usuário solicita à central telefônica uma conexão através de um protocolo de sinalização com a identificação do número do chamado. Caso haja disponibilidade de circuitos na rede para aquele endereço de destino, a solicitação é aceita e inicia-se a fase de transferência da informação. Durante essa fase, os recursos de rede são dedicados exclusivamente a essa comunicação, mesmo que o usuário não envie informação (nesse caso, mudo) e, após a transferência, os recursos são liberados.

Uma central telefônica típica com padrão E1 para agrupamento de canais de voz (padrão europeu utilizado no Brasil) é composta de troncos de entrada e saída em taxa de 2 *Mbits/s*,

que corresponde à agregação de trinta canais telefônicos de 64 *kbit/s* mais dois canais de controle e sinalização. Essa agregação é efetuada para facilitar a transmissão dos canais telefônicos até a central. Na entrada da central, os sinais de 2 *Mbits/s* são desmontados, os canais retirados e a informação é encaminhada para o tronco de saída relativo ao seu endereço destino. O núcleo da central telefônica é composto por uma matriz de comutação que faz a troca dos canais dos troncos de entrada para os troncos de saída.

A tecnologia de comutação de circuitos é bastante adequada para voz, mas também é utilizada para a transmissão de dados. Neste caso, utilizam-se múltiplos canais de 64 *kbits/s*, $N \times 64 \text{ kbits/s}$, para compor a taxa de sinal de dados mais adequada ao modem, por exemplo, 128 *kbits/s*, 256 *kbits/s* etc. Esses canais são transmitidos agrupados e comutados na central. Por serem dedicados, são também denominados de linhas privativas (LPs), pois o circuito permanece estabelecido mesmo se o usuário não envia informação, sendo, portanto, ineficiente para a transmissão de dados.

Em 1984 foi padronizada pelo *International Telecommunication Union Telecommunication Standardization Sector* (ITU-T) a rede digital de serviços integrados, ou *Integrated Services Digital Network* (ISDN), cujo objetivo era a formação de uma rede única, adequada ao transporte de voz, dados e vídeo. A ISDN é composta da telefonia digital e dos serviços de transporte de dados oferecidos por operadoras de telefonia regionais e envolve a digitalização da rede telefônica, que permite a transmissão de voz, dados, texto, gráficos, música, vídeo e outras fontes sobre os fios de telefone existentes.

A ISDN disponibiliza ao usuário três tipos de canais: o canal B de 64 *kbits/s* para a transmissão de voz ou dados, o canal D de 16 *kbits/s* ou 64 *kbits/s* para a transmissão da sinalização de chamadas e o canal H para transmissão de informação de alta velocidade. O acesso básico é composto de dois canais B e um canal D de 16 *kbits/s*, num total de 144 *kbits/s*. Entretanto, devido a adição de cabeçalhos de controle e *bits* de sincronismo, esta taxa é elevada para 192 *kbits/s*. A ISDN utiliza a tecnologia de comutação de circuitos tradicional e, para transmissão de dados, foi especificada, em 1988, na recomendação ITU-T I.122 a tecnologia de rede *Frame Relay*, que será melhor analisada adiante.

Com o surgimento de novos serviços de telecomunicações que incluem a transmissão de imagens simultaneamente aos sinais de voz e dados (*triple play*), foram encontradas restrições na capacidade de transmissão da rede telefônica convencional e redes *Frame Relay*. Esses serviços necessitam de elevadas capacidades de transmissão, com taxas superiores a 2 *Mbits/s*, impossibilitando o suporte na rede instalada. Assim, a ISDN evolui

para um novo conceito conhecido internacionalmente como rede digital de serviços integrados faixa larga, ou *Broadband ISDN* (BISDN), que se tornou viável pelo emprego das fibras ópticas. É neste contexto que surgiu a tecnologia ATM, que tem como umas das principais aplicações o transporte do IP, pois promove confiabilidade (QoS) durante a passagem pelo *backbone* [6].

Quando o ATM surgiu no início dos anos 90, pensou-se que uma nova era na área de redes estava surgindo, pois se pretendia que fosse empregado tanto no ambiente das redes locais (LAN) quanto no das redes de núcleo (WAN). Com sua característica de prover qualidade de serviço fim a fim na conexão que estabelece, era cogitado como o protocolo perfeito para o transporte multimídia. Seus idealizadores eram da área de telefonia, que provia comunicações da mais alta qualidade na época, e procuraram dotar essa técnica de transferência de informação com recursos que produzissem as mesmas características de qualidade da rede telefônica.

Contudo, o ATM não se tornou a solução fim a fim ideal. Os adaptadores de terminais ATM para computadores pessoais eram caros e as especificações de interconexão tardaram muito para serem emitidas, além de serem consideradas complexas. Sua aplicação principal passou a ser nos *backbones* dos provedores de serviço, operadoras de telefonia e nas redes corporativas de grandes empresas. Nas redes corporativas, contudo, o surgimento da tecnologia *Gigabit Ethernet*, que se estabeleceu naturalmente nas redes locais, afastou a utilização da tecnologia ATM.

A qualidade de serviço garantida pelo ATM é realizada através do estabelecimento manual de conexões virtuais permanentes denominadas *Permanent Virtual Connections* (PVCs), com roteadores interconectados. Dessa forma, é preciso o estabelecimento de uma rede de PVCs na qual todos os roteadores estejam conectados logicamente entre si. Sendo necessário um número muito grande de PVCs, observa-se um problema de escalabilidade para esse tipo de aplicação, pois as PVCs são configuradas manualmente e inviabiliza esse método para redes grandes.

Para resolver este problema, uma tecnologia híbrida começou a surgir para agregar as características boas do ATM ao IP e criar a tecnologia IP + ATM [6]. ao invés de IP/ATM. As vantagens dessa agregação são a gerência de uma rede única, ao invés da rede ATM e da rede IP, a resolução do problema de escalabilidade com a criação de mecanismos de estabelecimento automático de conexões e agregação de QoS ao IP tradicional. Essa tecnologia híbrida é denominada MPLS, considerada uma grande evolução que objetiva a

resolução de muitos desafios encontrados para a grande explosão da *Internet* e das comunicações de dados em alta velocidade em geral.

3.2 A tecnologia *Frame Relay*

O *Frame Relay* é um protocolo WAN de alto desempenho que opera na camada física e na camada de enlace de dados do modelo de referência OSI. Como resultado, a sobrecarga do protocolo é muito menor que outras tecnologias de encaminhamento de pacotes (*packet-switching*). Essa redução parte do pressuposto de que a camada física subjacente é relativamente livre de erros de transmissão, e que se os erros ocorrerem, os protocolos de camadas superiores ou os dispositivos do usuário final irão se recuperar de tais erros. Assim, o *Frame Relay* não fornece qualquer integridade dos dados e nenhum meio de controle de fluxo.

Nessa tecnologia também são provisionados circuitos virtuais permanentes (PVCs) para os usuários. Esses circuitos são canais lógicos entre o dispositivo de acesso *Frame Relay*, denominados *Frame Relay Access Devices* (FRAD), na rede. Dependendo do serviço oferecido, podem ser fornecidos circuitos virtuais comutados, os *Switched Virtual Circuits* (SVCs).

3.2.1 Identificador de conexão de enlace (DLCI)

O identificador de conexão de enlace, ou *Data-link Connection Identifier* (DLCI), identifica o circuito virtual *Frame Relay* pelos quais os quadros são roteados. Cada DLCI tem um caminho de comutação configurado permanentemente para um determinado destino. Dessa forma, em um sistema com muitos DLCIs configurados, é possível a comunicação simultânea com muitos lugares diferentes. A interface do usuário na rede, a *User-Network Interface* (UNI), fornece a demarcação entre o FRAD e a rede, sendo que a combinação entre a UNI e o DLCI especifica a extremidade de um circuito virtual, ou *virtual circuit* (VC), particular.

O DLCI tem um significado local e a numeração, armazenada em um campo específico no quadro, é normalmente decidida pelo usuário e atribuída pela operadora *Frame Relay*, sendo normalmente entre 1 e 1022.

Nos equipamentos comutadores *Frame Relay*, os quadros que chegam às portas de entrada são analisados e, de acordo com o valor do DLCI, são encaminhados a uma determinada porta de saída do comutador.

Nos comutadores há uma tabela de mapeamento portas/DLCIs de entrada para as portas/DLCIs de saída, encaminhando assim os quadros da origem ao seu destino. Esse caminho por fim percorrido é o chamado PVC. É denominado virtual uma vez que é lógico, ou seja, pode ser reconfigurado a qualquer momento sem a necessidade de alteração física nos equipamentos. A tabela de encaminhamento é configurada manualmente pelo operador da rede.

3.2.2 O quadro *Frame Relay*

A técnica utilizada pelo *Frame Relay* é a de encapsular em um quadro os dados que constituem a informação do usuário, adicionando informações de controle que compõem o cabeçalho e o sufixo, conforme ilustra a Figura 10. Essas informações são utilizadas pela rede para transportar os dados de um dispositivo a outro, realizando o encaminhamento, o controle de congestionamento e o controle de erros.

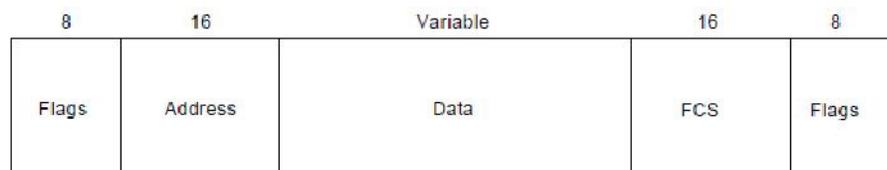


Figura 10 - Formato do quadro *Frame Relay* [1]

O cabeçalho é composto pelo *flag* (delimitador de quadro) de início de quadro e pelo *address* (endereço), que contém muitas outras informações. O sufixo é composto pelo *Frame Check Sequence* (FCS) e pelo *flag* de término de quadro. Os *flags*, portanto, delimitam o início e o final do quadro e têm sempre o mesmo valor, representado pelo hexadecimal 7E ou pelo binário 01111110.

O campo *data* tem tamanho variável (até 1.600 *bytes*) e contém a informação dos protocolos das camadas superiores e os dados do usuário. Este campo é transmitido transparentemente da origem ao destino e não é examinado por nenhum FRAD intermediário ou *switch Frame Relay*.

O campo FCS implementa uma verificação de redundância cíclica, a *cyclic redundancy check* (CRC), e fornece a detecção de erros para quadros com um comprimento de até 4096 *bytes*.

O campo *address* contém muitas informações, dentre a mais importante está o DLCI de 10 *bits*, utilizado para encaminhamento do quadro, conforme explicado anteriormente. As outras informações desse campo referem-se a informações de congestionamento da rede e se o quadro pode ou não ser descartado, em caso de congestionamento.

O *Frame Relay* pode operar com taxas desde 19.200 *bits/s* até velocidades nominais de circuitos E1 (2 *Mbits/s*) ou E3 (34 *Mbits/s*) e dota a rede de uma flexibilidade maior quando comparada as LPs (circuitos dedicados) citados no item 3.1. Por ser um processo de comutação executado quadro a quadro, se em determinado instante uma conexão não possuir quadros para serem enviados, quadros de outra conexão podem ser comutados e encaminhados naquele instante, otimizando a utilização dos recursos da rede. Para o usuário, tudo se passa como se o circuito fosse dedicado, mas o custo de um circuito *Frame Relay* é aproximadamente a metade do custo de uma linha privativa, devido a técnica de compartilhamento de banda.

3.3 Asynchronous Transfer Mode (ATM)

O ATM foi a tecnologia escolhida pelo ITU-T para a rede do tipo faixa larga (BISDN), mostrando-se eficiente para a transmissão de dados e sendo amplamente disseminada nas redes *backbones*, sendo frequentemente utilizada para transporte de outros protocolos de rede tais como IP (IP *over* ATM).

É um serviço orientado a conexão, na qual os dados transmitidos são organizados em células de comprimento fixo. Protocolos de camadas superiores e dados do usuário, como os pacotes IP, são segmentados em unidades de dados do protocolo, as *Protocol Data Units* (PDUs), de 48 *bytes*. Estas PDUs são precedidas por cabeçalho de 5 *bytes*, resultando em células de 53 *bytes* que são a entrada de *switches* ATM e podem ser multiplexadas juntas.

Cada cabeçalho da célula ATM contém um identificador de caminho virtual, o *Virtual Path Identifier* (VPI), e um identificador de canal virtual, o *Virtual Channel Identifier* (VCI), que juntos definem um circuito virtual ATM que as células devem utilizar como caminho até seu destino.

Na presença de tráfego no formato de células, estas são misturadas (agregadas), mapeadas na rede e inseridas na fibra óptica para serem transportadas para seu destino, como pode ser observado na Figura 11. Quando não há informação para ser transmitida, células vazias são inseridas para completar a taxa de transmissão de 155 Mbits /s ou 622 Mbits /s, de cada porta do comutador, uma vez que a rede é síncrona.

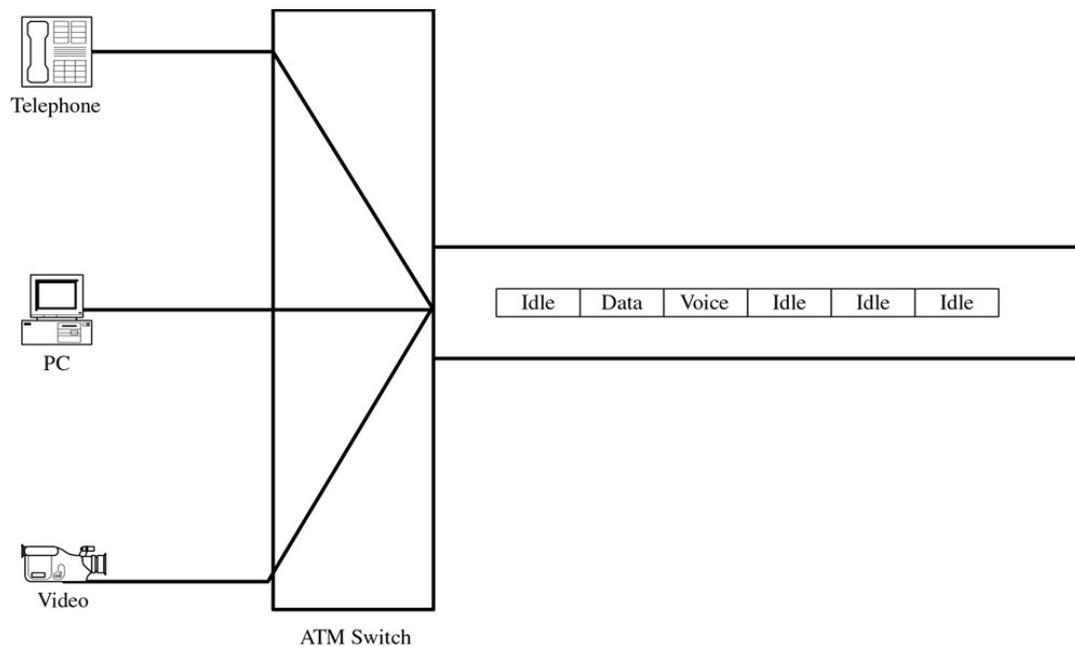


Figura 11 - Transmissão de células ATM vazias para completar a taxa de transmissão [4]

Portanto, pode-se auferir que a técnica ATM é uma combinação da Multiplexação por Divisão do Tempo, a *Time Division Multiplexing* (TDM), com células utilizando *slots* pré-atribuídos, e TDM estatístico, ou *Statistical TDM*, com células utilizando quantos *slots* estiverem disponíveis ou forem necessário para um determinado fluxo de tráfego.

Com os dois identificadores de conexão, VPI e VCI, inseridos no cabeçalho de cada célula, a fibra óptica (meio de transmissão) pode ser subdividida virtualmente em caminhos virtuais, os *Virtual Paths* (VPs), e canais virtuais, *Virtual Channels* (VCs), conforme ilustra a Figura 12. Um VP contém um grupo de VCs, sendo que cada VC contém um tipo de serviço, como por exemplo, um VC para voz, um para dados, um para vídeo etc.



Figura 12 - VCs concatenam-se para formarem VPs [1]

O encaminhamento das células é feito de acordo com os valores de VPI/VCI, conforme ilustra a Figura 13. A tabela indica que para cada valor de VPI/VCI de entrada, em determinada porta, há valores de VPI/VCI e portas de saídas. A tabela é configurada manualmente em todo comutador ATM e o resultado final de um encaminhamento fim a fim, passando por vários comutadores, corresponde a uma conexão ATM, denominada *Virtual Channel Connection (VCC)* ou *Virtual Path Connection (VPC)*.

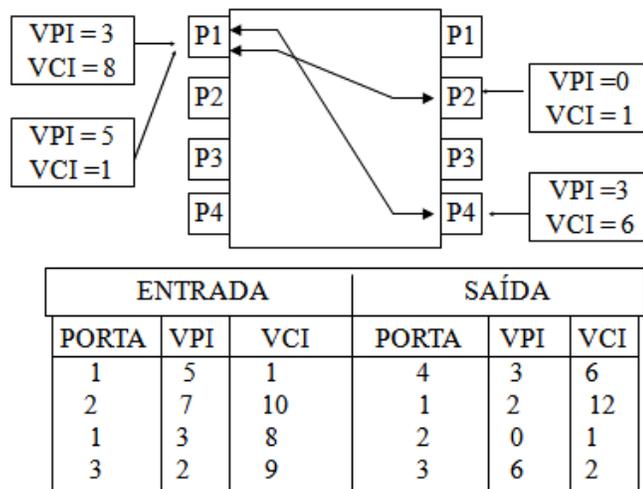


Figura 13 - Comutação ATM

Uma vez que esses identificadores têm apenas significado local em cada comutador e não correspondem a endereços fim a fim, a comutação ATM é muito rápida e implementada em hardware. Da mesma forma que as redes *Frame Relay*, as redes ATM utilizam multiplexação estatística para agregar o tráfego, permitindo o compartilhamento da banda por vários usuários de forma eficiente.

O tratamento de uma chamada que utiliza a rede ATM, de maneira sucinta, é uma declaração do usuário de valores de parâmetros de tráfego que descrevem o tipo da conexão requerida, além do desempenho desejado em relação a perdas de pacotes e atrasos. A rede utiliza essas informações para decidir se aceita ou rejeita a conexão. Se aceitar, a rede policia a conexão para garantir que os valores de tráfego declarados não

sejam ultrapassados. Esse processo previne congestionamentos na rede e garante que o desempenho requerido para cada conexão seja atingido.

3.3.1 Qualidade de serviço ATM (QoS ATM)

Gerenciamento de tráfego é a principal característica do ATM que o distingue de outros protocolos de rede e o torna adequado para implantação em redes de alta velocidade e para o fornecimento de garantias de desempenho em um ambiente integrado.

Os parâmetros de tráfego e desempenho, acordados pelo usuário e pela rede, formam um contrato de tráfego, no qual são especificados:

- Classe de QoS;
- Parâmetros de tráfego, tais como a taxa de pico de células, ou *Peak Cell Rate* (PCR), a taxa média de células, ou *Sustainable Cell Rate* (SCR), o tamanho máximo do surto de dados, ou *Maximum Burst Size* (MBS);
- Se a conexão é do tipo *Constant Bit Rate* (CBR), *Variable Bit Rate* (VBR), *Available Bit Rate* (ABR) ou *Unspecified Bit Rate* (UBR), dependendo da taxa de transmissão.

Cada classe de QoS contém uma combinação específica de limites para os parâmetros de desempenho, tais como: perda de células, atraso, células erradas etc. Estabelecida a conexão, o provedor da rede deve garantir o suporte da classe de QoS requerida e se a rede, em algum momento, não puder prover os objetivos acordados, deve desfazer a conexão, enviando uma mensagem apropriada ao usuário.

O serviço CBR provê uma taxa de *bit* sempre constante, sendo utilizado por aplicações que requerem uma alocação de banda fixa, continuamente disponível, enquanto durar a conexão.

O *real time* VBR (rt-VBR) é utilizado, como o CBR, por aplicações que requerem fortes restrições ao atraso, sendo apropriado para voz e vídeo com taxa variável (videoconferência). As fontes transmitem a uma taxa que varia com o tempo, sendo caracterizadas através da taxa de pico PCR, da taxa média SCR e do tamanho máximo do surto MBS, além da classe de QoS.

O serviço *non real time* VBR (nrt-VBR) é utilizado por aplicações com tráfego em surto e sem restrições ao atraso. O envio de um e-mail multimídia é um exemplo de serviço que

requer essa categoria. O nrt-VBR é também caracterizado pela taxa de pico PCR, taxa média SCR e do tamanho máximo do surto MBS e classe de QoS.

Na categoria ABR, a taxa de pico PCR e a taxa mínima SCR são os únicos parâmetros negociados durante o estabelecimento da conexão. A rede informa as condições de tráfego à fonte, através de realimentação por um mecanismo de controle de fluxo, permitindo que ela se adapte, aumentando ou diminuindo a taxa de pico PCR, conforme haja ou não banda disponível, minimizando a perda de células. Nenhum parâmetro de QoS é negociado.

Na categoria UBR, a banda disponível para as aplicações é a não utilizada pelas demais categorias de serviço. As células UBR são transmitidas quando há banda disponível (“melhor esforço”) e descartadas em caso de congestionamento. Não há um mecanismo de realimentação das condições de tráfego para o usuário e nenhum parâmetro de QoS é negociado. Essa categoria pode ser utilizada por aplicações que não possuem restrições de atraso e que fazem o próprio controle de erro e de fluxo. Transferência de arquivos e e-mail são exemplos dessas aplicações.

O ATM *Forum* e o ITU-T padronizaram também a categoria *Guaranteed Frame Rate* (GFR). A GFR também conhecida como UBR+, mantém a simplicidade do UBR, mas provê um serviço diferenciado ao usuário. Uma banda mínima é garantida para cada conexão e mecanismos de acesso “justo” à banda disponível são implementados.

3.4 A arquitetura da *Internet*

A *Internet* foi desenvolvida com investimentos da *Defense Advanced Research Projects Agency* (DARPA) em pesquisas da Universidade de Stanford e de Bolt, Beranek e Newman (BBN), quando ela se interessou no estabelecimento de uma rede de comutação de pacotes que facilitaria a comunicação entre sistemas de computadores diferentes em suas instituições de pesquisas, sendo primeiramente batizada de ARPANET, em meados de 1970.

Antes da ARPANET já existia outra rede que ligava esses departamentos e as bases militares, mas como os EUA estavam em plena guerra fria e toda a comunicação dessa rede passava por um computador central que se encontrava no Pentágono, ela era extremamente vulnerável.

Na década de 70, as universidades e outras instituições que faziam trabalhos relativos à defesa tiveram permissão para se conectarem a ARPANET, e em 1975 existiam aproximadamente cem sites. Os pesquisadores estudaram como o crescimento alterou o modo como as pessoas utilizavam a rede e presumiram que manter a velocidade alta seria o maior problema. Porém, na realidade a maior dificuldade se tornou a manutenção da comunicação entre os computadores (interoperabilidade).

No final de 1970 a ARPANET havia crescido de tal forma que seu protocolo de comunicação original, denominado *Network Control Protocol* (NCP), havia se tornado inadequado. Depois de algumas pesquisas, este foi alterado para o modelo TCP/IP, desenvolvido em UNIX. Sua maior vantagem era que permitia (o que parecia ser na época) o crescimento praticamente ilimitado da rede, além de possuir fácil implementação em uma grande diversidade de plataformas.

O IP pertence à camada de rede (*Layer 3*) e contém as informações de endereço e controle que permitem o roteamento dos pacotes. Juntamente com o TCP, da camada de transporte (*Layer 4*), o IP é o principal protocolo utilizado na *Internet*. O TCP/IP tornou-se a base de desenvolvimento da *Internet* e da *World Wide Web* (www).

Os padrões dessa arquitetura não são definidos pela ISO. As definições dos protocolos encontram-se nas *Request For Comments* (RFCs) distribuídas pelo IETF, que são publicadas e posteriormente analisadas pela comunidade da *Internet*, resultando em especificações semelhantes às recomendações do ITU-T [5].

3.4.1 Os protocolos IP e TCP

A arquitetura da *Internet* possui um conjunto de protocolos organizados em camadas, como ilustra a Figura 14 que compara com modelo OSI, sendo os dois principais o IP e o TCP. O IP é responsável pelo encaminhamento dos pacotes pela rede, entre as diversas sub-redes, desde a origem até o destino final, enquanto o TCP é responsável pela transmissão de forma segura no meio. Assim, o TCP/IP tem a tarefa do transporte fim a fim confiável das mensagens de dados entre os sistemas.

O IP é um protocolo não orientado a conexão (*connectionless*), sendo do tipo datagrama e não estabelecendo circuitos virtuais entre a origem e o destino, com as informações de transporte e controle utilizadas nas sub-redes contidas no próprio pacote. Já o conjunto TCP/IP forma um protocolo orientado a conexão, oferecendo um serviço de transporte de

alta confiabilidade para redes que exigem qualidade de serviço. Quando essa confiabilidade não é importante, utiliza-se como protocolo de transporte o *User Datagram Protocol* (UDP), que não é orientado a conexão.

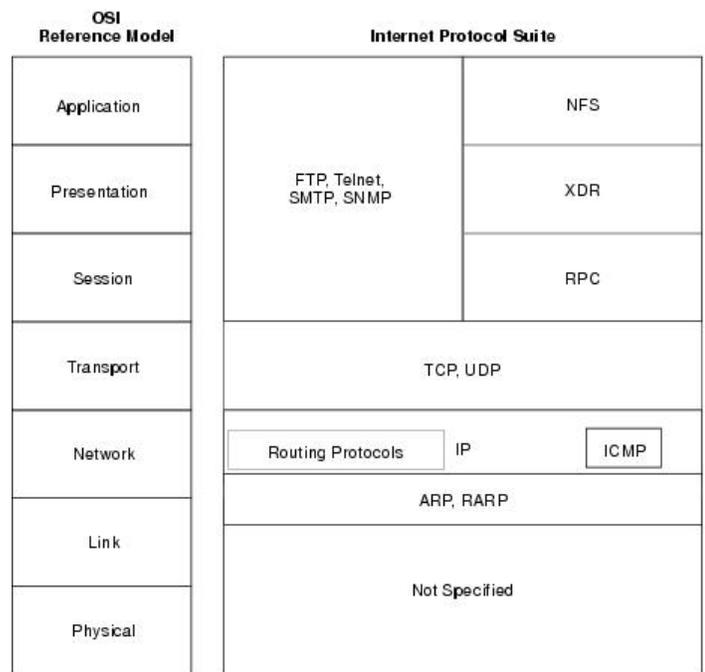


Figura 14 - Conjunto dos protocolos da *Internet* referenciados ao modelo OSI [1]

As portas são endereços associados às aplicações operantes em um sistema. Há endereços reservados, como os do correio eletrônico e transferência de arquivos, e outros endereços cujos valores são atribuídos de forma aleatória, de acordo com a gama dos que não foram utilizados, quando uma determinada aplicação solicita um endereço. Uma aplicação *Internet* é localizada por um *socket*, que corresponde ao conjunto do endereço IP do pacote com o endereço da porta [4].

Para executar as funções de envio e roteamento, as mensagens recebidas pelo TCP e UDP podem ser fragmentadas em tamanhos menores, de acordo com o tamanho máximo da sub-rede considerada. Os vários pacotes de uma mensagem podem seguir caminhos diferentes pelas sub-redes até o destino, caracterizando o modo de transporte do IP como de baixa confiabilidade, visto que os pacotes podem ser duplicados, perdidos e entregues fora de ordem sem que as camadas superiores sejam avisadas.

Assim, existem vários procedimentos que indicam a forma como os *gateways* devem tratar os pacotes, sinalizando quando uma mensagem de erro deve ser gerada e quando os pacotes devem ser descartados.

3.4.2 O endereçamento IP

O esquema de endereçamento IP é função integrante do processo de roteamento de datagramas IP através da rede. Cada endereço IP tem componentes específicos e segue um formato básico, sendo que esses endereços podem ser subdivididos e utilizados para criar endereços de sub-redes.

Cada nó em uma rede TCP/IP é atribuído a um único endereço lógico de 32 *bits* que é dividido em duas partes principais: o *network number* (número de rede) e o *host number* (número do nó), como mostrado na Figura 15. O número de rede identifica uma rede e deve ser atribuído pelo Centro de Informação da Rede Internet, ou *Internet Network Information Center* (InterNIC) se a rede for parte da *Internet*. Um provedor de serviços de *Internet*, ou *Internet Service Provider* (ISP), pode obter bloco de endereços de rede do InterNIC e atribuí-los conforme necessário. O número do nó identifica um nó em uma rede e é atribuído pelo administrador local da rede.

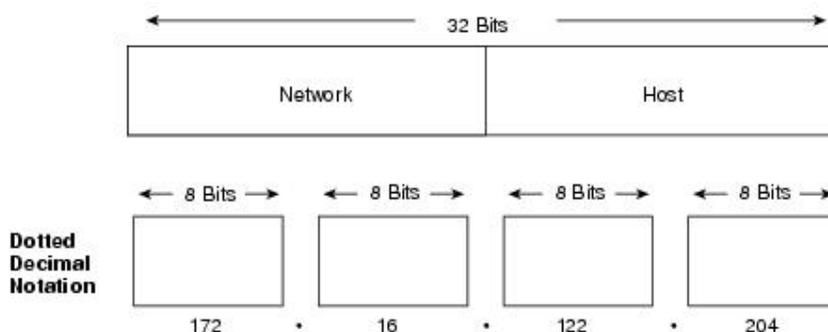


Figura 15 - Endereço IP composto de 32 *bits* separados em quatro octetos [4]

O endereço IP de 32 *bits* é ordenado em grupos de oito *bits*, formando um octeto, separados por pontos e representados em formato decimal. Cada *bit* no octeto tem um peso binário (128, 64, 32, 16, 8, 4, 2, 1) e o valor mínimo de um octeto é zero, enquanto o máximo é 255. A Figura 15 ilustra o formato básico de um endereço IP.

O endereçamento IP suporta cinco diferentes classes: A, B, C, D e E. Apenas as classes A, B e C são disponíveis para o uso comercial, cujas variações são:

- Classe A 1.0.0.0 a 127.255.255.255;
- Classe B 128.0.0.0 a 191.255.255.255;
- Classe C 192.0.0.0 a 223.255.255.255.

A divisão de endereços em classes tem a particularidade de facilitar a leitura do endereço das sub-redes pelos *gateways*. O simples teste dos *bits* iniciais identifica a classe utilizada, aumentando a eficiência no desempenho da rede.

Os endereços IP são colocados em um campo predeterminado no cabeçalho dos pacotes IP, ilustrado na Figura 16, quando transmitidos pela rede, indicando o endereço de origem (*source address*) e de destino (*destination address*).

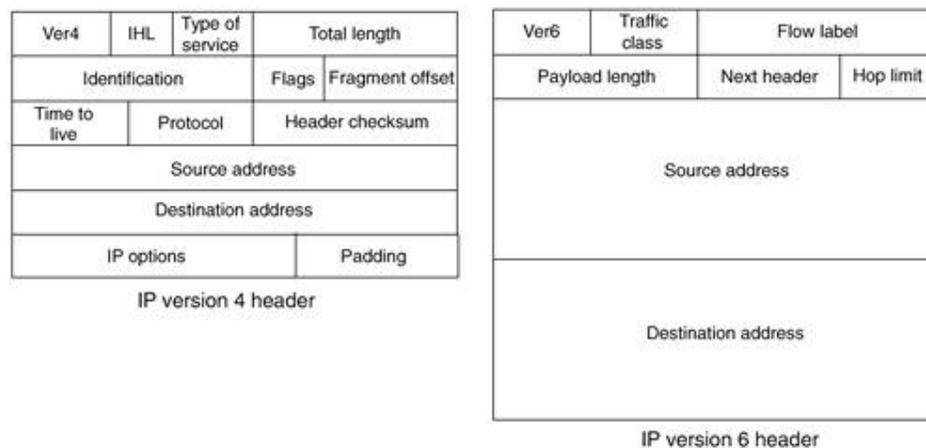


Figura 16 - Cabeçalhos IPv4 e IPv6 [1]

Além dos endereços de origem e destino, o cabeçalho IP também contém campos que indicam a versão do protocolo chamados Ver4 ou Ver6, como o IPv4 utilizado atualmente ou IPv6, especificado na RFC2460 de dezembro de 1998 e que irá substituir o IPv4 com a utilização de endereços de 128 *bits* [7]. Também indicam o nível de prioridade através do campo *Type of Service* (ToS), o comprimento total do pacote com o campo *Total length*, o controle de erros no cabeçalho com o campo *Header checksum* e o tempo de vida do pacote com o campo *Time to live*.

3.5 O modelo de sobreposição do IP e ATM

Com a proliferação do IP devido ao crescimento vertiginoso da *Internet*, o conceito de “IP sobre qualquer coisa” tomou precedência sobre o foco de forçar o ATM a se comportar como um protocolo de LAN legado. ATM sobre LAN, guiado pela LAN *emulation* (LANE), *Classical IP over ATM* e *Multiprotocol over ATM* (MPOA), registrou um crescimento limitado e foi ultrapassado pelas tecnologias *Fast Ethernet* (100 Mbps) e *Gigabit Ethernet* (1000 Mbps).

Porém, o ATM registrou um grande crescimento na utilização em redes WAN. QoS e garantias de classe de serviço oferecidas pelo ATM o conduziram para sua grande implantação nos provedores de serviço. A qualidade de serviço forneceu ao ATM a capacidade de multisserviço para oferecer classes de serviço separadas para voz, servidores e dados.

Os serviços *Frame Relay* são também oferecidos sobre os *backbones* ATM, utilizando a função ATM-to-Frame Relay Internetworking. Isso conduziu à extensiva implantação das redes virtuais privadas *Frame Relay*.

A relação entre o IP e ATM é fonte de muita controvérsia e debate. Ambas as tecnologias são amplamente implantadas, e cada uma tem a sua força. O IETF e o ATM *Forum* desenvolveram padrões para a integração do IP sobre ATM. O trabalho desses grupos foi focado principalmente em como as capacidades do ATM e do IP poderiam ser impulsionadas para fornecer uma solução, resultando na proliferação de redes IP sobrepostas na infraestrutura ATM.

O IP e o ATM são duas tecnologias completamente diferentes. O ATM é um serviço orientado a conexão e estabelece circuitos virtuais (PVC e SVC) antes de enviar o tráfego sobre um caminho predeterminado, utilizando células de comprimento fixo com a qualidade de serviço pré-determinada [8]. O ATM também tem seu próprio protocolo de roteamento na interface privada de rede a rede, ou *Private Network-to-Network Interface* (PNNI). PNNI é um protocolo hierárquico de estado de enlace no qual cada nó constrói uma visão topológica completa da rede e determina o melhor caminho, baseando-se em parâmetros de QoS inerentes ao ATM.

O IP, por outro lado, é uma tecnologia não orientada a conexão. Sua ampla aceitação é baseada em sua habilidade de utilizar qualquer mecanismo de Camada 2 e transporte físico.

Em cada nó (roteador) de uma rede IP realiza-se uma decisão independente para determinar o próximo salto do pacote até seu destino final.

O IP utiliza *Interior Gateway Protocols* (IGPs) para as decisões de roteamento dentro de uma rede ou dentro de um sistema autônomo de provedor de serviços de *Internet*. O *Open Shortest Path First* (OSPF) é um exemplo comum de IGP do tipo estado de enlace no qual cada roteador constrói tabelas com a topologia da rede e calcula o caminho mais curto para cada destino na rede, tipicamente utilizando o algoritmo de Dijkstra. Esses cálculos são armazenados em tabelas de roteamento utilizadas para determinar o próximo enlace do pacote, baseando-se em seu endereço de destino. O resultado é o mecanismo do tipo “melhor esforço” (*best effort mechanism*), que não tem nenhum conceito de QoS ou caminhos alternativos baseados em restrições.

Redes com roteamento IP são conectadas utilizando PVCs através de uma nuvem ATM. Isto cria um modelo de sobreposição que não é escalável ou gerenciável, como observado na topologia A da Figura 17, principalmente porque todos os roteadores IP na nuvem se tornam adjacentes. Este método também utiliza ineficientemente os recursos da rede, pois os *switches* ATM (Camada 2) são invisíveis no roteamento IP. Isso significa, por exemplo, que um PVC composto por muitos *switches* ATM de Camada 2 será utilizado pelo roteamento IP tão prontamente quanto um PVC com um único salto (um *switch* ATM), pois ambos os PVCs, de uma perspectiva de roteamento IP, são compostos por um único salto IP.

O modelo sobreposto requer que cada roteador tenha uma adjacência com todos os outros roteadores na rede. Devido às adjacências serem estabelecidas com circuitos virtuais ATM, a rede exige uma malha completa de circuitos virtuais para interconectarem os roteadores. Conforme o número de roteadores cresce, o número de circuitos virtuais completamente interconectados aumenta a uma taxa conforme (3.1), na qual n é o número de nós (roteadores),

$$\frac{n(n-1)}{2} \quad (3.1)$$

Como mostrado na topologia A da Figura 17, há oito roteadores, que conduzem a 28 circuitos virtuais a serem provisionados. O resultado em uma rede ATM com um grande número de VCs mostra o problema de escalabilidade, uma vez que o gerenciamento dos circuitos virtuais é uma árdua tarefa para os administradores da rede.

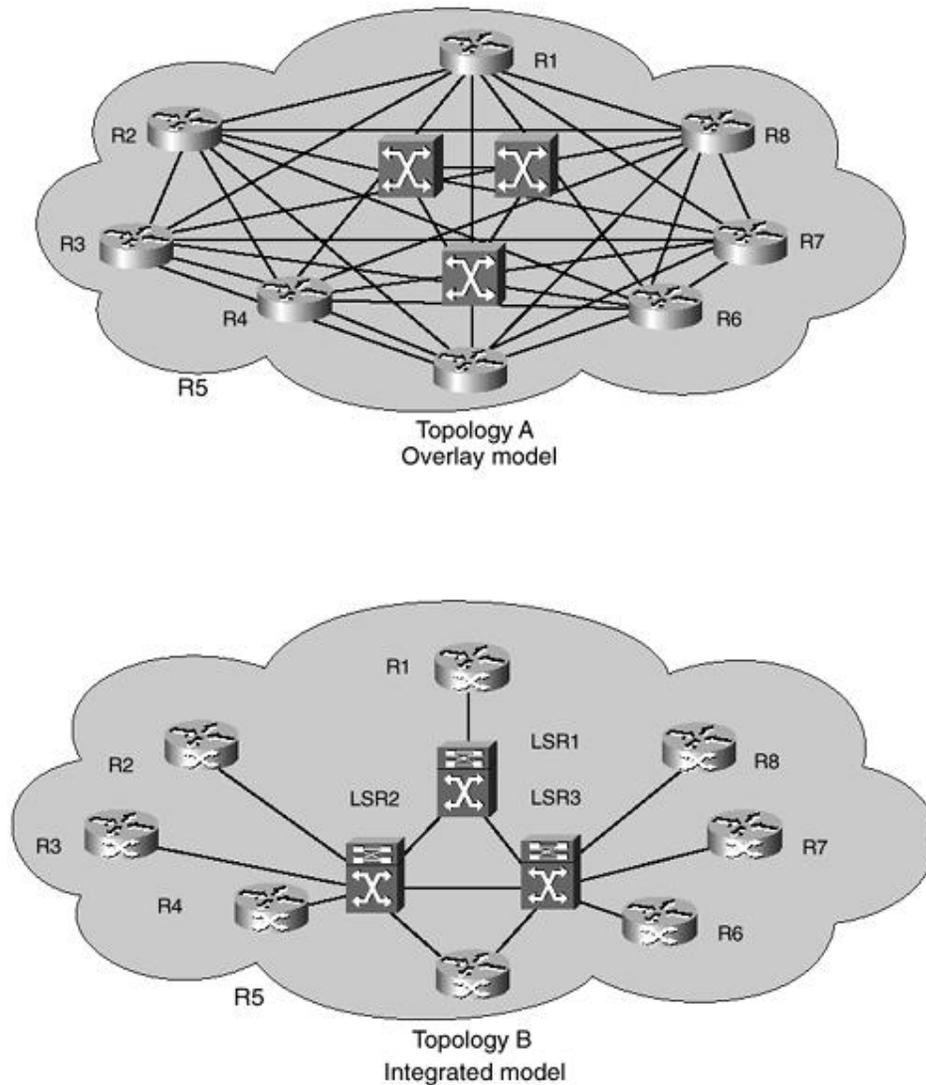


Figura 17 - Modelo sobreposto contra o modelo integrado [6]

Outro problema com as redes tradicionais resulta dos protocolos de roteamento, como o OSPF, que não desempenham bem em nuvens grandes e totalmente interligadas devido a duplicação da atualização do estado de enlace e ao grande número de adjacências que tem que manter. A oscilação de rota causada por falhas no circuito pode exceder o processamento do roteador e causar um comportamento de convergência de rota não determinística.

Na topologia A da Figura 17, o roteador R2 tem sete adjacências. A quantidade de informação de roteamento que é propagada nessa rede durante uma mudança na topologia devido à mudança de ligação ou nó pode ser maior que da ordem de n^4 , na qual n é o número de roteadores no núcleo. Conforme n aumenta, a quantidade de tráfego de

roteamento pode sobrecarregar os roteadores do núcleo, levando a comportamentos não determinísticos.

O MPLS resolve o problema do modelo sobreposto eliminando o conceito de nuvem ATM. Com o MPLS, tem-se um modelo integrando no qual os *switches* ATM são conhecedores do IP (*IP-aware*) e as ligações ATM são tratadas como ligações IP. Dessa forma, cada *switch* ATM pode se tornar um par de roteamento para o IP, como ilustrado na topologia B da Figura 17. O roteador R2 forma apenas três adjacências, com R3, R4 e LSR2. O máximo número de adjacências que qualquer roteador tem é amplamente reduzido e não mais aumenta com o tamanho da rede (número de roteadores), aumentando a escalabilidade no núcleo da rede.

Adicionalmente, esta integração de camadas resulta em um modelo de comutação e roteamento distribuídos que tira proveito das capacidades oferecidas por cada camada. A parte do roteador é necessária para fazer uso dos algoritmos de roteamento em Camada 3, para trocar informações de acessibilidade e calcular os caminhos. A parte MPLS é necessária para traduzir a informação de acessibilidade em um rótulo comutável que pode ser compreendido pelos *switches* (Camada 2), que possuem aptidões de *hardware* avançadas para comutar dados com máxima taxa de transmissão (*wire speed*).

3.6 A comutação de rótulos

A grande evolução nas redes de telecomunicações ocorreu com o surgimento da comutação de rótulos (*label switching*), caracterizada por dispositivos capazes de atribuir rótulos (*labels*) de comprimento fixo aos pacotes ou células. Entidades de comutação desempenham buscas em tabelas baseadas nesses rótulos simples para determinarem o destino dos dados, reduzindo significativamente o processamento nos nós pelos quais o pacote deve passar até seu destino final.

O rótulo resume a informação essencial sobre o encaminhamento do pacote ou célula. Esta informação inclui o destino, precedência, informação de QoS, rota de engenharia de tráfego e se o pacote pertence a algum grupo de rede virtual privada, ou *Virtual Private Network* (VPN).

Com a comutação de rótulos, a análise completa do cabeçalho de Camada 3 é desempenhada apenas uma vez: na entrada da rede (borda de entrada). Após isso, este cabeçalho é mapeado no rótulo que cada entidade de comutação de rótulos ou roteador na

rede utiliza para encaminhar as células ou pacotes até seus destinos. Na borda de saída do núcleo da rede, um comutador ou roteador de borda realiza a troca do rótulo pelo cabeçalho de Camada 3 associado.

Decisões de encaminhamento baseadas em algumas ou todas as diferentes fontes de informação podem ser feitas por meio de uma única tabela de pesquisa com rótulos. Por esta razão, a comutação de rótulos torna possível, para roteadores e *switches*, a tomada de decisões de encaminhamento baseada em múltiplos endereços de destinos.

A comutação de rótulos integra funções de roteamento e comutação, combinando informação de acessibilidade, fornecida pela função do roteador, com os benefícios da engenharia de tráfego, alcançados pela otimização das capacidades dos *switches*. A implementação nas redes WAN ocorre com a utilização do MPLS, atualmente muito difundido nas redes de núcleo devido à grande quantidade de benefícios.

Para compreender o benefício da integração de comutação e roteamento é necessário o entendimento do funcionamento do roteamento na camada de rede, bem como seus benefícios e problemas que fizeram com que a comutação de rótulos fosse a melhor solução no crescimento das redes de telecomunicações.

4 O roteamento na Camada de Rede

O roteamento na camada de rede é baseado na troca de informações de acessibilidade. Conforme um pacote percorre a rede, cada roteador extrai toda a informação relevante para o encaminhamento a partir de seu cabeçalho da Camada 3. Esta informação é então utilizada como um índice para uma procura na tabela de roteamento (*routing table*) e determinação do próximo nó de destino do pacote. Isto é repetido em cada roteador da rede até o destino final do pacote. Os protocolos de roteamento dinâmico ou a configuração estática constroem a base de dados necessária para analisar o endereço IP de destino (tabela de roteamento). Em cada nó, o próximo salto que o pacote deverá seguir é novamente definido, justificando a denominação do roteamento tradicional como *hop-by-hop*.

A informação nos pacotes, tal como a informação sobre qualidade de serviço, normalmente não é considerada para conseguir o máximo desempenho no encaminhamento. Tipicamente, apenas o endereço de destino ou prefixo é considerado. Contudo, a informação de QoS torna outros campos de um cabeçalho IP relevantes, como o de tipo de serviço, ou *Type of Service* (ToS), exigindo que uma análise complexa do cabeçalho deva ser realizada em cada roteador percorrido pelo pacote até a rede de destino.

A função de roteamento pode ser considerada duas componentes separadas:

- Componente de encaminhamento;
- Componente de controle.

4.1 Componente de encaminhamento

O componente de encaminhamento utiliza a informação mantida na tabela de roteamento do roteador e no cabeçalho da Camada 3 do pacote. Ele utiliza um conjunto de algoritmos, que

definem o tipo de informação extraída do cabeçalho e o procedimento que o roteador deverá seguir para encontrar uma entrada associada na tabela de roteamento. O roteador então encaminha o pacote baseado nesta informação. O encaminhamento do pacote pode ser como segue:

- Encaminhamento *unicast* ocorre quando o roteador utiliza o endereço de destino do cabeçalho da camada de rede e o algoritmo “de mais longa correspondência” (encontra o caminho mais próximo do destino final) sobre o endereço de destino para encontrar uma entrada associada na tabela de roteamento;
- Encaminhamento *unicast* com ToS ocorre quando o roteador utiliza o endereço de destino e o campo ToS do cabeçalho da Camada 3, além do algoritmo “de mais longa correspondência” sobre o endereço de destino, assim como a equivalência exata do valor de ToS para encontrar uma entrada associada na tabela de roteamento;
- Encaminhamento *multicast* ocorre quando o roteador utiliza os endereços de origem e destino do cabeçalho da Camada 3, bem como a interface de entrada que o pacote chega. O roteador utiliza o algoritmo de “de mais longa correspondência” sobre os endereços de origem e destino, assim como uma equivalência exata da interface de entrada para encontrar uma entrada associada na tabela de roteamento.

4.2 Componente de controle

O componente de controle é responsável pela construção e manutenção da tabela de roteamento. É implementado por protocolos de roteamento dinâmicos, como o *Open Shortest Path First (OSPF)* e *Border Gateway Protocol (BGP)*, que trocam informação de roteamento entre roteadores, assim como os algoritmos tipo o de Dijkstra utilizados para converter as tabelas de topologia em tabelas de roteamento.

4.2.1 Os algoritmos de roteamento

A informação sobre qual rota o pacote IP deve escolher para atingir seu destino é extraída de uma tabela de roteamento, que indica todas as possibilidades de rota. A geração dessas tabelas é, de forma geral, automática e utiliza um algoritmo que gera as informações baseadas em cálculos realizados sobre as métricas de um enlace. Há dois tipos principais de algoritmos de roteamento:

- Vetor de distância, que constrói a tabela com a menor distância conhecida para cada destino, indicando o caminho para atingi-lo; essa distância pode ser medida em número de saltos (*hops*) ou número de roteadores, para atingir a sub-rede. A atualização dessas tabelas é feita por troca de informação com os nós vizinhos;
- Estado de enlace (*link state*), que leva em consideração a banda disponível dos enlaces, além do número de saltos; pode considerar atraso, velocidade ou qualquer outra métrica que o administrador deseje utilizar;

No algoritmo de estado de enlace, cada roteador conhece a topologia completa da rede IP pela divulgação da informação dos *gateways* conectados entre si por enlaces. Um enlace existe entre dois *gateways* que estão conectados à mesma rede física.

São estabelecidas métricas denominadas “custos”, associadas aos enlaces, para proceder ao cálculo da menor distância (e conseqüentemente do menor custo) até o destino. O custo pode ser o número de saltos, atraso, banda ou um peso que o administrador arbitrariamente atribui. O algoritmo de Dijkstra é um exemplo de escolha do menor caminho, no qual são atribuídos pesos para cada aresta, como ilustra a Figura 18.

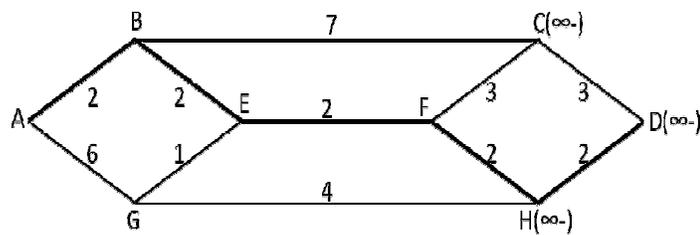


Figura 18 - Algoritmo de Dijkstra [3]

Supondo que os nós representam roteadores, o roteador A, para estabelecer uma rota até D, inicia o cômputo das métricas e elege os caminhos com os menores valores para atingir o destino. Assim, o caminho por B é escolhido, pois tem peso atribuído de 2, que pode ser representado por B(2,A) que indica a origem (A) e o peso do enlace (2). O nó seguinte, E, quando atingido por B, tem a entrada de E(4,B), sendo 4 a soma dos pesos AB+BE. Esse processo continua até o traçado final A, B, E, F, H e D, com as menores somas possíveis e indicando o caminho mais curto. Como regra geral, esse algoritmo é utilizado por vários protocolos que consideram outras métricas, além do número de saltos.

No início do funcionamento da *Internet*, o algoritmo de roteamento mais utilizado no seu núcleo era o de vetor de distância. A partir de 1988 passou-se a utilizar o algoritmo de estado de enlace devido ao seu crescimento e aos problemas que o tamanho da rede passou a apresentar para a convergência do roteamento.

4.2.2 Alguns protocolos de roteamento mais utilizados

Sistemas autônomos, ou *autonomous systems* (AS), são uma coleção de prefixos de roteamento IP sobre o controle de um ou mais operadores de rede que apresentam uma política de roteamento comum e claramente definida para a *Internet*. Dentro da rede *Internet* os roteadores são organizados hierarquicamente, sendo que aqueles utilizados para trocarem informações dentro dos sistemas autônomos utilizam os chamados *Interior Gateway Protocols* (IGPs) para alcançar seu propósito.

Um exemplo de IGP é o *Routing Information Protocol* (RIP), que emprega o algoritmo de vetor de distância Bellman-Ford. Ele utiliza o número de saltos como forma de métrica, sendo quinze o número máximo de saltos, e comunica-se com seus vizinhos a cada 30 segundos. Quando um deles não se re-anuncia por 180 segundos é removido de sua tabela de roteamento. Possui roteamento dinâmico, se ajustando automaticamente para manter a informação e encaminhamento para a topologia de tráfego ou de mudanças.

A vantagem do RIP é que possui um algoritmo de simples configuração, porém funciona apenas em redes de pequena escala, uma vez que sua métrica baseia-se no número de saltos. Além disso, se acontece uma falha de ligação entre os roteadores, seu algoritmo de roteamento converge muito lentamente.

A cada 30 segundos os roteadores enviam informações sobre suas rotas para todos os que estão ao seu lado. Ao receber a tabela, o roteador compara com a tabela já existente e somente faz a troca de algum destino se este ainda não constar na tabela ou se o número de saltos estiver inferior.

Outro exemplo de protocolo do tipo IGP é o *Open Shortest Path First* (OSPF), sendo um protocolo publicado em literatura aberta e não é proprietário de nenhum fabricante. Utiliza o algoritmo de Dijkstra e é do tipo estado de enlace, sendo que pode considerar outras formas para o caçulo do custo (métrica) além dos saltos. Com ele é possível realizar o balanço de carga em caminhos de igual custo, além de aprender rotas externas vindas de outros

sistemas autônomos. A distância máxima aceita é de 65.535 saltos, muito superior ao RIP, e também possui roteamento dinâmico.

Suas vantagens são que, frente ao RIP, permite que o administrador da rede atribua um custo (métrica) para cada rota da rede e pode ser utilizado em maior escala. Além disso, utiliza a largura de banda de cada rota para fazer o encaminhamento. Um problema é que, por ser mais complexo que o RIP e outros protocolos IGP, tende a exigir mais processamento quando em sua área existem muitos roteadores.

Para a atualização de sua tabela de roteamento, ele guarda as informações sobre o mesmo destino, ou seja, armazena rotas diferentes para o mesmo destino. Sempre que há uma atualização em um dos roteadores da rede, este encaminhada uma mensagem para os demais, caso contrário, a cada 30 minutos é enviada uma mensagem para avisar que está tudo sem problemas. Todos os roteadores guardam a tabela completa de roteamento dentro da rede, e por isso o alto processamento exigido.

Diferentemente do IGP, o *Border Gateway Protocol* (BGP) é utilizado para a troca de informações entre os sistemas autônomos, ou seja, entre sistemas que pertencem a domínios distintos e vizinhos entre si. No modelo TCP/IP, este protocolo pertence a camada de aplicação, e foi criado para substituir o *Exterior Gateway Protocol* (EGP) e utiliza o algoritmo de vetor de caminho (*path vector*), semelhante ao vetor de distância. As atualizações completas de roteamento são enviadas no início da sessão e as atualizações adicionais incrementais são enviadas em seguida. É um protocolo orientado a conexão e, portanto, é tido como confiável.

A utilização de atributos como métrica na escolha do melhor caminho permite-lhe ótima granularidade (capacidade de subdivisão da rede). Além disso, o uso de endereçamento hierárquico e a capacidade de manipular o fluxo de tráfego resultam numa rede projetada para crescer.

O BGP possui própria tabela de encaminhamento, apesar de ser capaz de compartilhar e pesquisar a tabela de roteamento IP interno. Também possui roteamento externo, utilizado para comunicação entre roteadores de um mesmo sistema independente.

Suas vantagens são que possibilita a troca de rotas, evita *loops* de roteamento em topologias arbitrárias e sua versão BGP4, foi projetada para suportar os problemas causados pelo grande crescimento da *Internet*.

Para a atualização de sua tabela, faz uso de algumas mensagens utilizadas para a comunicação entre roteadores BGP, nas quais cada uma tem importante função nas conexões:

- Mensagens tipo *open* são utilizadas para o estabelecimento de uma conexão BGP;
- Mensagens tipo *notification* reportam erros e servem para representar possíveis problemas nas conexões BGP;
- Mensagens tipo *update* são utilizadas para os anúncios propriamente ditos, incluindo rotas que devem ser incluídas na tabela e também rotas que devem ser removidos da tabela BGP;
- Mensagens tipo *keepalive* são utilizadas para manter a conexão entre roteadores BGP, caso não existam atualizações através de mensagens *update*.

Uma expressão utilizada para definir rotas que devem ser removidas da tabela BGP é *withdrawn*, do inglês e significa retirada.

4.3 Classes de equivalência

No roteamento, os pacotes são separados em classes de equivalência, *Forwarding Equivalency Classes* (FECs), que são conjuntos de pacotes de Camada 3 encaminhados da mesma maneira sobre o mesmo caminho e com o mesmo tratamento de encaminhamento. Enquanto atribui um pacote a uma FEC, o roteador deve verificar o cabeçalho IP e também alguma outra informação, tal como a interface na qual este pacote ingressou.

Alguns exemplos de FECs são:

- Prefixo do endereço de destino, e corresponde ao atual critério das redes IP (*destination based routing*);
- Um conjunto de pacotes *unicast* cujos endereços de destino equivalem a um prefixo endereço IP determinado com *bits* ToS parecidos;
- Um conjunto de pacotes *unicast* cujos endereços de destino equivalem a um prefixo de endereço IP determinado e têm o mesmo número da porta TCP de destino.

Um exemplo é apresentado na Figura 19, na qual o 200.15.45.9 e 200.15.45.126 estão na mesma FEC com um prefixo de endereço de 200.15.45.0/25 e porta de destino TCP 23.

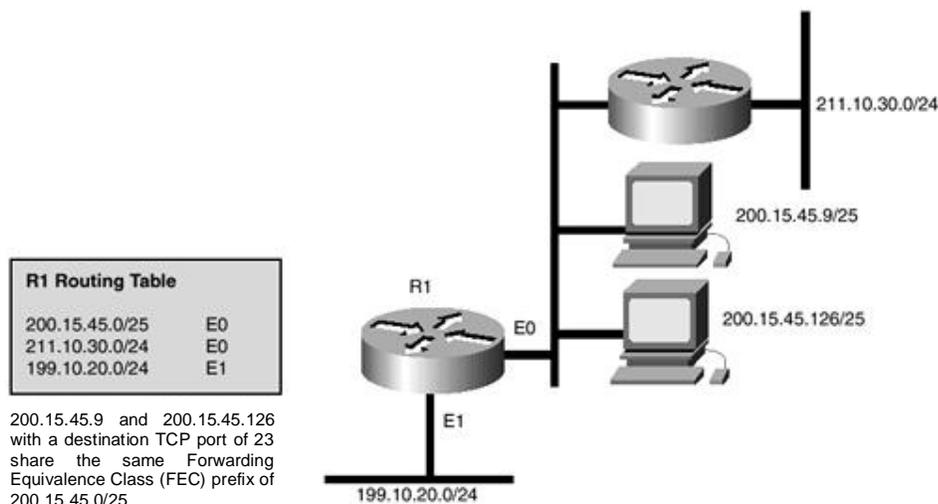


Figura 19 - Exemplo de classe de equivalência (FEC) [9]

Em cada nó da rede há uma classificação dos pacotes em classes de equivalência, que determinam o próximo nó, o que exige muito processamento do roteador. No MPLS, cada FEC é representada por um rótulo, cuja atribuição é feita no roteador de entrada enquanto os próximos nós apenas comutam os pacotes de uma interface à outra, baseado no rótulo. Isso significa que o núcleo (*core*) comuta os pacotes e a distribuição em classes de equivalências é feita apenas no roteador de entrada.

4.4 As restrições do roteamento tradicional

O tradicional encaminhamento de pacotes analisa o endereço IP de destino contido no cabeçalho da camada de rede de cada pacote conforme este viaja de sua origem até seu destino final. Um roteador analisa o endereço IP de destino independentemente em cada salto na rede, com a utilização de protocolos dinâmicos ou configuração estática para a construção da base de dados necessária (tabela de roteamento) para tal.

Embora bem sucedido, e obviamente amplamente implementado, existem algumas restrições nesse método de encaminhamento que diminuem sua flexibilidade. Novas técnicas são, portanto, desejadas para endereçar e expandir as funcionalidades de uma infraestrutura de rede baseada no IP. O MPLS fornece soluções para essas restrições, como será observado adiante.

4.4.1 O paradigma do roteamento na camada de rede

Conforme explicado no roteamento em Camada 3, todos os pacotes para um mesmo destino seguem o mesmo caminho pela rede, se não houver nenhum outro caminho de mesmo custo. Conforme explicado no item 4.2.1, por custo entende-se o número de saltos, atraso, banda ou um peso que o administrador arbitrariamente atribui ao caminho. Sempre que um roteador tem dois caminhos de mesmo custo para um destino, os pacotes rumo a esse destino seguem um deles ou ambos ao mesmo tempo, resultando em certo nível de balanceamento de carga.

Os roteadores são responsáveis pelo processo de decisão que seleciona o caminho a ser seguido pelos pacotes. Esses dispositivos Camada 3 participam na coleção e distribuição de informação de camada de rede, e desempenham a comutação com base no conteúdo dos cabeçalhos (de Camada 3) de cada pacote. É possível conectar os roteadores diretamente por ligações ponto a ponto ou redes locais (por exemplo, *hub* compartilhado), ou é possível conectá-los por *switches* LAN ou WAN (por exemplo, os *switches Frame Relay* ou ATM). Esses *switches* (LAN ou WAN) infelizmente não têm a capacidade de manter a informação de roteamento de Camada 3 ou de selecionar o caminho utilizado por um pacote através da análise de seu endereço de destino em Camada 3.

Deste modo, *switches* de Camada 2 (LAN ou WAN) não podem estar envolvidos no processo de decisão de encaminhamento em Camada 3. No caso de ambiente WAN, o projetista da rede tem que estabelecer caminhos virtuais em Camada 2 manualmente através da rede WAN. A partir destes pacotes de Camada 3 são encaminhados entre os roteadores que estão conectados fisicamente à rede de Camada 2.

O estabelecimento de caminhos WAN em Camada 2 é complexo. Estes são normalmente baseados em conexões ponto a ponto (circuitos virtuais na maioria das redes WAN) e são estabelecidos apenas mediante pedido e com configuração manual. Qualquer dispositivo de roteamento (roteador de entrada) na borda da rede de Camada 2 que queira encaminhar pacotes de Camada 3 para qualquer outro dispositivo de roteamento (roteador de saída), portanto, precisa estabelecer uma conexão direta na rede para o dispositivo de saída, ou enviar sua informação para um dispositivo diferente para transmissão até o destino final.

A rede ilustrada na Figura 20 é baseada em um núcleo ATM (Camada 2) rodeado por roteadores que desempenham o encaminhamento em camada de rede. Assumindo que as únicas conexões entre os roteadores são as apresentadas na figura, todos os pacotes enviado de São Paulo para, ou através, de Brasília devem ser enviados ao roteador de Curitiba, onde eles são analisados e enviados de volta sobre a mesma conexão ATM em

Curitiba para o roteador de Brasília. Este passo extra introduz atrasos na rede e sobrecarrega desnecessariamente o processamento do roteador de Curitiba, bem como a conexão ATM entre o roteador e o *switch* ATM adjacente em Curitiba.

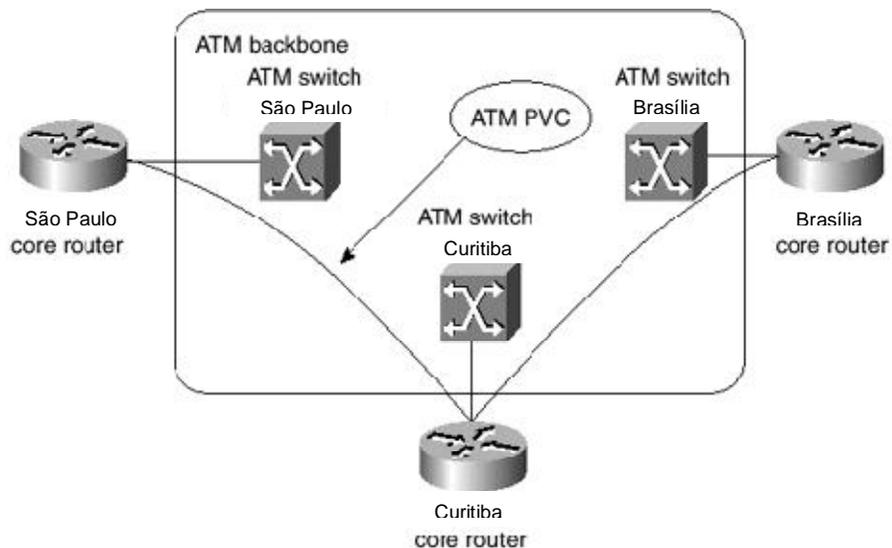


Figura 20 - Exemplo de rede IP baseada em núcleo ATM [9]

Para garantir o melhor encaminhamento de pacotes na rede, um circuito virtual ATM deve existir entre quaisquer dois roteadores conectados ao núcleo ATM. Embora isto possa ser fácil de ser atingido em redes pequenas, como as da figura, vários problemas de escalabilidade surgem em redes grandes, nas quais muitos roteadores se conectam ao mesmo núcleo WAN.

Os seguintes fatos ilustram os problemas de escalabilidade que podem ser encontrados [9]:

- Toda vez que um novo roteador é conectado ao núcleo WAN da rede, um circuito virtual deve ser estabelecido entre este roteador e quaisquer outros, para um roteamento ótimo;
- Com algumas configurações de protocolo de roteamento, todo roteador anexado ao núcleo WAN de Camada 2 (construído com *switches* ATM ou *Frame Relay*) precisa de um circuito virtual dedicado para qualquer outro roteador anexado ao mesmo núcleo. Para alcançar essa redundância, todo roteador também deve ter uma adjacência com cada roteador pertencente ao mesmo núcleo. Assim, se uma rede utiliza como protocolo de roteamento o OSPF, por exemplo, todos os roteadores transmitem cada mudança na topologia da rede para todos os outros roteadores

conectados ao mesmo *backbone* WAN, resultando em um tráfego proporcional ao quadrado do número de roteadores;

- O provisionamento dos circuitos virtuais entre os roteadores é complexo, pois é muito difícil prever a exata quantidade de tráfego entre quaisquer dois roteadores na rede. Para simplificar, alguns prestadores de serviço optam pela falta de garantia de serviço na rede, que nas conexões ATM é chamada taxa de bit não especificada, ou *Unspecified Bit Rate* (UBR);

A falta de troca de informação entre os roteadores e os *switches* WAN não foi um problema para os prestadores de serviços tradicionais que utilizavam *backbones* compostos apenas por roteadores, ou para aqueles que forneciam apenas serviços WAN (circuitos virtuais ATM ou *Frame Relay*). Porém, muitos outros fatores impulsionaram a utilização de *backbones* mistos:

- Prestadores de serviços tradicionais são convidados a oferecerem os serviços IP. Eles querem alavancar seus investimentos e basear estes novos serviços em sua infraestrutura WAN existente;
- Os prestadores de serviços foram solicitados a fornecer maiores garantias de qualidade de serviço, mais facilmente alcançadas com os *switches* ATM que com os tradicionais roteadores;
- As rápidas exigências em largura de banda antes da introdução de roteadores com interfaces ópticas forçaram algumas grandes operadoras a começarem a utilizar a tecnologia ATM devido às interfaces dos roteadores naquele tempo não fornecerem as velocidades oferecidas pelos *switches* ATM.

Assim, era evidente a necessidade de utilização de um mecanismo diferente para permitir a troca de informações da camada de rede entre os roteadores e os *switches* WAN, e tornar possível a participação dos *switches* nas decisões de encaminhamento dos pacotes de forma que a conexão direta entre os roteadores de borda não fosse mais necessária.

4.4.2 Serviços de pacote diferenciados

O método de destino de salto em salto (*hop-by-hop*), analisando apenas endereços de origem e destino da Camada 3, evita algumas abordagens inovadoras para o projeto da rede e otimização do fluxo de dados. Na Figura 21, por exemplo, a ligação direta entre os roteadores de núcleo de São Paulo e Brasília é responsável pela condução do tráfego que entra na rede por quaisquer roteadores de borda conectados em São Paulo, mesmo que

este *link* esteja congestionado e os *links* de São Paulo para Curitiba e de Curitiba para Brasília não.

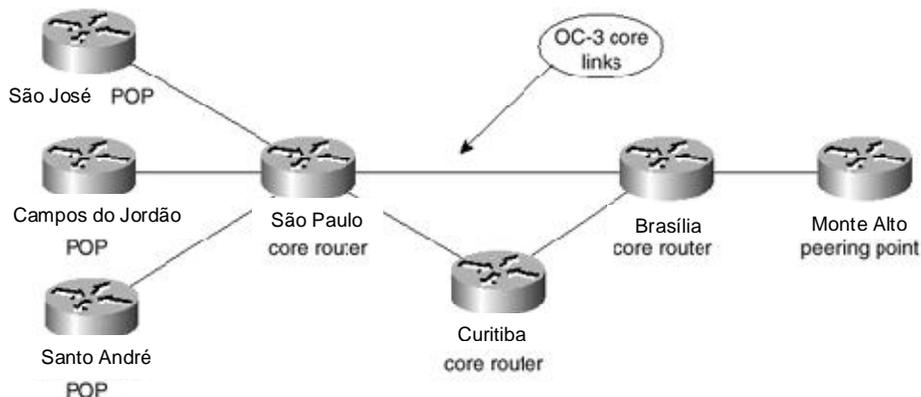


Figura 21 - Rede que se beneficiaria da engenharia de tráfego [9]

Embora algumas técnicas existam para interferir no processo de decisão, como o *Policy Based Routing* (PBR), nenhuma técnica escalável existe para decidir o caminho inteiro que o pacote deve seguir até seu destino final. Na rede da Figura 21, o PBR deveria ser implementado no roteador de núcleo de São Paulo para desviar um pouco do tráfego que chega para Brasília através do roteador de Curitiba. Porém, essa implantação em qualquer roteador de núcleo reduz consideravelmente o desempenho do roteador.

Uma vez que a maioria dos provedores de serviço implantam redes com caminhos redundantes, claramente existe uma necessidade para que o roteador de entrada seja capaz de decidir o encaminhamento de um pacote, o que afeta o caminho que o pacote segue através da rede, e aplique um rótulo (*label*) ao pacote, indicando aos outros dispositivos o caminho que este deveria seguir.

Essa necessidade também deveria permitir os pacotes que são destinados para a mesma rede IP seguirem caminhos diferentes dos determinados pelo protocolo de roteamento de Camada 3. Esta decisão também deveria ser baseada em fatores diferentes que o endereço IP de destino do pacote, como por exemplo, o nível de qualidade de serviço que o pacote necessita.

4.4.3 Controle e roteamento independentes

No roteamento IP tradicional, qualquer mudança na informação que controla o encaminhamento dos pacotes é comunicada a todos os dispositivos dentro do domínio de roteamento. Essas mudanças envolvem um período de convergência dentro do algoritmo de roteamento.

Um mecanismo que possa mudar como o pacote é encaminhado, sem afetar outros dispositivos dentro da rede, é certamente desejável. Para implementar tal mecanismo, os roteadores não podem se basear nas informações dos cabeçalhos IP para encaminhar o pacote, sendo necessário um rótulo adicional anexado ao pacote encaminhado para indicar seu comportamento desejado.

Com o encaminhamento dos pacotes sendo realizada com base em rótulos (*labels*) anexados aos pacotes IP originais, qualquer mudança dentro do processo de decisão pode ser comunicada para outros dispositivos através da distribuição de novos rótulos. Devido a esses dispositivos meramente encaminharem o tráfego baseados em rótulos anexos, mudanças podem ocorrer sem qualquer impacto em outros dispositivos que desempenham o encaminhamento de pacote.

4.4.4 Propagação de informações de roteamento externas

O encaminhamento convencional de pacotes dentro do núcleo de uma rede IP exige que informações externas de roteamento sejam anunciadas para todos os dispositivos de roteamento na rota. Isto é necessário para que os pacotes possam ser roteados baseados no endereço de destino contido dentro do cabeçalho da camada de rede do pacote. Continuando o exemplo da Figura 21, os roteadores de núcleo deveria ter que armazenar todas as rotas de *Internet* para poderem propagar os pacotes entre os roteadores de borda de entrada conectados em São Paulo até o roteador de borda de saída conectado em Brasília.

Este método tem grandes implicações na escalabilidade em termos de propagação de rota, utilização de memória e processamento dos roteadores de núcleo, e realmente não é uma função exigida se o desejado é transmitir um pacote de uma borda da rede a outra.

Assim, um mecanismo que permita dispositivos internos de roteamento comutarem os pacotes através da rede de um roteador de entrada para um roteador de saída sem analisar os endereços de destino da camada de rede é essencial.

5 Multiprotocol Label Switching (MPLS)

O MPLS é um multiprotocolo, uma vez que pode ser aplicável em qualquer protocolo da camada de rede, que com um método aperfeiçoado de encaminhamento de pacotes na rede, que utiliza a comutação de rótulos (*label switching*).

Os rótulos são valores de comprimento fixo, anexados aos pacotes ou células ATM, que resumem toda a informação necessária para o encaminhamento dos pacotes pela rede sem a necessidade de uma análise mais profunda dos dados transmitidos, como a análise do endereço de destino em cada nó. São inseridos entre o cabeçalho da Camada 3 e da Camada 2 no caso de tecnologias de roteamento baseadas em quadros (*frames*), ou estão contidos nos campos VPI/VCI no caso de tecnologias baseadas em comutação de células (*cells*), como o ATM. Ele combina a tecnologia de comutação da Camada 2 com a tecnologia de roteamento da Camada 3.

O primeiro objetivo do MPLS é criar uma estrutura de rede flexível que forneça aumento do desempenho e estabilidade. Isto inclui a capacidade de engenharia de tráfego e de realização de VPNs, que oferecem qualidade de serviço com múltiplas classes de serviço.

Em uma rede MPLS, como ilustrada na Figura 22, aos pacotes de entrada é atribuído um rótulo pelo roteador de borda de entrada, nomeado *Edge Label-Switching Router* (E-LSR). Os pacotes são encaminhados ao longo de um caminho determinado por um rótulo, denominado *Label-Switched Path* (LSP), onde roteadores capazes de analisar e comutar rótulos, denominados *Label-Switched Routers* (LSRs), tomam as decisões de encaminhamento baseados apenas no conteúdo do rótulo [10]. Em cada salto, o LSR analisa o rótulo, o remove e aplica um novo, com informações de como encaminhar o pacote para o próximo LSR e assim por diante. O rótulo é removido no E-LSR de saída, e a partir daí o pacote é encaminhado ao seu destino final pelo endereço IP de destino.

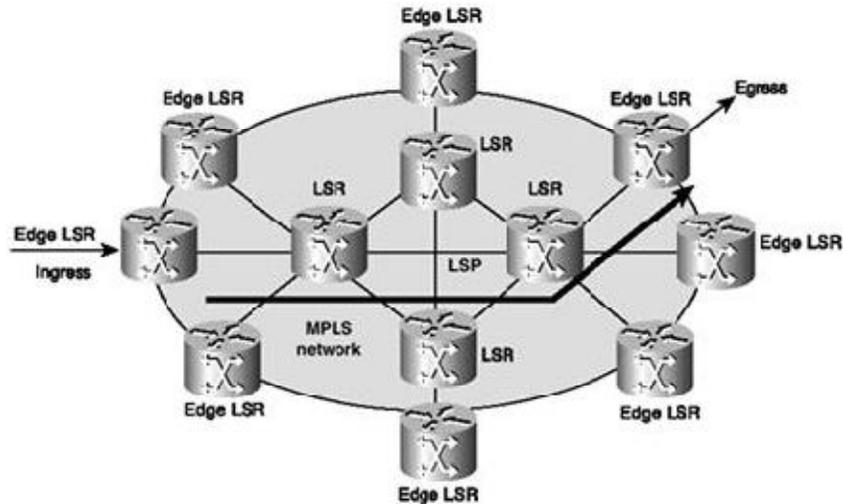


Figura 22 - Topologia de uma rede MPLS [9]

Sendo o MPLS baseado em troca de rótulos, assim como o ATM, essa tecnologia também pode ser aplicada em *switches* ATM através da implantação de componentes de controle. A informação necessária para a comutação dos rótulos pode ser transportada nos campos VPI/VCI do cabeçalho ATM. Esses valores são locais para uma determinada interface de entrada ou saída, ou seja, mesmos valores de VPI/VCI denotam diferentes circuitos virtuais quando utilizados interfaces diferentes. Os *switches* ATM utilizados com o MPLS recebem o nome de ATM LSR [9].

O MPLS integra o desempenho e a capacidade de gerenciamento de tráfego da Camada 2 com a escalabilidade e flexibilidade das redes de roteamento em Camada 3. Integrando o roteamento IP com a comutação ATM, oferece redes IP sobre ATM escaláveis, além da possibilidade de *switches* ATM ou roteadores existentes no núcleo da rede poderem encaminhar pacotes baseando-se apenas nos rótulos gerar mínima sobrecarga de pesquisa nos nós.

Está prática fornece uma alternativa escalável para sobrepor o modelo ATM e extinguir a necessidade do complexo endereçamento, roteamento e esquemas de sinalização do ATM.

O objetivo inicial da comutação baseada em rótulo era trazer a velocidade da comutação em Camada 2 para a Camada 3. Atualmente, este não é mais o principal benefício com o surgimento dos novos *switches* de Camada 3 (como o Cisco Catalyst 6500 proprietário da Cisco System, Inc.), que utilizam a tecnologia baseada em circuitos integrados de aplicação

específica, *application-specific integrated circuit* (ASIC), e desempenham pesquisas de rota em velocidade suficiente para suportar a maioria dos tipos de interface.

O amplo interesse na comutação de rótulos deu origem à formação do grupo de trabalho IETF MPLS em 1997. O MPLS evolui de muitas tecnologias anteriores, incluindo versões proprietárias (patenteadas) como *Tag Switching* da Cisco Systems, *Aggregate Route-Based IP Switching* (ARIS) da IBM, *Cell-Switched Router* (CRS) da Toshiba, *IP Switching* da Ipsilon e *IP Navigator* da Lucent.

Desde o princípio do *Tag Switching* a Cisco Systems tem trabalhado em conjunto com o IETF para desenvolver e ratificar o padrão MPLS, que incorporou a maioria das características e benefícios do *Tag Switching*.

5.1 Benefícios do MPLS

O método de comutação baseado em rótulo permite aos roteadores e *switches* ATM, que aceitam o MPLS, tomarem decisões de encaminhamento baseadas no conteúdo de um rótulo simples, ao invés de desempenharem uma complexa pesquisa da rota baseada no endereço IP de destino. Esta técnica traz muitos benefícios para as redes baseadas no IP, como:

- Redes virtuais privadas (VPNs). Utilizando o MPLS, provedores de serviço podem criar VPNs de Camada 3 através de sua rede *backbone* para múltiplos usuários, utilizando uma infraestrutura comum, sem a necessidade de criptografia;
- Engenharia de tráfego (TE). Ela fornece a habilidade de estabelecer explicitamente caminhos únicos ou múltiplos para fluxo de dados através da rede. Também fornece a habilidade de estabelecer características de desempenho para uma classe de tráfego. Essa característica otimiza a utilização de largura de banda para caminhos subutilizados;
- Qualidade de serviço. Utilizando a QoS do MPLS, prestadores de serviço podem fornecer múltiplas classes de serviço com garantias de qualidade de serviço para seus clientes de VPN;
- Integração do IP e ATM. A maioria das redes emprega um modelo de sobreposição no qual o ATM é utilizado na Camada 2 e o IP na Camada 3, como explicado anteriormente, que trazem grandes problemas de escalabilidade. Com o MPLS, é

possível a migração de muitas funções do ATM para a Camada 3, simplificando, assim, o provisionamento da rede, gerenciamento e a complexidade da rede.

Provedores de serviços têm percebido as vantagens do MPLS comparadas ao convencional IP sobre ATM. As grandes redes das grandes empresas que atualmente utilizam o ATM como infraestrutura de Camada 2 para o IP estão entre as beneficiárias dessa tecnologia.

Outra vantagem do MPLS é o fornecimento de redes virtuais privadas, com a utilização do IP, mais escaláveis e gerenciáveis, uma vez que uma rede MPLS consegue suportar muitas VPNs ao mesmo tempo. O serviço VPN é de grande importância por prover as empresas com redes IP privadas dentro de suas infraestruturas. Com um *backbone* MPLS, a informação da VPN pode ser processada apenas nos pontos de entrada e saída, com rótulos MPLS acompanhando os pacotes através do *backbone* compartilhado até seu correto ponto de saída.

Com o MPLS, é possível isolar as tabelas de roteamento da *Internet* dos núcleos da rede do provedor de serviço. Assim como os dados de uma VPN, o MPLS permite o acesso a tabela de roteamento apenas nos pontos de entrada e saída de uma rede fornecedora de serviço. Com o MPLS, pode-se atribuir um rótulo, associado com específicos pontos de saída, ao tráfego entrando pela borda do sistema autônomo do provedor. Como resultado, roteadores e *switches* internos precisam apenas processar a conectividade com os roteadores de borda dos provedores, blindando os dispositivos do núcleo do grande volume de roteamento trocado na *Internet*. Esta separação entre roteadores interiores das rotas de *Internet* também fornece melhor isolamento de erros, segurança e melhora a estabilidade.

O MPLS combina o desempenho e aptidões da comutação em Camada 2 (camada de enlace) com a comprovada escalabilidade do roteamento em Camada 3 (camada de rede), permitindo aos prestadores de serviços atender aos desafios do crescimento explosivo da utilização das redes enquanto fornecem a oportunidade de diferenciar os serviços sem sacrificar a infraestrutura de rede existente (legado). O MPLS é uma arquitetura flexível e pode ser implantada em qualquer combinação de tecnologias em Camada 2.

Uma característica importante também é a capacidade de engenharia de tráfego, ou *traffic engineering* (TE) necessárias para o uso eficiente dos recursos da rede. A engenharia de tráfego permite a troca da carga de tráfego de porções da rede superutilizadas para as subutilizadas, de acordo com o destino do tráfego, tipo, carga, hora do dia etc.

O suporte MPLS é oferecido para todos os protocolos de Camada 3, e o dimensionamento é possível além daquele tipicamente oferecido nas redes atuais. Ele eficientemente permite entregar serviços IP sobre uma rede comutada ATM, suportando também a criação de diferentes rotas entre a origem e o destino em um *backbone* Internet baseado puramente em roteadores. Através da incorporação do MPLS em suas arquiteturas de rede, muitos prestadores de serviços reduziram os custos, aumentaram renda e produtividade, forneceram serviços diferenciados (*DiffServ*) e adquiriram uma vantagem competitiva contra aqueles que não oferecem serviços MPLS, como as VPNs em Camada 3 (*Layer 3 VPNs*) ou engenharia de tráfego [11].

5.1.1 MPLS versus IP sobre ATM

Quando integrada aos *switches* ATM, a comutação de rótulos tira vantagem da comutação por *hardware* otimizada para se aproveitar do comprimento fixo das células ATM e comutá-las em alta velocidade. Para as rede de multisserviços, a comutação de rótulos permite aos *switches* fornecerem serviços ATM, *Frame Relay* e IP em uma plataforma única, fornecendo redução de custos operacionais e simplificação de provisionamento.

Para os provedores de serviços que utilizam *switches* ATM no núcleo de suas redes, a comutação de rótulos permite uma solução mais escalável e gerenciável que a sobreposição do IP sobre uma rede ATM. A comutação de rótulos evita problemas de escalabilidade da adjacência de muitos roteadores e fornece suporte para uma estrutura hierárquica dentro de uma rede de *Internet*.

Quando aplicado ao ATM, o MPLS integra as funcionalidade IP e ATM ao invés de sobrepor o IP no ATM. Isto torna a infraestrutura ATM visível ao roteamento IP e retira a necessidade de mapeamentos aproximados entre os recursos IP e ATM. O MPLS não precisa das técnicas de endereçamento e roteamento do ATM, como o PNNI, embora possa ser utilizado simultaneamente se necessário.

Em WANs com infraestrutura ATM, o MPLS é uma solução simples de integração de protocolos de roteamento com o ATM. O tradicional IP sobre ATM envolve o estabelecimento de uma rede de PVCs entre roteadores ao redor do núcleo ATM, porém, devido ao método no qual os circuitos virtuais são sobrepostos na rede ATM, a infraestrutura ATM se torna invisível aos roteadores. Assim, uma única falha em uma ligação ATM poderia fazer várias ligações entre roteadores falharem, criando problemas com grandes quantidades de atualizações de roteamento e processamento subsequente.

Sem o ajuste dos pesos da rota, todos os PVCs são vistos pelo roteamento IP como caminhos com de salto únicos e com o mesmo custo. Isso pode conduzir a um roteamento ineficiente na rede ATM.

5.1.2 MPLS *versus* Roteamento em Camada 3

Conforme os pacotes de Camada 3 são encaminhados de um roteador para outro, cada um toma uma decisão independente, analisando o endereço de destino de Camada 3 no cabeçalho do pacote e encaminhando o pacote para o próximo salto baseado em um algoritmo de roteamento. Sendo assim, cada roteador escolhe independentemente o próximo salto para o pacote, com base na análise do cabeçalho e no tipo de algoritmo de roteamento.

As decisões de encaminhamento são resultados de duas funções:

- Classificação dos pacotes de Camada 3 em classes de encaminhamento, FECs, baseada nos prefixo de endereço;
- Mapeamento das FECs para o próximo salto.

Todos os pacotes que pertencem a uma FEC particular com um mesmo nó de destino percorrem o mesmo caminho. Se o roteamento de múltiplo caminho é utilizado, os pacotes seguem um do conjunto de caminhos associados a FEC. Conforme o pacote percorre a rede, em cada nó (roteador) seu cabeçalho é reanalisado e ele é atribuído a uma FEC.

No MPLS, a atribuição de um pacote par um FEC determinada é feita apenas uma vez, quando o pacote entra na rede. A classe de equivalência para qual o pacote é atribuído é codificada em um valor de comprimento fixo denominado, o rótulo, que é anexado ao pacote e o acompanha pela rede. Nos saltos subsequentes, não há mais análise do cabeçalho de Camada 3, apenas o rótulo é utilizado como entrada em uma tabela que especifica o próximo salto. Não é o mesmo rótulo que acompanha o pacote durante o percurso pela rede, mas em cada nó há uma substituição do rótulo de acordo com o destino do próximo salto.

No paradigma do encaminhamento MPLS, tão breve o pacote é atribuído a uma FEC, não há mais análise do cabeçalho nos roteadores subsequentes. Todo o encaminhamento é guiado pelos rótulos.

O encaminhamento MPLS pode ser realizado por *switches* Camada 2, que podem realizar a pesquisa dos rótulos e suas substituições mesmo sem a capacidade de analisar os cabeçalhos em Camada 3 ou fazê-lo em velocidade adequada.

Os roteadores MPLS podem atribuir pacotes que chegam de diferentes portas em classes de equivalência diferentes. Isto forma a base de construção das redes virtuais privadas MPLS (MPLS VPNs). O roteamento tradicional, por outro lado, pode considerar apenas a informação que acompanha o cabeçalho de Camada 3 do pacote.

Um pacote que ingressa na rede por um determinado roteador pode ser rotulado diferentemente de um mesmo pacote que ingressa por outro roteador. Como resultado, pode ser realizar facilmente decisões de encaminhamento que dependem do roteador de entrada. Isto não pode ser realizando no roteamento tradicional, pois essa informação não acompanha o pacote.

A engenharia de tráfego obriga o pacote a seguir rotas particulares, de forma a balancear o tráfego sobre ligações pouco utilizadas. No MPLS, um rótulo pode ser utilizado para representar a rota de forma que a identidade de uma rota explícita não precisa acompanhar o pacote. No roteamento tradicional, isto requer que o pacote carregue uma codificação de rota consigo. Os roteadores convencionais analisam o cabeçalho do pacote não apenas para escolher seu próximo salto, mas também para determinar sua precedência ou classe de serviço. Eles podem, então, aplicar limiares de descarte diferentes aos pacotes.

5.2 Operação MPLS

As maiores diferenças entre MPLS e as tecnologias WAN tradicionais são a forma como os rótulos são atribuídos e a capacidade de carregar uma pilha de rótulos (*label stack*) anexada ao pacote [12]. O conceito de pilha de rótulos permite as aplicações como engenharia de tráfego e redes virtuais privadas.

A FEC na qual o pacote é atribuído é codificada em um valor de comprimento fixo chamado rótulo, sendo os pacotes rotulados antes de serem encaminhados. Nos nós subsequentes, não há mais análise do cabeçalho da camada de rede, o rótulo é utilizado como entrada em uma tabela que especifica o próximo salto e um novo rótulo. O rótulo antigo é substituído por um novo, e o pacote é encaminhado para o nó seguinte. Nas redes MPLS, os rótulos guiam todo o encaminhamento do pacote. Isto traz uma série de vantagens sobre o encaminhamento tradicional na camada de rede.

O encaminhamento MPLS pode ser feito por *switches*, que realizam a pesquisa na tabela e substituição do rótulo, mas não analisam os cabeçalhos da camada de rede. Os *switches* ATM realizam uma função similar através da comutação de células baseada nos valores de VPI/VCI encontrados no cabeçalho ATM. Se os valores VPI/VCI forem substituídos por valores de rótulos, os *switches* ATM podem encaminhar células baseadas nesses valores. Os *switches* ATM precisam ser controlados por um elemento de controle MPLS baseado em IP, como um *Label Switch Controller* (LSC), formando a base de integração do IP com o ATM utilizando MPLS.

Uma FEC é atribuída a um pacote quando ele entra na rede. O roteador de entrada pode utilizar qualquer informação que ele tenha sobre o pacote, assim como a interface ou porta de entrada, mesmo se a informação não possa ser obtida do cabeçalho da camada de rede. Um pacote que entra na rede por um roteador específico pode ser rotulado diferentemente do mesmo pacote entrando por outro roteador, sendo as decisões de encaminhamento que dependem do roteador de entrada facilmente realizáveis. Isto não pode ser feito com o encaminhamento convencional, pois a identidade do roteador de entrada de um pacote não o acompanha.

Redes com engenharia de tráfego forçam os pacotes a seguirem um caminho particular, como por exemplo um caminho pouco utilizado. Este caminho é explicitamente selecionado quando o pacote entra na rede ou mesmo antes, ao invés de ser selecionado pelo algoritmo de roteamento dinâmico normal conforme o pacote percorre a rede. No MPLS, um rótulo pode ser utilizado para representar uma rota, de forma que a identidade explícita da rota não precisa acompanhar o pacote.

5.2.1 Operação em modo quadro

O MPLS em modo de quadro (*frame-mode* MPLS) é assim chamado, pois os pacotes rotulados são trocados como quadros na Camada 2. Além disso, é realizado com dispositivos de comutação de Camada 3 (roteadores).

No IP, um determinado roteador irá considerar dois pacotes na mesma FEC se houver algum prefixo de endereço "X" em sua tabela de roteamento, tal que esse endereço seja o "de mais longa correspondência" para cada endereço de destino dos pacotes. Sendo assim, pode acontecer do pacote ser encaminhado para diversos nós intermediários até seu destino final.

A Figura 23 ilustra uma série de LSRs (de borda e de núcleo) interconectados, formando um caminho físico entre dois elementos, a *Station A* e a *Station B*.

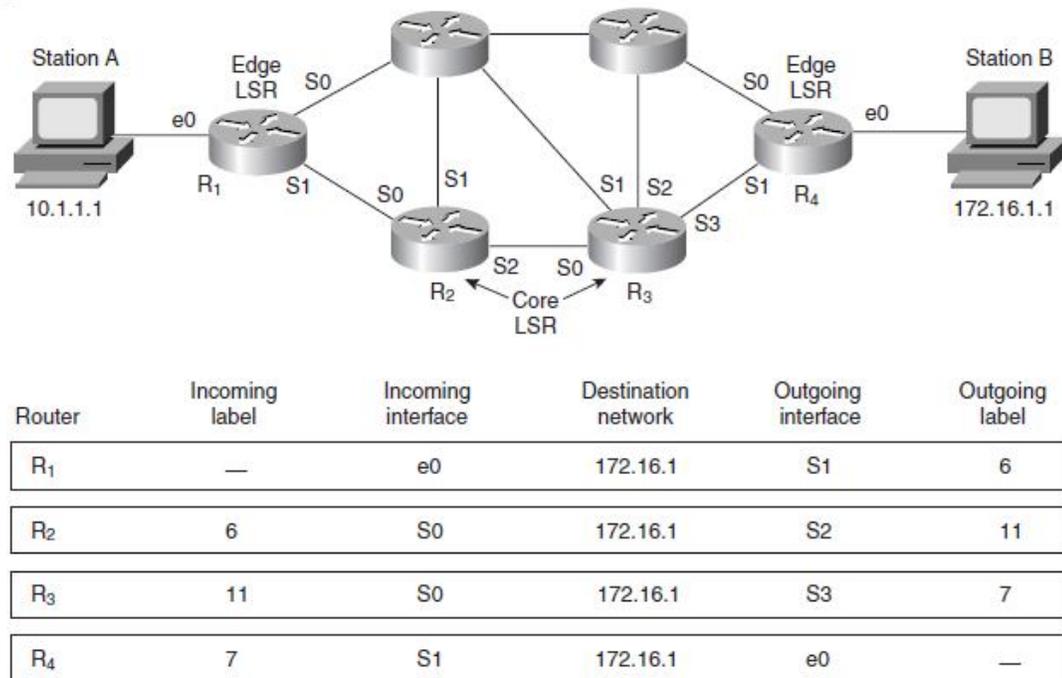


Figura 23 - Pacotes IP em uma rede MPLS [1]

O quadro gerado pela *Station A* segue o formato do padrão *Ethernet* com um cabeçalho normal de Camada 2 seguido por um cabeçalho de Camada 3. Uma vez que o endereço de destino encontra-se em uma rede diferente, a *Station A* deseja ter como alvo o cabeçalho de Camada 2 para seu *gateway* padrão. Neste caso, o *gateway* padrão também serve como E-LSR de entrada. O LSR de entrada referencia sua tabela interna e determina que o pacote precisa ser encaminhado pela interface S₁ para o próximo LSR.

Além disso, o LSR de entrada deve inserir um rótulo entre os cabeçalhos da Camada 2 e da Camada 3 para indicar o caminho pelo qual o quadro deve percorrer até a *Station B*. O R₂ observa que o quadro recebido pela interface S₀ e determina que há um rótulo incorporado. Portanto, o roteador trata o quadro de acordo com sua tabela de comutação, que informa que o quadro deve sair pela interface S₁ e substitui o rótulo por um novo valor. Todos os LSR subsequentes lidam com os pacotes de maneira similar até que alcance o LSR de saída, que retira toda a informação do rótulo.

A operação de comutação de rótulos é realizada da mesma maneira indiferentemente se o pacote rotulado possui apenas um ou uma pilha de rótulos. Em ambos os casos, o LSR age apenas no topo da pilha, ignorando outros rótulos abaixo. Esta função permite uma variedade de aplicações MPLS nas quais roteadores de borda podem concordar nas regras de classificação e rótulos associados sem o conhecimento dos roteadores de núcleo [12].

O *Penultimate Hop Popping (PHP)* consiste em retirar o rótulo do pacote antes que este chegue ao roteador E-LRS de saída, uma vez que o rótulo não tem mais utilidade e o encaminhamento será realizado pelo endereço IP de destino (Camada 3).

Sem essa prática, o E-LRS de saída deve inspecionar o rótulo na pilha de rótulos e realizar a pesquisa, que irá mostrar que o rótulo deve ser retirado e o restante do pacote IP deve ser inspecionado. Assim, uma pesquisa na Camada 3 deve ser realizada no pacote IP para encaminhá-lo ao destino final. Essa pesquisa dupla em tabelas reduz o desempenho no nó.

Com o PHP, o E-LRS pode pedir uma operação de retirada de rótulo para seu *upstream LSR*, que enviam um pacote IP pro para o E-LRS de saída que realiza apenas a pesquisa na Camada 3 e encaminha o pacote para seu destino final.

5.2.2 Operação em modo célula

Sendo o MPLS baseado em troca de rótulos, assim como o ATM, essa tecnologia pode ser aplicada em *switches* ATM através da implantação de componentes de controle. A informação necessária para a comutação dos rótulos pode ser transportada no campo VCI da célula ATM. Se dois níveis de rotulamento forem necessários, então o campo VPI também pode ser utilizado, embora o tamanho do cabeçalho VPI limite o tamanho das redes nas quais é utilizado. Porém, a utilização apenas do VCI é adequada para a maioria das aplicações. Esses valores são locais para uma determinada interface de entrada ou saída: mesmos valores para VPI/VCI denotam diferentes VCs quando utilizados em outras interfaces.

Está prática fornece uma alternativa escalável para superpor o modelo ATM e extinguir a necessidade do complexo endereçamento, roteamento e esquemas de sinalização do ATM. Essa implementação utiliza alguns termos específicos do ATM, como:

- *Label Switching Controlled ATM interface (LC-ATM interface)* é a interface ATM controlada por comutação de rótulo. Trata-se de uma interface em um roteador ou *switch* ATM na qual o valor VPI/VCI é atribuído através dos protocolos de controle MPLS (LDP);
- ATM-LSR é um *switch* ATM que executa protocolos no plano de controle e o encaminhamento MPLS entre as interfaces LC-ATM no plano de dados pelos meios tradicionais de comutação da célula ATM;
- LSR baseado em quadros (*Frame-based LSR*) é um LSR que encaminha quadros completos entre suas interfaces, como o tradicional roteador;
- Domínio ATM-LSR é um conjunto de ATM-LSRs interconectados pelas interfaces LC-ATM;
- ATM-LSR de borda é um LSR baseado em quadros com pelo menos uma interface LC-ATM.

O ATM-LSR de borda de entrada recebe um pacote rotulado ou não, realiza uma busca em sua base de informações de rótulos e encontra o valor VPI/VCI de saída, utilizado como rótulo de saída. O pacote rotulado é segmentado em células ATM e enviado para o próximo ATM-LSR, sendo valor VPI/VCI obtido colocado no cabeçalho ATM de cada célula ATM.

A partir desse momento, até que o pacote rotulado deixei o domínio ATM-LSR, a pesquisa de rótulo é feita puramente baseada nos valores VPI/VCI e não mais no cabeçalho de rótulo MPLS, que ainda está presente no pacote rotulado devido a necessidade de manutenção de alguns campos para utilização posterior.

Os ATM-LSRs comutam células baseadas no valor VPI/VCI do cabeçalho da célula ATM, com o mesmo mecanismo da comutação tradicional de células ATM, e os mecanismos de alocação e distribuição do MPLS são responsáveis por estabelecerem o mapeamentos corretos de VPI/VCI de entrada/saída.

O ATM-LSR de borda de saída remonta as células novamente no formato de um pacote rotulado, realiza a pesquisa de rótulo na base de dados e encaminha o pacote para o próximo LSR. A pesquisa de rótulo na base de dados é baseada nos valores VPI/VCI das células de entrada, não na pilha de rótulos do cabeçalho MPLS. Isto porque os ATM-LSRs alteraram apenas os valores VPI/VCI, não os rótulos dentro das células ATM.

Na distribuição dos rótulos sob demanda acontece um problema na reutilização de rótulos iguais para destinos iguais chamado mescla de circuito virtual (VC merge). Se um ATM-LSR

tiver recebido um rótulo para um determinado destino de seu vizinho *downstream* (aquele que envia o rótulo), este deveria ser reutilizado caso outro *upstream* LSR solicitasse um rótulo para o mesmo destino. Porém, se essa otimização for feita, o roteador que enviou o rótulo teria problemas quando as células ATM chegassem simultaneamente de ambos os roteadores que solicitaram rótulos e receberam o mesmo valor, as células ATM terão o mesmo valor de VPI/VCI, conteúdos diferentes, e se intercalarão. A Figura 24 ilustra esse problema.

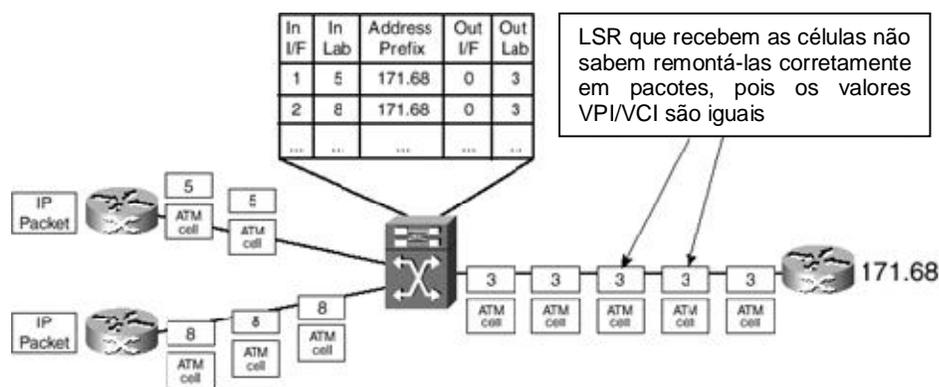


Figura 24 - Problema de intercalação de células no domínio ATM-LSR [9]

Dessa forma, o ATM-LSR deverá pedir um novo rótulo todas as vezes que precisar encaminhar células, mesmo que para um mesmo destino. Com rótulos diferentes, os campos VPI/VCI serão diferentes e o LSR consegue remontar as células novamente em pacotes, como ilustra a Figura 25.

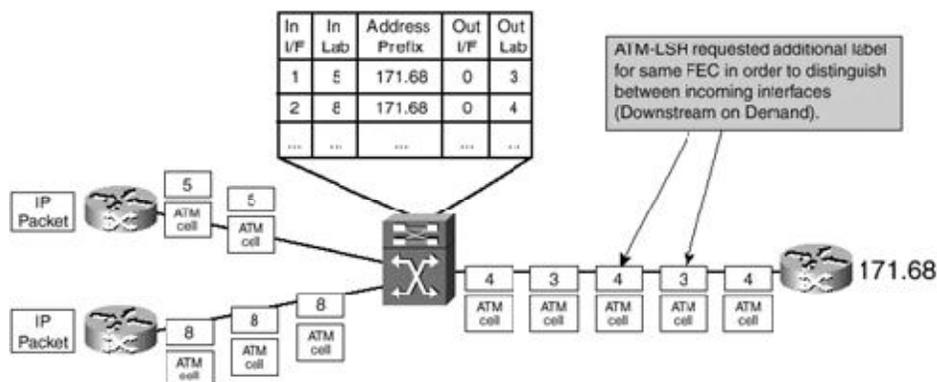


Figura 25 - Rótulos diferentes para células com um mesmo destino [9]

Para utilizar a otimização do mesmo rótulo, modificações no *hardware* são realizadas de forma que o *switch* ATM armazene (*buffer*) as células ATM de entrada até que todo o pacote seja recebido (isto pode ser determinado através da análise do indicador AAL5 de final de quadro) [13].

5.3 Arquitetura do nó MPLS

Os nós MPLS têm dois planos de arquitetura: o plano de encaminhamento MPLS (*Forwarding plane*) e o plano de controle MPLS (*Control plane*). Nós MPLS podem desempenhar roteamento Camada 3 ou comutação Camada 2 somados à comutação de pacotes rotulados. A Figura 26 ilustra a arquitetura básica de um nó MPLS.

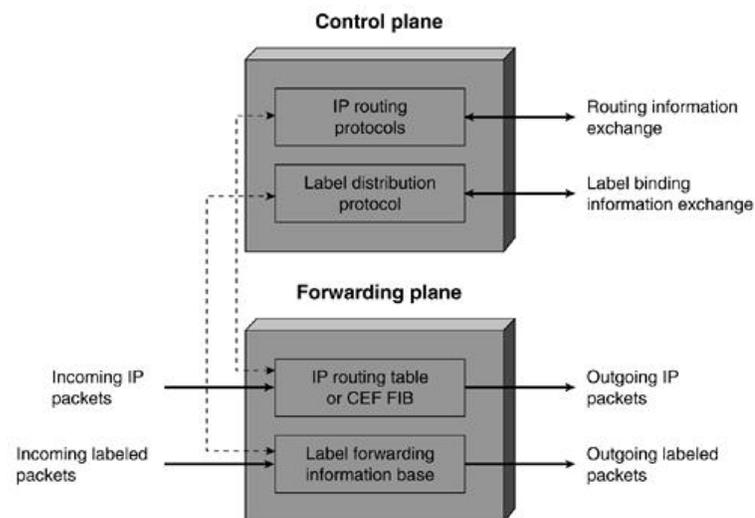


Figura 26 - Arquitetura de um nó MPLS [6]

5.3.1 Plano de encaminhamento

O plano de encaminhamento MPLS é responsável por encaminhar os pacotes baseados nos valores dos rótulos, utilizando uma base de informação de rótulos de encaminhamento, a *label forwarding information base* (LFIB), mantida pelo nó MPLS. O algoritmo utilizado pelo componente de encaminhamento utiliza informação contida nessa tabela para e a informação contida no valor do rótulo. Cada nó MPLS mantém duas tabelas relevantes para o encaminhamento MPLS: a base de informações de rótulo, ou *label information base* (LIB) e a LFIB. A LIB contém todos os rótulos atribuídos pelo nó local MPLS e o mapeamento desses rótulos para rótulos recebidos de seus vizinhos MPLS. A LFIB utiliza um subconjunto de rótulos contidos na LIB para o encaminhamento local dos pacotes.

O rótulo MPLS é um identificador de tamanho fixo igual a 32 *bits*, utilizado para identificar uma FEC e normalmente tem significado local. O rótulo representa a classe de equivalência para qual o pacote é atribuído. No caso do ATM, o rótulo é alocado no campo VCI ou VPI do cabeçalho ATM e, se o quadro é do tipo *Frame Relay*, o rótulo ocupa o campo DLCI do cabeçalho.

As tecnologias de Camada 2 como a *Ethernet*, *Token Ring*, *FDDI* e ligações ponto a ponto não podem utilizar seus campos de endereço em Camada 2 para transportar os rótulos. Estas tecnologias carregam os rótulos em cabeçalhos chamados *shim headers*. Esse cabeçalho é inserido entre a camada de enlace e a camada de rede (Camada 2 e Camada 3), como ilustra a Figura 27. O uso do rótulo *shim header* permite o suporte MPLS na maioria das tecnologias de Camada 2.

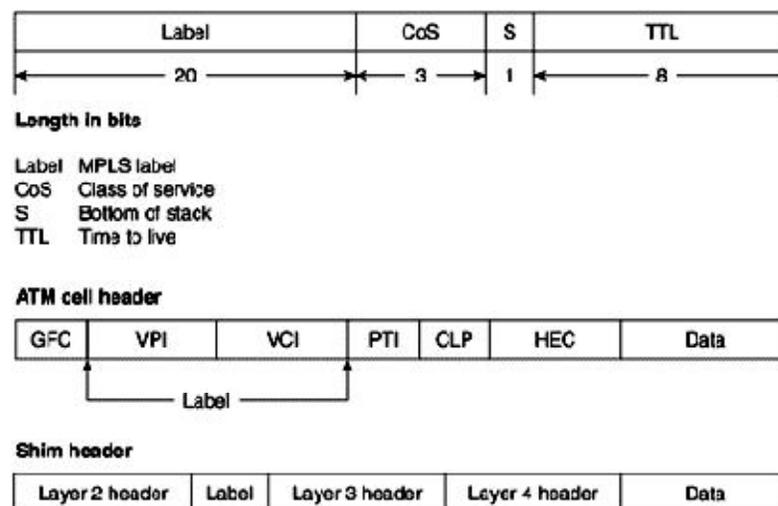


Figura 27 - Formatos de rótulo MPLS [6]

O suporte aos *shim headers* exige que o roteador remetente tenha uma forma de indicar ao roteador destinatário que o quadro contém esse tipo de cabeçalho. Isto é facilitado diferentemente em várias tecnologias. A *Ethernet* utiliza os valores *ethertype* (campo de um pacote Ethernet que define qual protocolo de camada superior o pacote transporta) 0x8847 e 0x8848 para indicar a presença de um *shim header*. Alguns valores de rótulos são reservados para determinadas funções, como mostra a Tabela 1

Tabela 1 - Valores MPLS reservados

Rótulo	Descrição
0	Rótulo nulo explícito no IPv4. Este valor de rótulo é válido apenas no final da pilha de rótulos. Ele indica que a pilha de rótulos deve ser removida e o encaminhamento do pacote deve ser baseado no cabeçalho IPv4.
1	Rótulo de alerta para o roteador. É análogo ao uso da opção “alerta de roteador” nos pacotes IP. Este valor é válido em qualquer lugar da pilha de rótulos, exceto no final.
2	Rótulo nulo explícito no IPv6. Este valor é válido apenas no final da pilha de rótulos. Isto indica que a pilha de rótulos deve ser removida e o encaminhamento do pacote deve ser feito baseado no cabeçalho IPv6.
3	Rótulo nulo implícito. Este é um rótulo que um nó MPLS pode atribuir e distribuir, mas que na verdade nunca aparece no encapsulamento. Ele é usado no <i>penultimate hop popping</i> .
4 – 15	Reservado para uso futuro.

O rótulo MPLS, observado na Figura 27, contém os seguintes campos:

- *Label*, ou rótulo, possui 20 *bits* e carrega o valor atual do rótulo MPLS;
- *CoS* possui 3 *bits* e afeta o enfileiramento e algoritmos de descarte aplicados ao pacote conforme este percorre a rede. Define a classe de serviço a que um pacote pertence, indicando a prioridade do pacote;
- *S*, chamado de *bottom of stack*, ou pilha, possui 1 *bit* e suporta a pilha de rótulos hierárquicos. Esse campo tem valor binário, sendo definido para 1 na última entrada da pilha de rótulo, permitindo múltiplos rótulos, utilizado em aplicações VPN;
- *TTL*, chamado *Time to Live*, ou tempo de sobrevivência, possui 8 bits e tem o papel de contar por quantos roteadores o pacote passou, sendo que se esse número foi maior que 255 o pacote é descartado para evitar *loops*.

O *bit* do campo *S*, conforme a Figura 27, implementa a pilha de rótulos em que mais de um cabeçalho de rótulo pode ser anexado a um único pacote IP. Ele é definido para 1 para indicar o final da pilha de rótulos. Todos os outros bits da pilha são definidos para zero. No MPLS baseado em pacotes, o topo da pilha aparece logo após o cabeçalho da camada de enlace, e o final da pilha antes do cabeçalho da camada de rede. O encaminhamento de pacotes é realizado utilizando os valores de rótulo do topo da pilha, e esta função é muito importante para VPNs MPLS e engenharia de tráfego [12].

A LFIB mantida pelo nó MPLS consiste de uma sequencia de entradas. Como ilustrado na Figura 28, cada entrada é formada de um rótulo de entrada e uma ou mais subentradas. Ela é indexada pelo valor contido no rótulo de entrada.

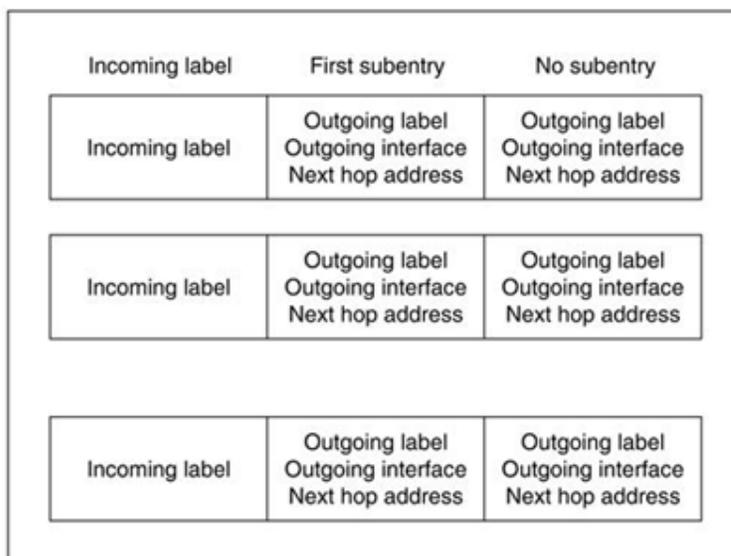


Figura 28 - Label Forwarding Information Base (LFIB) [6]

Cada subentrada consiste de um rótulo de saída (*outgoing label*), interface de saída (*outgoing interface*) e endereço do próximo salto (*next hot address*). Subentradas dentro de uma entrada individual podem ter rótulos de saída diferentes ou iguais. O encaminhamento *multicast* exige subentradas com múltiplos rótulos de saída, pois um pacote de entrada chegando por uma interface precisa ser enviado para múltiplas interfaces de saída. Além dessas informações, o pacote pode ter em sua tabela de encaminhamento informação relacionada aos recursos do pacote, assim como uma fila de saída (*outgoing queue*) que o pacote deve ser colocado.

Um nó MPLS pode manter uma única tabela de encaminhamento, uma tabela de encaminhamento para cada uma de suas interfaces, ou uma combinação das duas. No caso de múltiplas tabelas de encaminhamento, são considerados o valor do rótulo de entrada e a interface de entrada na qual o pacote foi recebido.

A comutação de rótulos utiliza um algoritmo de encaminhamento baseado na troca de rótulos (*label swapping*). Nós MPLS que mantém uma única LFIB extraem os valores dos rótulos dos pacotes recebidos e os utilizam como entrada na LFIB. Feita a correspondência, o nó MPLS substitui o rótulo de entrada no pacote pelo valor do rótulo de saída encontrado

na subentrada correspondente. Assim, envia o pacote pela interface especificada pela subentrada e para o próximo salto, também especificado pela subentrada da LFIB. Se a subentrada especifica uma fila de saída, o nó MPLS coloca o pacote na fila.

Se o nó MPLS mantém múltiplas LFIBs para cada interface, ele utiliza a interface física na qual o pacote foi recebido e seleciona uma LFIB particular, que será usada para encaminhar o pacote conforme procedimento anterior.

Um nó MPLS pode obter toda a informação que ele necessita para encaminhar um pacote assim como determinar a reserva de recursos necessária pelo pacote através de um único acesso a memória. Esta pesquisa em alta velocidade e a habilidade de encaminhamento faz a comutação de rótulo uma tecnologia de comutação de alta velocidade. MPLS pode também ser utilizado para transportar outros protocolos Camada 3, tornando-o atrativo para a migração da tecnologia IPv4 para IPv6.

5.3.2 Plano de controle

O plano de controle MPLS é responsável por preencher e manter a LFIB. Todo nó MPLS deve processar um protocolo de roteamento IP para trocar informações de roteamento IP com outros nós MPLS na rede. Nós MPLS com ATM habilitado utilizariam um LSC externo, como um roteador de grande porte, para participar do processo de roteamento IP.

Protocolos de estado de enlace, como o OSPF, são os escolhidos, pois fornecem em todos os nós MPLS uma visão inteira da rede. Nos roteadores convencionais, a tabela de roteamento IP é utilizada para construir a base de informação de encaminhamento, a *Forwarding Information Base* (FIB). Porém, no MPLS, as tabelas de roteamento IP fornecem informação sobre a rede de destino e prefixos de sub-rede utilizados para o rótulo associado.

A informação do rótulo associado pode ser distribuída utilizando o protocolo de distribuição de rótulos, o *Label Distribution Protocol* (LDP). Os protocolos de estado de enlace espalham a informação de roteamento entre um conjunto de roteadores que não são necessariamente adjacentes, visto que a informação dos rótulos associados é distribuída apenas entre roteadores adjacentes. Isto torna os protocolos de roteamento de estado de enlace inadequados para distribuir a informação dos rótulos. Porém, extensões dos protocolos de roteamento, como BGP, podem ser utilizadas.

Isto torna a distribuição da informação de associação de rótulos consistente com a distribuição da informação de roteamento e evita uma condição rara em que um nó MPLS recebe a informação sobre o rótulo, mas não tem a informação de roteamento apropriada. Também simplifica a operação total do sistema, pois previne a necessidade para um protocolo separado, como o LDP, para distribuir a informação de associação de rótulos.

Os rótulos trocados com nós MPLS adjacentes são utilizados para construir a LFIB. O MPLS utiliza um paradigma de encaminhamento baseado na troca de rótulo que pode ser combinado com uma série de módulos de controle diferentes. Cada módulo de controle é responsável pela atribuição e distribuição de um conjunto de rótulos, assim como a manutenção de outras informações importantes de controle. IGPs são utilizados para definir o alcance, ligação e mapeamento entre FEC e endereços do próximo salto.

5.4 Elementos MPLS

Os elementos MPLS que ajudam no entendimento da interação com os vários protocolos Camada 2 e Camada 3 são:

- Roteadores de comutação de rótulo, ou *Label-Switched Router* (LSR);
- Caminho de rótulo comutado, *Label-Switched Path* (LSP);
- Protocolo de distribuição de rótulo, ou *Label Distribution Protocol* (LDP).

5.4.1 *Label-Switched Router* (LSR)

O LSR é um dispositivo que implementa o controle MPLS e encaminhamento de componentes. Ele encaminha um pacote baseado no valor do rótulo encapsulado e também encaminha pacotes Camada 3 nativos.

Os LSRs são roteadores ou *switches* ATM com MPLS habilitado que utilizam rótulos para encaminhar o tráfego. Os *switches* ATM que permitem o uso do MPLS (ATM MPLS LSR) são construídos pela integração de um *software* MPLS ou adicionando-se uma funcionalidade MPLS através de um LSC externo. Um passo fundamental na comutação de rótulos é que os LSRs concordem com o rótulo que eles deveriam utilizar para encaminhar o tráfego. Eles chegam a esse entendimento utilizando o protocolo de distribuição de rótulos.

LSRs de borda são localizados nos limites do ponto de presença, ou *point of presence* (POP), de uma rede MPLS e aplicam rótulos (ou uma pilha de rótulos) nos pacotes. A imposição de rótulos também é chamada de *push label*. Os LSRs de borda também desempenham a retirada do rótulo na saída do domínio MPLS, ação referida como *pop label*, e a função de roteamento IP convencional. As várias ações que podem ser desempenhadas com os pacotes rotulados por um LSR são:

- *Aggregate* (agregar), remove o rótulo do topo da pilha e desempenha uma pesquisa em Camada 3;
- *Pop* (remoção), remove o rótulo do topo da pilha e transmite o *payload* restante como um pacote ainda rotulado (se S estiver definido como 1) ou como um pacote IP tradicional;
- *Push* (empurrar), substitui o rótulo do topo da pilha por um conjunto de rótulos;
- *Swap* (trocar), substitui o rótulo do topo da pilha por outro;
- *Untag*, remove o rótulo do topo da pilha e encaminha o pacote IP para o IP especificado do próximo salto.

5.4.2 *Label-Switched Path (LSP)*

O LSP é uma conexão configurada entre dois LSRs na qual as técnicas de comutação de rótulo são utilizadas para o encaminhamento de pacotes. Um LSP é um caminho de tráfego específico na rede MPLS, sendo provisionado utilizando, por exemplo, o LDP.

Pode ser considerado um caminho sobre um conjunto de LSRs em nos quais pacotes pertencentes a uma determinada FEC percorrem para alcançar seu destino.

O MPLS permite uma hierarquia de rótulos conhecidos como pilha de rótulos. É, portanto, possível ter diferentes caminhos em diferentes níveis de rótulos para um pacote alcançar seu destino. LSPs são unidirecionais, significando que um pacote poderia tomar um caminho diferente para seu retorno à origem.

Na Figura 29, LSR1 e LSR6 são LSRs de borda, e LSR2, LSR3, LSR4 e LSR5 são LSRs de núcleo. Para o propósito do encaminhamento com rótulo, LSR1 e LSR6 são pares no nível de *gateway* de borda, e LSR2, LSR3, LSR4 e LSR5 são pares no nível de *gateway* interior. Esta ilustração mostra dois LSPs: um fim a fim de nível 1 do LSR1 para o LSR6 e um nível 2 através de LSR4 e LSR5.

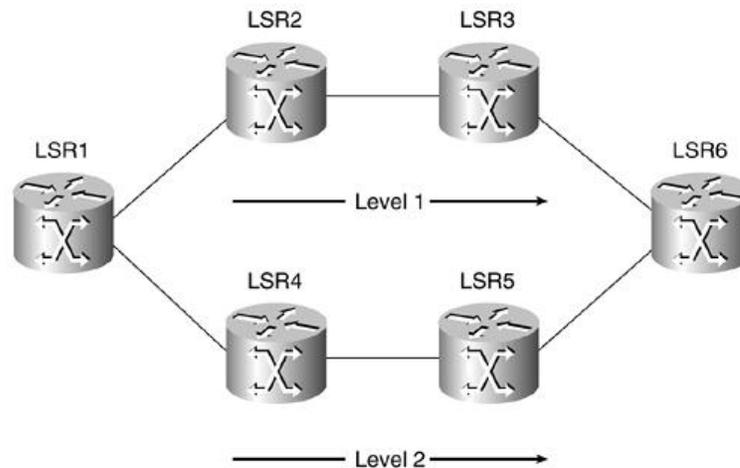


Figura 29 - Níveis de *Label-Switched Path* [6]

O estabelecimento do LSP pode ser realizado de duas formas:

- Controle independente;
- Controle ordenado.

Os controles independente e ordenado para estabelecimento de LSP podem coexistir na mesma rede sem problemas na arquitetura ou interoperabilidade. O método de controle independente fornece convergência e estabelecimento de LSPs mais rápidos, pois o LSR pode estabelecer e enviar informações sobre associações de rótulos a qualquer momento, sem o atraso de esperar por mensagens para propagar em ordem de um lado da rede para o outro. O estabelecimento de LSP imediatamente segue a convergência dos protocolos de roteamento. No método de controle ordenado, as informações de associação de rótulos propagam pela rede antes do estabelecimento do LSP. Porém, o método do controle ordenado fornece melhores capacidades de prevenção de *loop*.

5.4.3 *Label Distribution Protocol (LDP)*

O protocolo de distribuição de rótulo, LDP, é utilizado em conjunto com os protocolos de roteamento de Camada 3 padrão para distribuir informação de associação de rótulos entre os dispositivos LSR em uma rede de comutação de rótulos. LDP permite um LSR distribuir rótulos para seus pares LDP utilizando a porta TCP 646. O uso do TCP como protocolo da camada de transporte resulta em confiabilidade na entrega da informação LDP com controle de fluxo robusto e mecanismos de controle de congestionamento.

Quando um LSR atribui um rótulo a uma FEC, é preciso permitir que seus pares relevantes saibam deste rótulo e de seu significado. O LDP é utilizado para este propósito. Um conjunto de rótulos desde os LSR de entrada até os LSR de saída, em um domínio MPLS, define um LSP. Rótulos são mapas do roteamento da camada de rede para os caminhos comutados da camada de enlace. LDP ajuda no estabelecimento de um LSP utilizando um conjunto de procedimentos para distribuir os rótulos entre os pares LSR.

O LDP fornece um mecanismo de descoberta de LSR para permitir que os roteadores pares localizem uns aos outros e estabeleçam comunicação. Isto define quatro classes de mensagens:

- Mensagens de descoberta (*discovery messages*) utilizam o protocolo de transporte UDP e mensagens multiponto do tipo *hello messages* para aprenderem sobre outros LSRs para os quais LDP tem uma conexão direta. Só então se estabelece uma conexão TCP e uma eventual sessão LDP com seus pares. As sessões LDP são bidirecionais. Estas mensagens anunciam e mantêm a presença de um LSR na rede;
- Mensagens de adjacência (*adjacency messages*) utilizam o TCP e fornecem inicialização de sessão utilizando mensagens de inicialização (*initialization messages*) no início da negociação da sessão LDP. Esta informação inclui o modo de alocação de rótulo, valores de tempo de sobrevivência (*keepalive timer values*) e a faixa de rótulos para ser utilizada entre dois LSRs. LDP *keepalives* são enviados periodicamente utilizando as mensagens *keepalives* (*keepalives messages*). As sessões LDP são finalizadas entre pares LSRs se estas mensagens não são recebidas dentro de um intervalo de tempo;
- Mensagens de anúncio de rótulo (*label advertisement messages*) fornecem anúncios da associação de rótulos entre as FECs e os rótulos. Mensagens de retirada de rótulo (*label withdrawal messages*) são utilizadas para reverter o processo de associação. Mensagens de liberação de rótulo (*label release messages*) são utilizadas pelos LSRs que receberam a informação de mapeamento de rótulo e querem liberar o rótulo que não tem mais utilidade para eles;
- Mensagens de notificação (*notification messages*) fornecem informações de consulta e de sinal de erro entre os pares LSRs que tem uma sessão LDP estabelecida entre eles.

O LDP utiliza o TCP para garantir a entrega das mensagens, com exceção das mensagens de descoberta, como mencionado anteriormente, que utilizam o UDP. As mensagens LDP

são especificadas como um conjunto de objetos do tipo, comprimento, valor, ou *type, length, value* (TLV). A distribuição e atribuição de rótulo no LDP podem acontecer de dois modos: sob demanda ou não solicitado.

A arquitetura MPLS permite um LSR pedir explicitamente, de seu próximo salto para uma FEC particular, um rótulo para associar a uma FEC. Isto é conhecido como distribuição de rótulo sob demanda, ou *downstream-on-demand*. Este modo utiliza mensagens de solicitação de rótulo (*label request messages*) para solicitar mapeamentos de rótulos dos *downstream* LSRs (sabendo que os pacotes partem do *upstream* LSR para o *downstream* LSR).

A arquitetura MPLS também permite que um LSR distribua informações de associação de rótulos para LSRs que não tenha explicitamente solicitado. Isto é conhecido como distribuição de rótulo não solicitada, ou *unsolicited downstream*. Ambas as técnicas de distribuição podem ser utilizadas na mesma rede ao mesmo tempo. Em ambas, o *upstream* LSR e o *downstream* LSR devem entrar em acordo com a utilização de uma técnica. Isto é negociado entre os LSRs durante a negociação da sessão LDP através da troca de mensagens de inicialização.

5.5 Sobrevivência, detecção, e prevenção de *loop* no MPLS

LSPs são construídos utilizando LDP, que faz uso da informação adquirida pelos protocolos de roteamento em Camada 3 e por isso é suscetível aos *loops* de roteamento, a menos que o protocolo de Camada 3 possa, por si próprio, evitá-los. Apesar de tentativas pelos protocolos de roteamento no estabelecimento de rotas sem *loop*, quase todos os protocolos podem levar a situação de *looping* em condições transientes.

Nos protocolos de roteamento do tipo vetor de distância, os nós não tem uma visão completa da topologia da rede. Nessas redes, a “contagem até o infinito” pode ocorrer, onde a métrica de uma rota em *loop* é incrementada e anunciada entre os nós até ela alcance um valor máximo.

Em protocolos de roteamento de estado de enlace, como o OSPF, cada nó mantém toda a base de dados topológica de sua área de roteamento na rede. Nessas redes, há a possibilidade de *loops* de roteamento transitórios se a sincronização da atualização da base de dados não é alcançada sobre a mudança topológica, especialmente durante o período imediatamente após a falha da ligação.

Se os *loops* de roteamento não são prevenidos, o tráfego MPLS é afetado da seguinte maneira:

- Os pacotes usados para o estabelecimento do LSP são encaminhados em um *loop* de roteamento sem fim, e o LSP nunca é apropriadamente estabelecido fim a fim. Isto continua até que o *loop* de roteamento seja quebrado;
- Pacotes de dados injetados em um *looping* LSP estabelecido continuam a ser comutados no LSP até que o *loop* de roteamento seja interrompido.

Existem três maneiras básicas de controlar os *loops* no MPLS:

- Sobrevivência de *loop*;
- Detecção de *loop*;
- Prevenção de *loop*.

5.5.1 Sobrevivência de *loop*

Neste método de controle de *loop*, os LSPs são permitidos a formarem *loops*. Porém, os pacotes em *looping* não são autorizados a impactarem na transmissão de pacotes que não estão em *looping*. Nós MPLS que tem a habilidade de desempenhar o decremento TTL para os LSPs tem esta capacidade. Segmentos não-TTL como as ligações ATM utilizam a alocação de espaço do *buffer* por VC nos *switches* ATM como um mecanismo de controle de *loop*.

O encaminhamento IP utiliza o campo TTL em um pacote, que é decrementado a cada salto IP. Se ele alcança zero, o pacote é assumido estar em *looping* e então é descartado. Isto permite aos roteadores conservar recursos e focar no encaminhamento de pacotes que não estão em *looping* e na atualização das tabelas de roteamento. Assim que as tabelas de roteamento tenham convergido e estão estáveis, o *loop* é normalmente interrompido, a menos que haja um erro de configuração em um dos roteadores.

O MPLS utiliza uma aproximação similar para lidar com *loops*. Os pacotes MPLS rotulados utilizando um *shim header* colocado entre os cabeçalhos da Camada 2 e da Camada 3 carregam o campo TTL que opera como o IP TTL para permitir que pacotes em *looping* sejam descartados. A Figura 30 mostra uma situação de *looping* no LSP. Porém, como

mostrado na Figura 31, o campo TTL é decrementado até zero, e o *loop* é eventualmente interrompido.

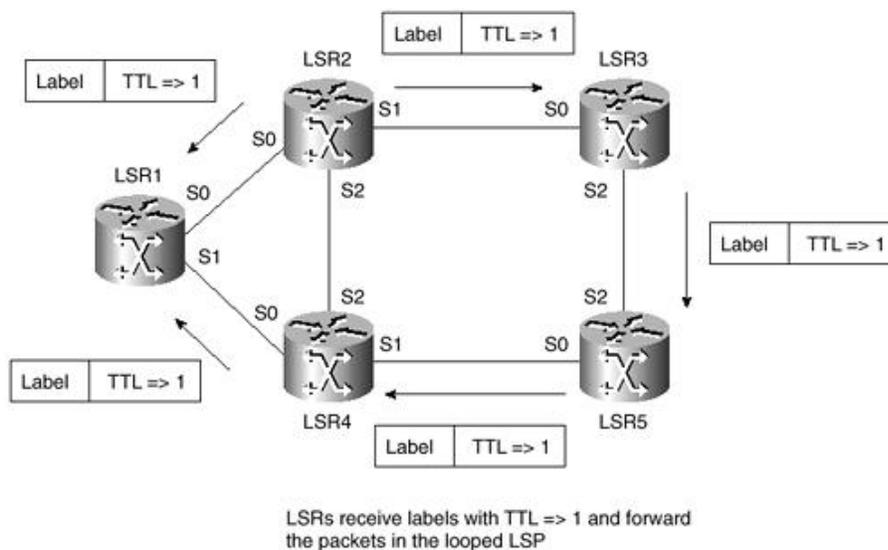


Figura 30 - LSP com *loop* [6]

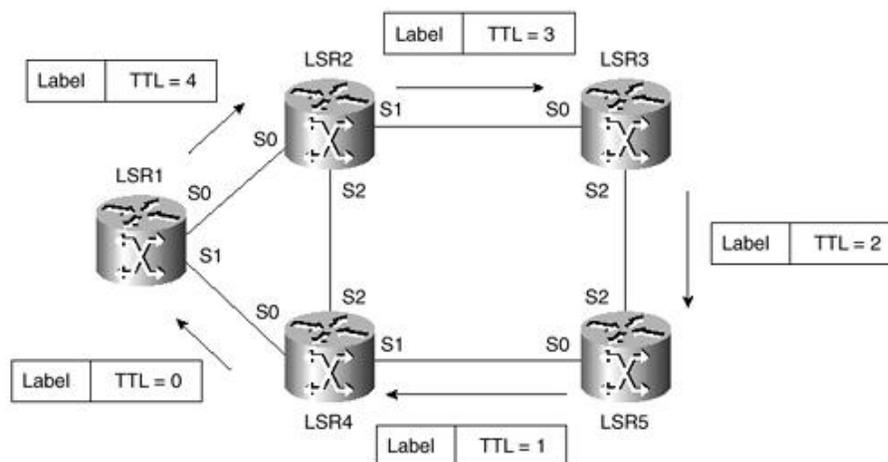


Figura 31 - Detecção do *loop* com o decremento do campo TTL [6]

O campo TTL não é disponível nas conexões ATM, referenciadas como segmentos não-TTL pela arquitetura MPLS. ATM LSRs são *switches* ATM como capacidades de roteamento em Camada 3. O espaço do *buffer* consumido por um único circuito virtual pode ser limitado através de configuração. Este método de alocação de *buffer* controlado é utilizado para controlar os *loops*.

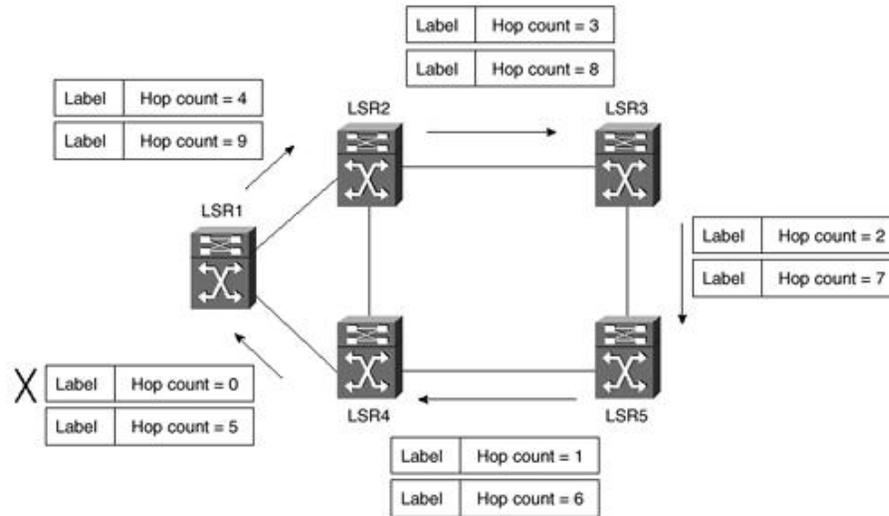
Desta maneira, pacotes presos em *loops* podem consumir apenas uma quantidade limitada de espaço do *buffer* do *switch* e não sobrecarregará o ATM LSR. O *switch* deveria também estar apto a encaminhar pacotes de atualização de roteamento, que garante que o roteamento eventualmente convirja e que *loops* transitórios sejam interrompidos. *Loops* não transitórios criados devido a erros de configuração podem também ser suprimidos, pois o ATM LSR pode ainda encaminhar pacotes de controle e pacotes que não estão em *looping*. Apenas pacotes em *looping* sofrem restrição de recursos em termos de espaço de *buffer*.

5.5.2 Detecção de *loop*

Este método de detecção de controle de *loop* permite a um LSP formar *loop*, mas ele pode detectar e desconectar o *loop* em um curto período de tempo. O LDP é um exemplo de protocolo que suporta esse método. A detecção de *loop* faz uso de técnicas utilizadas na sobrevivência de *loops* como o decremento do campo TTL e a alocação de espaço de *buffer* por VC em *switches* ATM como mecanismo de controle de *loop*.

Em adição as técnicas de sobrevivência de *loop*, uma técnica conhecida como contagem de saltos, ou *hop count*, é utilizada para controle. Este método trabalha como o TTL. Porém, a informação da contagem de salto é transportada no pedido LDP e mensagens de resposta. Esse mecanismo é ilustrado na Figura 32.

Na atribuição de rótulos sob demanda força os pedidos a se propagarem do ponto de mudança da topologia para a saída da rede MPLS. Assumindo que a contagem de salto se inicie como valor 7, este é decrementado em cada salto no ATM LSR que participa do *loop* até alcançar zero. Quando isto ocorre, a atribuição de rótulos falha, e o LSP é derrubado. Depois que o roteamento converge e estabiliza, um novo pedido de atribuição é emitido, resultando na formação de um novo LSP.



ATM LSRs decrement the *Hop count* value by 1 and forward the packets in the looped LSP. The LSP loop is broken when *Hop count* = 0

Figura 32 - Mecanismo de contagem de saltos para detecção de *loops* [6]

5.5.3 Prevenção de *loop*

A prevenção de *loop* é um método de controle que previne a formação de um caminho com *loops* antes que qualquer pacote seja enviado. Para um único estado de enlace de entrada existente para cada estado de enlace de saída, cada mensagem de controle utilizada para solicitar um rótulo contém uma lista de endereços de nós LSR, nos quais cada endereço é adicionado por um LSR quando o LSR encaminha a mensagem. Se um LSR encontra seu próprio endereço na mensagem recebida, ele detecta um *loop* e previne o LSP com *loop* de ser construído.

Quando múltiplos estados de enlace de entrada existem para cada estado de enlace de saída, existe um tipo de prevenção de *loop* chamada vetor de difusão de caminho, ou *path vector diffusion* (PD).

No algoritmo PD, um *loop* LP pode ser prevenido de se formar utilizando-se uma lista de endereços LSR, referidos como um vetor de caminhos. O vetor de caminho é uma lista de LSRs nos quais tenham passado uma mensagem de solicitação de rótulo ou de mapeamento de rótulo tenha.

Uma mensagem de solicitação de rótulo enviada por um LSR para seu vizinho contém um vetor de caminho apenas com o endereço de LSR solicitante. O LSR receptor adiciona seu próprio endereço no vetor antes de emitir uma solicitação de rótulo para esta FEC ao

próximo salto. Se a mensagem de solicitação ou de mapeamento de rótulo passar por um *loop* de roteamento, eventualmente o LSR verá seu próprio endereço na mensagem e então detectará o *loop*. Neste caso, previne a construção de um LSP com *loop*.

5.6 Algumas aplicações do MPLS

5.6.1 Engenharia de tráfego

A engenharia de tráfego, ou *Traffic Engineering* (TE), é um processo que melhora a utilização da rede de maneira geral e abrange vários aspectos relacionados ao desempenho da rede. Ela inclui a tentativa de melhora na utilização dos recursos da rede pela reordenação uniforme do tráfego e provimento de funcionalidade para rápida recuperação em caso de falha de um nó ou enlace [16]. Um importante resultado desse processo é a prevenção de congestionamentos sobre qualquer caminho.

É importante notar que a TE não necessariamente seleciona o caminho mais curto entre dois dispositivos. É possível que, para dois fluxos de dados, os pacotes atravessem a rede por caminhos completamente diferentes, mesmo que o nó de origem e o destino final sejam os mesmos. Assim, o segmento menos exposto ou menos utilizado pode ser utilizado e serviços diferenciados podem ser fornecidos.

No MPLS, engenharia de tráfego é inerentemente fornecida utilizando-se caminhos roteados explicitamente. Os LSPs são criados independentemente, especificando diferentes caminhos que são baseados em políticas definidas pelo usuário. No entanto, isto pode exigir grande intervenção do operador. O *Resource Reservation Protocol* (RSVP) e *Constraint-based Routed Label Distribution Protocol* (CR-LDP) são duas abordagens possíveis para fornecimento de engenharia de tráfego dinâmico e QoS [14].

O RSVP é um protocolo da camada de transporte e pode ser utilizado por roteadores para solicitar ou entregar níveis específicos de QoS para aplicação de dados. Ele define como os aplicativos alocam recursos e como renunciam recursos reservados se não há mais necessidade para tal. A operação RSVP geralmente resultará em recursos sendo reservados em cada nó ao longo do caminho. Sua extensão RSVP-TE, conforme a RFC 3209, é mais utilizada para a aplicação de engenharia de tráfego.

O protocolo de sinalização CR-LDP pode prover, entre outras capacidades, a configuração de rotas explicitamente roteadas através de alocação de roteadores individuais (*strict explicit*

routes) ou grupos de roteadores (*loose explicit route*) e realocação de recursos para os caminhos através de prioridades (*path preemption with setup and holding priority*). Este mecanismo permite ao CR-LDP realocar recursos requeridos por uma determinada rota. Se uma rota com recursos suficientes não pode ser encontrada, caminhos existentes podem ser reconfigurados para realocar recursos para o novo caminho, processo este denominado de *path preemption*.

5.6.2 Qualidade de serviço

Atualmente, com a grande variedade de tipos de informação a serem transmitidas através das redes, é evidente que estas redes precisem da habilidade de diferenciar entre os diversos tipos de tráfego e tratá-los de formas distintas. Cada tipo de informação exige requisitos diferentes para a sua transmissão. Aplicações de vídeo são sensíveis à variação de retardo e a mídia texto é bastante tolerante a esta característica, por exemplo.

Para possibilitar a transferência de dados multimídia, as redes devem garantir qualidade de serviço. QoS (*Quality of Service*) consiste em um conjunto de parâmetros de desempenho da rede que pode ser negociado pelo usuário durante o estabelecimento de uma conexão. Estes parâmetros podem ser:

- O retardo no estabelecimento da conexão;
- O retardo no encerramento da conexão;
- A probabilidade de falha no estabelecimento da conexão;
- A probabilidade de falha na liberação da conexão;
- A vazão em cada sentido da conexão;
- O retardo de transferência médio em cada sentido;
- O retardo de transferência máximo em cada sentido;
- A variação estatística do retardo;
- A taxa de erro;
- A prioridade de serviços;
- A probabilidade de queda de uma conexão.

Devido ao estabelecimento de caminhos virtuais (LSP), diferenciação dos pacotes em classes de serviço (FECs) e a distribuição de rótulos, o gerenciamento de QoS em redes MPLS torna-se simples e prático [15].

5.6.3 Redes virtuais privadas

As redes virtuais privadas, ou *Virtual Private Networks* (VPNs), são redes privadas por onde podem trafegar informações de forma segura, construídas sobre a infraestrutura de uma rede pública como a internet.

Utilizando-se a técnica conhecida como “tunelamento”, pacotes são transmitidos através da rede pública em um túnel privado que simula uma conexão ponto a ponto. 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, uma VPN utiliza os serviços das redes IP.

A utilização de redes públicas tende a apresentar custos muito menores que os obtidos com a implantação de redes privadas, sendo este, justamente, o grande estímulo para o uso de VPNs. No entanto, para que esta abordagem se torne efetiva, a VPN deve prover um conjunto de funções que garantam confidencialidade, integridade e autenticidade. O MPLS, atuando como mecanismo de encaminhamento dentro de um cenário de VPN, provê agilidade, facilidade de gerenciamento para grandes redes e suporte a QoS, bem como suporte a segurança.

O fundamento utilizado em VPNs é a comutação de rótulos habitual do MPLS com a utilização da pilha de rótulos. Uma pilha de rótulos é acrescentada ao pacote e, apenas quando estiver no destino final haverá correspondência para o último rótulo da pilha.

6 Conclusão

A interligação de redes locais com tecnologias diferentes abriu espaço para a criação das grandes redes nas quais todos estão interligados, porém a interoperabilidade entre os sistemas é um problema a ser resolvido. Para isso foram criados modelos de referências, com padrões especificados por organizações internacionais (ISO e ITU-T), como o modelo de referência OSI e o modelo de referência TCP/IP, atualmente o mais difundido.

Ambos baseiam-se em camadas, cada qual com seus protocolos e funções, que se comunicamos entre si, como se apenas as camadas pares fossem visíveis umas as outras, para a realização da comunicação entre os sistemas. Mesmo com ambos os modelos sendo arquiteturas abertas (não proprietárias), o modelo TCP/IP foi o mais aceito devido a seus protocolos eficazes, mas o modelo OSI é utilizado para referenciar as camadas da arquitetura TCP/IP.

Para o encaminhamento dos pacotes existem os serviços orientados a conexão e os não orientados a conexão, cada qual utilizado em determinada aplicação dependendo das exigências. Diferentemente dos serviços sem conexão, as conexões orientadas estabelecem, primeiramente, uma conexão com o serviço desejado antes de transmitir qualquer informação, oferecendo certa garantia de entrega. Apesar do custo dos serviços orientados a conexões ser maior, muitas aplicações exigem a garantia de entrega, sendo portanto muito utilizado.

O transporte de dados começou a ser realizado em redes projetadas para o serviço de voz, sendo apenas posteriormente implantadas redes de dados independentes e paralelas às redes de voz, com sistemas de gerência próprios e alto custo operacional devido à falta de integração com as tecnologias já instaladas e a utilização do legado. Com o aumento do tráfego de dados, surgiu a necessidade por parte das operadoras e provedores de serviço,

de uma infraestrutura *backbone* com tecnologia WAN eficiente, flexível para futuras ampliações (escalável) e com baixo custo operacional.

Esse contexto foi favorável para o surgimento de plataformas que possibilitassem a união das redes de telecomunicações tradicionais, orientadas a conexão, e das redes *Internet*. Foram desenvolvidas, então, a tecnologia *Frame Relay* e ATM, com maior capacidade de transmissão que o *Frame Relay*. Ambos utilizam a comutação em Camada 2 para encaminhar os pacotes, de maneira rápida, eficiente e oferecendo a qualidade de serviço exigida pelas aplicações.

O problema do ATM era a interoperabilidade com o IP, protocolo de rede mais difundido do mundo. Isso porque o conceito de sobreposição do IP sobre o ATM geravam sérios problemas de escalabilidade, uma vez que os *switches* ATM de Camada 2 são invisíveis aos roteadores Camada 3 que utilizam o IP, fazendo o número de adjacências ser muito grande. Além disso, os circuitos virtuais devem ser gerenciados pelos administradores da rede, sendo inviável se o número de roteadores anexos à nuvem ATM for muito grande.

Com os princípios do roteamento em Camada 3 sendo simples e bem definidos, com protocolos de roteamento criando automaticamente tabelas de roteamento (*routing tables*) para encaminhar os pacotes, o ideal seria uma tecnologia que integrasse o IP e o ATM, ao invés de sobrepô-lo. Isso unificaria as vantagens do roteamento em Camada 3 com a confiabilidade do encaminhamento em Camada 2. Nesse contexto surgiu o MPLS, que resolve o problema do modelo sobreposto eliminando o conceito de nuvem ATM. Porém, não necessariamente necessita-se da utilização de comutadores ATM para a implementação do MPLS, podendo-se utilizar roteadores ou switches comuns no núcleo da rede.

Com o MPLS, tem-se um modelo integrado no qual os switches ATM são conhecedores do IP e as ligações ATM são tratadas como ligações IP. Dessa forma, cada switch ATM pode se tornar um par de roteamento para o IP, diminuindo drasticamente o número de adjacências, que não mais aumentam com o tamanho da rede.

O MPLS é um multiprotocolo, pois pode ser aplicável em qualquer protocolo da camada de rede. Ele resume toda a informação necessária para o encaminhamento de pacotes pela rede em um rótulo simples de comprimento fixo, que nos pacotes é inserido entre os cabeçalhos de Camada 2 e de Camada 3, e nas células ATM é inserido no campo VPI/VCI.

Tabelas de encaminhamento nas entidades comutadoras de rótulos são responsáveis pelas substituições dos rótulos pelos diversos nós da rede até o destino final do pacote.

O MPLS integra o desempenho e a capacidade de gerenciamento de tráfego da Camada 2 com a escalabilidade e flexibilidade das redes de roteamento em Camada 3. Integrando o roteamento IP com a comutação ATM, oferece redes IP sobre ATM escaláveis, além da possibilidade de *switches* ATM ou roteadores existentes no núcleo da rede poder encaminhar pacotes baseando-se apenas nos rótulos gerar mínima sobrecarga de pesquisa nos nós.

A grande importância da integração IP + ATM é a utilização do legado dos *backbones* das operadoras, pois quando o ATM foi desenvolvido acreditava-se que seria a tecnologia fim a fim, abrangendo redes WAN e LAN, e grandes investimentos foram feitos. Porém, o alto custo dos adaptadores ATM para PCs inibiram sua vasta implantação, abrindo espaço para o IP, simples e escalável, que teve um crescimento vertiginoso e se tornou o mais difundido protocolo de rede. Porém, mesmo com as vantagens da comutação em Camada 2 (como o ATM), com taxas de transmissão que permitem aplicações de vídeo, era inviável a integração dessas tecnologias devido ao problema da escalabilidade. Assim, o MPLS possibilitou que, sem a troca dos equipamentos existentes a integração pudesse ser feita, trazendo escalabilidade e grandes taxas de transmissão na rede.

As vantagens mais relevantes na utilização do MPLS são a facilidade de engenharia de tráfego, permitindo direcionar fluxos de dados específicos para caminhos menos congestionados, e a implementação de redes virtuais privadas (VPNs), que através da pilha de rótulo se torna bastante simplificada.

Referências Bibliográficas

- [1] **Cisco Systems, Inc.** *Internetworking Technologies Handbook*. 4th ed. Indianapolis : Cisco Press, 2003.

- [2] **TANENBAUM, Andrew S.** *Redes de Computadores*. 4. ed. Rio de Janeiro : Campus Elsevier, 2003.

- [3] **TRONCO, Tania R.** *Redes de Nova Geração*. 1. ed. São Paulo : Érica, 2006.

- [4] **BATES, Regis J.** *Broadband Telecommunications Handbook*. 2nd ed. Phoenix : McGraw-Hill telecommunications, 2002.

- [5] **SOCOLOFSKY, Theodore e KALE, Claudia.** A TCP/IP Tutorial. *RFC 1180*. 1991.

- [6] **ALWAYN, Vivek.** *Advanced MPLS Design and Implementation*. Indianapolis : Cisco Press, 2002.

- [7] **DEERING, Stephen E. e HINDEN, Robert M.** Internet Protocol, Version 6 (IPv6) Specification. *RFC 2460*. 1998.

- [8] **LAUBACH, Mark.** Classical IP and ARP over ATM. *RFC 1577*. 1994.

- [9] **GUICHARD, Jim e PEPELNJAK, Ivan.** *MPLS and VPN Architectures*. Indianapolis : Cisco Press, 2001.

- [10] **ROSEN, Eric C., VISWANATHAN, Arun e CALLON, Ross.** Multiprotocol Label Switching Architecture. *RFC 3031*. 2001.

- [11] **Cisco Systems, Inc.** *DiffServ - The Scalable end-to-end Quality of Service Model*. San Jose : Cisco Press, 2005.
- [12] **ROSEN, Eric C., et al.** MPLS Label Stack Encoding. *RFC 3032*. 2001.
- [13] **HEINANEN, Juha.** Multiprotocol Encapsulation over ATM Adaptation Layer 5. *RFC 1483*. 1993.
- [14] **Cisco Systems, Inc.** *Advanced Topics in MPLS-TE Deployment*. San Jose : Cisco Press, 2002.
- [15] **Cisco Systems, Inc.** *Cisco IOS MPLS Quality of Service*. San Jose : Cisco Press, 2001.
- [16] **ALOIA, Eduardo José.** Sistematização Crítica das Tendências de Padronização de Arquitetura de Protocolos em Redes Ópticas. São Carlos : s.n., 2003.