

Redes de Comunicação de Dados



PUC -Rio
Departamento de Informática

Luiz Fernando Gomes Soares
lfgs@inf.puc-rio.br



PUC-Rio / DI

TeleMídia

Camada de Enlace



- ▶▶ Delimitação e transmissão de quadros
- ▶▶ Detecção de erros
- ▶▶ Controle de Acesso
- ▶▶ Correção de erros que por ventura ocorram no nível físico (opcional)
- ▶▶ Controle de fluxo (opcional)
- ▶▶ Multiplexação (opcional)



PUC-Rio / DI

TeleMídia

Correção de Erro

- ➔ Uma vez detectado o erro, o quadro é descartado e, opcionalmente, é enviado um aviso ao sistema que o transmitiu. A recuperação do erros, se desejada, pode ser realizada por retransmissão.
- ➔ Alternativamente o nível de enlace pode utilizar bits de redundância que lhe permitam não só detectar erros, mas também corrigi-los, sem a necessidade de retransmissão. A correção de erros sem retransmissão é particularmente útil em enlaces onde o custo de comunicação e o retardo de transferência são elevados, como por exemplo, em enlaces de satélites.



Correção de Erro

Os três procedimentos mais utilizados para controlar erros são:

- ➔ algoritmo de bit alternado (stop-and-wait),
- ➔ janela n com retransmissão integral (go-back-n) e
- ➔ janela n com retransmissão seletiva (selective repeat).



Bit alternado

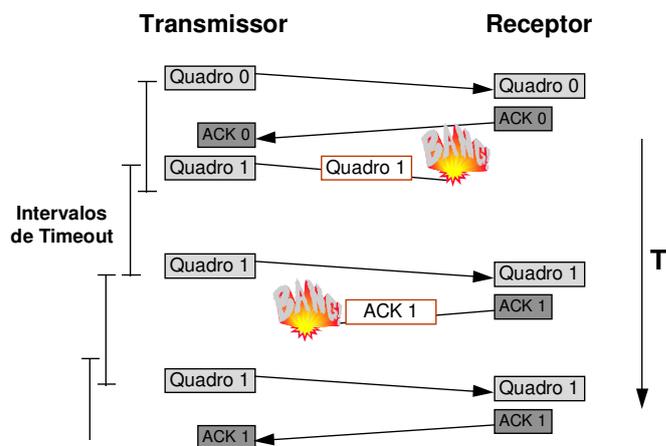
- No *algoritmo de bit alternado* o transmissor só envia um novo quadro quando recebe o reconhecimento do quadro enviado anteriormente.
- Após a transmissão, o transmissor espera um certo tempo pelo reconhecimento. Caso este tempo se esgote (timeout), o quadro é retransmitido.
- Considerando que os quadros podem ser transmitidos mais de uma vez, é necessário numerá-los para que o receptor possa distinguir quadros originais de retransmissões. Como o transmissor só envia um novo quadro depois do anterior ser reconhecido, só é preciso um bit para diferenciar quadros sucessivos. O primeiro quadro é numerado com o bit 0, o segundo com o bit 1, o terceiro com o bit 0, e assim sucessivamente.



PUC-Rio / DI

TeleMídia

Bit Alternado



PUC-Rio / DI

TeleMídia

Janela de Transmissão

- ➔ A técnica de bit alternado oferece uma solução simples porém ineficiente para o controle de erro, pois, enquanto o transmissor espera por reconhecimentos o canal de comunicação não é utilizado.
- ➔ Para aumentar a eficiência na utilização dos canais de comunicação foram elaborados protocolos que permitem que o transmissor envie diversos quadros mesmo sem ter recebido reconhecimentos dos quadros anteriormente enviados. O número máximo de quadros, devidamente numerados, que podem ser enviados sem que tenha chegado um reconhecimento define a largura da *janela de transmissão*.

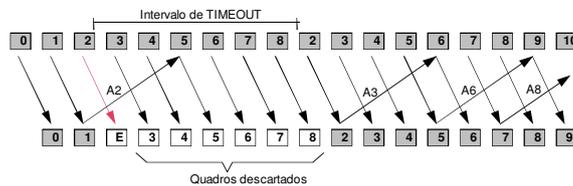


Controle com Janelas

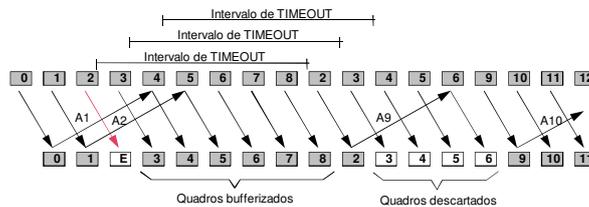
- ➔ Como no protocolo de bit alternado, o transmissor fica sabendo que ocorreu um erro em um quadro por ele enviado quando seu reconhecimento não chega, após decorrido um intervalo de tempo suficiente para tal. Nesse caso, dois procedimentos podem ser implementados para recuperar o erro:
 - *Retransmissão integral*: todos os quadros a partir do que não foi reconhecido são retransmitidos.
 - *Retransmissão seletiva*: apenas o quadro que não foi reconhecido é retransmitido.
- ➔ Para aumentar ainda mais a eficiência na utilização do canal de transmissão, em ambos os casos o receptor não precisa enviar um reconhecimento para cada quadro que recebe. O transmissor ao receber o reconhecimento do quadro n conclui que os quadros enviados antes dele, foram recebidos corretamente.



Controle de Erro com Janela Deslizante



a) Protocolo com retransmissão integral



b) Protocolo com retransmissão seletiva



Controle de Fluxo

- ➔ Outro problema tratado pelo nível de enlace diz respeito a situações onde o transmissor sistematicamente envia quadros mais depressa que o destinatário pode receber.
- ➔ A solução para essa situação é introduzir um mecanismo de controle de fluxo para compatibilizar as velocidades do transmissor e do receptor.



Controle de Fluxo

- ➔ No protocolo de bit alternado, o próprio mecanismo de retransmissão de quadros controla o fluxo, pois um novo quadro só é enviado depois do receptor ter processado o quadro anterior e enviado um reconhecimento.



Janela Deslizante

- ➔ Nos protocolos com janela n maior que 1, o controle do fluxo é feito com base em quadros especiais ou em janelas de transmissão e recepção.
- ➔ O número máximo de quadros que o transmissor pode enviar é determinado pela largura (T) de sua janela de controle de fluxo. Após enviar T quadros sem receber nenhuma atualização da janela, o transmissor suspende o envio de dados, só voltando a fazê-lo após receber uma resposta de controle que indica que o receptor está pronto para processar novos quadros.



Camada de Enlace



- Delimitação e transmissão de quadros
- Detecção de erros
- Controle de Acesso
- Correção de erros que por ventura ocorram no nível físico (opcional)
- Controle de fluxo (opcional)
- Multiplexação (opcional)



Multiplexação da Conexão Física

- Finalmente, dado que uma estação tem o controle do acesso ao meio físico, ela pode multiplexar entre vários usuários do nível de rede (que interagem com o nível de enlace através de SAPs de enlace), o acesso à conexão física, ora enviando um quadro provindo de um usuário do nível de rede, ora de outro.



Tipos de Serviço

- Como usual em todas as camadas do RM-OSI, a camada de enlace pode oferecer ao nível de rede serviços orientados a conexão e serviços não orientados a conexão.
- Dentre os serviços mais usuais podem ser destacados:
 - Serviço sem conexão e sem reconhecimento (apenas detecção de erros no nível de enlace): também conhecido como serviço de datagrama não confiável.
 - Serviço sem conexão e com reconhecimento (detecção e correção de erros no nível de enlace): também conhecido como serviço de datagrama confiável.
 - Serviço com conexão (usualmente com controle de erro e fluxo no enlace).



PUC-Rio / DI

TeleMídia

Camada de Enlace

- Orientados a Caracter
 - BSC - Binary Synchronous Communication
- Orientados a Bit
 - HDLC - High Level Data-Link Control
 - LAP-B
 - LAP-D
 - LAP-M
 - LLC
 - SDLC - Synchronous Data Link Control (IBM)



PUC-Rio / DI

TeleMídia

High Level Data-Link Control

- Transmissão Síncrona de Quadros
- Bit Stuffing
- Controle de Erro
 - Detecção: CRC
 - Correção: Retransmissão
- Controle de Fluxo
- Vários modos de operação



Padrão IEEE 802.2



IEEE 802.2

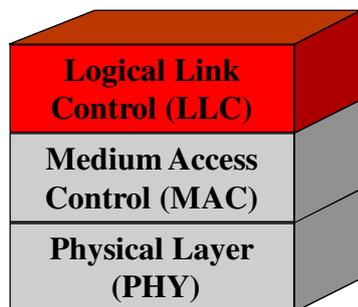
- ➔ O padrão IEEE 802.2 é um exemplo de protocolo que, juntamente à subcamada MAC, implementa as funções necessárias da camada de enlace.
- ➔ O padrão define como é realizado o controle de erro, de fluxo e a multiplexação, de forma a realizar os tipos de serviço orientados a conexão e sem conexão.



PUC-Rio / DI

TeleMídia

Camada de Controle de Enlace Lógico IEEE 802.2



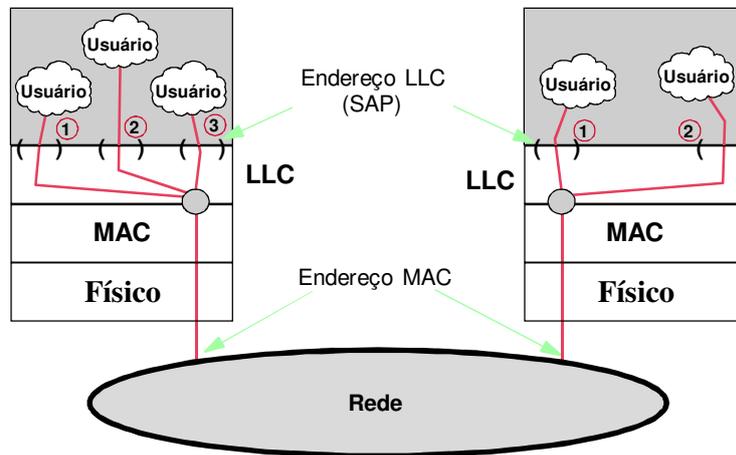
- ➔ Independência da camada MAC
- ➔ LSAPs
- ➔ Multiplexação
- ➔ Controle de erros e de fluxo
- ➔ Tipos de operação
- ➔ Classes de procedimentos



PUC-Rio / DI

TeleMídia

IEEE 802.2 - Multiplexação



PUC-Rio / DI

TeleMídia

Formato da Unidade de Dados do Protocolo LLC (PDU LLC)

DSAP	SSAP	Controle	Dados
8 Bits	8 Bits	8 ou 16 Bits	N x 8 Bits

DSAP: endereço do ponto de acesso ao serviço LLC destino

SSAP: endereço do ponto de acesso ao serviço LLC origem



PUC-Rio / DI

TeleMídia

Especificação da Interface LLC/Nível Superior

➤ Operação Tipo 1

- serviço sem conexão e sem reconhecimento
- transferências de dados ponto a ponto, entre grupos, ou por difusão

➤ Operação Tipo 2

- serviço orientado a conexão (controle de fluxo, sequenciação e recuperação de erros)
- conexões ponto a ponto

➤ Operação Tipo 3 (Addendum 2)

- serviço sem conexão e com reconhecimento (sequenciação e recuperação de erros)
- conexões ponto a ponto



PUC-Rio / DI

TeleMídia

Classes de Serviço IEEE 802.2

➤ Classe I

- Operação Tipo 1

➤ Classe II

- Operação Tipo 1 e Tipo 2

➤ Classe III

- Operação Tipo 1 e Tipo 3

➤ Classe IV

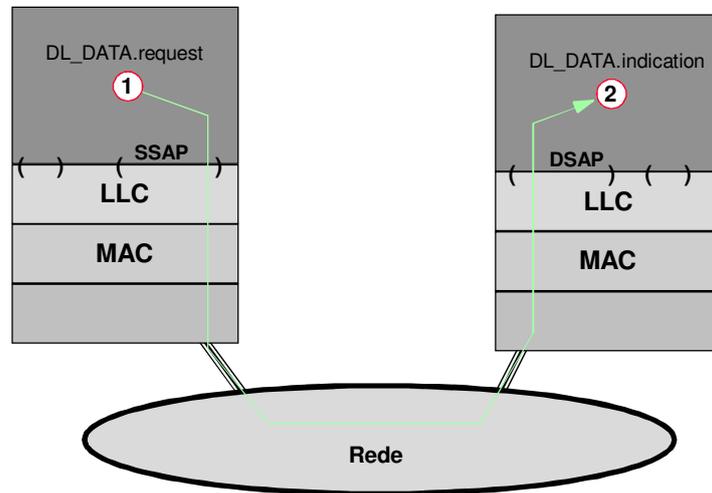
- Operação Tipo 1, 2 e 3



PUC-Rio / DI

TeleMídia

Serviço Sem Conexão e Sem Reconhecimento



PUC-Rio / DI

TeleMídia

Formato da Unidade de Dados do Protocolo LLC (PDU LLC)

DSAP	SSAP	Controle	Dados ⚡
8 Bits	8 Bits	8 ou 16 Bits	N x 8 Bits

DSAP: endereço do ponto de acesso ao serviço LLC destino

SSAP: endereço do ponto de acesso ao serviço LLC origem



PUC-Rio / DI

TeleMídia

Campo de Controle da PDU LLC Operação Tipo 1

Formato Não-Numerado (U)



- M - bits identificadores de comando não-numerado

- P/F - (P = 1) solicitação de resposta imediata e
(F = 1) indicador de resposta de solicitação imediata



PUC-Rio / DI

TeleMídia

Comandos e Respostas LLC Operação Tipo 1

Tipo 1	
Comandos	Respostas
UI	
XID	XID
TEST	TEST

- UI - transmissão de informação não-numerada

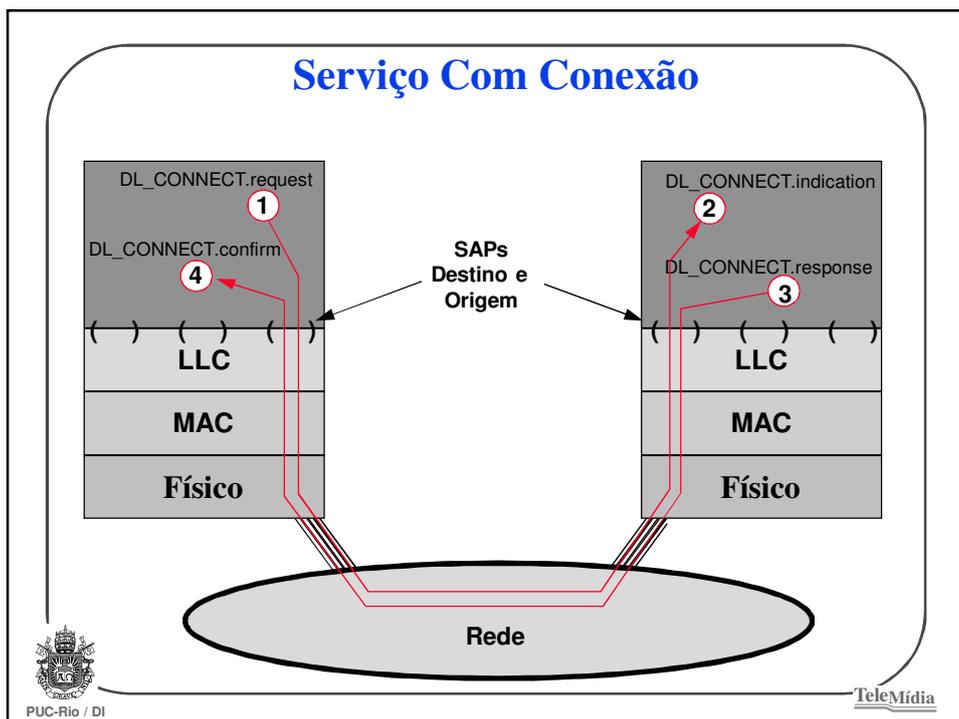
- XID - troca de identificação

- TEST - comando usado para testar uma conexão LLC-LLC



PUC-Rio / DI

TeleMídia



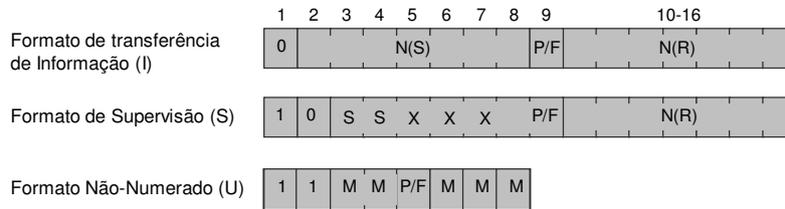
Comandos e Respostas LLC Operação Tipo 2

Tipo 2	
Comandos	Respostas
I	I
RR	RR
RNR	RNR
REJ	REJ
SABME	UA
DISC	DM
	FRMR

- I transmissão de informação numerada
- RR indica que a LLC está pronta para receber PDUs I
- RNR indica impossibilidade de recepção de PDUs I
- REJ solicita a retransmissão das PDUs I a partir de N(R)
- SABME estabelece uma conexão com LLC destino no modo balanceado assíncrono
- DISC encerra conexão de enlace no modo balanceado assíncrono
- UA reconhecimento de recebimento e aceitação de comandos SABME e DISC
- DM indica que a estação está desconectada da conexão de enlace
- FRMR indica rejeição de PDU que não pode ser corrigida por retransmissão

PUC-Rio / DI TeleMídia

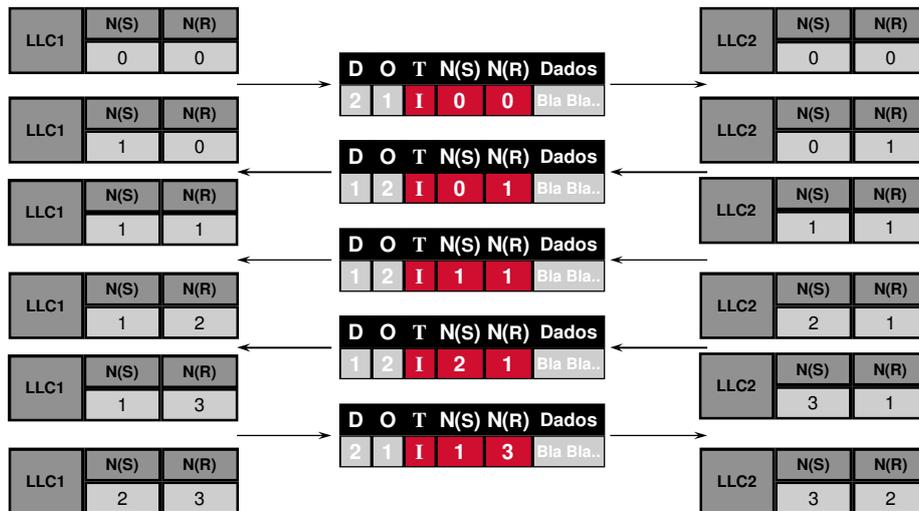
Formato do Campo de Controle



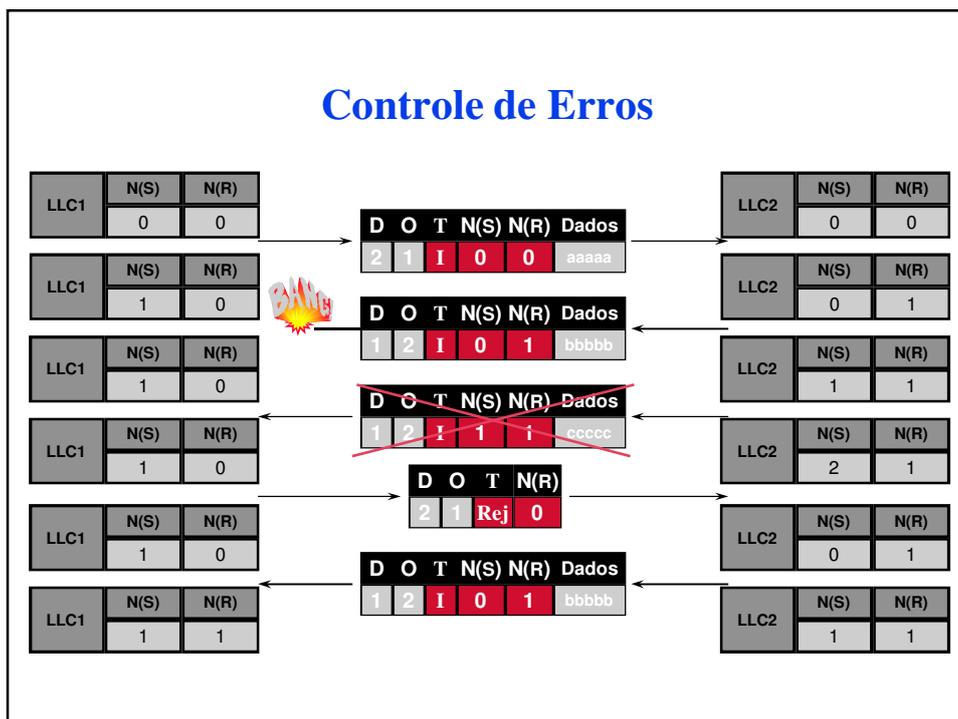
- N(S) - número de sequência da PDU transmitida
- N(R) - número de sequência da PDU esperada
- S - bits de função de supervisão
- M - bits identificadores de comando não-numerado
- X - bits reservados
- P/F - (P = 1) solicitação de resposta imediata e
(F = 1) indicador de resposta de solicitação imediata



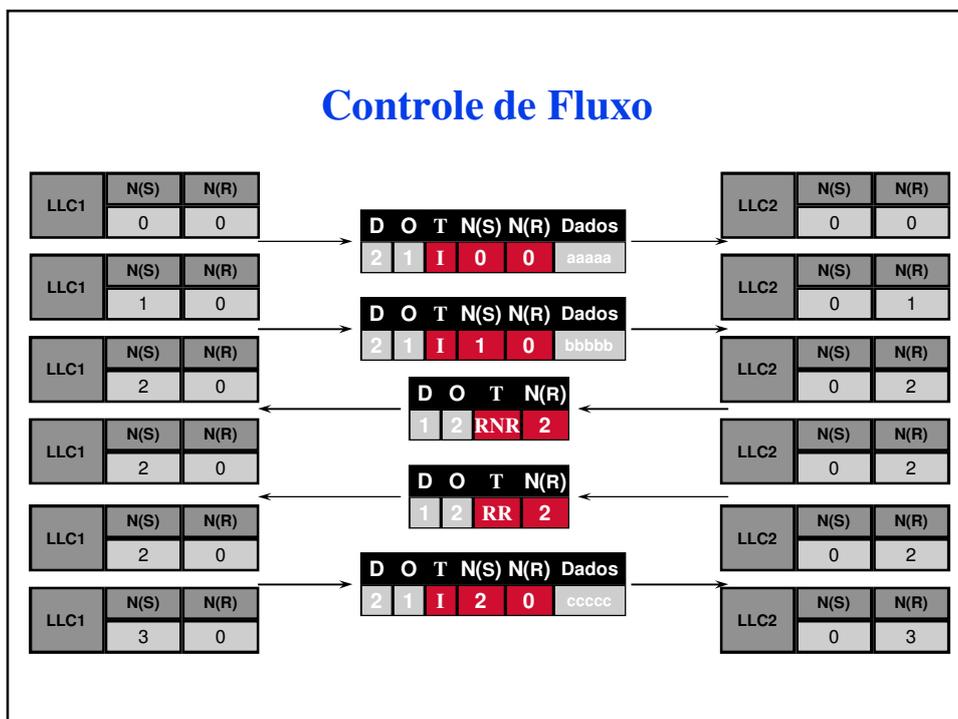
Troca de Quadros I



Controle de Erros



Controle de Fluxo



Drivers da Placa de Interface de Rede



PUC-Rio / DI

TeleMídia

Interação entre Driver de Protocolo e Placa de Interface de Rede

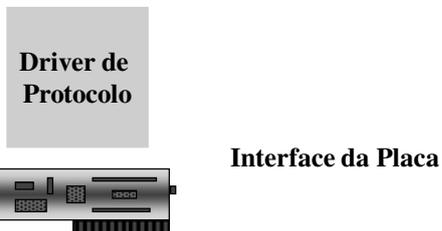
- ▶ Interação direta
- ▶ Interação através de um driver de placa não padrão
- ▶ Interação através de um driver de placa padrão



PUC-Rio / DI

TeleMídia

Interação Direta



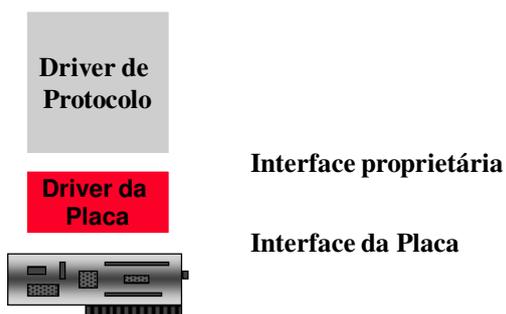
- Driver de protocolo comunicando diretamente com o hardware da placa
- Driver de Protocolo muito eficiente
- Driver de Protocolo totalmente dependente das características da placa (Hardware, Protocolo MAC,...)



PUC-Rio / DI

TeleMídia

Driver de Placa não Padrão



- Driver de protocolo comunicando com a placa através de um driver não padrão
- Driver de Protocolo independente das características da placa
- Driver de Protocolo totalmente dependente da interface proprietária oferecida pelo Driver da Placa



PUC-Rio / DI

TeleMídia

Drivers de Placa

- Normalmente fornecido juntamente com a Placa de Interface de Rede
- Esconde detalhes do hardware da placa
- Interface de acesso aos serviços de comunicação oferecidos pela placa
- Interfaces padronizadas
 - PD Packet Driver
 - NDIS Network Driver Interface Specification
 - ODI Open Data-Link Interface
 - ASI Adapter Support Interface
- Conversores de interfaces (Shim Drivers)



PUC-Rio / DI

TeleMídia

Network Driver Interface Specification

- Desenvolvido pela Microsoft e 3COM
- Oferece independência das Camada MAC e Física
- Oferece suporte a várias placas de interface de rede
- Oferece suporte a várias pilhas de protocolo (Arquitetura Multiprotocolar)
- Controla as interações entre as placas de rede e as pilhas de protocolos através da subcamada VECTOR
- Entrega de pacotes sequencial às pilhas de protocolo

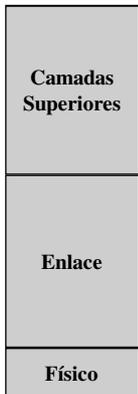


PUC-Rio / DI

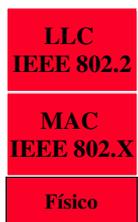
TeleMídia

Arquitetura Multiprotocolar NDIS

Modelo OSI



Modelo IEEE



PROTOCOL.INI

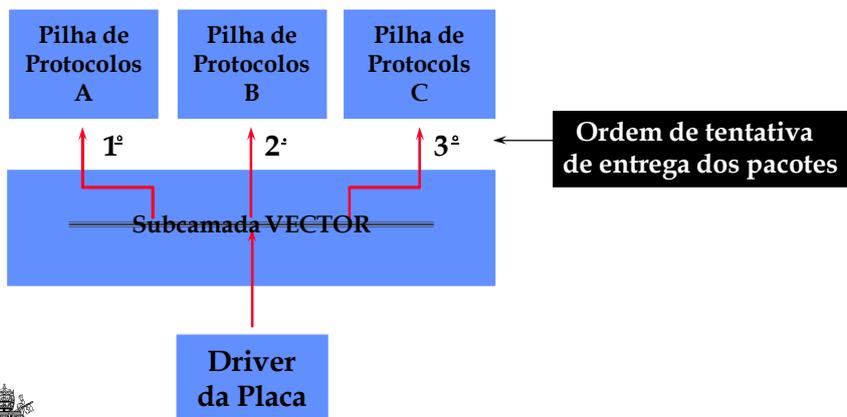
```
[zka]
sxs
ssa
[asa]
xs
sd
```



PUC-Rio / DI

TeleMídia

Funcionamento da Subcamada VECTOR



PUC-Rio / DI

TeleMídia

Open Data-Link Interface

- Desenvolvido pela Novell e Apple
- Oferece independência das Camada MAC e Física
- Reconhece vários tipos de quadros ethernet e token-ring
- Oferece suporte a várias placas de interface de rede (Board Number)
- Oferece suporte a várias pilhas de protocolo (Arquitetura Multiprotocolar)
- Controla as interações entre as placas de rede e as pilhas de protocolos, através da subcamada LSL
- Entrega de pacotes diretamente à pilha de protocolos correta



PUC-Rio / DI

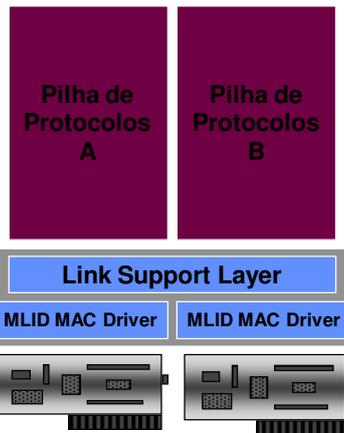
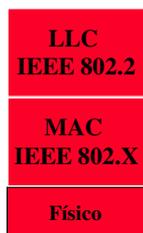
TeleMídia

Arquitetura Multiprotocolar ODI

Modelo OSI



Modelo IEEE



NET.CGF

[zka]
sxs
ssa

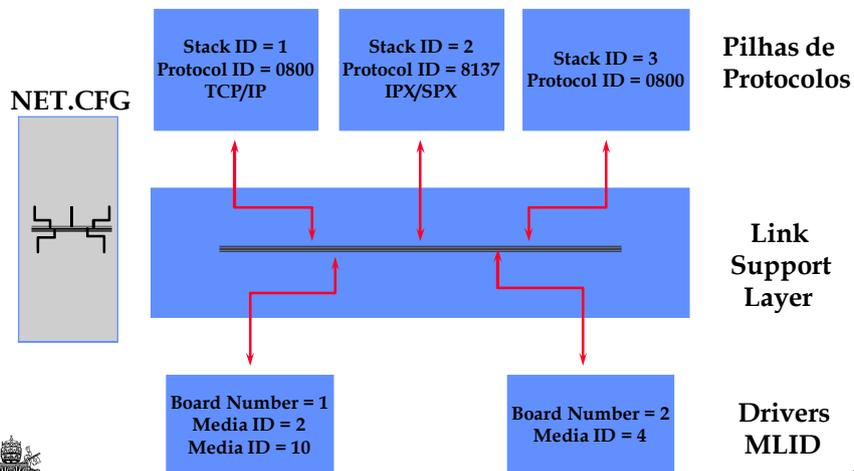
[asa]
xs
sd



PUC-Rio / DI

TeleMídia

Funcionamento da Subcamada LSL



PUC-Rio / DI

TeleMídia

Perguntas

- Quando é mais eficiente realizar a correção de erros por retransmissão e quando por redundância na própria informação?
- Explique o procedimento de controle de erro "bit alternado". Discuta a eficiência do procedimento.
- O que é janela de retransmissão? Em uma janela n , porque é necessário termos $n+1$ numerações diferentes?
- Explique a diferença entre os procedimentos de controle de erro janela n com retransmissão integral (go-back- n) e janela n com retransmissão seletiva (selective repeat).
- Cite duas formas de realizar o controle de fluxo.
- Como funciona o controle de fluxo por quadros especiais?
- Como funciona o controle de fluxo por janela deslizante?
- Por que não se deve misturar a janela de controle de fluxo com a janela de controle de erro?



PUC-Rio / DI

TeleMídia

Perguntas

- ➔ Para que serve a multiplexação na camada de enlace? Como ela pode ser realizada?
- ➔ O que é um SAP? Como é composto o endereço de um SAP de enlace de forma hierárquica e como é feito no caso de endereçamento horizontal? Qual a vantagem de se ter um endereço hierárquico?
- ➔ Quais as funcionalidades existentes no serviço LLC sem conexão e sem reconhecimento?
- ➔ Quais as funcionalidades existentes no serviço LLC com conexão?
- ➔ Como é realizado o controle de erro no protocolo LLC? Como são enviados os reconhecimentos? O protocolo é de janela deslizante? Se sim, é com repetição seletiva ou retransmissão integral?
- ➔ O protocolo LLC com conexão pode funcionar sem o quadro REJ? Para que serve este quadro?



Perguntas

- ➔ O que vem a ser driver de protocolo e driver de placa?
- ➔ Quais as formas de integração entre um driver de protocolo e um driver de placa? Discuta a eficiência de cada forma versus a independência de plataforma.



Para as próximas questões siga as seguintes orientações:

- Considere dois pontos A e B ligados através de um enlace. Considere também apenas os quadros do tipo informação (I) e três tipos de quadros supervisor: RECEIVE READY (RR), RECEIVE NOT READY (RNR) e REJECT (REJ). A janela do protocolo é de tamanho 3. Considere também que o controle de erro utiliza o método "repetição seletiva (selective repeat)".
- Todas as mensagens levam exatamente 1s para serem transmitidas e chegarem ao destino, e todas as transmissões acontecem em múltiplos de 1s, isto é, nunca acontecem em uma fração de segundo. Ao receber um quadro de informação, o receptor verifica se existe algum quadro de informação para ser enviado naquele instante (todo o processamento nas estações podem ser considerados instantâneos, isto é, levam 0s). Se houver, ele manda a confirmação (N(R)= número do próximo quadro esperado) de carona no quadro de informação. Se não, ele espera no máximo 4s para mandar o N(R). Quadros do tipo RNR são enviados imediatamente, quando uma mensagem chega e não encontra mais buffers disponíveis. O "timeout" é de 8s. Nenhuma mensagem, por motivo algum, espera mais do que o "timeout" para ser retransmitida. Quando é necessário a retransmissão de várias mensagens, elas são retransmitidas uma após outra, com intervalo de 1s.
- Para cada situação, desenhe um diagrama ilustrando a situação (um eixo de tempo para A e outro para B, como exemplificado na figura) usando a seguinte convenção:

• I, N(S), N(R)	0s +	1,0,0	+ 0s
• RR, N(R)	+		+ 1s
• RNR, N(R),	+		+ 2s
• REJ, N(R)	+		+ 3s
	+		+ 4s
	+		+ 5s



PUC-Rio / DI

TeleMídia

Questões

- Situação 1: iniciado o sistema, suponha que mensagens cheguem para transmissão em A em 0s; 1s; 2s; 3s e 7s ; e em B em 8s. Considere ainda que o nó B consome todos os seus buffers de uma só vez no instante $t=5,5s$.
- Situação 2: iniciado o sistema, suponha que mensagens cheguem para transmissão em A em 0s; 1s; 6s; e 7s. Considere ainda que o nó B consome todos os seus buffers de uma só vez no instante $t=8,5s$.
- Situação 3: iniciado o sistema, suponha que mensagens cheguem para transmissão em A em 0s; 1s; 6s; e 8s. Considere que as mensagens emitidas pelo nó B entre 8,5s e 10s chegam com erro no nó A.



PUC-Rio / DI

TeleMídia

Questões

- ▶ Situação 4: iniciado o sistema, suponha que mensagens cheguem para transmissão em A em 0s; 1s; 2s; e 3s. Considere que as mensagens emitidas pelo nó A entre 0,5s e 1,5s chegam com erro no nó B.
- ▶ Situação 5: iniciado o sistema, suponha que mensagens cheguem para transmissão em A em 1s; 2s; 4s. Considere que as mensagens emitidas pelo nó A entre 1,5s e 3,5s chegam com erro no nó B. Considere também que as mensagens emitidas pelo nó B entre 4s e 7s chegam com erro no nó A.
- ▶ Situação n: Imagine vários cenários como os exemplificados anteriormente e trace os diagramas de mensagens correspondentes.



Gateways



Roteadores e Gateways

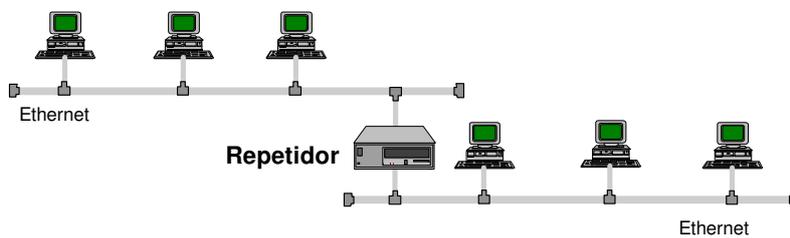
- ➔ Quando mensagens são deslocadas de uma rede para outra, conversões de protocolo se fazem necessárias. Esse procedimento é realizado pelos Gateways.
- ➔ Os gateways são classificados conforme o maior nível de protocolo convertido.



PUC-Rio / DI

TeleMídia

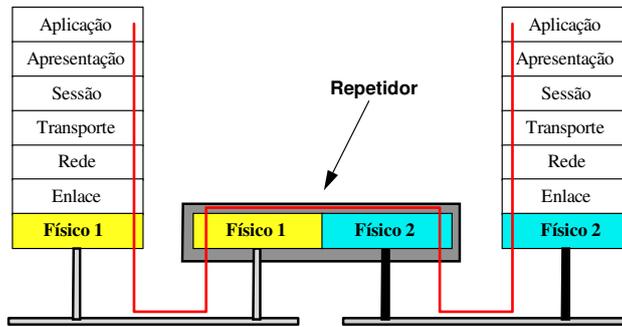
Repetidores



PUC-Rio / DI

TeleMídia

Repetidores



PUC-Rio / DI

TeleMídia

Repetidores

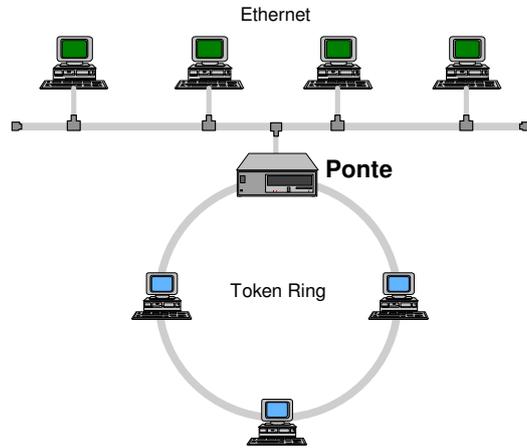
- Em redes em anel onde a estação é a responsável pela retirada dos próprios quadros, caberá ao repetidor a retirada dos quadros.
- Em redes que utilizam protocolos baseados em contenção, caberá ao repetidor também a função de detecção de colisão em um segmento, e a sinalização, no(s) outro(s) segmento(s), da ocorrência da colisão.
- Não pode haver um caminho fechado entre dois repetidores quaisquer da rede.
- Um tráfego extra inútil é gerado pelo repetidor quando os pacotes repetidos não se destinam às redes que interligam.



PUC-Rio / DI

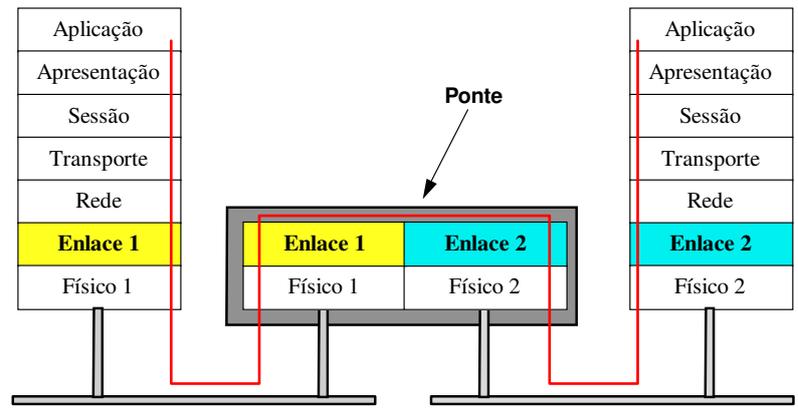
TeleMídia

Pontes



TeleMídia

Pontes



TeleMídia

Pontes

- ➔ Em redes em anel onde a estação é a responsável pela retirada dos próprios quadros, caberá à ponte a retirada dos quadros.
- ➔ No caso de haver caminhos fechados na rede, o número de duplicações de quadros é finito e igual ao número de caminhos fechados existentes entre as estações de origem e de destino, e por serem finitas, tais duplicações podem ser tratadas.
- ➔ Segmentação e remontagem. Não há solução; quadros grandes demais para serem encaminhados devem ser descartados.
- ➔ Ligação de redes que suportam diferentes esquemas de prioridade: o esquema é perdido.



PUC-Rio / DI

TeleMídia

O Padrão IEEE 802.1D - Pontes Transparentes

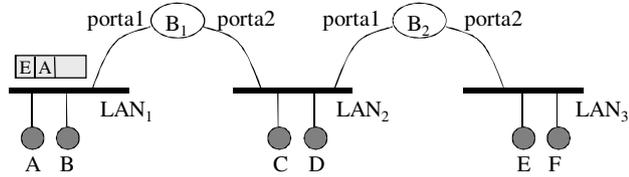
- ➔ A denominação transparente deve-se ao fato das LANs não sofrerem nenhuma modificação ao serem interconectadas por esse tipo de ponte.
- ➔ As pontes transparentes operam com base em uma *tabela de rotas* com pares: endereço de destino/porta de saída.
- ➔ As pontes transparentes operam em modo promíscuo. Ao receber um quadro, a ponte verifica na tabela de rotas se o endereço de destino está associado a uma porta diferente da porta de entrada.
 - Se o endereço não for encontrado na tabela, o quadro é retransmitido através de todas as portas exceto aquela através da qual ele entrou.
 - Se for encontrada na tabela de rotas uma associação do endereço de destino com uma porta diferente da de entrada, o quadro é retransmitido por essa porta.



PUC-Rio / DI

TeleMídia

Pontes Transparentes - Exemplo



ED EO
E|A Dados

Estação	Porta
A	1

B₁

Estação	Porta

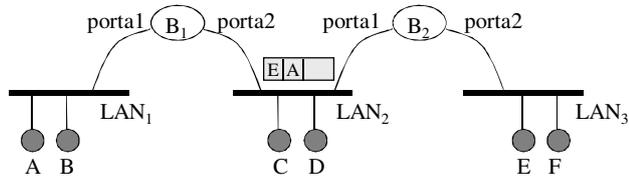
B₂



PUC-Rio / DI

TeleMídia

Pontes Transparentes - Exemplo



ED EO
E|A Dados

Estação	Porta
A	1

B₁

Estação	Porta
A	1

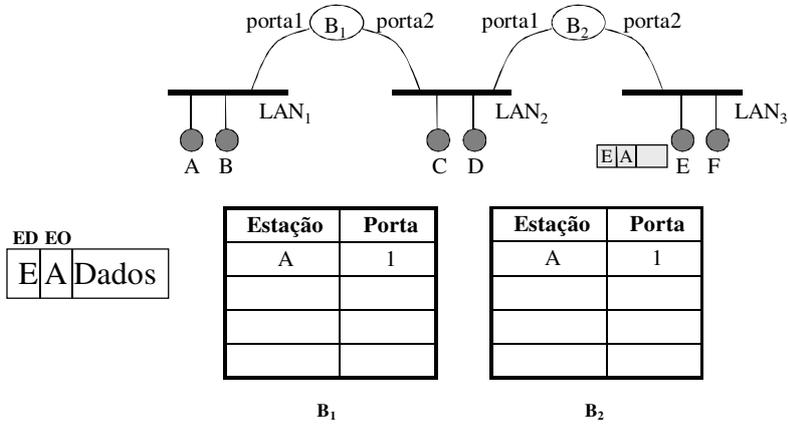
B₂



PUC-Rio / DI

TeleMídia

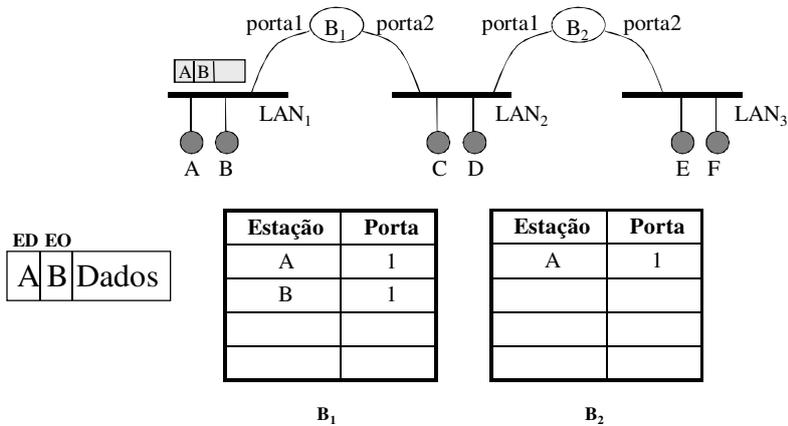
Pontes Transparentes - Exemplo



PUC-Rio / DI

TeleMídia

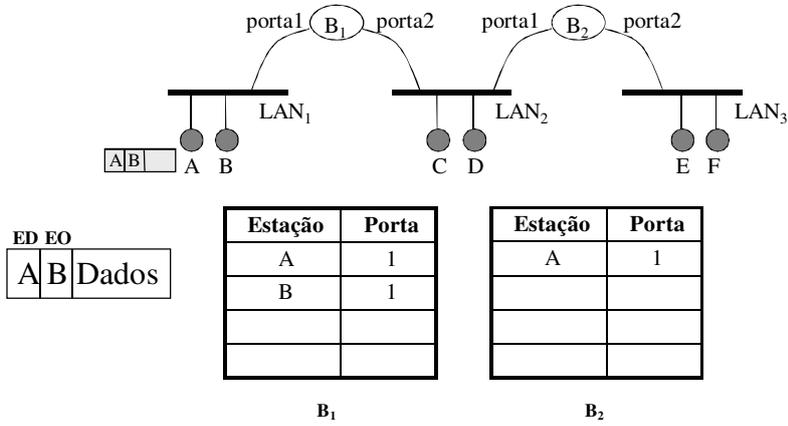
Pontes Transparentes - Exemplo



PUC-Rio / DI

TeleMídia

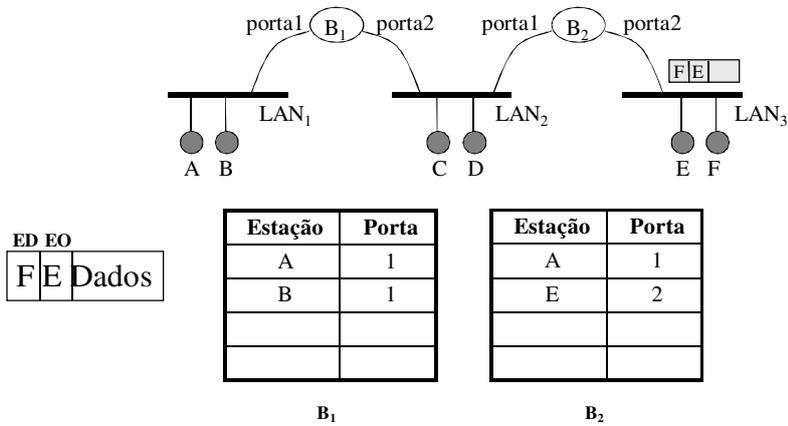
Pontes Transparentes - Exemplo



PUC-Rio / DI

TeleMídia

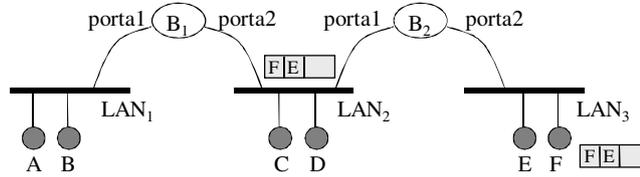
Pontes Transparentes - Exemplo



PUC-Rio / DI

TeleMídia

Pontes Transparentes - Exemplo



ED EO
F E Dados

Estação	Porta
A	1
B	1
E	2

B₁

Estação	Porta
A	1
E	2

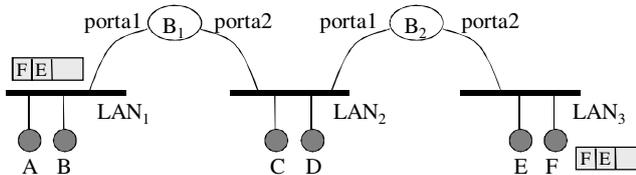
B₂



PUC-Rio / DI

TeleMídia

Pontes Transparentes - Exemplo



ED EO
F E Dados

Estação	Porta
A	1
B	1
E	2

B₁

Estação	Porta
A	1
E	2

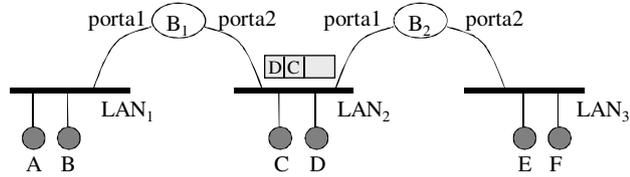
B₂



PUC-Rio / DI

TeleMídia

Pontes Transparentes - Exemplo



ED EO
D|C Dados

Estação	Porta
A	1
B	1
E	2
C	2

B₁

Estação	Porta
A	1
E	2
C	1

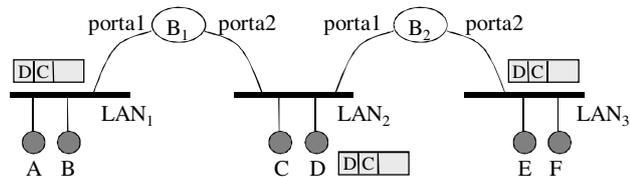
B₂



PUC-Rio / DI

TeleMídia

Pontes Transparentes - Exemplo



ED EO
D|C Dados

Estação	Porta
A	1
B	1
E	2
C	2

B₁

Estação	Porta
A	1
E	2
C	1

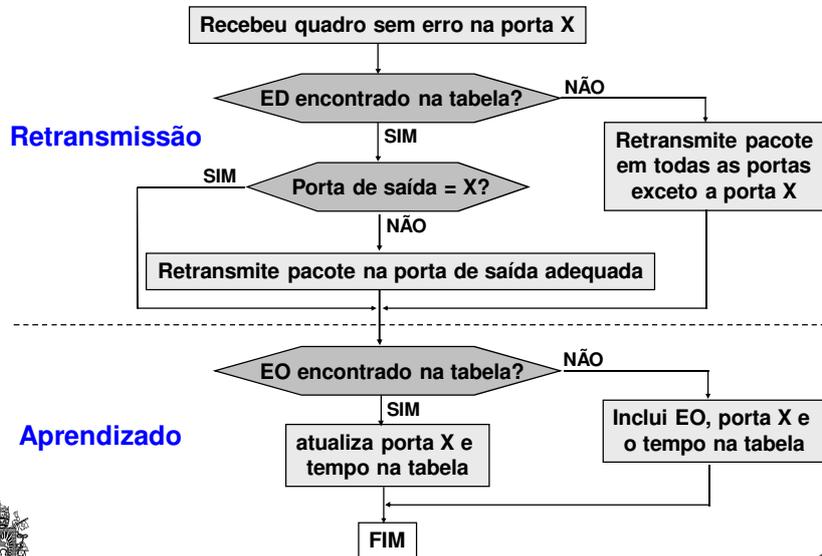
B₂



PUC-Rio / DI

TeleMídia

Pontes Transparentes



PUC-Rio / DI

TeleMídia

O Padrão IEEE 802.1D - Pontes Transparentes

- ➔ No início a tabela de rotas está vazia. Quando um quadro chega através de uma porta, seu endereço de origem é usado para atualizar a tabela de rotas. Se não existir nenhuma entrada na tabela com esse endereço, é criada uma entrada associando-o à porta através da qual ele entrou.
- ➔ Periodicamente, um processo varre a tabela e remove todas as entradas que não foram acessadas recentemente (devido a reconfiguração).
- ➔ O algoritmo usado pelas pontes é o de Baran, e funciona bem quando não existem rotas alternativas na inter-rede.



PUC-Rio / DI

TeleMídia

O Padrão IEEE 802.1D - Pontes Transparentes

- ➔ O problema é contornado com a utilização do seguinte resultado da teoria dos grafos: “para todo grafo conectado, consistindo em nós e arcos conectando pares de nós (as redes são os nós, e as pontes, os arcos do grafo), existe uma árvore de arcos que estende-se sobre o grafo (*spanning tree*) que mantém a conectividade do grafo, porém não contém caminhos fechados”.
- ➔ O padrão IEEE 802.1D define um algoritmo que deriva automática e dinamicamente a *spanning tree* de uma inter-rede.
- ➔ Periodicamente, cada ponte transmite uma mensagem por difusão em todas as redes às quais está conectada, com sua identificação. Um algoritmo distribuído seleciona a ponte que vai ser a raiz da árvore (por exemplo, a que possui o menor número de identificação). Uma vez determinada a raiz, a árvore é construída com cada ponte determinando o menor caminho para a raiz. Em caso de empate, é escolhido o caminho cuja ponte tiver o menor identificador.



PUC-Rio / DI

TeleMídia

Pontes com Roteamento na Origem

- ➔ A estação de origem escolhe o caminho que o quadro deve seguir e inclui a informação de roteamento no cabeçalho do quadro.
 - A informação de roteamento é construída da seguinte forma: cada LAN possui um identificador único, e cada ponte possui um identificador único no contexto das redes às quais está conectada.
 - Uma rota é uma seqüência de pares (identificador de rede, identificador de ponte). O primeiro bit do endereço de origem dos quadros cujo destino não está na mesma rede da estação de origem é igual a 1.
- ➔ Ao escutar um quadro cujo primeiro bit do endereço de origem é igual a 1, a ponte analisa a informação de roteamento do quadro. Se o identificador da LAN através da qual ele chegou é seguido pelo identificador da ponte em questão, ela retransmite o quadro na rede cujo endereço vem depois do seu identificador na informação de roteamento do quadro.



PUC-Rio / DI

TeleMídia

Gateways de Nível 3

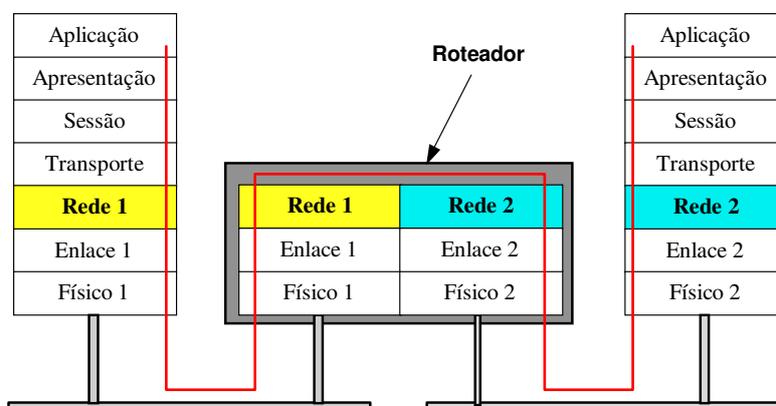
- Gateways conversores de meio: suas funções resumem-se em receber um pacote do nível inferior, tratar o cabeçalho inter-redes do pacote, descobrindo o roteamento necessário, construir novo pacote com novo cabeçalho inter-redes, se necessário, e enviar esse novo pacote ao próximo destino, segundo o protocolo da rede local em que este se encontra. Esse tipo de gateway da camada de rede é também chamado de *roteador*.
- Gateways tradutores de protocolos atuam traduzindo mensagens de uma rede, em mensagens da outra rede, com a mesma semântica de protocolo. Por exemplo, o *open* em uma rede poderia ser traduzido por um *call request* em outra ao passar pelo gateway.



PUC-Rio / DI

TeleMídia

Roteador

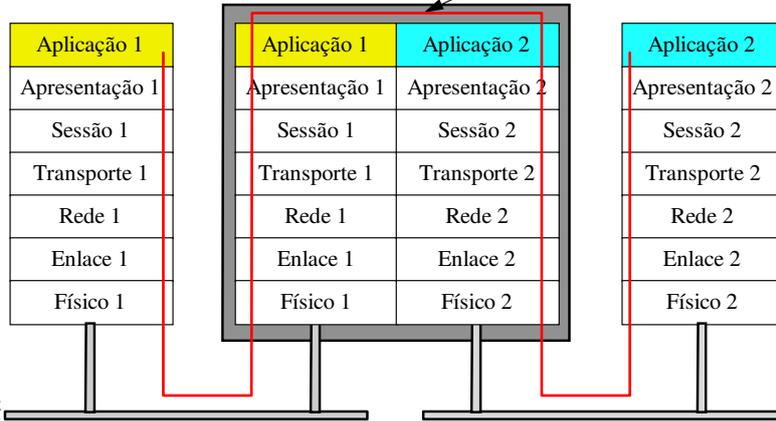


PUC-Rio / DI

TeleMídia

Gateway de Aplicação

Gateway converte o protocolo de Aplicação 1 no protocolo de Aplicação 2 e vice-versa



PUC-Rio / DI

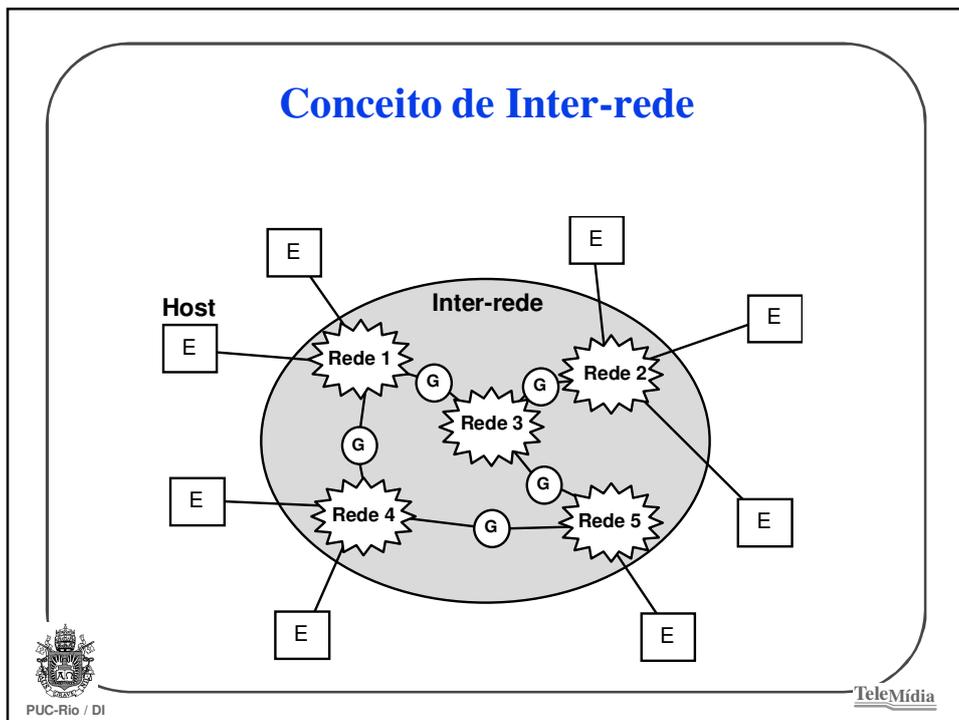
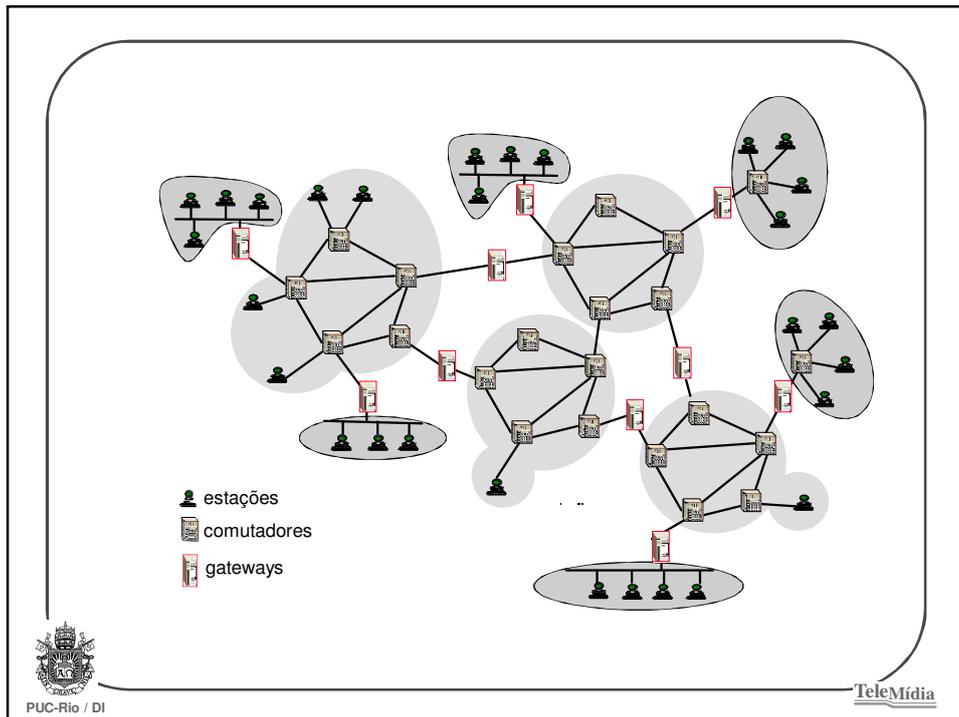
TeleMídia

Nível de Rede

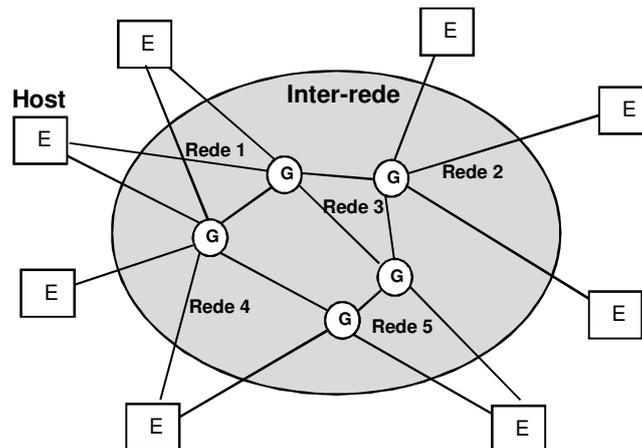


PUC-Rio / DI

TeleMídia



Conceito de Inter-rede



PUC-Rio / DI

TeleMídia

Organização Interna do Nível de Rede

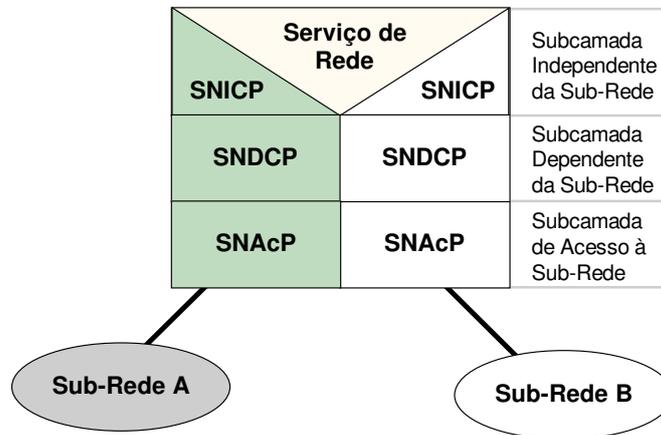
- ➔ Uma das principais funções do nível de rede é o roteamento, desde a estação origem até a estação de destino.
- ➔ Quando as funções do nível de rede são executadas através da combinação de várias sub-redes distintas, a especificação das funções de roteamento é facilitada pelo uso de *subcamadas*. As subcamadas são empregadas para separar as funções de roteamento das sub-redes individuais das funções de roteamento entre sub-redes (roteamento inter-redes).
- ➔ Por esse motivo, o nível de rede do RM-OSI [ISO 87a] foi dividido em três subcamadas.



PUC-Rio / DI

TeleMídia

Nível de Rede



PUC-Rio / DI

TeleMídia

Subcamada de Acesso a Sub-Rede

- A *subcamada de acesso à sub-rede* é dependente da tecnologia de sub-rede específica. As funções dessa subcamada incluem a transferência de dados dentro da sub-rede, incluindo as funções de roteamento e retransmissão de informações que trafegam dentro da sub-rede.
- Os protocolos de acesso à sub-rede (*Subnetwork Access Protocol* SNAcP) podem variar muito de uma rede para outra.
- Exemplos:
 - o protocolo X.25 de uma rede pública.
 - Em uma rede local o SNAcP é geralmente inexistente.



PUC-Rio / DI

TeleMídia

Subcamada Independente da Sub-Rede

- ➔ A *subcamada independente da sub-rede* é responsável pelo fornecimento do serviço de rede entre dois sistemas finais.
- ➔ O protocolo de convergência independente da sub-rede (*Subnetwork Independent Convergence Protocol* SNICP) só é usado quando os sistemas finais estão localizados em sub-redes distintas.
- ➔ A principal tarefa da subcamada independente da sub-rede é o roteamento entre redes.
- ➔ Exemplo de protocolo inter-rede: o protocolo IP da Internet.



PUC-Rio / DI

TeleMídia

Subcamada Dependente da Sub-Rede

- ➔ Resumindo, o SNACp torna possível a transferência de dados dentro de uma sub-rede e o SNICP, com base no serviço de rede OSI fornecido em sub-redes interligadas, realiza as adaptações necessárias à conexão de sistemas finais localizados em sub-redes distintas.
- ➔ A função da *subcamada dependente da sub-rede* é harmonizar sub-redes que oferecem serviços diferentes apresentando uma interface única para o SNICP.



PUC-Rio / DI

TeleMídia

Funções do Nível de Rede

- Multiplexação.
- Endereçamento.
- Mapeamento entre endereços.
- Roteamento.
- Segmentação e remontagem.
- Controle de erros .
- Controle de fluxo.
- Controle de congestionamento.



PUC-Rio / DI

TeleMídia

Funções do Nível de Rede

- As funções de controle de erro e de fluxo são similares às realizadas no nível de enlace, só que o controle é realizado em todo o caminho da subrede (no SNAcP) ou em todo o caminho inter-rede (no SNICP).
- A multiplexação também é semelhante ao realizado no nível de enlace e visa permitir que mais de um usuário de transporte utilize do mesmo ponto de acesso (SAP) de enlace.



PUC-Rio / DI

TeleMídia

Endereçamento

- ➔ Não importa a camada de protocolo que estamos falando, o endereçamento dos pontos de acesso a serviços de uma camada K (SAPs da camada K) deve ser completamente independente dos demais endereçamentos dos outros níveis de protocolo.
- ➔ Basicamente dois tipos de endereçamento são possíveis: o hierárquico e o horizontal.



Endereçamento Hierárquico

- ➔ No *endereçamento hierárquico* o endereço de uma entidade é constituído de acordo com os endereços correspondentes aos vários níveis de hierarquia de que ela faz parte.
- ➔ Um exemplo comum, no nível de rede, é um endereço de SAP de rede formado pelo número da sub-rede a que pertence, pelo número da estação dentro dessa sub-rede e pelo número da porta associada (SAP de sub-rede).
 - O IP (*Internet Protocol*) da Arpanet é um exemplo de utilização de endereço hierárquico, onde a identificação de um SAP de rede (único por estação) é formada pelo endereço da rede (sub-rede) e pelo endereço da estação.
 - O endereço hierárquico é também o método sugerido pelo ITU-T, através da recomendação X.121, para interconexão de redes públicas de pacotes. Nessa recomendação os endereços são números decimais formados por três campos, um código do país (três dígitos), um código para a rede (um dígito no máximo 10 redes) e um campo para o endereçamento dentro da rede (10 dígitos).



Endereçamento Horizontal

- ➔ No *endereçamento horizontal*, o endereço de uma entidade de protocolo não têm relação alguma com os vários níveis de hierarquia de que ela faz parte.
- ➔ Um exemplo comum desse tipo de endereçamento seriam os endereços globalmente administrados, constituídos pelo número de assinatura do usuário, como os utilizados pelo padrão IEEE 802 , para o nível de enlace.



Endereçamento

- ➔ Considerações sobre o roteamento parecem indicar vantagens na utilização de endereços hierárquicos, uma vez que estes contêm informações explícitas sobre o local onde se localizam as entidades, informações que podem ser usadas quando necessário.
- ➔ Já o endereço horizontal, por ser independente da localização, vai facilitar os esquemas de reconfiguração por permitir uma mobilidade das entidades sem renumeração das mesmas.



Mapeamento de Endereço

- ➔ O mapeamento de endereço é a função que determina o endereço do SAP de um nível a partir do endereço do SAP do nível superior a quem oferece seus serviços.
- ➔ O mapeamento pode ser direto através de uma tabela de translação, ou através de um conjunto de regras executadas por um protocolo de resolução de endereços.
- ➔ O protocolo ARP da rede Internet é um exemplo de protocolo de resolução de endereços do nível de rede (sub-rede) para o nível de enlace.



PUC-Rio / DI

TeleMídia

Tabela de Rotas

- ➔ A implementação do roteamento exige uma estrutura de dados que informe os possíveis caminhos e seus custos, a fim de que se possa decidir qual o melhor. Diversos métodos têm sido utilizados para a manutenção da estrutura de dados.
- ➔ No encaminhamento por rota fixa, a tabela (estrutura de dados) uma vez criada, não é mais alterada.
- ➔ No encaminhamento adaptativo, a rota é escolhida de acordo com a carga na rede. Nas tabelas de rotas, são mantidas informações sobre o tráfego (como por exemplo o retardo sofrido em um determinado caminho), que são consultadas para a escolha do caminho mais curto (por exemplo, o de menor atraso). As tabelas devem ser periodicamente atualizadas.



PUC-Rio / DI

TeleMídia

Atualização das Tabelas

As tabelas devem ser periodicamente atualizadas, podendo tal atualização ser realizada de vários modos:

- ➔ No modo centralizado cada nó envia a um ponto central da rede as informações locais sobre a carga. Essas informações são utilizadas pelo ponto central para o cálculo das novas tabelas, que são então enviadas aos demais nós.
- ➔ No modo isolado a atualização é realizada com base nas filas de mensagens para os diversos caminhos e outras informações locais.
- ➔ No modo distribuído cada nó envia periodicamente aos outros nós as informações locais sobre a carga na rede. Essas informações são utilizadas para o cálculo da nova tabela.



PUC-Rio / DI

TeleMídia

Atualização das Tabelas

- ➔ Existem diversos algoritmos para a computação do “caminho mais curto” entre dois nós da rede. A definição de “caminho mais curto”, depende da qualidade de serviço que se deseja.
- ➔ O comprimento de um caminho pode ser medido pelo número de saltos, isto é, o número de nós intermediários pelos quais deve passar o pacote até chegar ao destino. Outra medida é a distância geográfica em metros. Ainda outra medida poderia ser o retardo de transferência do pacote.
- ➔ Os vários algoritmos são implementados nos diversos protocolos de atualização da tabela de rotas, entre eles o RIP, o OSPF e o DVRP (da Internet).



PUC-Rio / DI

TeleMídia

Roteamento Centralizado

- ➔ No roteamento centralizado existe, em algum lugar da rede, um Centro de Controle de Roteamento (CCR) responsável pelo cálculo das tabelas de rotas.
- ➔ Problemas:
 - Para se adaptar a um tráfego mutável, o cálculo das tabelas deve ser realizado com muita frequência, o que pode exigir um processamento muito grande.
 - Tráfego elevado nas linhas que levam ao CCR.
 - Confiabilidade: uma falha no CCR é crítica.
 - Inconsistências: os nós receberão suas tabelas em tempos diferentes, devido a retardos diferentes sofridos pelas mensagens que as transportaram.



Roteamento Isolado

- ➔ Algoritmo de Baran: um nó ao receber um pacote tenta se livrar dele imediatamente pelo enlace que possui a fila mais curta no momento.
- ➔ Baran combinado com roteamento de rota fixa: até um certo limiar do tamanho das filas, o pacote deve ser encaminhado para a menor fila; depois de um certo limiar, deve ser utilizada a rota estática.



Roteamento Isolado - Baran

- ➔ Leva em conta que cada pacote deve incluir o nó de origem e um contador de saltos (nós intermediários por onde trafegou).
- ➔ Ao receber um pacote, o nó sabe a que distância o nó de origem está, a partir do enlace de chegada. Se na sua tabela o nó de origem está a uma distância maior, ele deve atualizar a tabela.
- ➔ Passado um determinado tempo, cada nó possuirá o caminho mais curto para qualquer outro nó.
- ➔ Como cada nó só registra a troca para melhor, de tempos em tempos ele deve reiniciar o processo, de forma que enlaces que caíam ou fiquem sobrecarregados não afetem a confiabilidade da tabela.
- ➔ Se as tabelas forem reiniciadas em um período pequeno, pacotes podem ter de ser transferidos por rotas desconhecidas. Se forem reiniciadas em um período longo, os pacotes podem ser transmitidos em rotas congestionadas ou com enlaces em falha.



PUC-Rio / DI

TeleMídia

Roteamento Isolado

- ➔ Ainda um outro algoritmo de roteamento isolado, simples porém de grande custo devido à quantidade de pacotes gerados, exige que um pacote ao chegar em um nó da rede, seja enviado por todos os enlaces de saída do nó, exceto aquele por onde chegou (inundação).



PUC-Rio / DI

TeleMídia

Roteamento Distribuído

- Quando um nó quer descobrir o caminho mais curto (aqui exemplificado pelo caminho de menor retardo) até um nó de destino, ele envia um pacote de requisição por difusão (em todos os seus enlaces) na rede, com um campo contendo o endereço do nó de destino e um campo contendo um caminho, inicialmente com seu endereço.
- Cada nó intermediário ao receber o pacote de difusão verifica se o campo caminho contém seu endereço. Se não contiver, o nó acrescenta ao caminho seu endereço e o difunde por todos os seus enlaces. Em caso contrário, simplesmente descarta o pacote.
- O primeiro pacote recebido pelo nó de destino contém no campo caminho a rota de menor retardo. O nó de destino envia então, como resposta ao pacote recebido, um outro pacote contendo essa rota (por exemplo, através da rota reversa ou uma rota que no momento lhe pareça a de menor retardo entre o destino/nova origem e a origem/novo destino).
- Um nó deve periodicamente enviar pacotes de difusão para descobrimento das rotas. Quanto menor o período, mais adaptável às flutuações de tráfego é o algoritmo. Quanto maior o período, menor é o tráfego que o algoritmo gera na rede. Uma solução de compromisso deve ser encontrada.



PUC-Rio / DI

TeleMídia

Roteamento Hierárquico

- No roteamento hierárquico os nós são divididos em regiões, com cada nó capaz de manter as informações de rotas das regiões a que pertence.
- Devemos notar que a subdivisão da camada de rede no RM-OSI torna natural separar cada rede como uma região.
- Para muitas redes, no entanto, essa hierarquia em dois níveis é insuficiente, podendo ser necessário agrupar regiões em super-regiões e assim sucessivamente.



PUC-Rio / DI

TeleMídia

Segmentação e Remontagem

- ➔ Os protocolos inter-redes foram projetados para interligar redes das mais diversas tecnologias. Infelizmente, o tamanho máximo permitido para os pacotes varia de uma tecnologia de sub-rede para outra. Dessa forma pode se fazer necessário a quebra (segmentação) e remontagem de pacotes ao se passar de uma rede para outra.



PUC-Rio / DI

TeleMídia

Controle de Congestionamento

- ➔ Quando temos pacotes em excesso em uma sub-rede, ou parte de uma sub-rede, o desempenho da rede se degrada e dizemos que temos um *congestionamento*.
- ➔ Os congestionamentos podem ser causados por vários fatores: se os nós da rede são lentos, se o tráfego de entrada exceder a capacidade das linhas de saída em um nó, etc.
- ➔ Existem vários algoritmos de controle de congestionamento que podem ser implementados em uma sub-rede. Eles são subdivididos em procedimentos preventivos (aqueles que tentam evitar o congestionamento) e reativos (aqueles que operam uma vez detectada uma condição de congestionamento).



PUC-Rio / DI

TeleMídia

Controle de Congestionamento versus Controle de Fluxo

- ➔ O controle de congestionamento é um controle global envolvendo todos os recursos da rede. Já o controle de fluxo se relaciona com o tráfego entre um transmissor e um determinado receptor.
- ➔ Algumas redes tentam evitar o congestionamento pelo controle de fluxo, embora seja impossível controlar a quantidade total de tráfego na rede usando regras de controle de fluxo fim a fim.
- ➔ O problema com o controle do fluxo é que ele não pode ser ajustado pela taxa média do tráfego gerado, pois isso limitaria uma aplicação cujo tráfego possui uma taxa variável, resultando em um serviço ruim nos momentos de pico do tráfego. Por outro lado, se o controle de fluxo estiver ajustado para permitir que a taxa de pico do tráfego seja aproximada, ele tem pouco valor como controle de congestionamento.



PUC-Rio / DI

TeleMídia

Controle de Congestionamento - Descarte de Pacotes

- ➔ Em uma rede com serviço de circuito virtual uma cópia do pacote deve ser mantida em algum lugar de modo a poder ser retransmitida. Neste caso uma situação de impasse pode ocorrer.
- ➔ O descarte indiscriminado de pacotes pode levar a situações onde estamos descartando exatamente uma confirmação que liberaria um buffer do nó. Uma melhora do algoritmo pode ser conseguida reservando permanentemente um buffer por linha de entrada, a fim de permitir que todos os pacotes que cheguem sejam inspecionados.
- ➔ Buffer na entrada: *bloqueio de head of line*.
- ➔ De forma a minimizar a banda passante desperdiçada no descarte de pacotes, podemos descartar o pacote da fila que viajou uma distância menor. Esta poderia ser uma outra variante no algoritmo de descarte.



PUC-Rio / DI

TeleMídia

Controle de Congestionamento - Controle Isorrítmico

- Evita o congestionamento pela limitação do número de pacotes em trânsito na sub-rede.
- Nesse método existem permissões que circulam pela sub-rede. Sempre que um nó deseja introduzir um novo pacote na sub-rede ele deve primeiro capturar uma permissão e destruí-la. A permissão é regenerada pelo nó que recebe o pacote no destino.
- Problemas:
 - Distribuição das permissões depende das aplicações na rede.
 - O passeio aleatório das permissões causa um tráfego extra na sub-rede, roubando seu desempenho.
 - A perda de uma permissão por uma falha qualquer deve ser recuperada, sob pena de diminuirmos a capacidade de transporte da rede.



PUC-Rio / DI

TeleMídia

Controle de Congestionamento - Controle da Taxa

- FECN
- BECN



PUC-Rio / DI

TeleMídia

Tipos de Serviço

- ➔ O serviço oferecido pela inter-rede, bem como o serviço oferecido pelas sub-redes pode ser com conexão ou sem conexão, com características similares aos serviços citados para o nível de enlace.
- ➔ É importante salientarmos que o serviço oferecido pela inter-rede é um assunto separado do serviço de cada rede.



Serviço Sem Conexão

- ➔ Nos *serviços sem conexão* (serviços de datagrama) uma mensagem é tratada de forma individual e entregue ao destino através do caminho mais conveniente, definido pelos algoritmos de roteamento.
- ➔ A inter-rede não dá garantias de entrega em seqüência dos pacotes e, muitas vezes, nem garantia de chegada de um pacote ao destino final (datagrama não confiável).
- ➔ A principal vantagem do serviço é sua simplicidade, deixando para níveis superiores o controle de erro, de seqüência e de fluxo, caso sejam necessários. O exemplo mais popular de protocolo que implementa este serviço é o IP da Internet.



Serviço Com Conexão

- ▶ Em um serviço com conexão (também chamado de serviço de circuito virtual), não importa se a nível de sub-rede ou inter-rede, todos os pacotes podem seguir a mesma rota, que é escolhida no momento do estabelecimento da conexão. Nesse caso a conexão é usualmente chamada de um circuito virtual. Temos então um serviço de circuito virtual com conexão de circuito virtual. Conexões de circuito virtual são utilizadas geralmente em sub-redes. O exemplo mais popular de protocolo que implementa este serviço é o nível de rede do X-25.
- ▶ Em uma outra opção, os pacotes de um serviço com conexão (serviço de circuito virtual) podem seguir rotas diferentes até o destino.



PUC-Rio / DI

TeleMídia

Serviço Com Conexão x Sem Conexão

- ▶ A defesa do serviço sem conexão é fortemente realizada pela comunidade da rede Internet.
- ▶ De qualquer forma, as estações deverão realizar em níveis superiores o controle de erro e de fluxo, havendo pouca vantagem em realizar essas funções duas vezes.
- ▶ Nem todas as aplicações necessitam da confiabilidade oferecida pelos serviços com conexão, pelo contrário, o atraso introduzido pela realização desses serviços pode ser até prejudicial, como no caso de transmissão de voz em tempo real.



PUC-Rio / DI

TeleMídia

Serviço Com Conexão x Sem Conexão

- A maior parte dos usuários não está disposta a executar protocolos de transporte complexos em suas máquinas, desejando um serviço confiável, suprido pelo serviço com conexão.
- Serviços com conexão é possível se fazer a pré-alocação de recursos (por exemplo, recursos de armazenamento nos gateways e nós), garantindo um melhor controle de congestionamento a nível inter-redes.



PUC-Rio / DI

TeleMídia

Perguntas

- Quando é necessário o uso de gateways?
- O que vem a ser um gateway de nível n?
- Como opera um gateway de nível 1? Como é usualmente chamado? Quais são suas principais características para os diversos tipos de redes?
- Como opera um gateway de nível 2? Como é usualmente chamado? Quais são suas principais características para os diversos tipos de redes?
- Como as pontes transparentes encaminham as mensagens recebidas?
- Como funciona o algoritmo de Baran? O que acontece se existem rotas alternativas para um mesmo destino? Como o problema pode ser resolvido?
- Como funciona uma ponte com roteamento na origem?
- Quais os dois tipos de gateways de nível 3?



PUC-Rio / DI

TeleMídia

Perguntas

- ▶ Dê quatro razões porque as redes são divididas em domínio? Como surge o conceito inter-rede?
- ▶ Porque o modelo OSI divide o nível de rede nas camadas inter-redes e sub-redes?
- ▶ Quando a camada sub-rede não existe?
- ▶ Qual a diferença entre os protocolos para inter-rede e sub-rede? Existe alguma diferença em termos de algoritmos de roteamento e encaminhamento?
- ▶ Quais as principais funções do nível de rede?
- ▶ O que é a multiplexação no nível de rede?
- ▶ O que é resolução de endereço? Quais as diversas formas de realizá-la?
- ▶ O endereço IP (SAP-inter-rede) é horizontal ou hierárquico? Se é hierárquico, por que é necessário a resolução de endereço para o nível de enlace?



PUC-Rio / DI

TeleMídia

Perguntas

- ▶ Quais as vantagens e desvantagens do endereçamento horizontal e hierárquico?
- ▶ Qual a diferença entre um encaminhamento por rota fixa e adaptativo?
- ▶ Cite as várias formas de roteamento adaptativo, comentando suas vantagens e desvantagens.
- ▶ O que vem a ser “caminho mais curto”?
- ▶ Descreva o algoritmo de Baran isolado.
- ▶ Dê um exemplo de algoritmo para roteamento distribuído.
- ▶ O que vem a ser roteamento hierárquico? Há limitação no número de níveis hierárquicos? Dê um exemplo.



PUC-Rio / DI

TeleMídia

Perguntas

- ▶ O que vem a ser congestionamento em uma rede?
- ▶ Qual a diferença do controle de congestionamento e o controle de fluxo?
- ▶ O que é o controle de congestionamento preventivo e reativo? Dê exemplos.
- ▶ Descarte de pacotes pode ser usado em controle de congestionamento preventivo?
- ▶ Quais os cuidados a serem tomados com descartes de pacotes?
- ▶ O que é bloqueio de “head of line”?
- ▶ Como funciona o controle isorrítmico?
- ▶ Descreva o funcionamento do FECN e do BECN.
- ▶ Faça uma discussão sobre os tipos de serviços a serem utilizados nas camadas de inter-rede e sub-rede.



PUC-Rio / DI

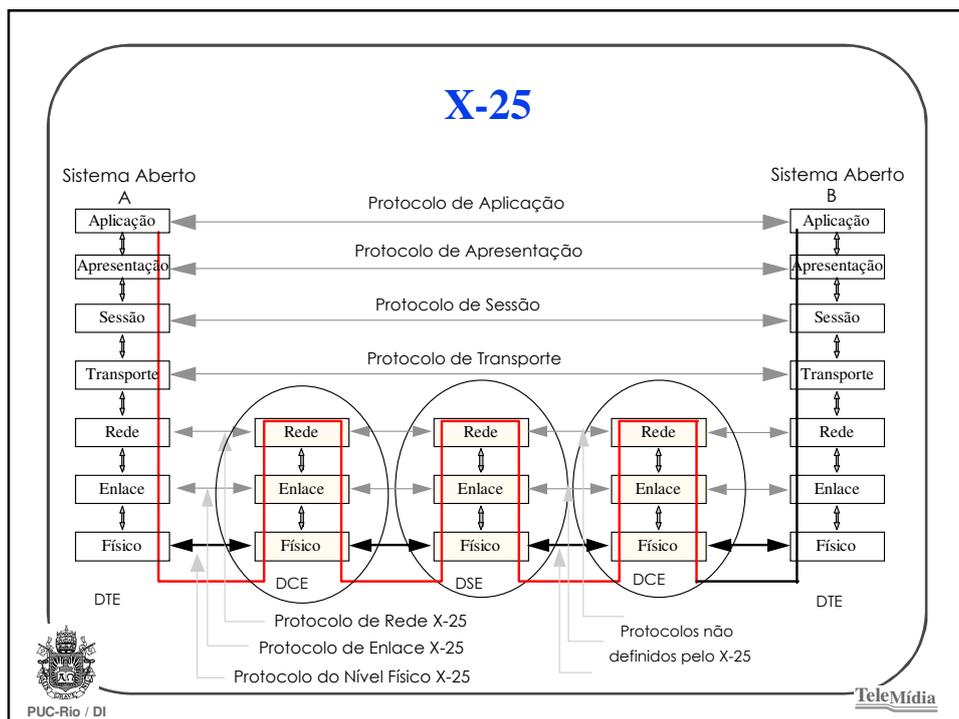
TeleMídia

X-25



PUC-Rio / DI

TeleMídia



- ## X-25
- No nível de enlace os protocolos utilizados são o LAP e o LAPB.
 - No nível de rede, é utilizado o serviço com conexão com circuito virtual.
 - Os dados trocados em uma conexão sofrem controle de fluxo e de erro, usando uma janela deslizante.
 - A confirmação de recebimento de um pacote pode ser de dois tipos, selecionados por um bit (bit D) do pacote de dados. Se D=0, a confirmação significa que o DCE recebeu o pacote e não o DTE remoto. Se D=1, a confirmação indica que o pacote foi entregue ao DTE remoto, representando, portanto, uma confirmação fim a fim.
- PUC-Rio / DI TeleMídia

Arquitetura Internet



PUC-Rio / DI

TeleMídia

Arquitetura Internet

- Pesquisa financiada pela DARPA (Defense Advanced Research Projects Agency) no início dos anos 70
- Arquitetura aberta que tornou-se um padrão de fato
- Evolução da ARPANET
- Arquitetura baseada no conceito de **inter-rede**

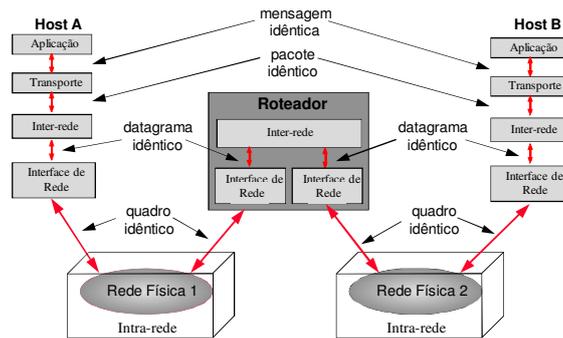


PUC-Rio / DI

TeleMídia

Arquitetura TCP/IP

- ➔ A arquitetura Internet TCP/IP é organizada em quatro camadas conceituais construídas sobre uma quinta camada que não faz parte do modelo, a camada intra-rede. A Figura abaixo mostra as camadas e o tipo de dados que é passado de uma para outra.



PUC-Rio / DI

TeleMídia

RM-OSI x Arquitetura TCP/IP

Arquitetura OSI



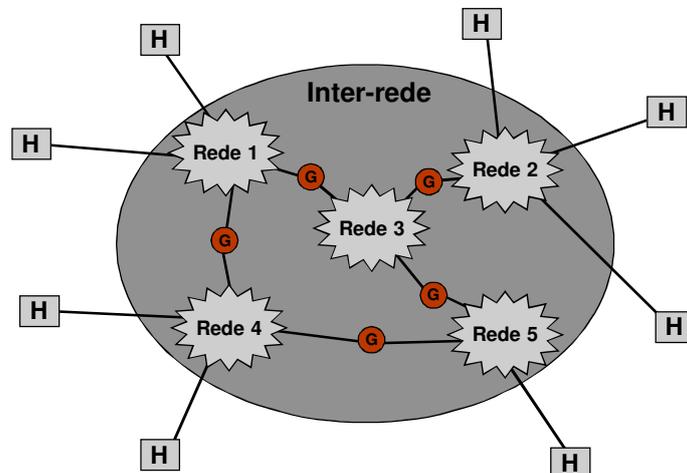
Arquitetura Internet



PUC-Rio / DI

TeleMídia

Conceito de Inter-rede



PUC-Rio / DI

TeleMídia

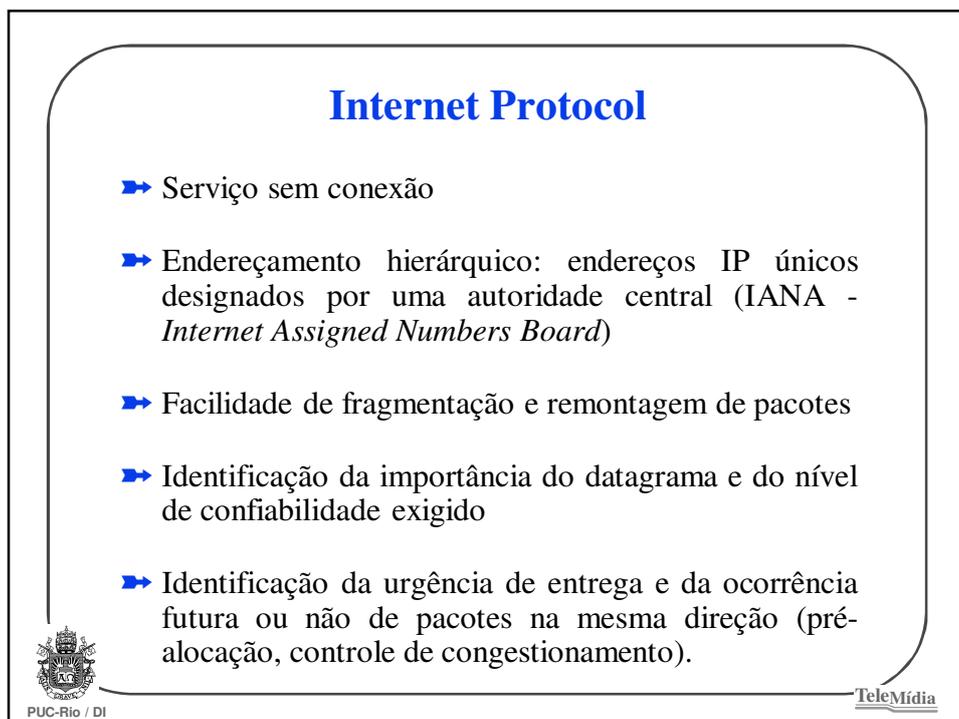
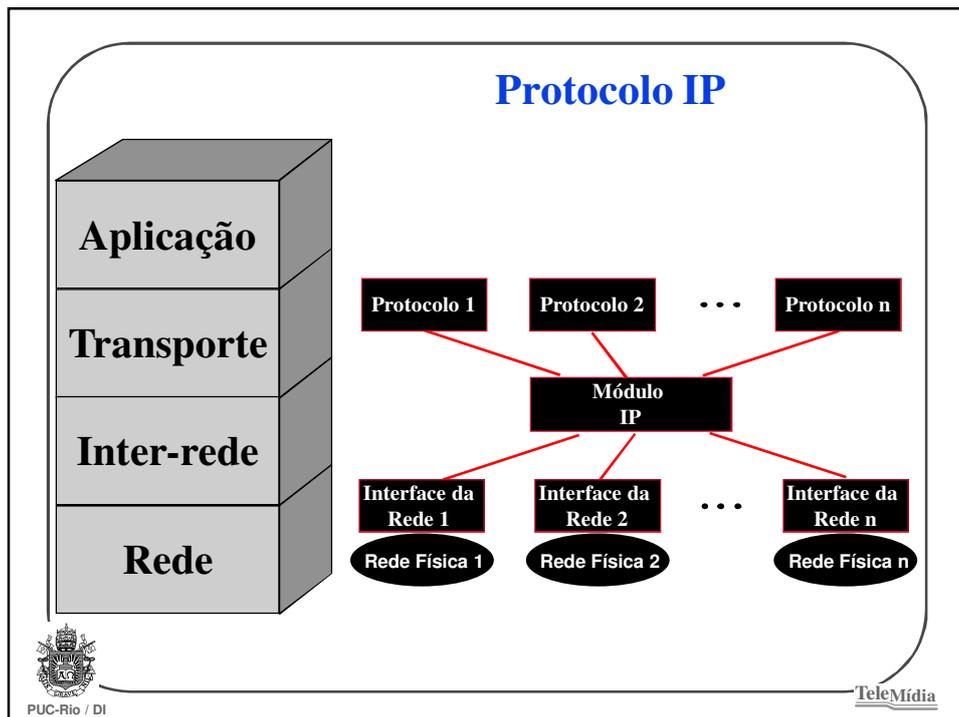
Internet Protocol

- Esquema de roteamento entre redes: roteamento adaptativo distribuído nos gateways
- Interconexão e roteamento através de *Gateways*
- Roteamento baseado no endereço da rede e não do *host*
- Estrutura interna da rede transparente
- Descarte e controle de tempo de vida dos pacotes inter-redes no gateway.
- Campo especial indicando qual o protocolo de transporte a ser utilizado no nível superior.



PUC-Rio / DI

TeleMídia



Esquema de Endereçamento IP v4



A.B.C.D

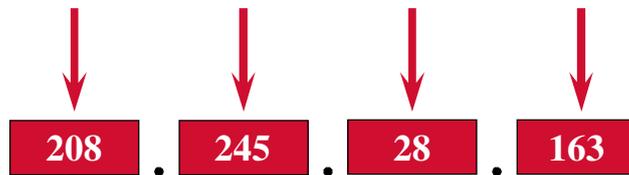


PUC-Rio / DI

TeleMídia

Notação de ponto decimal e hexadecimal

11010000 11110101 00011100 10100011



208.245.28.163

D0.F5.1C.A3



PUC-Rio / DI

TeleMídia

Esquema de Endereçamento IP v4

- Os endereços IP podem ser usados para nos referirmos tanto a sub-redes quanto a um host individual.
- Por convenção, um endereço de sub-rede tem o campo identificador de host com todos os bits iguais a 0.
- Podemos também nos referir a todos os hosts de uma sub-rede através de um endereço por difusão, quando, por convenção, o campo identificador de host deve ter todos os bits iguais a 1.
- Um endereço com todos os 32 bits iguais a 1 é considerado um endereço por difusão para a rede do host origem do datagrama.
- O endereço 127.0.0.0 é reservado para teste (loopback) e comunicação entre processos da mesma máquina.



PUC-Rio / DI

TeleMídia

Esquema de Endereçamento IP v4



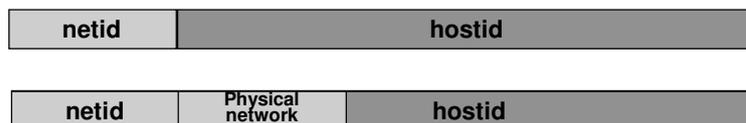
A.B.C.D



PUC-Rio / DI

TeleMídia

Hierarquização do Endereço



- ➔ A implementação é feita através de máscaras de 32 bits.



PUC-Rio / DI

TeleMídia

Problemas do Endereçamento IP v4

- ➔ O endereço se refere a conexão de rede e não à estação: se uma estação se mover de uma rede para outra, o endereço IP deve mudar.
- ➔ A mudança de classe de endereço (por exemplo pelo crescimento de uma rede) acarreta em grande problema administrativo.
- ➔ Pacotes podem viajar em rotas diferentes para uma estação com múltiplos endereços IP, dependendo do endereçamento usado.



PUC-Rio / DI

TeleMídia

Mapeamento de Endereços

➔ O mapeamento do endereço de um SAP de rede em um endereço de sub-rede (muitas vezes o endereço no nível de enlace se a rede não possui a subcamada de acesso à sub-rede do RM-OSI) pode ser feito através de:

- resolução através de mapeamento direto
- resolução através de vinculação dinâmica.



PUC-Rio / DI

TeleMídia

Mapeamento Endereço IP - MAC

Endereço de Classe D: 234.138.8.5 (EA-8A-08-05)

Endereço IP de Classe D	E	A	8	A	0	8	0	5
	1	1	1	0	1	0	1	0
	1	0	0	0	1	0	0	0
	0	0	0	0	0	1	0	0
	0	0	0	0	0	1	0	1

23 bits menos significativos mapeados ↓

Endereço Multicast MAC IEEE-802	0	1	0	0	5	E	0	A	0	8	0	5
	0	0	0	0	0	0	0	1	1	1	1	0
	0	0	0	0	0	0	0	0	1	0	1	0
	0	0	0	0	0	0	0	0	1	0	0	0
	0	0	0	0	0	0	0	0	1	0	0	1

■ Não mapeados



PUC-Rio / DI

TeleMídia

Mapeamento de Endereços

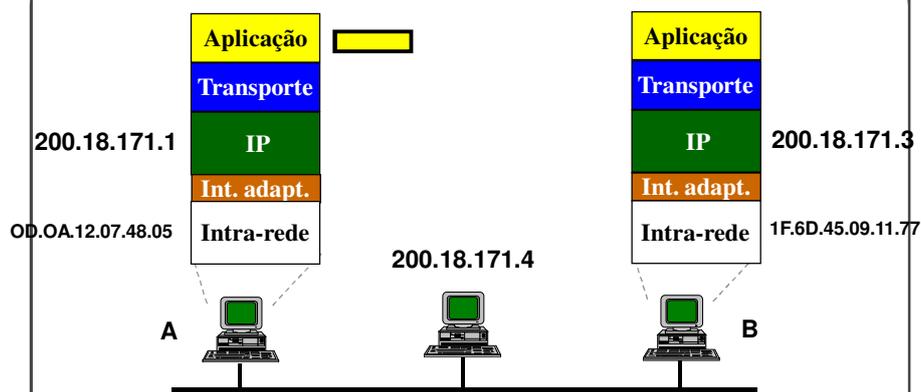
- ➔ A resolução através de vinculação dinâmica pode ser realizada através de um protocolo denominado ARP (*Address Resolution Protocol*).
- ➔ O ARP permite a um host encontrar o endereço de sub-rede de outro host na mesma sub-rede, através do envio de um datagrama por difusão na sub-rede, como acontece com as redes IEEE 802.
- ➔ A resolução pode ser feita também por um servidor de resolução, como nas redes IPOA.



PUC-Rio / DI

TeleMídia

Funcionamento do ARP

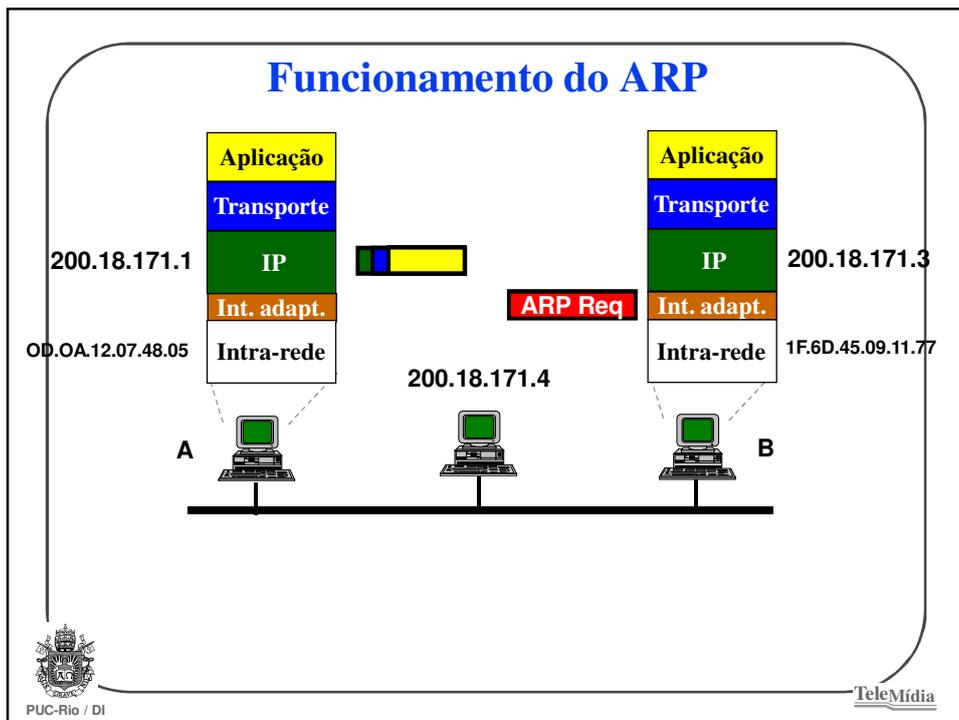
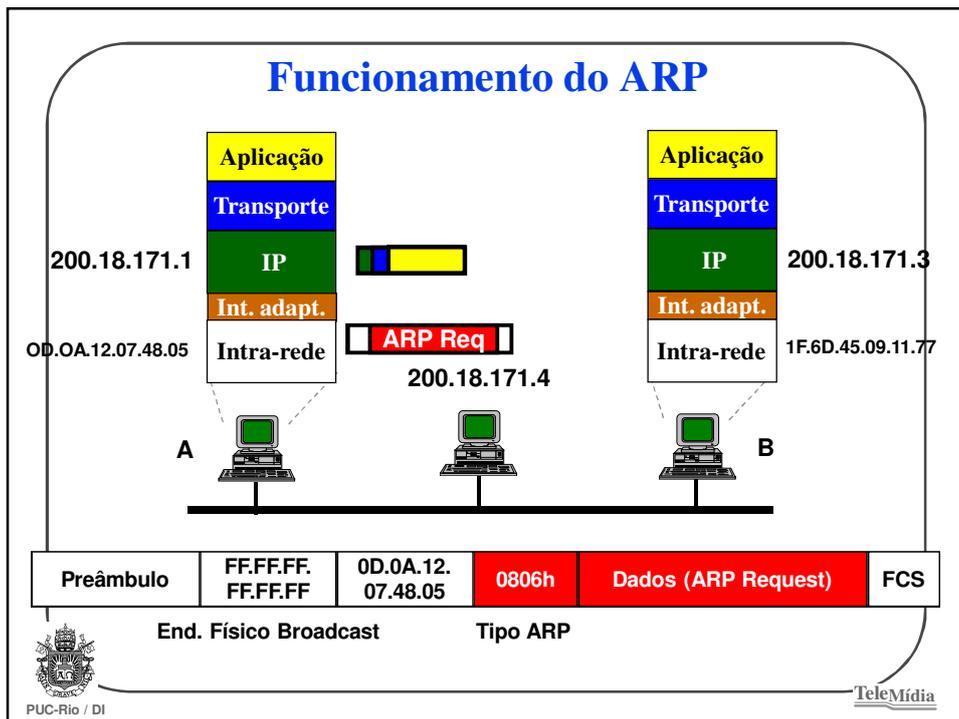


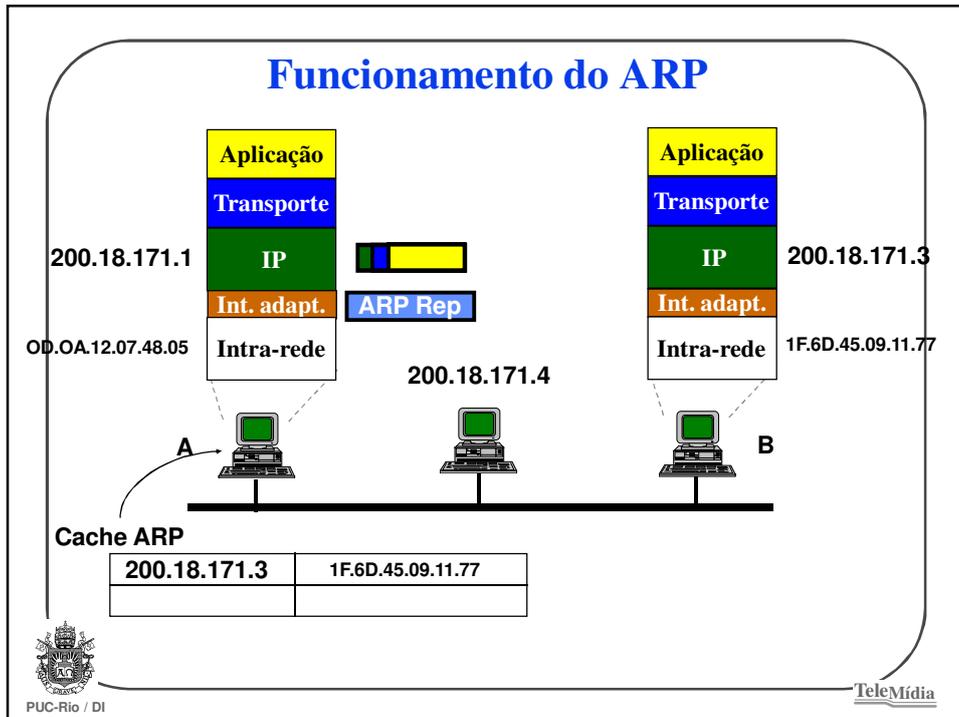
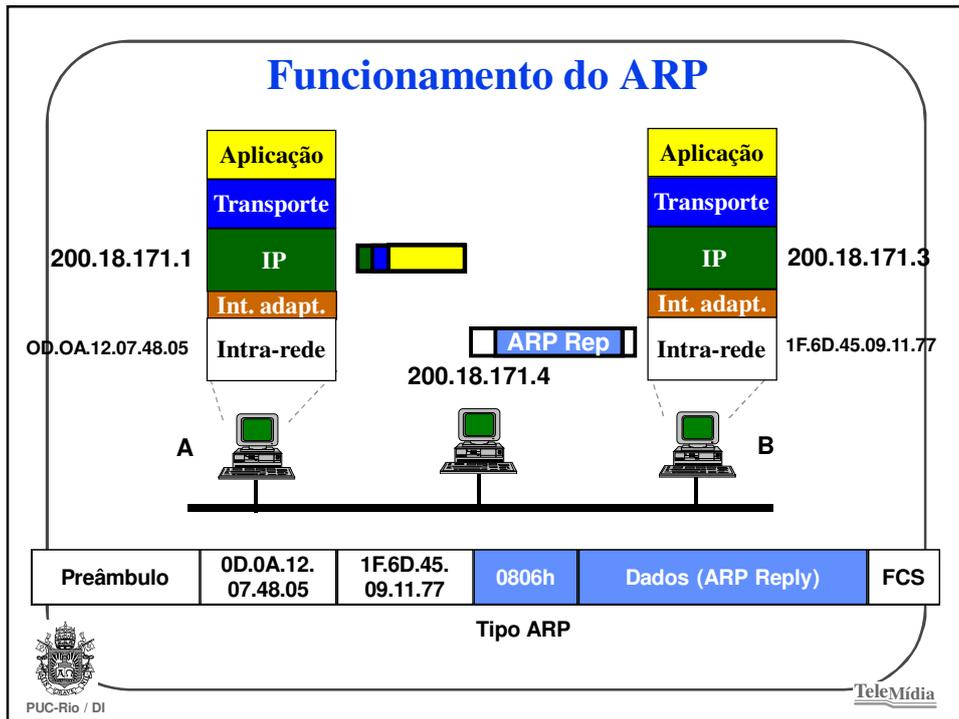
- ➔ Aplicação no host A quer enviar dados para aplicação no host B

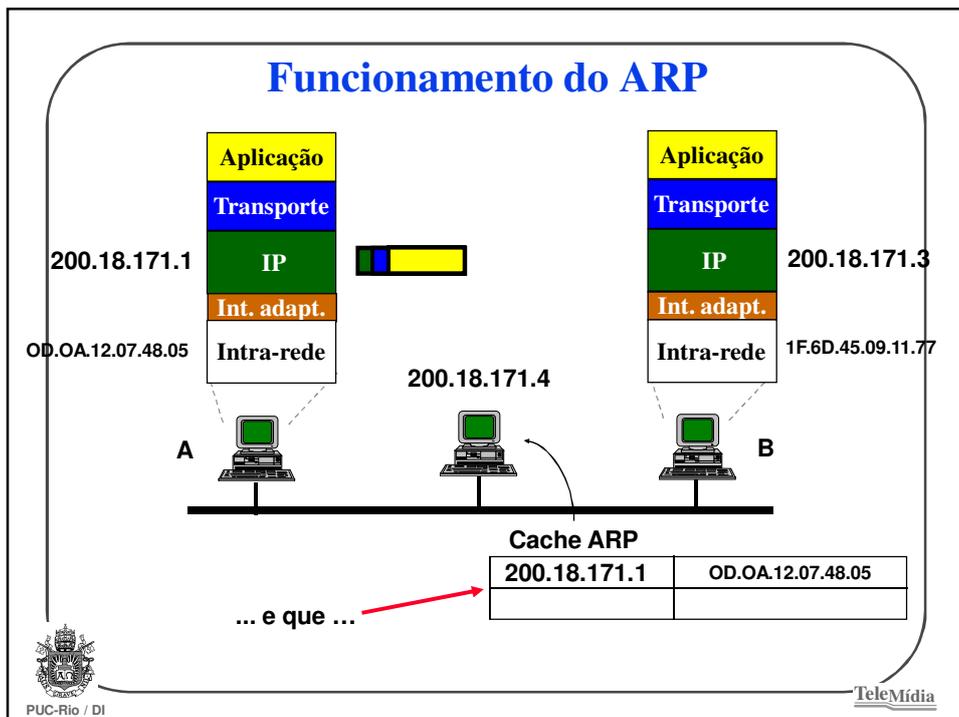
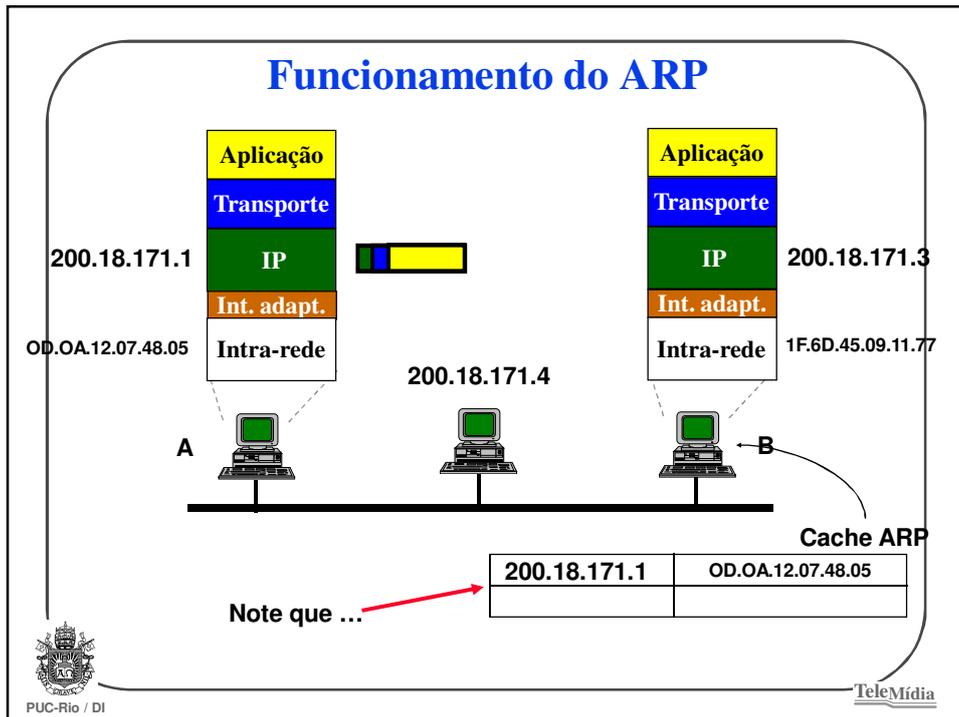


PUC-Rio / DI

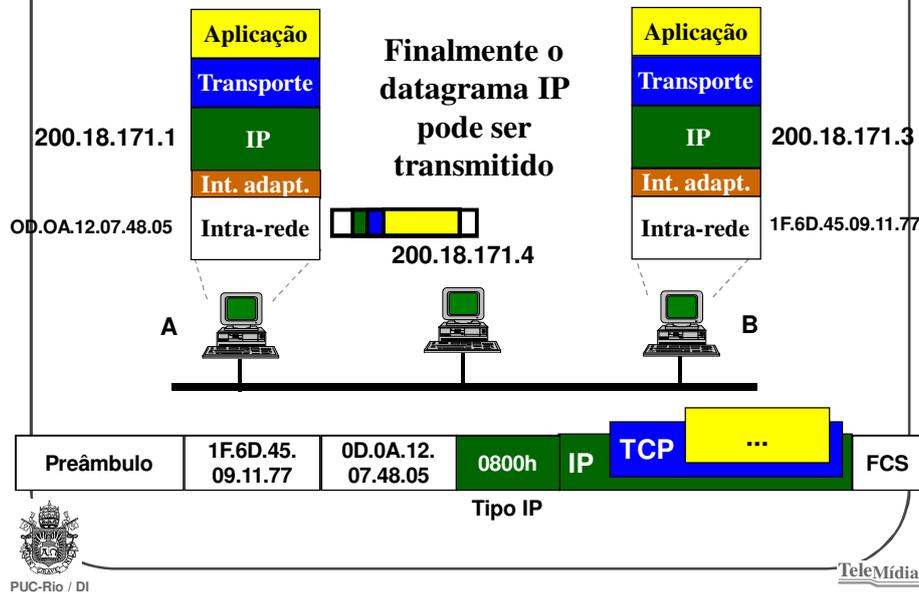
TeleMídia







Funcionamento do ARP



Mapeamento de Endereços

- ▶ Tomemos agora o problema inverso. Suponha que por algum motivo uma máquina não saiba seu endereço IP. Como ela poderia descobrir esse endereço?
 - O protocolo RARP (*Reverse Address Resolution Protocol*), adaptado do protocolo ARP, tem a resposta.
- ▶ A solução seria termos um servidor de endereços IP na rede.
 - Quando uma máquina quisesse saber seu endereço IP ela mandaria um quadro por difusão dentro de sua rede, que seria reconhecido como uma requisição pelo servidor. A partir do endereço de sub-rede, o servidor descobriria o endereço IP e o enviaria à máquina requisitante.
 - A comunicação poderia também se dar diretamente com o servidor.

Campos do Datagrama IP

VERS	HLEN	SERVICE TYPE	TOTAL LENGTH
IDENTIFICATION		FLAGS	FRAGMENT OFFSET
TIME TO LIVE	PROTOCOL	HEADER CHECKSUM	
SOURCE IP ADDRESS			
DESTINATION IP ADDRESS			
IP OPTIONS (IF ANY)		PADDING	
DATA			
...			

VERS = 4
HLEN = N x 4 octetos
(20 a 64 bytes)
TLEN = até 65535 octetos

PRECEDENCE	D	T	R	C	UNUSED
------------	---	---	---	---	--------

→ Baixo Custo (Low Monetary Cost)
→ Alta Confiabilidade (High Reliability)
→ Alta Vazão (High Throughput)
→ Baixo Retardo (Low Delay)
→ Prioridade do Datagrama: 0 - 7



PUC-Rio / DI

TeleMídia

Campos do datagrama IP

VERS	HLEN	SERVICE TYPE	TOTAL LENGTH
IDENTIFICATION		FLAGS	FRAGMENT OFFSET
TIME TO LIVE	PROTOCOL	HEADER CHECKSUM	
SOURCE IP ADDRESS			
DESTINATION IP ADDRESS			
IP OPTIONS (IF ANY)		PADDING	
DATA			
...			

TTL decrementado
em cada roteador
de 1 ou mais

TCP = 6
UDP = 17
ICMP = 1
EGP = 8
OSPF = 89

Alterado a cada
roteador



PUC-Rio / DI

TeleMídia

Campo FLAGS do Datagrama IP

VERS	HLEN	SERVICE TYPE	TOTAL LENGTH	
IDENTIFICATION		FLAGS	FRAGMENT OFFSET	
TIME TO LIVE	PROTOCOL	HEADER CHECKSUM		
SOURCE IP ADDRESS				
DESTINATION IP ADDRESS				
IP OPTIONS (IF ANY)			PADDING	
DATA				
...				

13 bits
Múltiplos de 8 octetos

TLEN informa tamanho do fragmento. Só último fragmento determina tamanho do datagrama original, através de TLEN e FRAGMENT OFFSET.

Identificador da Mensagem (não do datagrama)



More Fragments Bit

Do Not Fragment Bit (datagrama descartado e erro enviado para origem)

Não utilizado



PUC-Rio / DI

TeleMídia

Exemplo de Segmentação e Remontagem

➔ Datagrama original D_0 é dividido em dois novos datagramas: D_1 e D_2 .

➔ D_1 e D_2 têm conteúdo do cabeçalho idêntica ao do datagrama original.

- O valor do campo *identification* permanece inalterado. Esse campo, junto com os endereços de origem e destino, e com a identificação do protocolo de transporte, serve para especificar todos os fragmentos de um mesmo datagrama IP.
- O número de identificação só deve ser reutilizado depois de decorrido um tempo suficiente para o descarte do datagrama que utilizou anteriormente a mesma identificação.



PUC-Rio / DI

TeleMídia

Exemplo de Segmentação e Remontagem

- Os dados carregados pelo datagrama original, D_0 , são divididos em blocos de 8 octetos (64 bits).
- O primeiro datagrama, D_1 , é montado com um número inteiro de blocos de dados, denominado NFB (número de blocos de dados do fragmento).
- No cabeçalho de D_1 , o campo *total length*, que contém o valor copiado do datagrama original, é atualizado passando a carregar o comprimento do datagrama que acabou de ser criado.
- Ainda no datagrama D_1 , o *flag MF* do cabeçalho recebe o valor 1, especificando *more-fragments*.



PUC-Rio / DI

TeleMídia

Exemplo de Segmentação e Remontagem

- A segunda parte dos dados é colocada no outro datagrama, D_2 .
- O campo *total length* desse datagrama também é atualizado, passando a carregar o número de octetos desse datagrama (cabeçalho + dados).
- O *flag more fragments* mantém o valor copiado do datagrama D_0 , pois ele próprio pode ter sido anteriormente fragmentado.
- O campo *fragment-offset* recebe a soma do valor que tinha em D_0 com o valor NFB (que indica o número de blocos de dados transportado no fragmento anterior do datagrama D_1).
 - O valor do campo *fragment-offset* do datagrama original é utilizado na soma porque, novamente, o datagrama original pode já ter passado por um processo de fragmentação anterior.
 - Se fossem necessários N datagramas, o NFB usado deveria indicar o número de blocos carregados nos $N-1$ datagramas anteriores.



PUC-Rio / DI

TeleMídia

Exemplo de Segmentação e Remontagem

- Para remontar o datagrama original, o módulo IP destinatário combina os datagramas IP que possuem o mesmo valor para os campos *identification*, *protocol*, *source address* e *destination address*.
- A recombinação é feita através da colocação da porção de dados de cada fragmento na posição relativa indicada pelo valor do campo *fragment-offset* de seus cabeçalhos.
- O primeiro fragmento possui o *fragment-offset* igual a 0, e o último fragmento tem o *flag more fragments* igual a 0.



PUC-Rio / DI

TeleMídia

Campo IP OPTIONS

VERS	HLEN	SERVICE TYPE	TOTAL LENGTH	
IDENTIFICATION		FLAGS	FRAGMENT OFFSET	
TIME TO LIVE	PROTOCOL	HEADER CHECKSUM		
SOURCE IP ADDRESS				
DESTINATION IP ADDRESS				
IP OPTIONS (IF ANY)			PADDING	
DATA				
...				

1 octeto	1 octeto	1 octeto	1 octeto	1 octeto
Option Code	Length	Data 1	...	Data n



PUC-Rio / DI

TeleMídia

Opções IP

Código	Comp.	Descrição
0	-	Fim da lista de opções.
1	-	No operation .
2	11	Restrições de segurança e manipulação (militar).
7	variável	Record Route (Utilizado para rastrear a rota que o datagrama percorreu).
8	4	SATNET stream identifier (Obsoleto).
68	variável	Internet Timestamp (Utilizado para obter de cada roteador a hora de transmissão do datagrama).
131	variável	Loose Source Routing (Especifica a rota que um datagrama deve seguir).
137	variável	Strict Source Routing (Especifica a rota que um datagrama deve seguir).



PUC-Rio / DI

TeleMídia

Roteamento IP

- Cada computador possui uma tabela cujas entradas são pares: endereço de sub-rede/endereço de roteador. Essa tabela é denominada *tabela de roteamento IP*.
- Quando um módulo IP, por exemplo, executando em um roteador, tem que encaminhar um datagrama, ele inicialmente verifica se o destino do datagrama é um host conectado à mesma rede que seu hospedeiro.
- Se este for o caso, o datagrama é entregue à interface da rede que se encarrega de mapear o endereço IP no endereço físico do host, encapsular o datagrama IP em um quadro da rede, e finalmente transmiti-lo ao destinatário.



PUC-Rio / DI

TeleMídia

Roteamento IP

- ▶ Se a rede identificada no endereço de destino do datagrama for diferente da rede onde está o módulo IP, ele procura em sua tabela de roteamento uma entrada com o endereço de rede igual ao do endereço de destino do datagrama, recuperando assim o endereço do roteador que deve ser usado para alcançar a rede onde está conectado o destinatário do datagrama.
- ▶ O roteador recuperado da tabela pode não estar conectado diretamente à rede de destino, porém, se este for o caso, ele deve fazer parte do caminho a ser percorrido para alcançá-la.



PUC-Rio / DI

TeleMídia

Roteamento IP

- ▶ Uma desvantagem do esquema de roteamento apresentado é o tamanho das tabelas de roteamento, quando a inter-rede interliga um número muito grande de redes individuais..
- ▶ Uma alternativa é a utilização de *roteadores default*. Por exemplo, em redes que são ligadas à inter-rede por um único roteador, não é necessário ter uma entrada separada na tabela de rotas para cada uma das redes distintas da inter-rede. Simplesmente define-se o roteador como caminho default.
- ▶ O conceito de roteador default também se aplica a redes que possuem mais de um roteador. Nesse caso, quando não for encontrada uma rota específica para um datagrama, ele é enviado ao roteador default.



PUC-Rio / DI

TeleMídia

Roteamento IP

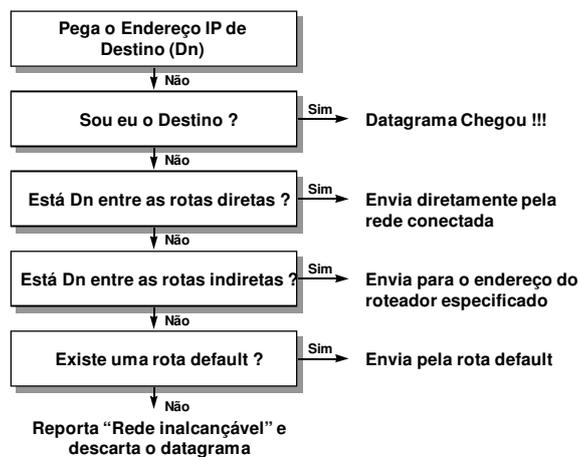
- ▶ Através de um protocolo específico ICMP (*Internet Control Message Protocol*) o roteador informará ao módulo IP se ele não é a melhor escolha para alcançar uma determinada rede.
- ▶ Essa mensagem, chamada *redirect* no ICMP, carrega como parâmetro o endereço do roteador que é a escolha correta.
- ▶ O módulo IP, ao receber uma mensagem *redirect*, adiciona uma entrada em sua tabela de roteamento associando a rede de destino ao endereço do roteador recebido na mensagem.



PUC-Rio / DI

TeleMídia

Encaminhamento IP



PUC-Rio / DI

TeleMídia

IP v6



Necessidades e Motivações

- ▶ Rápido crescimento da Internet (tamanho e carga)
 - Exaustão do espaço de endereçamento de 32 bits
- ▶ Segurança
 - Demanda crescente por comunicação segura
- ▶ Novas tecnologias de comunicação
- ▶ Novas aplicações
 - Necessidade de transmissão de dados para múltiplos usuários
 - Demanda por protocolos eficientes na comunicação de dados multimídia



Características Principais do IPv6

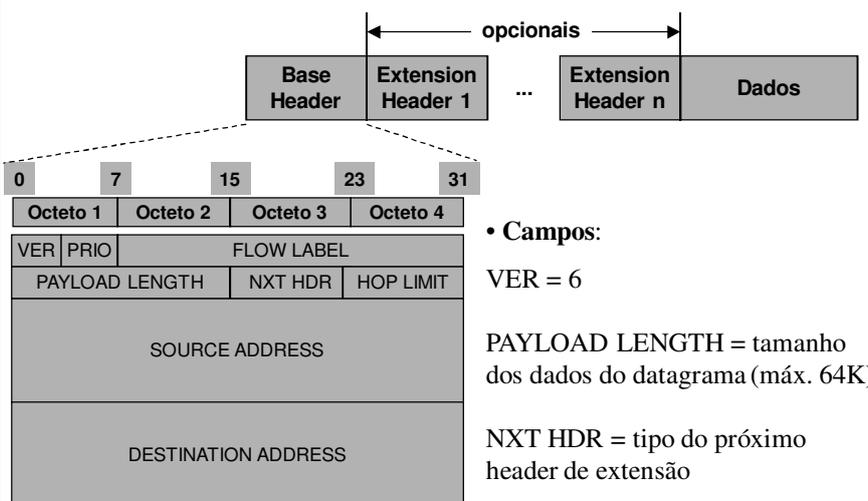
- ▶ Aumento do espaço de endereços de 32 para 128 bits
- ▶ Header básico mais simples para melhorar o desempenho do roteamento
- ▶ Melhoria na especificação de opções (headers de extensão)
- ▶ Suporte a rótulos de tipo de fluxo (Flow Labels)
 - Tratamento de requisitos especiais de QoS (áudio, vídeo, etc...)



PUC-Rio / DI

TeleMídia

Formato do Datagrama IPv6



• Campos:

VER = 6

PAYLOAD LENGTH = tamanho dos dados do datagrama (máx. 64K)

NXT HDR = tipo do próximo header de extensão

HOP LIMIT \equiv TTL em IPv4



PUC-Rio / DI

TeleMídia

Formato do Datagrama IPv6

➔ Campo PRIO: classe de prioridade

- 0 - não caracterizado (normal)
- 1 - baixíssima prioridade
- 2 - Transferência de dados não interativa (ex. SMTP)
- 4 - Transferência de dados interativa (ex. FTP, HTTP)
- 6 - Tráfego interativo (ex. TELNET, X-Windows)
- 7 - ICMP, protocolos de roteamento, SNMP
- 8 a 15 - tráfego sem controle de congestionamento (áudio, vídeo, ...)

➔ Campo FLOW LABEL

- Valor que determina a forma de tratamento do datagrama IP.
 - Especificado pelo protocolo que está controlando a QoS.
- A associação {FLOW LABEL, QoS} pode ser feita por um outro protocolo (ex: RSVP) ou datagrama a datagrama através de um header de extensão



PUC-Rio / DI

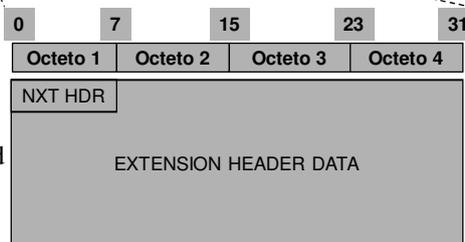
TeleMídia

Formato do Datagrama IPv6



NXT HDR:

- 0 = Hop-by-Hop Option Header
- 43 = Routing Extension Header
- 44 = Fragment Header
- 50 = Encapsulating Security Payload
- 51 = Authentication Header
- 60 = Destination Option Header
- 59 = No Next Header



PUC-Rio / DI

TeleMídia

Endereçamento IPv6

0110 1000	1110 0110	1000 1100	1111 1111	0000 0000	...	0000 0000	0000 1010
68	EA	8C	FF	00	...	00	0A

68EA:8CFF:00:00:00:00:0A
68EA:8CFF:0:0:0:0:A
68EA:8CFF::A

} Números de 16 bits

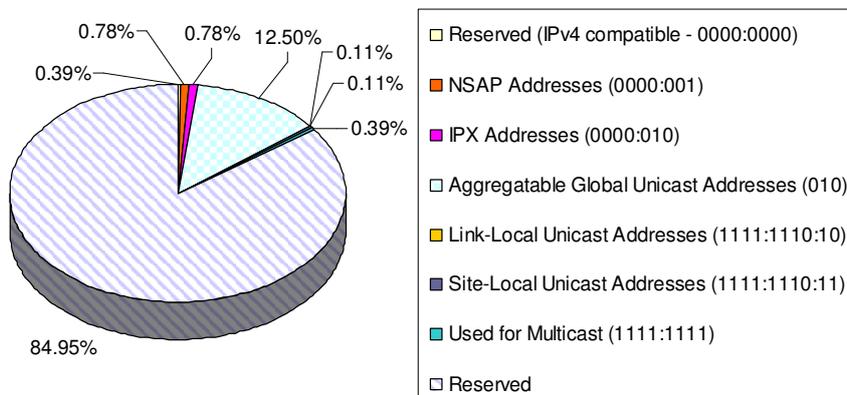
0:0:0:0:0:128.10.2.1
::128.10.2.1



PUC-Rio / DI

TeleMídia

Classes de Endereçamento

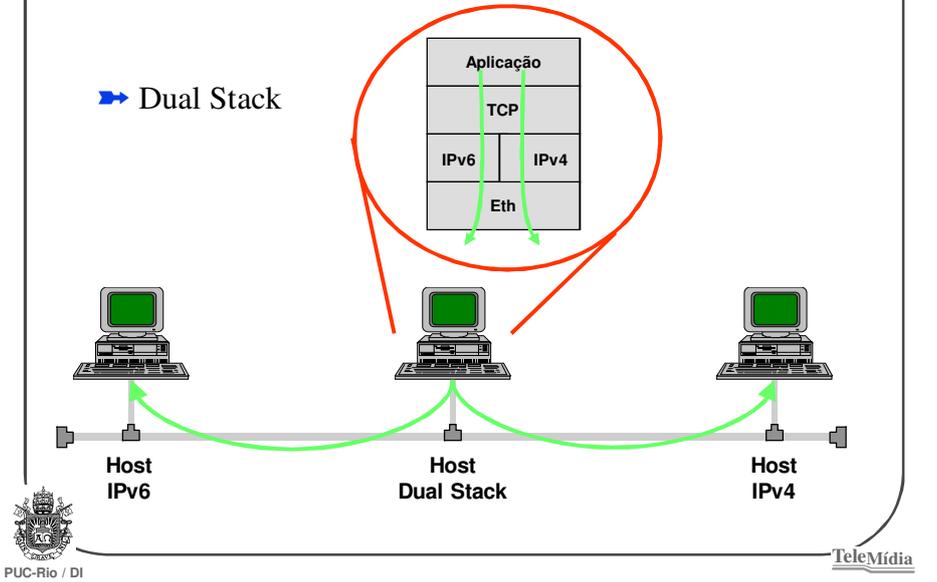


PUC-Rio / DI

TeleMídia

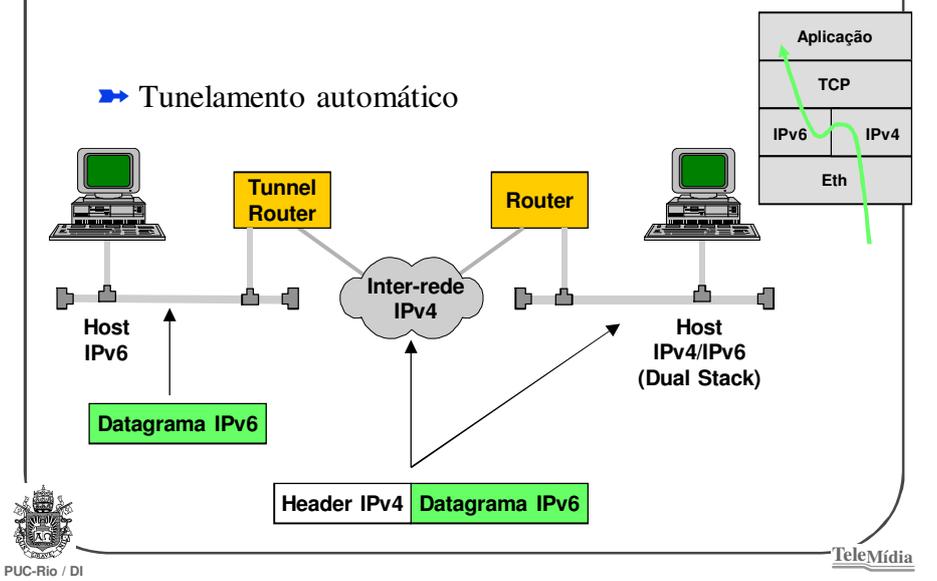
Estratégias de Migração IPv4 → IPv6

➔ Dual Stack



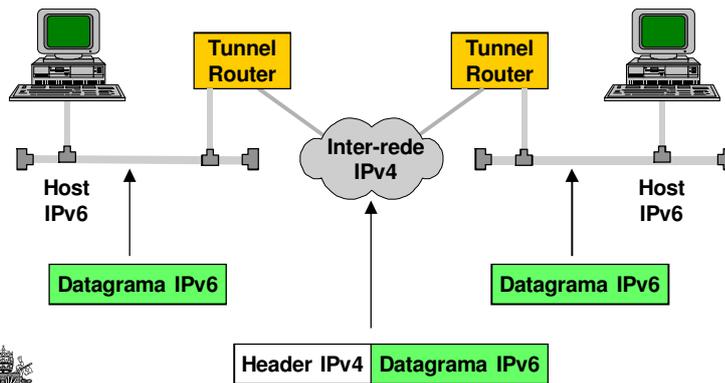
Estrat gias de Migra o IPv4 → IPv6

➔ Tunelamento autom tico



Estratégias de Migração IPv4 → IPv6

➔ Tunelamento configurado

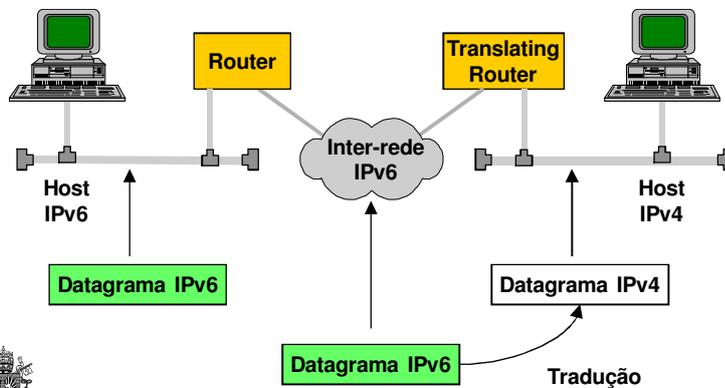


PUC-Rio / DI

TeleMídia

Estratégias de Migração IPv4 → IPv6

➔ Tradução de cabeçalhos



PUC-Rio / DI

TeleMídia

Construção da Tabela de Rotas



PUC-Rio / DI

TeleMídia

Encaminhamento

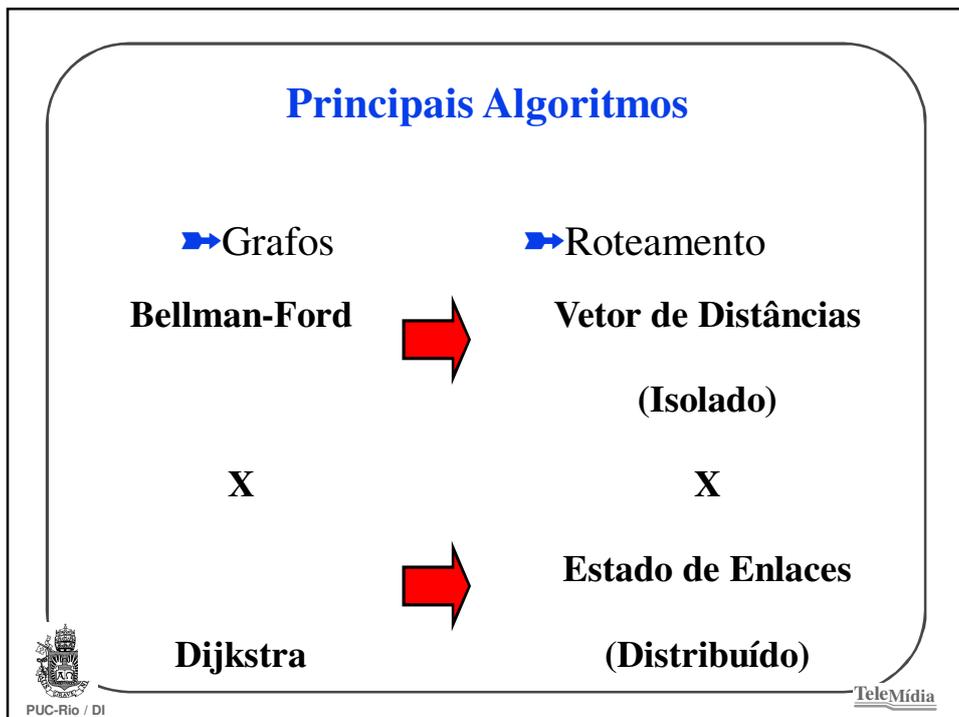
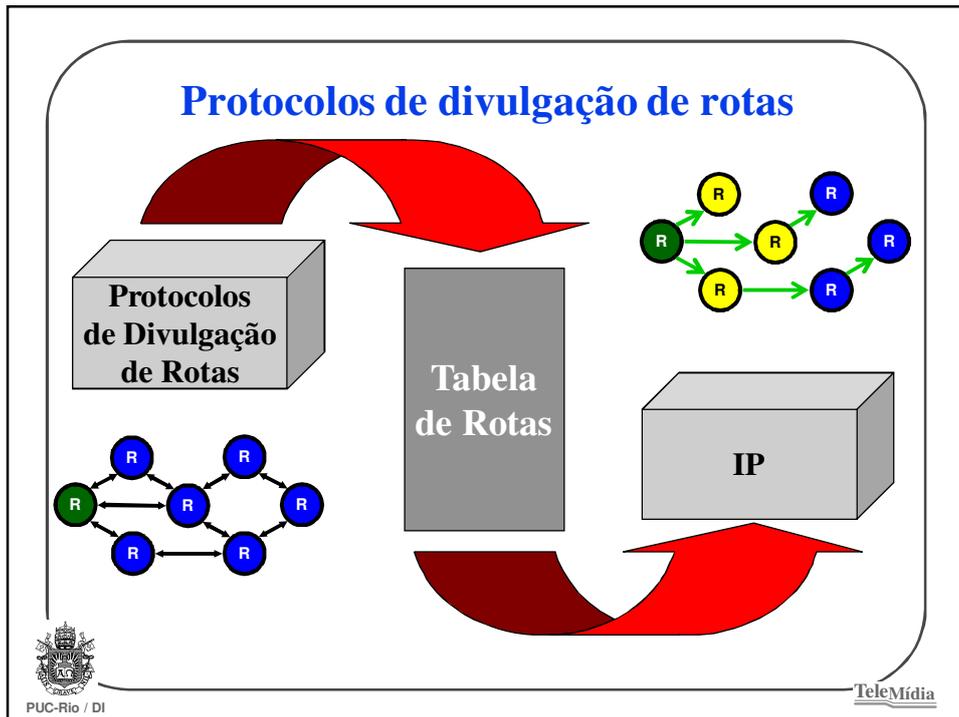
- ➔ Os roteadores possuem *tabelas de rotas* para realizarem o *encaminhamento*

- ➔ Essas tabelas contêm, a princípio, entradas que relacionam uma *sub-rede de destino* a uma *métrica* (ou *custo*)
 - Podem apresentar outras informações



PUC-Rio / DI

TeleMídia



Algoritmo de Vetor de Distâncias

- ➔ Inicialmente cada roteador possui em sua tabela de rotas uma entrada para cada sub-rede ao qual está ligado diretamente
- ➔ Periodicamente o roteador difunde para seus vizinhos um *vetor de distâncias* e recebe deles outros vetores
 - Vetor = { (Endereço de sub-rede, próx. roteador, custo) }
- ➔ Se um desses vetores revela uma rota nova, ou uma rota de menor custo do que uma presente atualmente na tabela de rotas local, esta rota será incorporada na tabela local



PUC-Rio / DI

TeleMídia

Algoritmo de Estado de Enlaces

- ➔ Cada roteador mantém um estado local (*router local state*)
 - Contém informações sobre interfaces operacionais do roteador e vizinhos alcançáveis por meio dessas interfaces
- ➔ Roteador difunde estado local a todos os outros roteadores
 - Distribuição periódica ou devido a mudanças no estado local
- ➔ Cada roteador constrói uma base de dados descrevendo a topologia da inter-rede (*link-state database*)
 - Base de dados = {estado local do roteador i}
- ➔ Cada roteador executa o algoritmo de Dijkstra sobre sua base
 - Constrói árvore de caminho mais curto onde o roteador é a raiz



PUC-Rio / DI

TeleMídia

Internet e Arquitetura TCP/IP

Sistemas autônomos



PUC-Rio / DI

TeleMídia

Sistemas Autônomos (*Autonomous Systems*)

- ➔ Divisão da Internet em Sistemas Autônomos (AS), cada AS consistindo em um conjunto de roteadores e sub-redes sob mesma administração
- ➔ Não existe um limite teórico para o tamanho de um AS:
 - pode ser uma simples sub-rede
 - pode ser uma inter-rede corporativa inteira
 - pode ser um conjunto de sub-redes de clientes de um provedor de acesso Internet

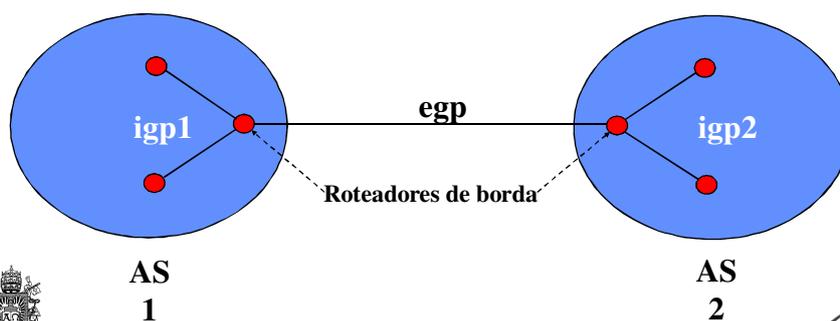


PUC-Rio / DI

TeleMídia

Sistemas Autônomos

- AS pode usar igps - protocolos de roteamento interno - para calcular rotas internas
- AS usa um egp - protocolo de roteamento externo - para trocar informações sobre rotas com outros ASs



PUC-Rio / DI

TeleMídia

Tipos de AS

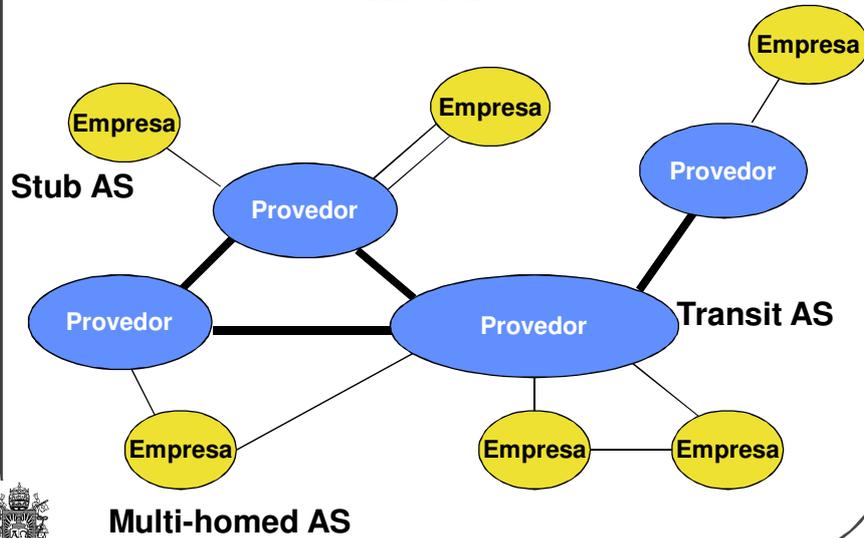
- Stub AS
 - AS com conexão para apenas um outro AS
- Multihomed AS
 - AS com várias conexões para outros Ass, mas que não aceita tráfego não-local
 - Tráfego não-local = Tráfego que não se originou no AS e nem se destina ao AS
- Transit AS
 - AS com várias conexões para outros ASs e que aceita tráfego não-local, freqüentemente sob certas condições (policy restrictions)



PUC-Rio / DI

TeleMídia

Arquitetura Topológica Atual da Internet



PUC-Rio / DI

TeleMídia

Identificação de AS

- ➔ ASN: número de 16 bits (~64K possíveis)
- ➔ ASs (e ASNs) no Brasil:
 - PUC-Rio (14553), Fapesp (2715), RNP (1916), Embratel (4230), Uninet (5772), Inside (7063), Telerj/Telemar (7298), Intelig (17379) ...
- ➔ Consulta a ASs e ASNs, AS traces:
 - <http://www.fixedorbit.com/search.htm>



PUC-Rio / DI

TeleMídia

Protocolos igp

RIP



PUC-Rio / DI

TeleMídia

Protocolo RIP (Routing Information Protocol)

➔ Utiliza Algoritmo de Vetor de Distâncias

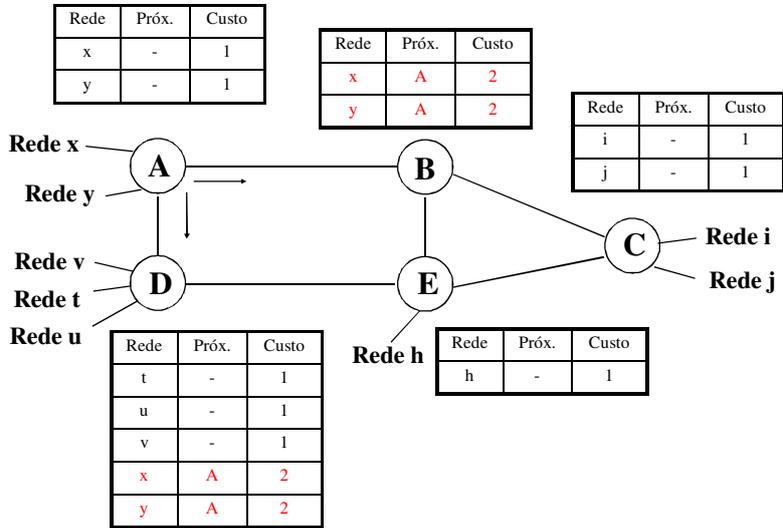
- Relembrando...
 - Inicialmente, cada roteador possui em sua tabela de rotas uma entrada para cada sub-rede ao qual está ligado diretamente
 - Periodicamente o roteador difunde para seus vizinhos um vetor de distâncias e recebe deles outros vetores
 - Vetor = { (Endereço de sub-rede, próx. roteador, custo) }
 - Se um desses vetores revela uma rota nova, ou uma rota de menor custo do que uma presente atualmente na tabela de rotas local, esta rota será incorporada na tabela local



PUC-Rio / DI

TeleMídia

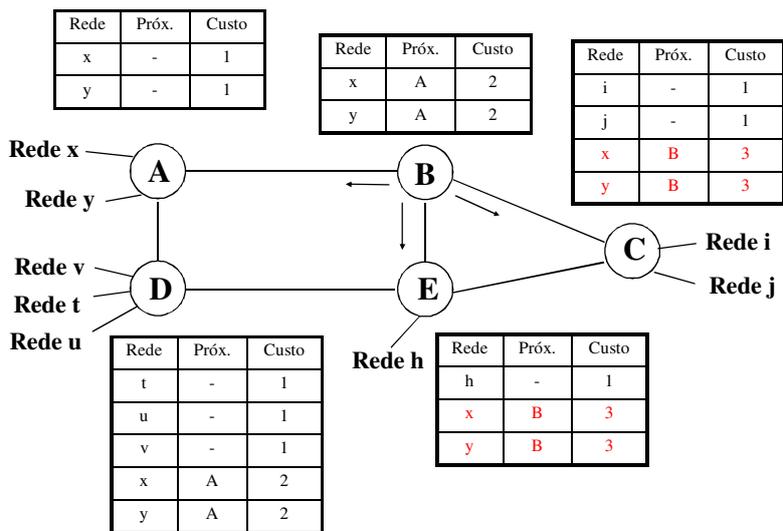
Protocolo RIP



PUC-Rio / DI

TeleMídia

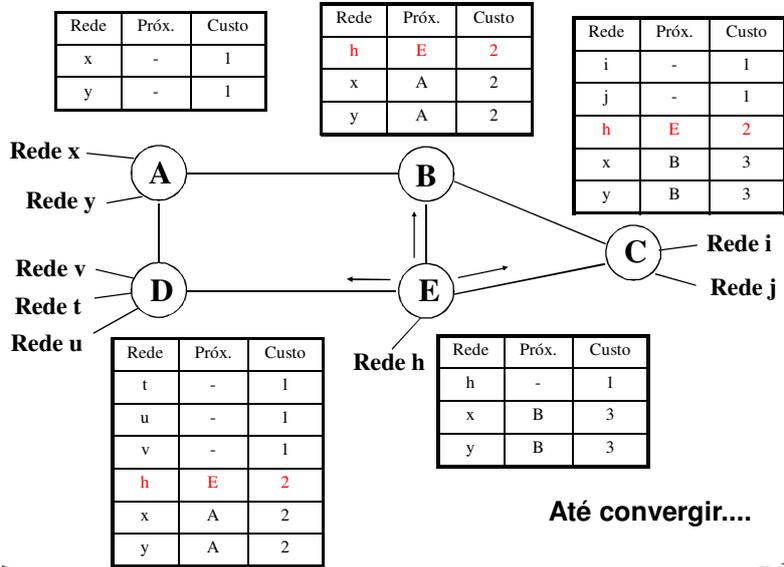
Protocolo RIP



PUC-Rio / DI

TeleMídia

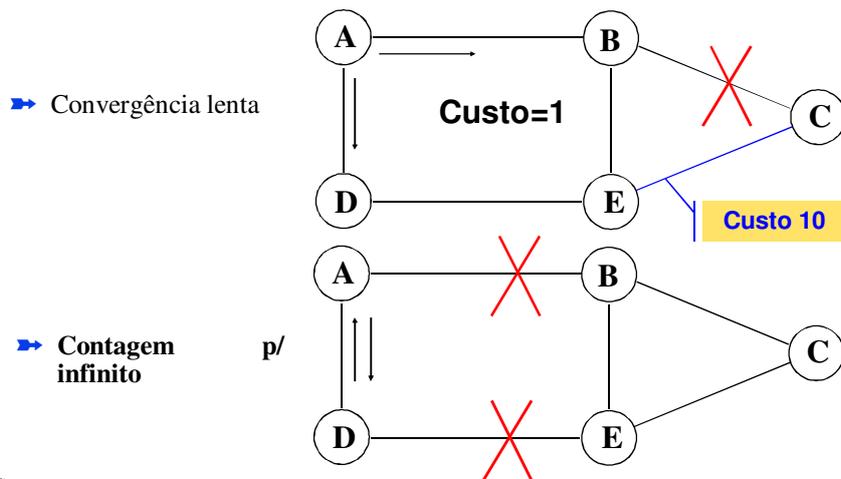
Protocolo RIP



PUC-Rio / DI

TeleMídia

Problemas bem-conhecidos de Bellman-Ford



PUC-Rio / DI

TeleMídia

Soluções para os problemas de Bellman-Ford

- ➔ Máximo de saltos: 16 (“infinito” do RIP)
 - Diâmetro máximo da rede == 15 saltos
- ➔ “Simple split horizon”
 - Omite rotas aprendidas por um vizinho em divulgações enviadas a esse vizinho
- ➔ “Split horizon with poisoned reverse”
 - Inclui rotas aprendidas por um vizinho em divulgações enviadas a esse vizinho, mas associa a elas custo “infinito”
- ➔ “Triggered updates”
 - Gerados quando o custo de uma rota muda
 - Opera independentemente da divulgação periódica



PUC-Rio / DI

TeleMídia

Limitações do RIP

- ➔ Uso de custo fixo
 - Número de saltos
- ➔ Baixa escalabilidade
 - “Diâmetro” limitado do AS (15 saltos)
 - Grandes tabelas de rotas



PUC-Rio / DI

TeleMídia

Protocolos igp

OSPF



PUC-Rio / DI

TeleMídia

Protocolo OSPF (Open Shortest Path First)

➔ Utiliza Algoritmo de Estado de Enlaces

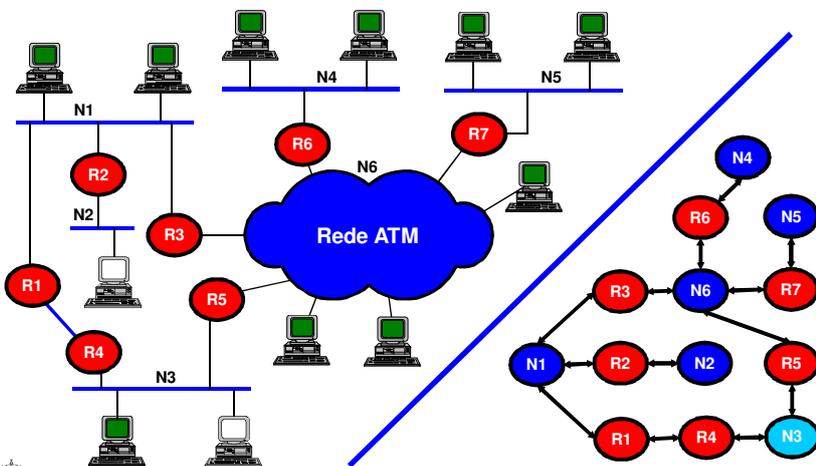
- Relembrando...
 - Cada roteador mantém um estado local (router local state)
 - Contém informações sobre interfaces operacionais do roteador e vizinhos alcançáveis por meio dessas interfaces
 - Roteador difunde estado local a todos os outros roteadores
 - Distribuição periódica ou devido a mudanças no estado local
 - Cada roteador constrói uma base de dados descrevendo a topologia da inter-rede (link-state database)
 - Base de dados = {estado local do roteador i }, p/ todo $i \in AS$
 - Cada roteador executa o alg. de Dijkstra sobre sua base
 - Constrói árvore de caminho mais curto onde o roteador é a raiz



PUC-Rio / DI

TeleMídia

Alg. de Roteamento X Alg. de Grafo



PUC-Rio / DI

TeleMídia

OSPF x RIP

- ➔ Responde mais rapidamente a mudanças na topologia
- ➔ Quando a rede está “estável”, os roteadores trocam pequenos volumes de informação entre si
- ➔ Métrica mais flexível
- ➔ Armazena múltiplas rotas para um mesmo destino
 - Balanceamento de carga
- ➔ Agrupamento de redes em áreas
 - Topologia da área é escondida do resto do AS
 - Redução no tráfego de informações sobre rotas



PUC-Rio / DI

TeleMídia

OSPF - Definições Básicas

➔ Tipos de sub-redes

- Sub-redes ponto-a-ponto
 - PPP, SLIP, HDLC, ...
- Sub-redes de múltiplo acesso
 - X.25, ATM, ...
- Sub-redes com broadcast
 - Ethernet, Token Ring, ...



OSPF - Definições Básicas (continuação)

➔ Link-State Advertisement (LSA)

- Unidade de informação que descreve o estado local de um roteador ou de uma sub-rede
 - Estado local de um roteador: estado de suas interfaces
 - Estado local de uma sub-rede: roteadores conectados àquela sub-rede

➔ Link-State Database (LSDB)

- Conjunto de LSAs



Representação de Redes e Roteadores

→ LSAs e LSDB

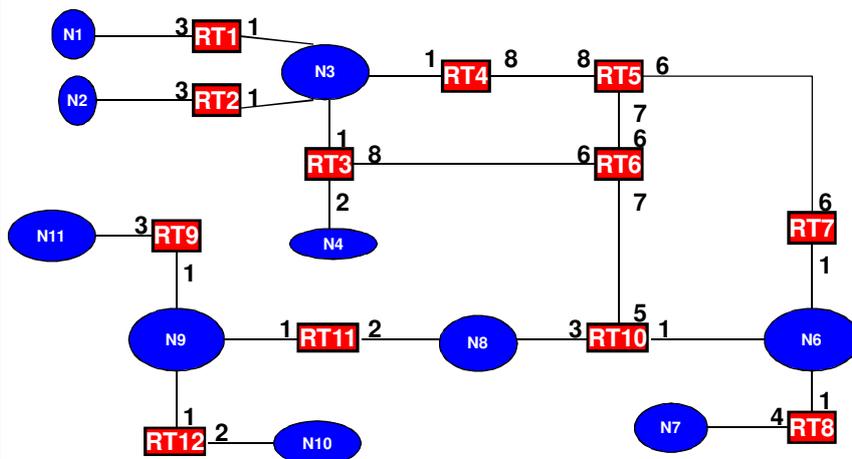
- Permitem construção de grafo direcionado
- Vértices: roteadores e sub-redes
 - Retângulos representam roteadores, círculos representam sub-redes
 - Sub-redes ponto-a-ponto não são representadas como vértices
- Arestas
 - Conectam dois roteadores se eles são ligados por uma sub-rede ponto-a-ponto
 - Conectam um roteador e uma sub-rede de múltiplo acesso ou com broadcast se o roteador possui uma interface ligada à sub-rede



PUC-Rio / DI

TeleMídia

Exemplo de Grafo Representando um AS



PUC-Rio / DI

TeleMídia

Exemplos de LSAs

➤ Representação da LSA de RT12

		De RT12	
Para	RT12		
	N9	1	
	N10	2	

➤ Representação da LSA de N9

		De N9	
Para	RT9	0	
	RT11	0	
	RT12	0	
	N9		

➤ Representação da LSA de RT6

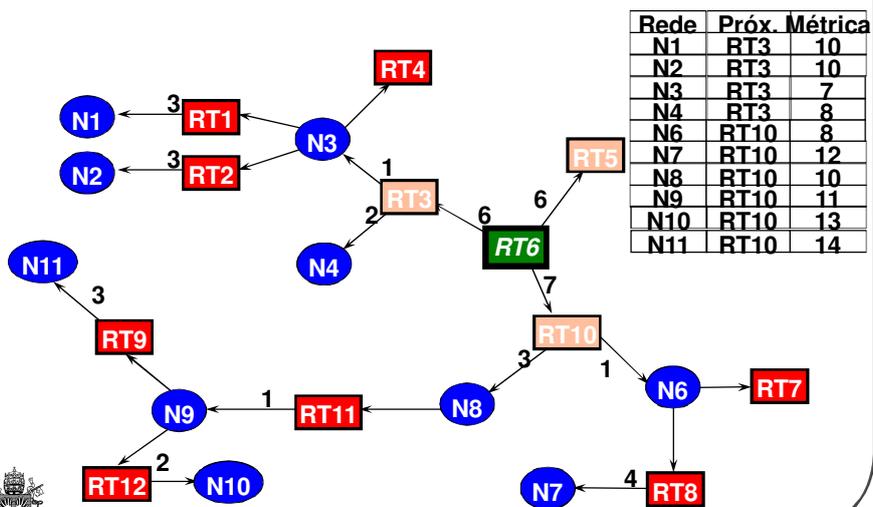
		De RT6	
Para	RT3	6	
	RT5	6	
	RT10	7	



PUC-Rio / DI

TeleMídia

Árvore SPF e Tabela de Rotas para RT6



PUC-Rio / DI

TeleMídia

OSPF: Múltiplas Áreas

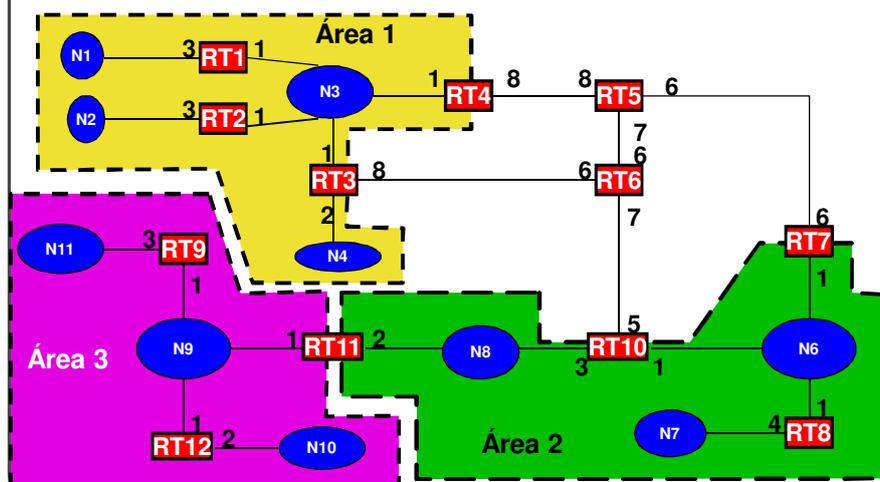
- ➔ Com o aumento de tamanho de um AS, começa a ser oneroso utilizar o OSPF em cada roteador
- ➔ Solução: roteamento em dois níveis
 - Rede dividida em áreas interligadas por um backbone
 - Cada área se comporta como um "AS" separado
 - O backbone também é uma área, e há diversos roteadores de fronteira de área (ABR) que o interligam às outras áreas
 - Para a área, ABR injeta rotas externas à área
 - Para o backbone, ABR injeta um sumário da área
- ➔ Divisão em áreas permite bases de dados menores, redução das tabelas de roteamento e diminuição do tempo de cálculo do algoritmo



PUC-Rio / DI

TeleMídia

AS Dividido em Áreas



PUC-Rio / DI

TeleMídia

LSAs das Áreas 1 e 2 Difundidos pelo Backbone

- Representação da área 1 por RT3 ➤ Representação da área 1 por RT4

		De	
		RT3	
Para	N1	4	
	N2	4	
	N3	1	
	N4	2	

		De	
		RT4	
Para	N1	4	
	N2	4	
	N3	1	
	N4	3	

- Representação da área 2 por RT7 ➤ Representação da área 2 por RT10

		De	
		RT7	
Para	N6	1	
	N7	5	
	N8	4	

		De	
		RT10	
Para	N6	1	
	N7	5	
	N8	3	



PUC-Rio / DI

TeleMídia

LSAs do Backbone Difundidos pela Área 1

- Representação do backbone injetada na área 1 por RT3 ➤ Representação do backbone injetada na área 1 por RT4

		De	
		RT3	
Para	N6	16	
	N7	20	
	N8	18	
	N9-N11	22	

		De	
		RT4	
Para	N6	15	
	N7	19	
	N8	18	
	N9-N11	29	



PUC-Rio / DI

TeleMídia

LSDB da área 1

		DE				
		RT 1	RT 2	RT 3	RT 4	N3
PARA	RT 1					0
	RT 2					0
	RT 3					0
	RT 4					0
	N1	3				
	N2		3			
	N3	1	1	1	1	
	N4			2		
	N6			16	15	
	N7			20	19	
	N8			18	18	
	N9-N11			22	29	



Perguntas

- ▶ Como funciona o protocolo de divulgação de rotas “Vetor de Distâncias”?
- ▶ Como funciona o protocolo de divulgação de rotas “Estado de Enlaces”?
- ▶ O que vem a ser o problema de convergência lenta e de contagem para o infinito no algoritmo de Bellman-Ford? Como podem ser resolvidos?
- ▶ Faça uma comparação entre o OSPF e o RIP.
- ▶ Quais os problemas da comutação de pacotes? Quando eles são preponderantes? Como eles são resolvidos na comutação rápida de pacotes?
- ▶ Quando a comutação rápida de pacotes tem um desempenho pior do que a comutação de pacotes?
- ▶ Por que são necessárias duas pilhas de protocolo na comutação rápida de pacotes?



Redes com Comutação por Rótulo



PUC-Rio / DI

TeleMídia

Comutação de Pacotes

- Retardos de transferência muito grandes.
- Jitter (variação do retardo) muito grande.
- **Carga de processamento elevada no comutador (roteador)**

Comutação Rápida de Pacotes



PUC-Rio / DI

TeleMídia

Comutação Rápida de Pacotes

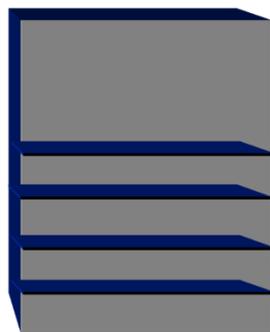
- ➔ Relaxar (ou eliminar) o controle de erro nas camadas de enlace e de rede.
- ➔ Relaxar (ou eliminar) o controle de fluxo nas camadas de enlace e de rede.
- ➔ Separar os canais de sinalização e de dados: pacotes de dados devem propagar na rede sem a necessidade do processamento da camada de rede (label switch).



PUC-Rio / DI

TeleMídia

Comutação de Rótulos



Pilha de Controle



Pilha de Dados



PUC-Rio / DI

TeleMídia

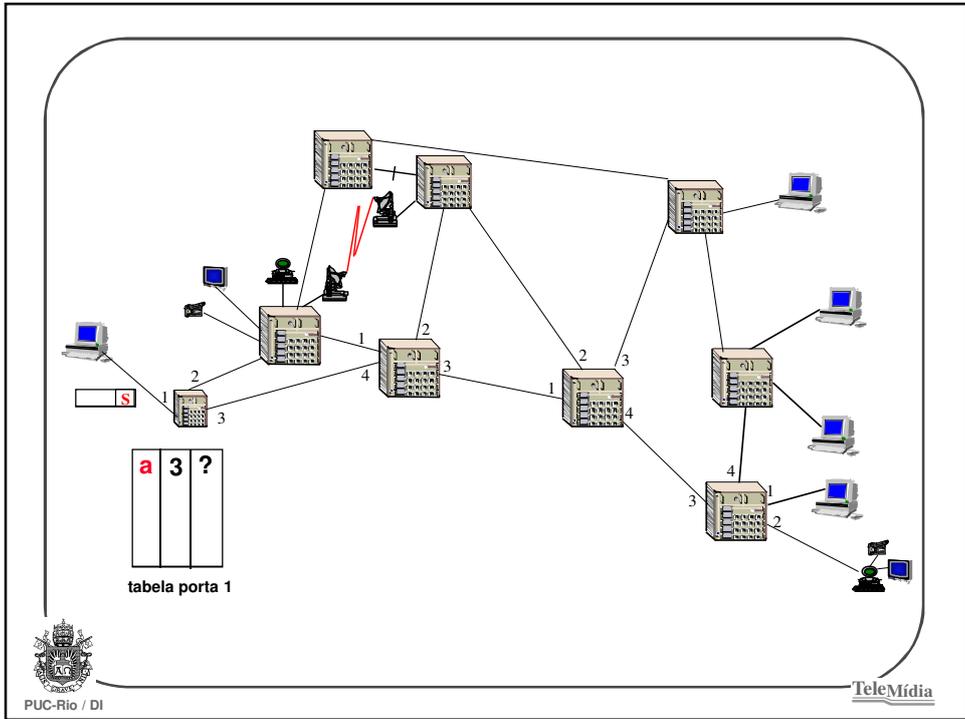
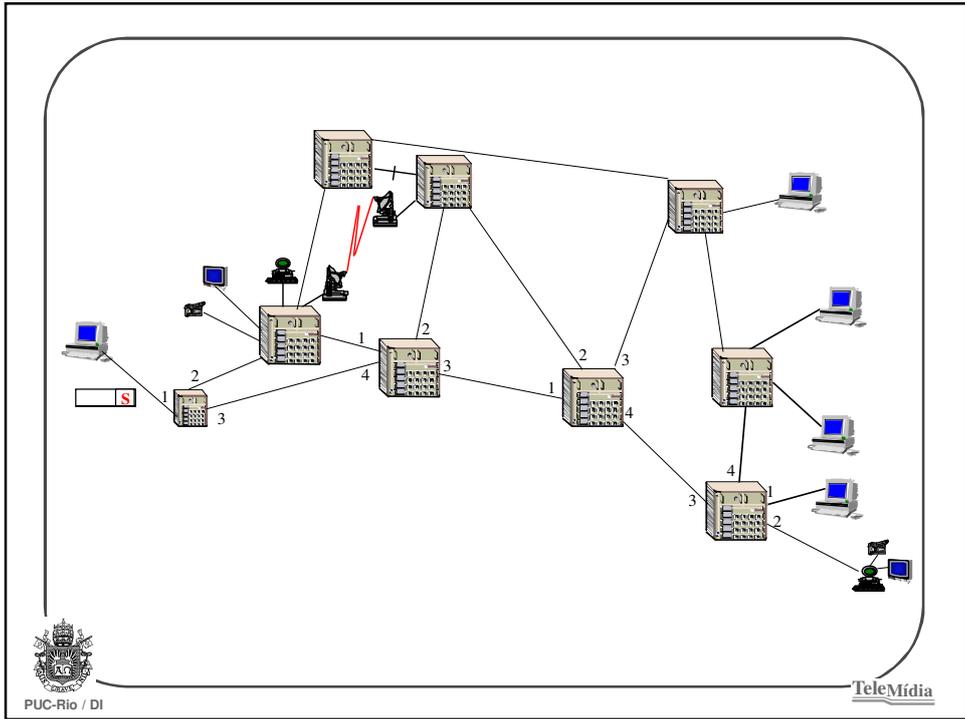
Comutação por Rótulos

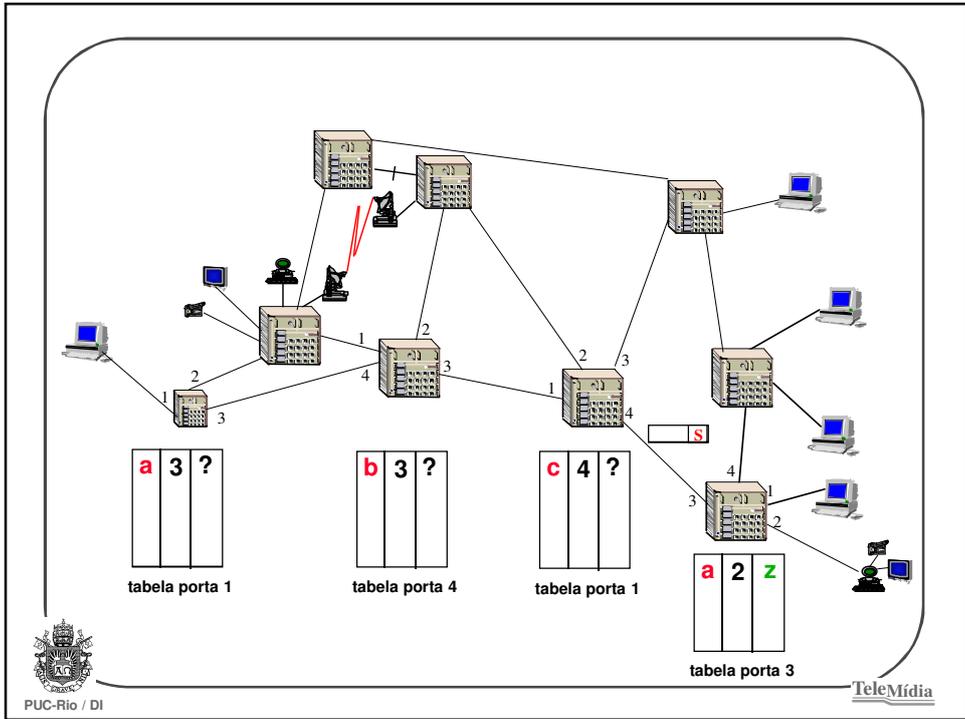
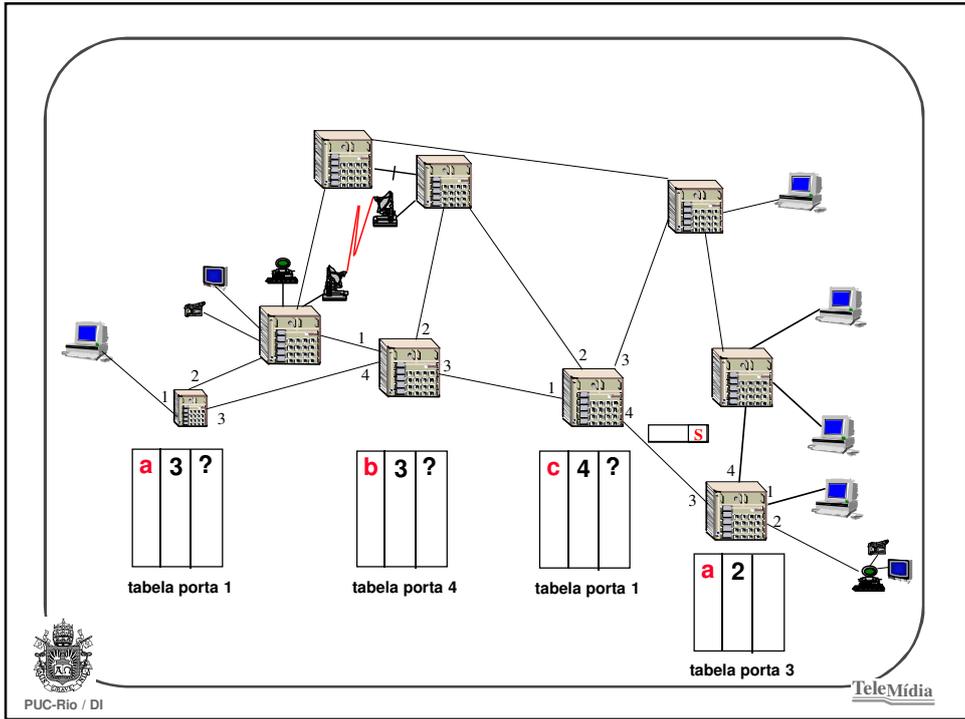
- ➔ A comutação por rótulos pode ser realizada em comutação orientada a conexão ou não orientada a conexão.
- ➔ São exemplos de redes orientadas a conexão com comutação de rótulos:
 - Frame Relay
 - ATM
- ➔ São exemplos de redes não orientadas a conexão com comutação de rótulos:
 - Comutação IP (MPLS)

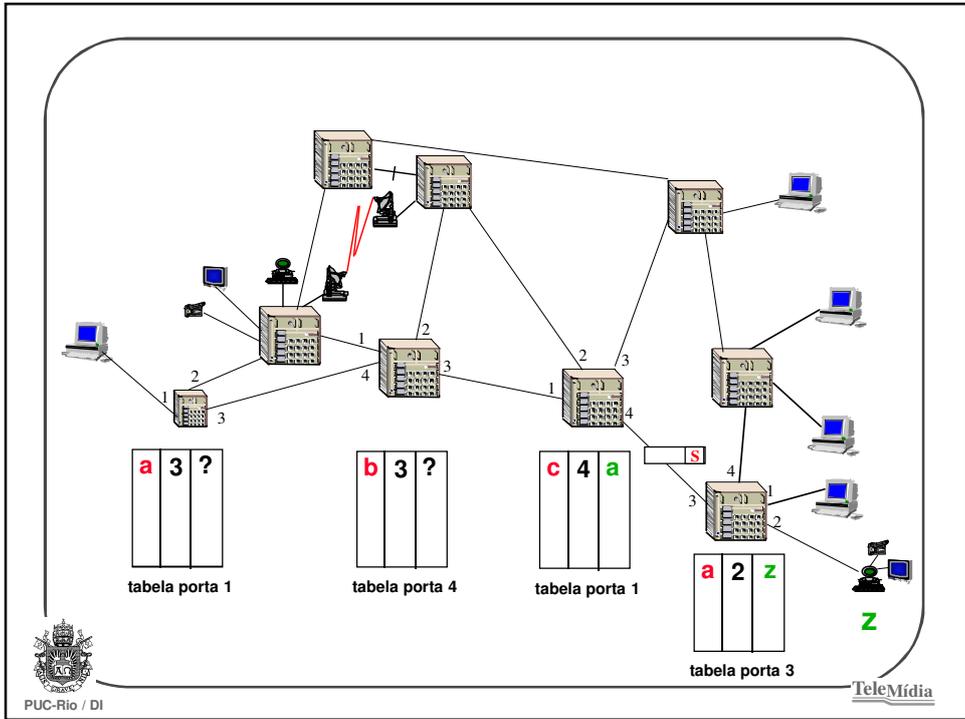
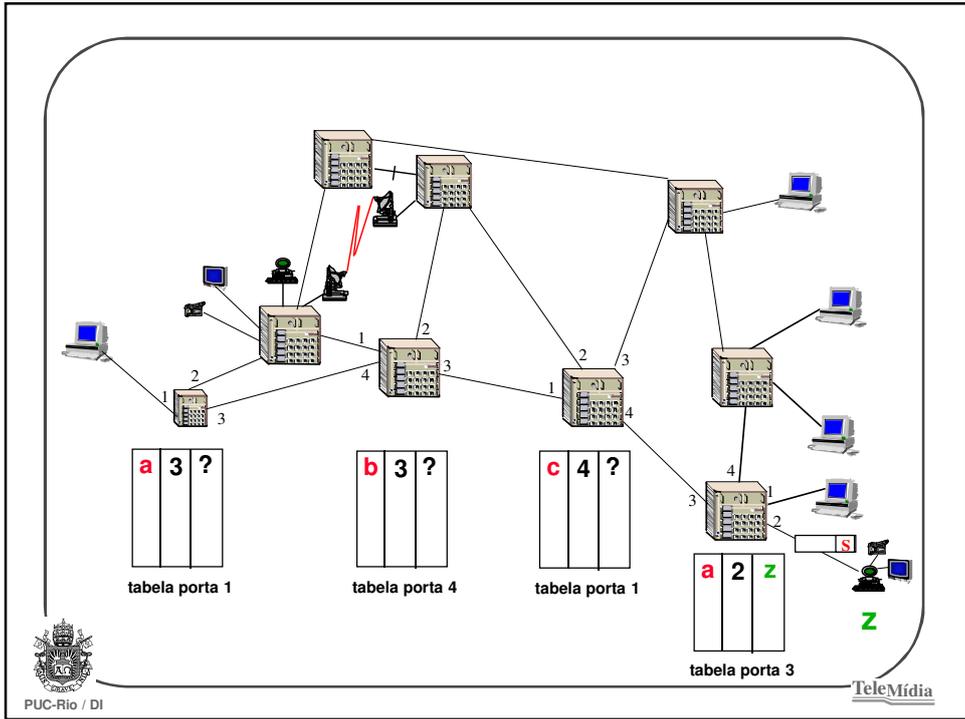


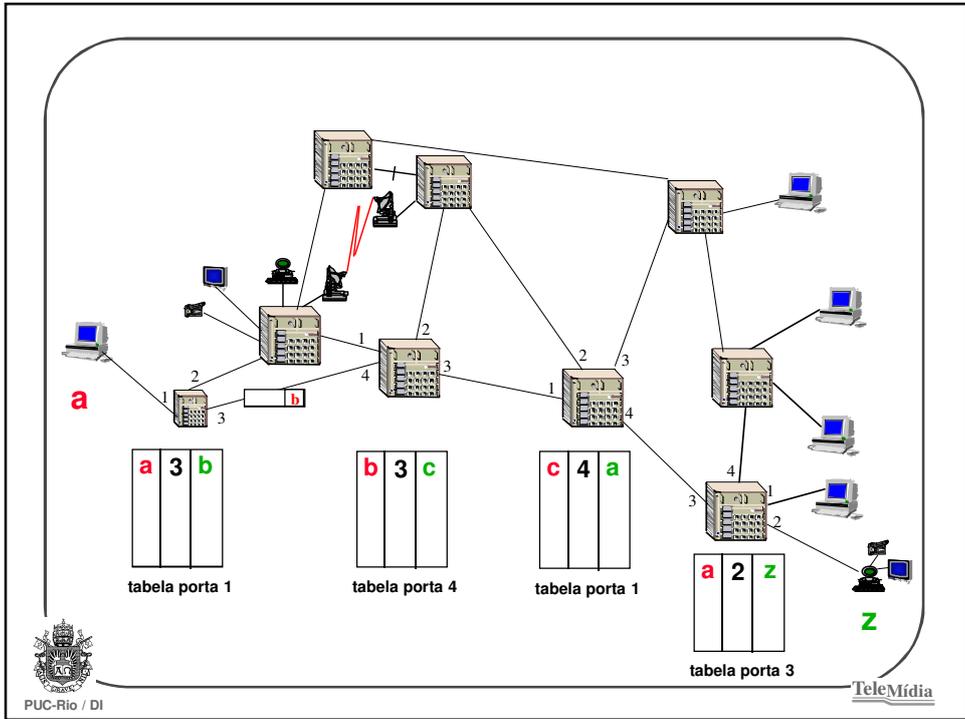
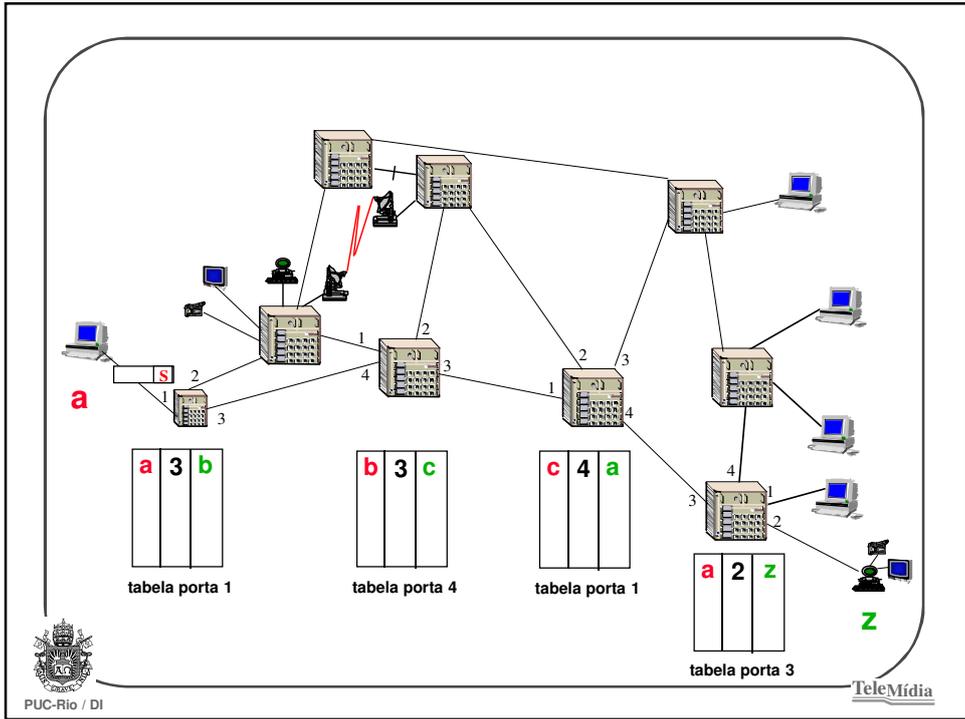
Comutação por Rótulo Com Conexão

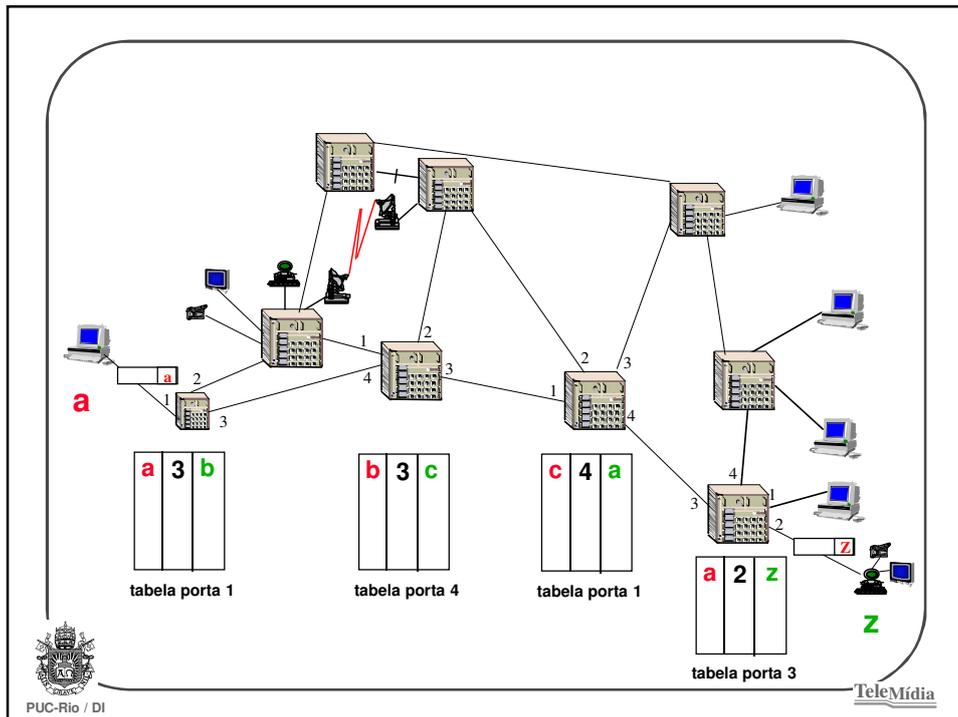












Comutação Por Rótulos

- ➔ A tabela de comutação de rótulos pode ser construída por procedimentos de estabelecimento de conexão, como descrito anteriormente, e neste caso temos *conexões dinâmicas (ou chaveadas)*.
- ➔ Procedimentos de gerenciamento podem também ser usados para escrever diretamente na tabela de comutação. Neste caso temos conexões semi-permanentes ou permanentes.

Frame Relay



PUC-Rio / DI

TeleMídia

Redes Frame Relay

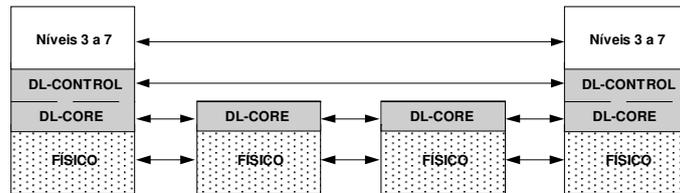
- ➔ Redes frame relay são redes com comutação de rótulos orientadas a conexão.
- ➔ Os pacotes que trafegam na rede, denominados quadros, são de tamanho variáveis.
- ➔ Na comutação dos quadros, o nível de enlace nos comutadores realiza a detecção de erro e a comutação de rótulos. Funções de correção de erro e controle de fluxo não são realizadas.



PUC-Rio / DI

TeleMídia

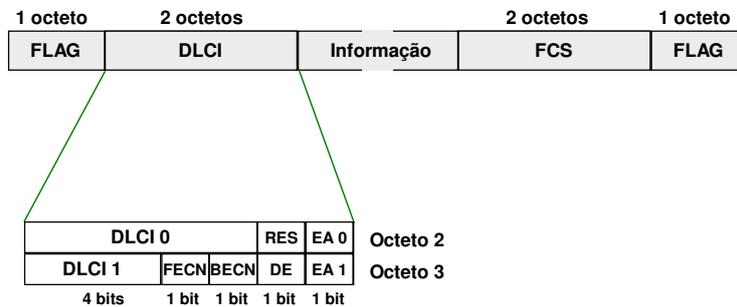
Serviço de Frame Relaying



PUC-Rio / DI

TeleMídia

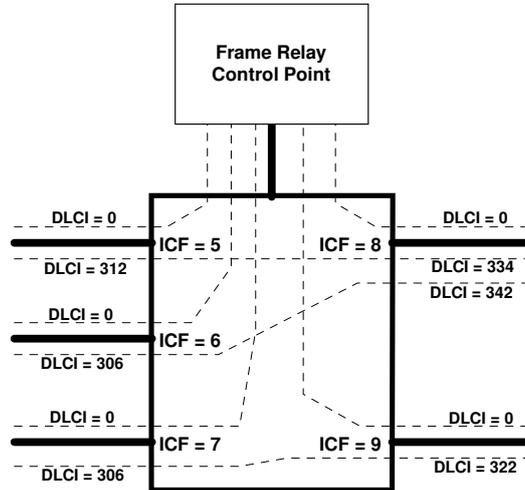
Quadros Frame Relay



PUC-Rio / DI

TeleMídia

Frame Handler



PUC-Rio / DI

TeleMídia

Redes ATM



PUC-Rio / DI

TeleMídia

Redes ATM

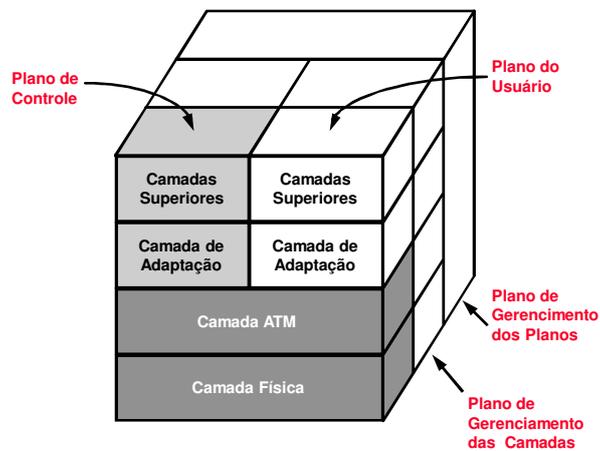
- Redes ATM são redes com comutação de rótulos orientadas a conexão.
- Os pacotes que trafegam na rede, denominados células, são de tamanho fixos (53 octetos).
- Na rede ATM, o nível de enlace, denominado camada ATM, realiza apenas a comutação de rótulos. Nenhuma função de controle de erro e controle de fluxo de dados é realizada.
- No estabelecimento da conexão, requisitos de qualidade de serviço (QoS), tais como retardo máximo, banda passante, etc., podem ser requeridos, e podem ser garantidos pela rede.



PUC-Rio / DI

TeleMídia

Modelo de Referência



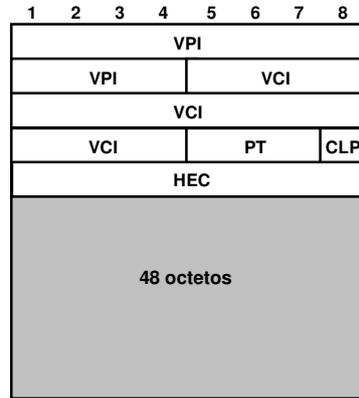
PUC-Rio / DI

TeleMídia

Células ATM



UNI



NNI



PUC-Rio / DI

TeleMídia

Comutação por Rótulo Sem Conexão



PUC-Rio / DI

TeleMídia

Comutação por Rótulo

➔ Mapeamento Baseado em Topologia

- Decisão de encaminhamento tomada baseada no endereço de destino, ou na composição endereços fonte e destino



PUC-Rio / DI

TeleMídia

Comutação por Rótulo

➔ Mapeamento Baseado em Topologia

➔ Mapeamento Baseado em Fluxo

- Decisão de encaminhamento tomada baseada no tipo de fluxo. Fluxos mais duradouros são comutados, os outros roteados normalmente.



PUC-Rio / DI

TeleMídia

Comutação Por Rótulos em Redes Não Orientadas a Conexão

- ➔ Em redes não orientadas a conexão, o LSP precisa ser refrescado de tempos em tempos. Diz-se que o LSP estabelece uma conexão “soft state”.
- ➔ No modo orientado a tráfego, um intervalo grande entre pacotes do mesmo fluxo gera a retirada do LSP das tabelas de comutação do caminho. O próximo pacote do fluxo vai então estabelecer um novo LSP.
- ➔ No modo orientado a topologia, cada alteração na tabela de roteamento pode gerar a destruição do LSP correspondente, e a geração de um novo LSP.



PUC-Rio / DI

TeleMídia

Comutação por Rótulo

- ➔ Mapeamento Baseado em Topologia
- ➔ Mapeamento Baseado em Fluxo

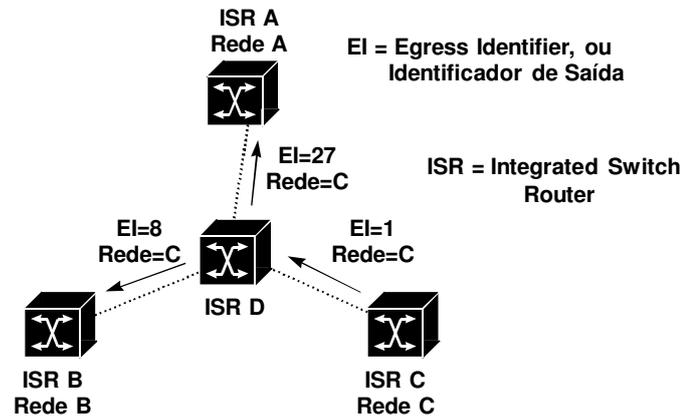


PUC-Rio / DI

TeleMídia

ARIS - AGGREGATE ROUTE-BASED IP SWITCHING

Tecnologia desenvolvida pela IBM.



PUC-Rio / DI

TeleMídia

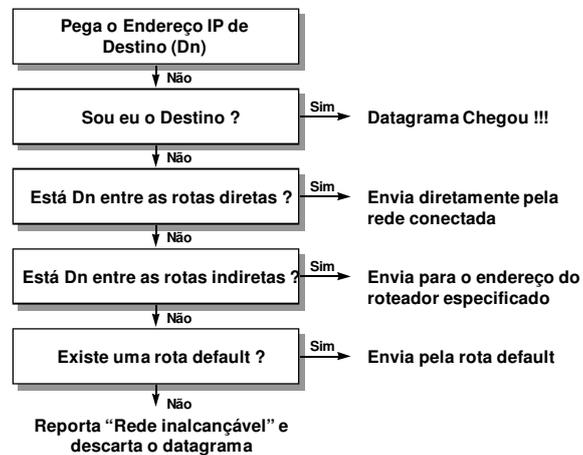
Comutação IP



PUC-Rio / DI

TeleMídia

Roteamento IP



PUC-Rio / DI

TeleMídia

Comutação IP

➔ Preocupação: compatibilidade

- Nem todas as soluções permitem que sejam utilizados os comutadores já existentes, geralmente se amarrando a hardwares específicos.



PUC-Rio / DI

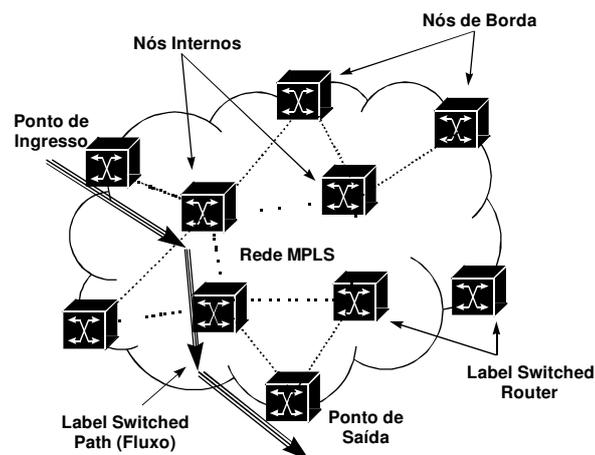
TeleMídia

MPLS - MultiProtocol Label Switching

- ➔ Tentativa de padronização: formado um grupo de trabalho do IETF, sob o nome de MultiProtocol Label Switching, ou simplesmente MPLS.
- ➔ Responsável por padronizar uma tecnologia base para uso em label switching e roteamento a nível de camada de rede, bem como a sua implementação sobre diversas outras tecnologias de conexão, que incluem: POS (Packet-over-SONET), Frame Relay, Ethernet, Token-Ring, ATM, etc.



MPLS



LDP - Label Distribution Protocol

- ➔ LSRs devem ser capazes de trocar mensagens de controle com seus vizinhos de forma a compartilhar um acordo comum do relacionamento de cada um dos rótulos associados a um fluxo qualquer. Este procedimento é conhecido como Protocolo de Distribuição de Rótulos (LDP).
- ➔ Dois grandes subgrupos:
 - Edge Control Operation (abordagem orientada à topologia), ou Operação Controlada na Borda, representado pelos padrões ARIS\$PEC e TDP;
 - Distributed Operation (abordagem orientada ao tráfego), ou Operação Distribuída, representado pelo FANP e IFMP.



PUC-Rio / DI

TeleMídia

LDP

- ➔ Modelo *downstream-on-demand* .
- ➔ É iniciado por um nó de borda (um ponto de ingresso ou saída).
- ➔ Mensagens explícitas do LDP carregam informações de gerenciamento de rótulos que são escolhidos pelo nó na direção do fluxo (rótulo de entrada) e comunicado ao nó vizinho no sentido inverso (que o usará como rótulo de saída).



PUC-Rio / DI

TeleMídia

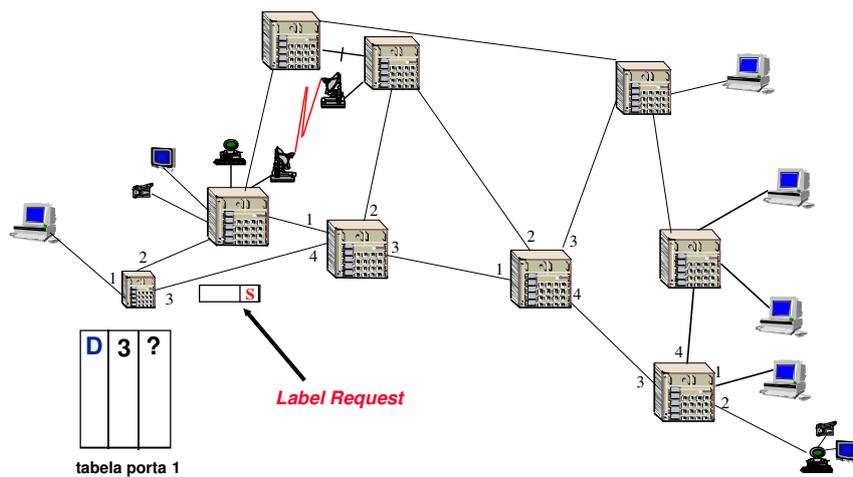
LDP

- ▶ Ao receber uma mensagem LDP *Hello*, o LSR inicia o estabelecimento de uma conexão TCP com o LSR que originou a mensagem.
- ▶ Uma vez estabelecida a conexão, parâmetros são trocados, por intermédio de mensagens LDP *Initialization*, o que inicia uma sessão LDP entre os LSRs pares. Estes LSRs devem concordar sobre a versão do protocolo LDP, o método de distribuição de rótulos a ser utilizado (*unsolicited downstream* ou *downstream-on-demand*), se o controle da distribuição será ordenado ou independente, se o modo de retenção de rótulos será liberal ou conservativo, valores de *timeout*, dentre outros parâmetros.
- ▶ Uma vez estabelecida uma sessão LDP, LSRs vizinhos podem trocar associações de rótulos a FECs por intermédio das mensagens de anúncio *Label Request* e *Label Mapping*.



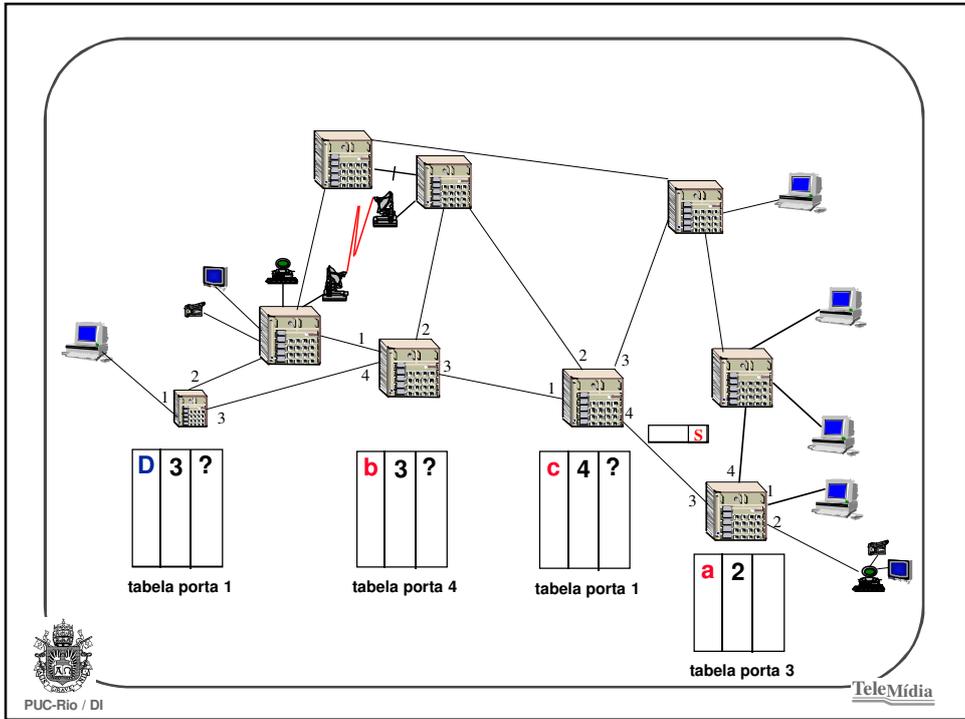
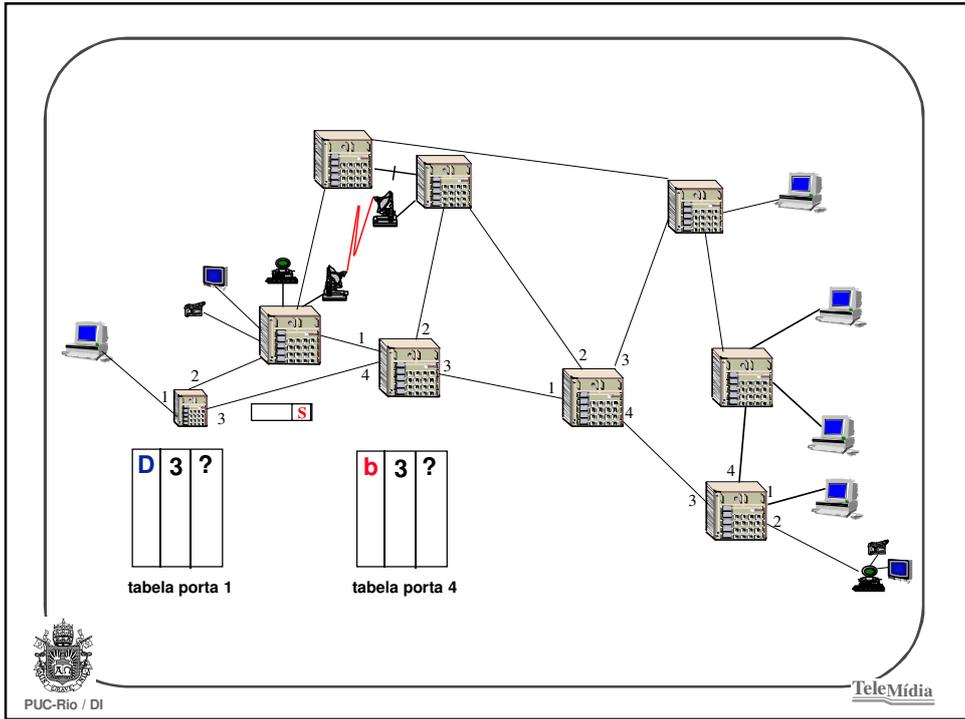
PUC-Rio / DI

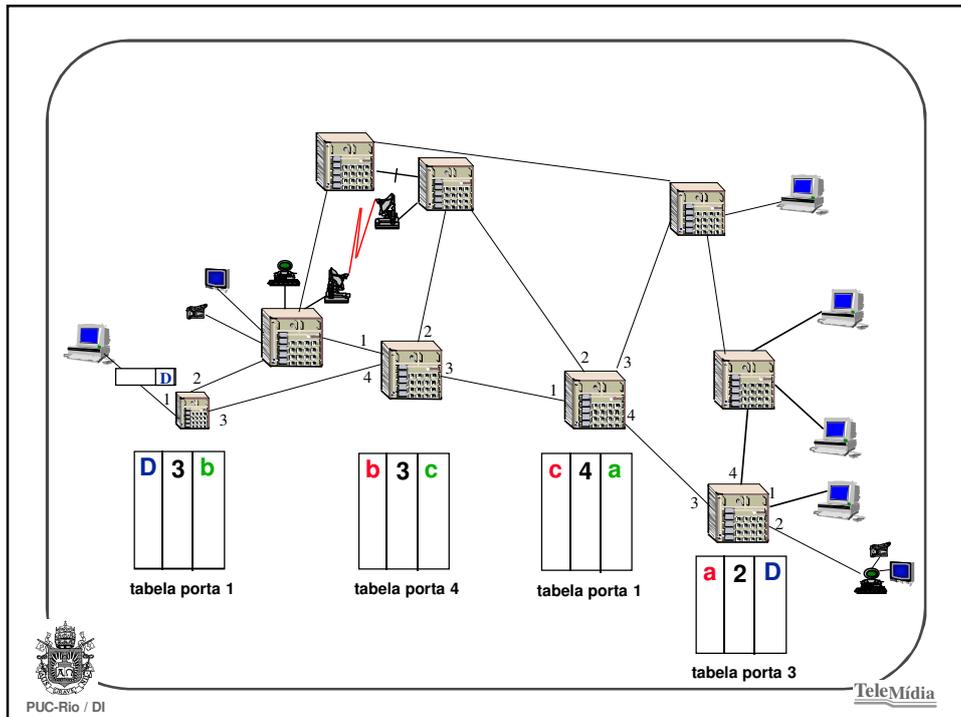
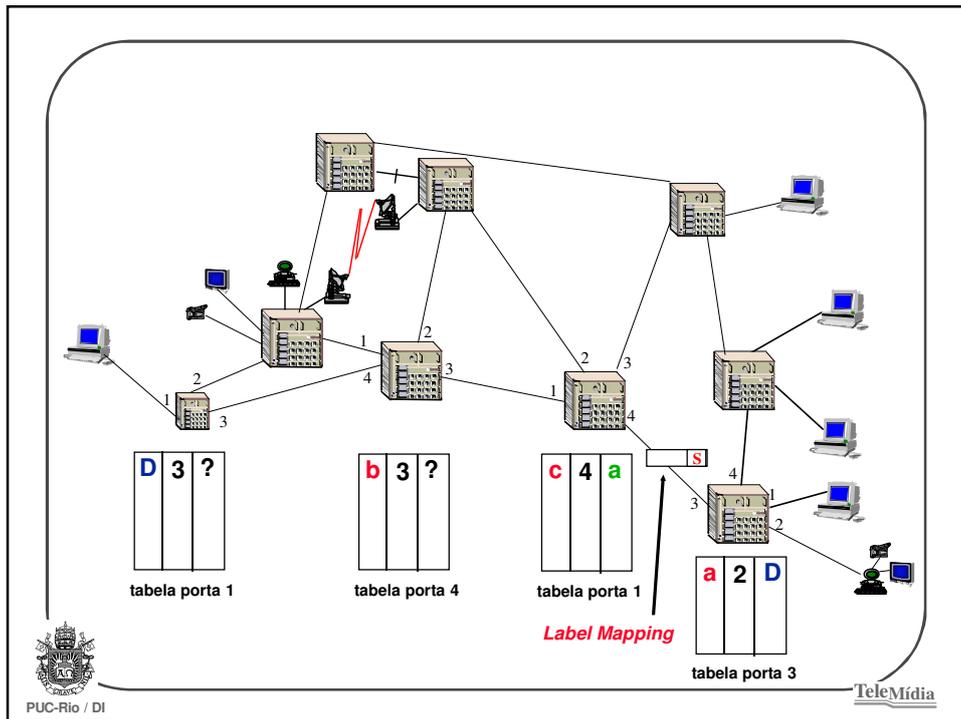
TeleMídia



PUC-Rio / DI

TeleMídia





Mensagens LDP

LDP Messages	
Discovery Messages	
Hello	Enviada por um LSR aos LSRs vizinhos de forma a anunciar a sua presença.
Address	Identifica o endereço de uma interface de um LSR.
Address Withdraw	Elimina a identificação de endereço de uma interface de um LSR.
Session Messages	
Initialization	Utilizada no estabelecimento de uma sessão LDP.
KeepAlive	Mantém a sessão LDP ativa.
Advertisement Messages	
Label Request	Solicita uma associação de um rótulo a uma FEC em particular.
Label Mapping	Informa a associação de um rótulo a uma FEC.
Label Abort Request	Utilizado para tornar sem efeito a solicitação de uma associação.
Label Withdraw	Elimina a associação rótulo a FEC informada previamente.
Label Release	Permite a um LSR sinalizar que a associação solicitada não é mais necessária.
Notification Messages	
Notification	Condições de erro ou status são enviadas através desta mensagem.



PUC-Rio / DI

TeleMídia

Perguntas

- Em que tipo de rede a comutação rápida de pacotes tem um bom desempenho?
- O que vem a ser comutação rápida por pacotes? Qual a diferença para a comutação de pacotes?
- Por que na comutação rápida de pacotes são necessárias duas pilhas de protocolo?
- Faça passo a passo o procedimento de uma comutação rápida por pacotes em redes com conexão?
- Faça passo a passo o procedimento de uma comutação rápida por pacotes em redes sem conexão?
- Descreva o funcionamento da rede *frame relay*.
- Descreva o funcionamento de uma rede ATM.
- Você poderia usar o mesmo procedimento de comutação rápida de pacotes das redes com conexão em redes sem conexão? Se sim, como? O que é conexão *soft state*?
- Como funciona o padrão MPLS?



PUC-Rio / DI

TeleMídia