

PTC 3450 - Aula 20

5.3 Roteamento intra-AS na Internet: OSPF

5.4 Roteamento entre os ISPs: BGP

5.5 MPLS

(Kurose, Seções 5.3, 5.4 e 6.5)

04/06/2024

Capítulo 5: Roteiro

5.1 Introdução

5.2 Protocolos de roteamento

- Estado de enlace
- Vetor de distâncias

5.3 Roteamento intra-AS na Internet: OSPF

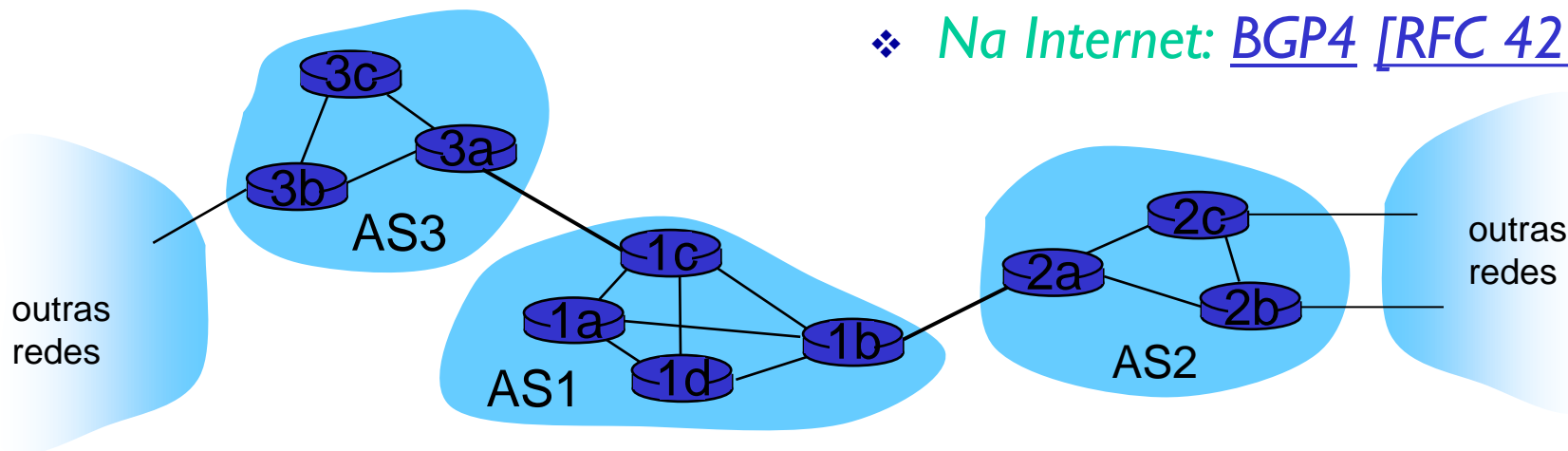
5.4 roteamento entre os ISPs: BGP

5.5 MPLS

5.6 Redes numa perspectiva mais ampla: introdução a redes complexas

Tarefa Inter-AS

- ❖ Suponha que roteador em AS1 recebe datagrama destinado para fora de AS1:
 - Roteador deve rotear pacote para roteador *gateway*, mas qual?



ASI precisa:

1. Aprender quais destinos são alcançáveis via AS2 e quais via AS3
 2. Propagar essa informação de alcançabilidade para todos roteadores em ASI
- ❖ *Trabalho do roteamento inter-AS!*
 - ❖ *Todas as ASs precisam rodar mesmo protocolo inter-AS!*
 - ❖ *Na Internet: BGP4 [RFC 4271]*

Roteamento Intra-AS

- ❖ Também conhecidos como *protocolos de roteadores internos (IGP – Interior Gate Protocols)*
- ❖ Protocolos de roteamento intra-AS mais comuns:
 - **RIP: Routing Information Protocol** [[RFC 2453 ... RFC4822](#)] – *ISPs de níveis mais baixos e redes corporativas*
 - **OSPF: Open Shortest Path First** [[RFC 2178](#)] – *ISPs de níveis mais altos (mais utilizado hoje)*
 - **EIGRP: Enhanced Interior Gateway Routing Protocol** (proprietário - Cisco)

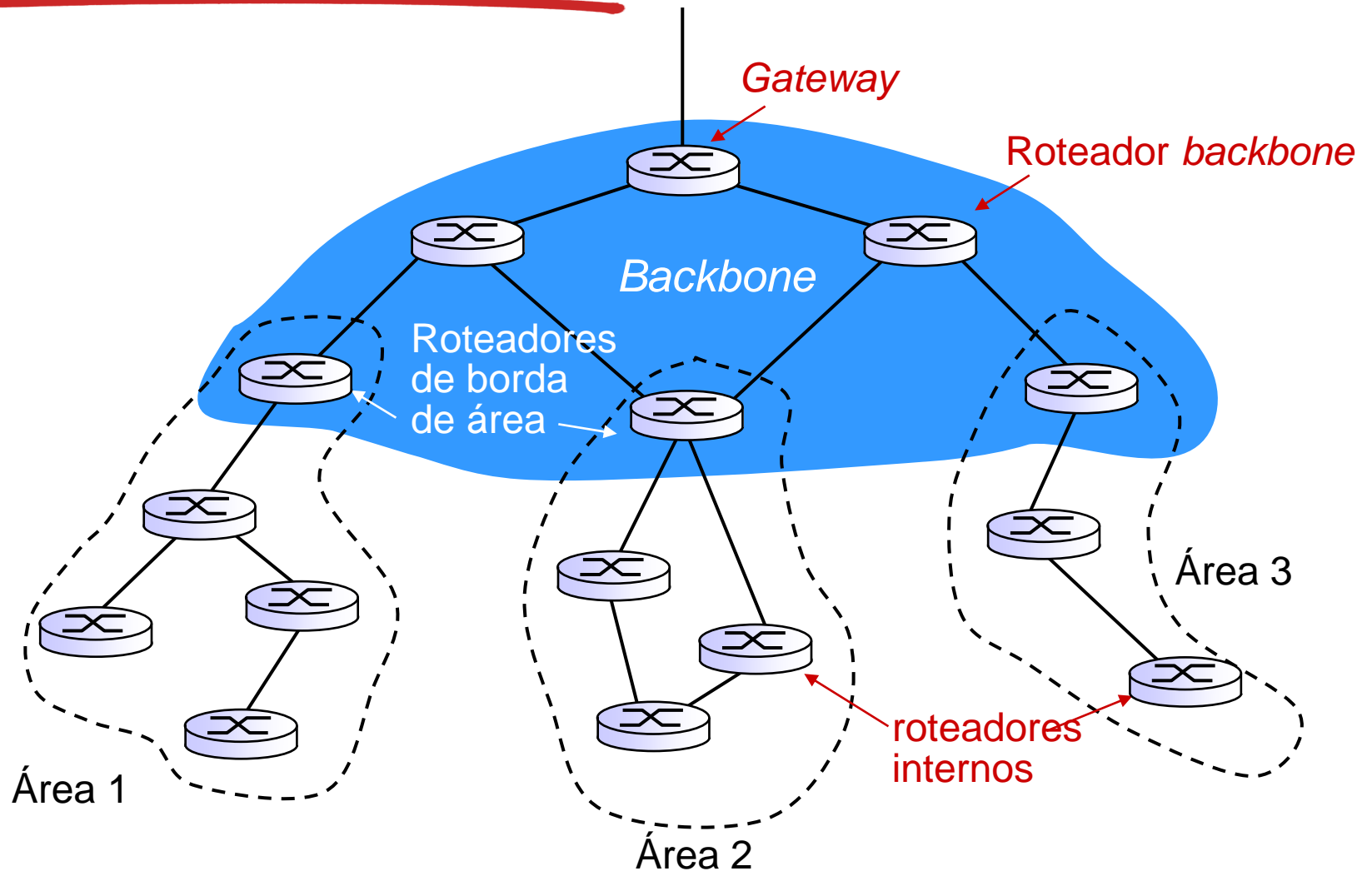
OSPF (*Open Shortest Path First*)

- ❖ “*Open*”: disponível publicamente [[RFC 1131](#) (1989) ... [RFC 7474](#) (2015)]
- ❖ “Sucessor” do RIP – características mais avançadas
- ❖ Usa algoritmo de estado de enlace
 - Inundação de pacotes para todos os roteadores do AS
 - Roteador constrói mapa da topologia (grafo) do AS inteiro
 - Cálculo de rota usando algoritmo de Dijkstra
 - Custos configurados pelo administrador
- ❖ Anúncio OSPF carrega uma entrada por vizinho
- ❖ Anúncios difundidos para AS *inteira* – quando há mudança no estado do enlace e periodicamente (30 min)
 - Carregados em mensagens OSPF diretamente sobre IP (em vez de sobre TCP ou UDP)

OSPF ferramentas “avançadas” (não estão no RIP)

- ❖ **Segurança:** todas mensagens OSPF são autenticadas (para evitar intrusões maliciosas)
- ❖ **Múltiplos caminhos** de mesmo custo permitidos (apenas um caminho no RIP)
- ❖ Para cada enlace, **múltiplas métricas de custo para diferentes ToS** (e.g., custo de enlace de satélite definido como “baixo” para ToS melhor esforço e alto para ToS tempo real)
- ❖ **Permite roteamento hierárquico** dentro de uma AS

OSPF Hierárquico



OSPF Hierárquico

- ❖ **Hierarquia de dois níveis:** área local e *backbone*
 - Cada área roda seu próprio algoritmo de roteamento LS; anúncios LS apenas na área
 - Cada nó tem topologia de área detalhada; apenas conhecem direção (menor caminho) para redes em outras áreas.
- ❖ **Roteadores de borda de área:** “resumem” distâncias para roteadores na própria área, anunciam para outros roteadores de borda de área; *todos pertencem ao backbone*
- ❖ **Roteadores backbone :** rodam roteamento OSPF limitado ao *backbone*.
- ❖ **Roteadores de borda (gateways):** conectam a outras AS's.

Capítulo 5: Roteiro

5.1 Introdução

5.2 Protocolos de roteamento

- Estado de enlace
- Vetor de distâncias

5.3 Roteamento intra-AS na Internet: OSPF

5.4 Roteamento entre os ISPs: BGP

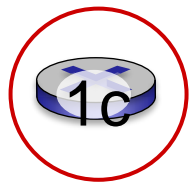
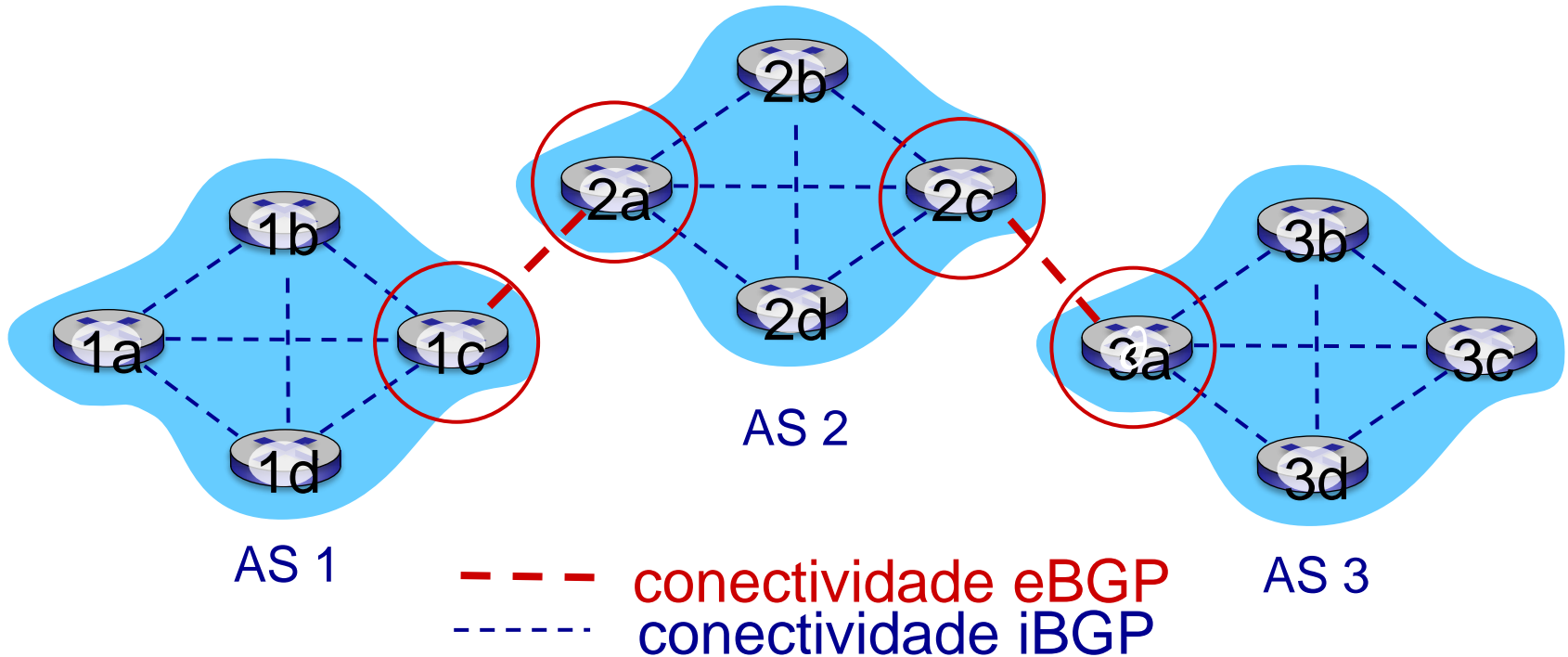
5.5 MPLS

5.6 Redes numa perspectiva mais ampla: introdução a redes complexas

Roteamento inter-AS na Internet: BGP

- ❖ **BGP** (*Border Gateway Protocol* – versão 4) [[RFC1105 \(1989\)](#) ... [RFC9072 \(2021\)](#)]: o protocolo de roteamento interdomínios *de fato*
 - “*Cola que mantém a Internet junta*”
- ❖ BGP provê a cada AS uma forma de:
 - Obter informação de alcançabilidade de sub-rede de ASs vizinhas (**eBGP**)
 - Propagar informação de alcançabilidade para todos roteadores internos à AS (**iBGP**)
 - Determinar “boas” rotas para sub-redes baseando-se nas informações de alcançabilidade e política da AS.
- ❖ Permite a uma sub-rede anunciar sua existência ao resto da Internet: “*Eu estou aqui!*”

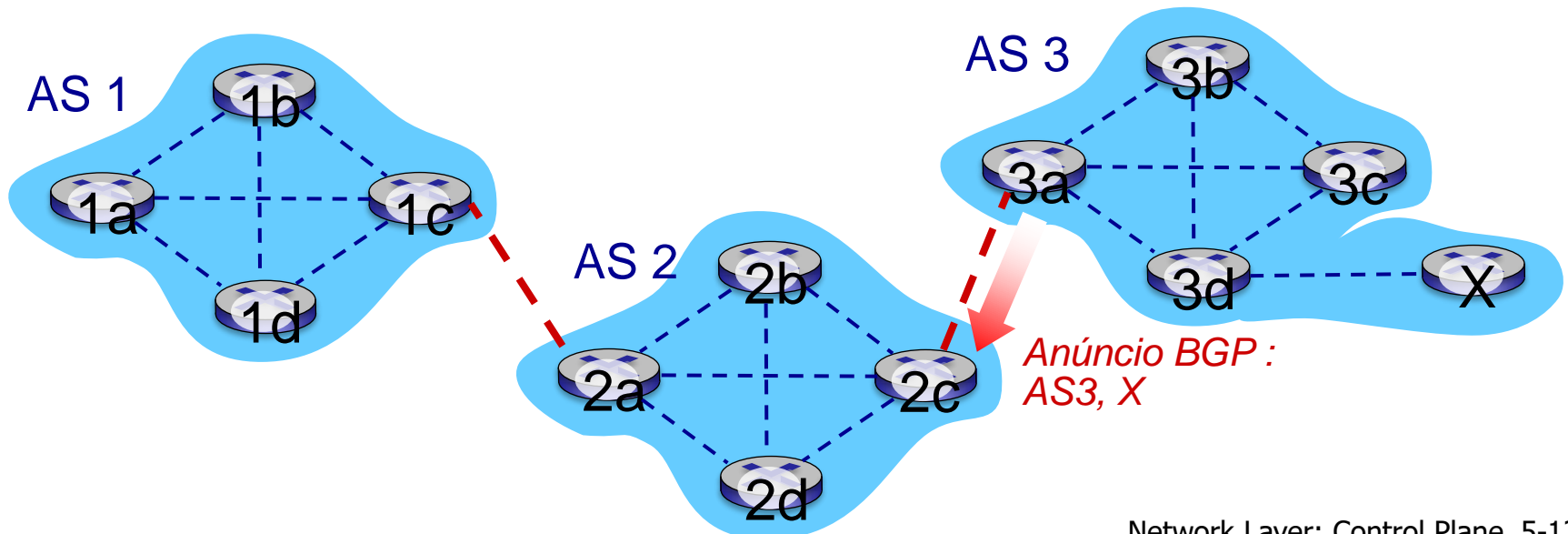
Conexões eBGP, iBGP



roteadores *gateway* rodam ambos protocolos eBGP e iBGP

Básico do BGP

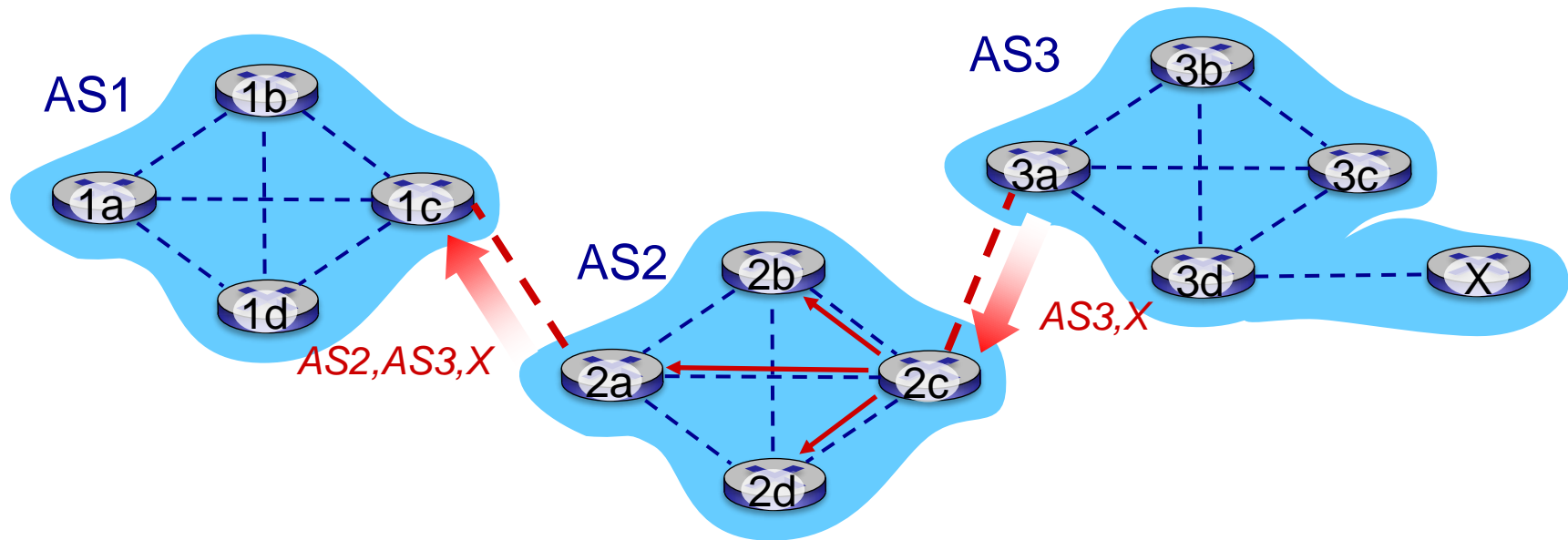
- ❖ Extremamente complicado ☹️
- ❖ **Sessão BGP**: 2 roteadores BGP (“peers”) trocam mensagens BGP:
 - Anunciando *caminhos* para diferentes *prefixos* de rede destinos (protocolo de “vetor de caminhos”)
 - Trocados por conexões TCP (porta 179)
- ❖ Quando *Gateway 3a* de *AS3* anuncia caminho **AS3,X** para o *Gateway 2c* de *AS2*:
 - *AS3 promete* que irá repassar datagramas até esse prefixo



Atributos de caminhos e rotas BGP

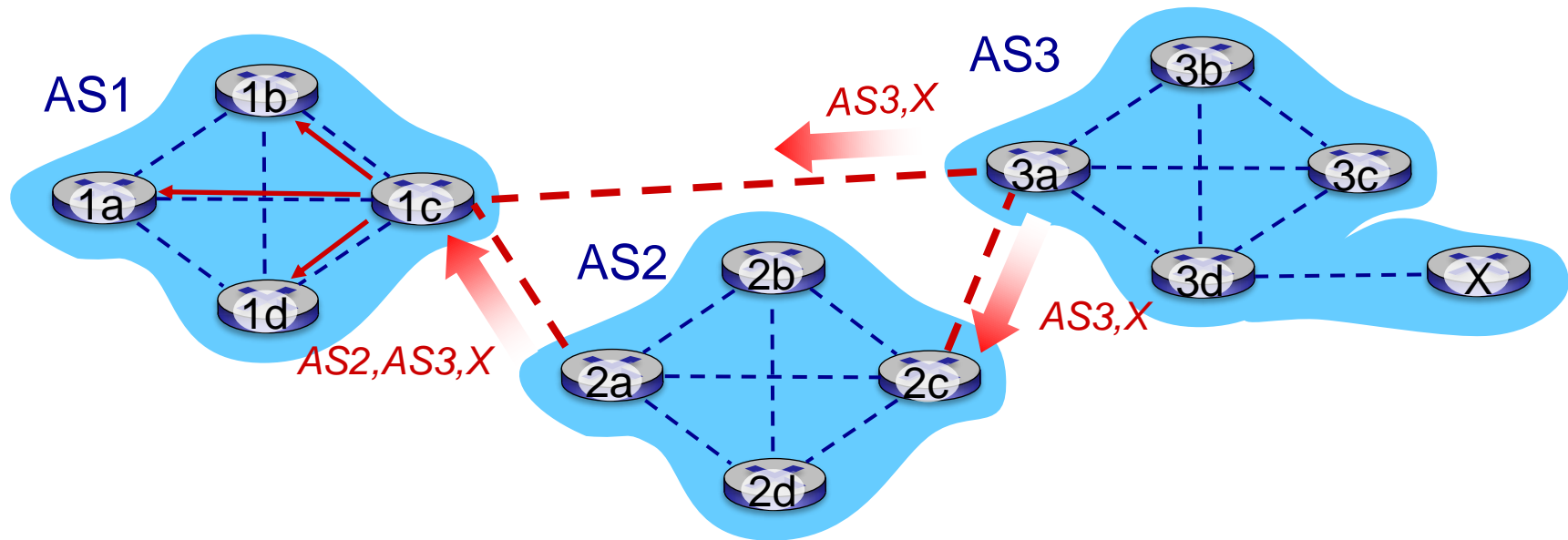
- ❖ Prefixo anunciado inclui atributos BGP
 - prefixo + atributos = “rota”
- ❖ 2 atributos importantes:
 - **AS-PATH**: contém ASN das ASs através dos quais anúncio de prefixo foi repassada: e.g., AS 67, AS 17
 - **NEXT-HOP**: indica roteador interno a AS específico na AS de próximo salto. (Podem haver múltiplos enlaces do AS corrente para o AS de próximo salto)
- ❖ Roteador *gateway* recebendo anúncio de rota usa **políticas de importação** para aceitar/declinar
 - por exemplo, nunca rotear através da AS x
 - Roteamento *baseado em política*

Anúncio de rota BGP



- Roteador 2c de AS2 recebe anúncio de caminho **AS3,X** (via eBGP) do Roteador 3a de AS3
- Baseado na política de AS2, o Roteador 2c de AS2 aceita caminho **AS3,X** e propaga (via iBGP) para todos os Roteadores de AS2
- Baseado na política de AS2, roteador 2a de AS2 anuncia (via eBGP) caminho **AS2,AS3,X** para Roteador 1c de AS1

Anúncio de rota BGP



Roteador *gateway* pode aprender **múltiplos** caminhos para um destino:

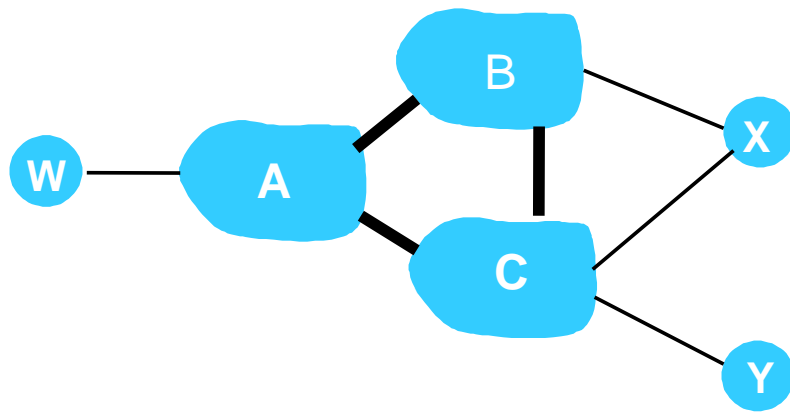
- Roteador *gateway* 1c de AS1 aprende caminho **AS2,AS3,X** de 2a
- Roteador *gateway* 1c de AS1 aprende caminho **AS3,X** de 3a
- Baseado na política de AS1, *gateway* 1c escolhe caminho **AS3,X**, e **anuncia caminho para AS1 via iBGP**

Seleção de rota BGP

❖ Roteador pode aprender sobre mais do que 1 rota até AS destino; seleciona rota baseado em:

1. Valor do atributo de preferência local: decisão política
2. AS-PATH mais curto
3. Roteador NEXT-HOP mais próximo: roteamento batata quente
4. Critérios adicionais

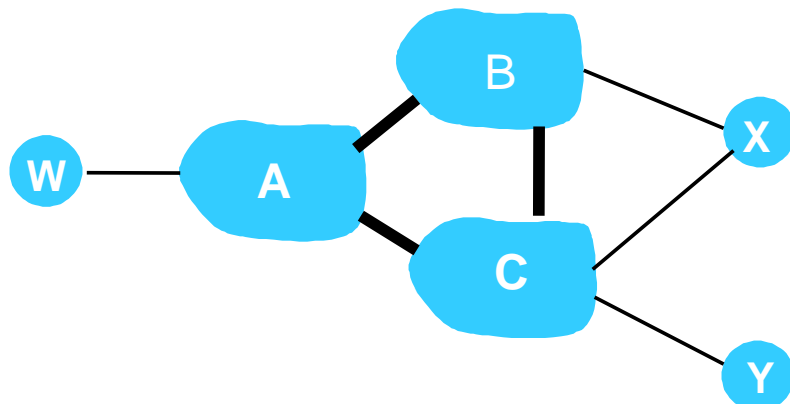
Política de roteamento BGP





legenda:  rede de provedor (AS)
 rede de cliente (AS)

- ❖ A,B,C são *redes de provedores*
- ❖ X,W,Y são clientes (de redes provedoras)
- ❖ X é *dual-homed*: ligada a duas redes
 - X não quer rotear de B via X para C
 - .. assim X não vai anunciar para B uma rota para C

Política de roteamento BGP (2)



legenda:  rede de provedor
 rede de cliente

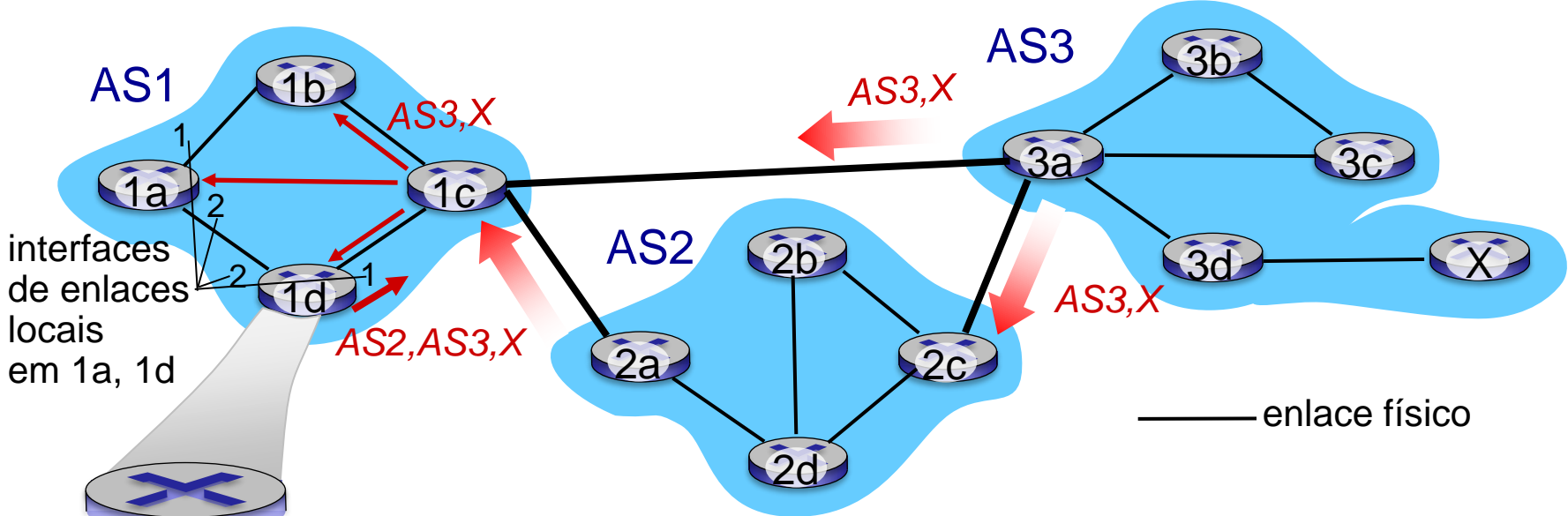
- ❖ A anuncia caminho AW para B
- ❖ B anuncia caminho BAW para X
- ❖ B deve anunciar BAW para C?
 - De forma alguma! B não tem nenhum ganho por rotear CBAW já que nem W nem C são clientes de B
 - B quer forçar C a rotear para w via A
 - B quer rotear *apenas* de/para seus clientes! (regra prática adotada por ISPs)
- ❖ Mais exemplos em <http://www.routeviews.org/> e <http://www.caida.org/tools/>

Mensagens BGP

- ❖ Mensagens BGP trocadas entre *peers* sobre conexão TCP
- ❖ Mensagens BGP :
 - **OPEN**: abre conexão TCP com *peer* e autentica remetente
 - **UPDATE**: anuncia novo caminho (ou arquiva antigo)
 - **KEEPALIVE**: mantém conexão ativa na ausência de UPDATES; também serve para ACKs pedido OPEN
 - **NOTIFICATION**: reporta erros em mensagens anteriores; também usado para fechar conexão

Entradas de tabela de repasse: BGP e OSPF

Q: Como roteador define tabela de repasse para prefixo distante?

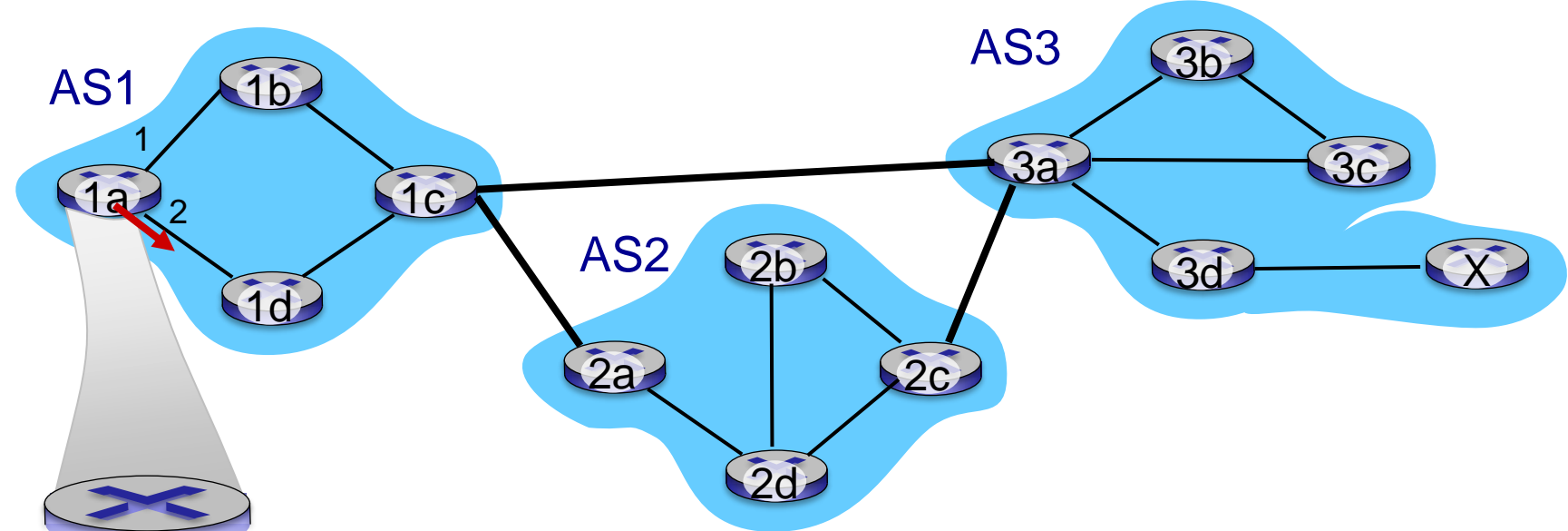


Dest	Interface
...	...
X	1
...	...

- Lembrando: 1a, 1b e 1d aprendem sobre destino X via iBGP por meio de 1c: “caminho para X passa por 1c”
- 1d: roteamento intradomínio OSPF: para chegar a 1c, repasse através da interface local de saída 1

Entradas de tabela de repasse: BGP e OSPF

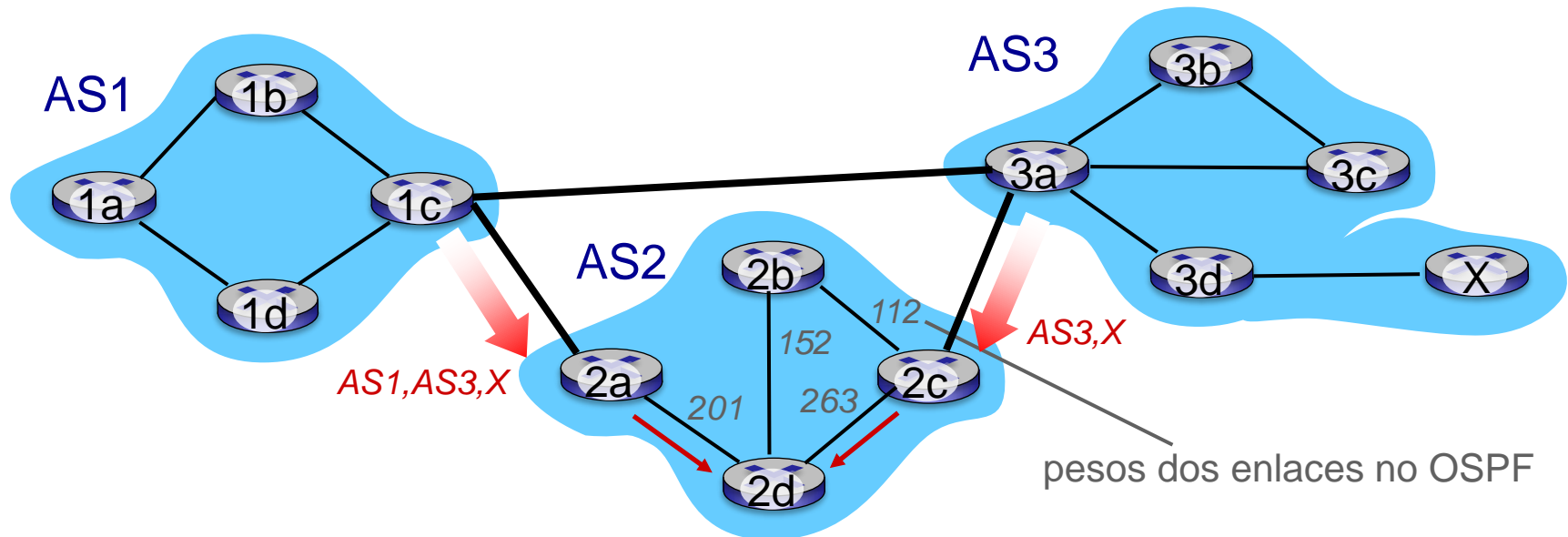
Q: Como roteador define tabela de repasse para prefixo distante?



dest	interface
...	...
X	2
...	...

- lembrando: 1a, 1b e 1c e 1d aprendem sobre destino X via iBGP por meio de 1c
- 1d: roteamento intradomínio OSPF: para chegar a 1c, repasse através da interface local de saída 1
- 1a: roteamento intradomínio OSPF: para chegar a 1c, repasse através da interface local de saída 2

Roteamento “Batata Quente”



- 2d aprende (via iBGP) que pode rotear para X via 2a or 2c
- **Roteamento “batata quente”**: escolhe gateway local que tem mínimo custo intradomínio (por exemplo, 2d escolhe 2a, apesar de mais saltos entre ASs para chegar a X): não se preocupa com custos interdomínios!

Por que roteamentos Intra-, Inter-AS diferentes?

Política:

- ❖ **inter-AS**: administrador quer controle sobre como tráfego é roteado, quem roteia sobre sua rede.
- ❖ **intra-AS**: administrador único, assim não são necessárias decisões políticas

Escala:

- ❖ roteamento hierárquico economiza tamanho de tabela, reduz tráfego de atualizações

Desempenho:

- ❖ **intra-AS**: pode focar em desempenho
- ❖ **inter-AS**: política pode dominar sobre desempenho
(nem noção de custo é usada...)

Exemplos de número de AS (ASN)

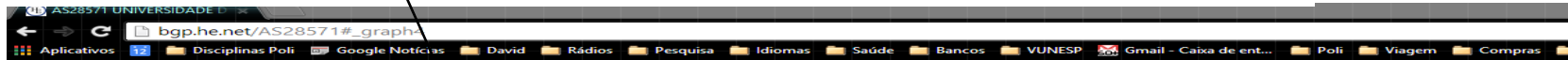


AS Numbers Query

[AS28571](#) Universidade de Sao Paulo - USP

[AS30515](#) USP - USP

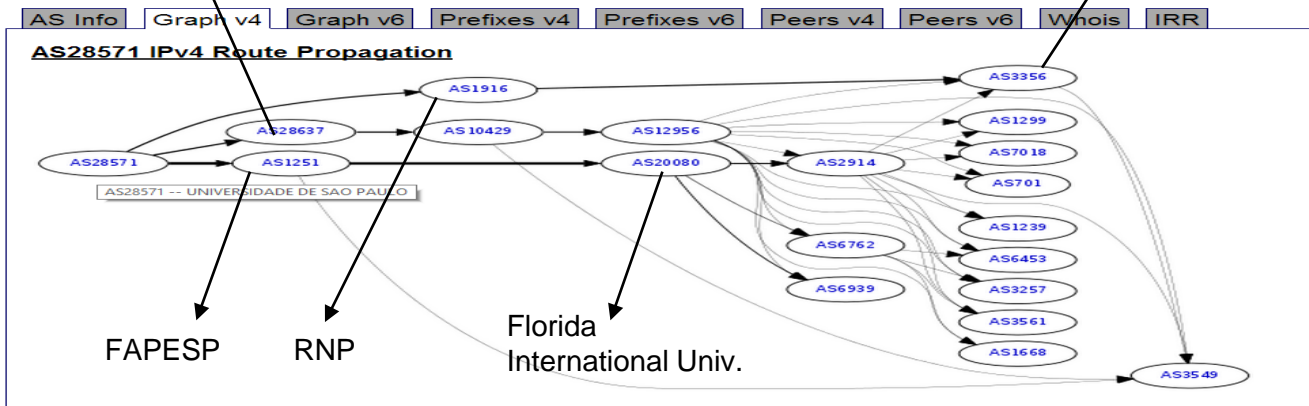
PRODESP



HURRICANE ELECTRIC
INTERNET SERVICES

AS28571 UNIVERSIDADE DE SAO PAULO

- Quick Links**
- BGP Toolkit Home
 - BGP Prefix Report
 - BGP Peer Report
 - Bogon Routes
 - World Report
 - Multi Origin Routes
 - DNS Report
 - Top Host Report
 - Internet Statistics
 - Looking Glass
 - Free IPv6 Tunnel
 - IPv6 Certification
 - IPv6 Progress
 - Going Native
 - Contact Us



Capítulo 5: Roteiro

5.1 Introdução

5.2 Protocolos de roteamento

- Estado de enlace
- Vetor de distâncias

5.3 Roteamento intra-AS na Internet: OSPF

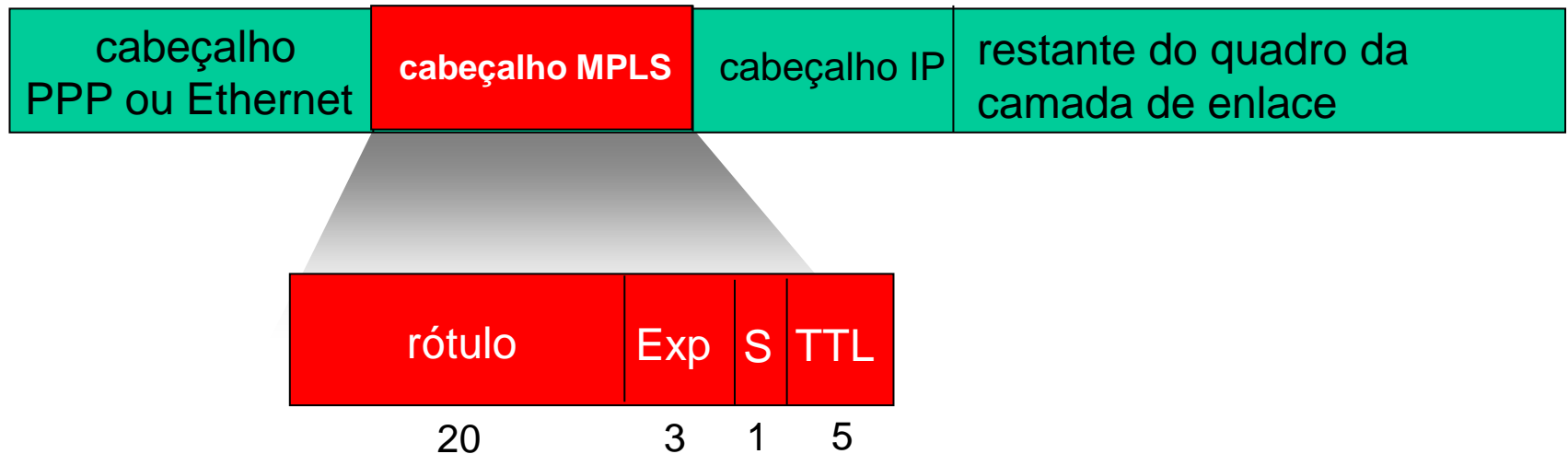
5.4 roteamento entre os ISPs: BGP

5.5 MPLS

5.6 Redes numa perspectiva mais ampla: introdução a redes complexas

Multiprotocol label switching (MPLS) [RFC3032 (2001)... RFC9017 (2021)]

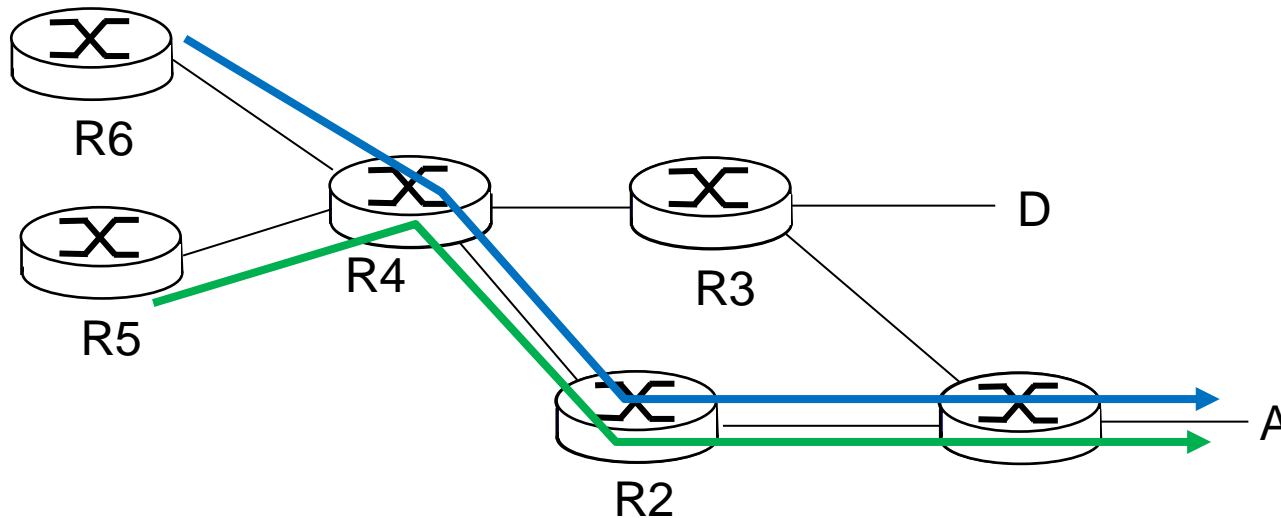
- ❖ Objetivo inicial : repasse IP em alta velocidade usando rótulo de comprimento fixo (em vez de endereço IP)
 - Pesquisa rápida usando identificador de comprimento fixo (ao invés de casamento com menor prefixo)
 - Empréstimo de ideias da abordagem de comutação de circuito
 - Mas datagramas IP ainda mantêm endereços IP!
 - Somente entre roteadores com capacidade MPLS (ou *roteador com comutação por rótulos* – label-switched router)



Roteadores com capacidade MPLS

- ❖ Repassa pacotes para interface de saída baseado apenas no valor do rótulo (*não inspeciona endereço IP*)
 - Tabelas de repasse MPLS distintas das tabelas de repasse IP
- ❖ **Flexibilidade:** decisões de repasse MPLS podem *diferir* daquelas do IP
 - Usa endereços de fonte e destino para rotear fluxos para mesmo destino de forma diferente (engenharia de tráfego)
 - Altera roteamento de fluxos rapidamente se enlace falha: caminhos de *backup* pré-computados (útil para VoIP)

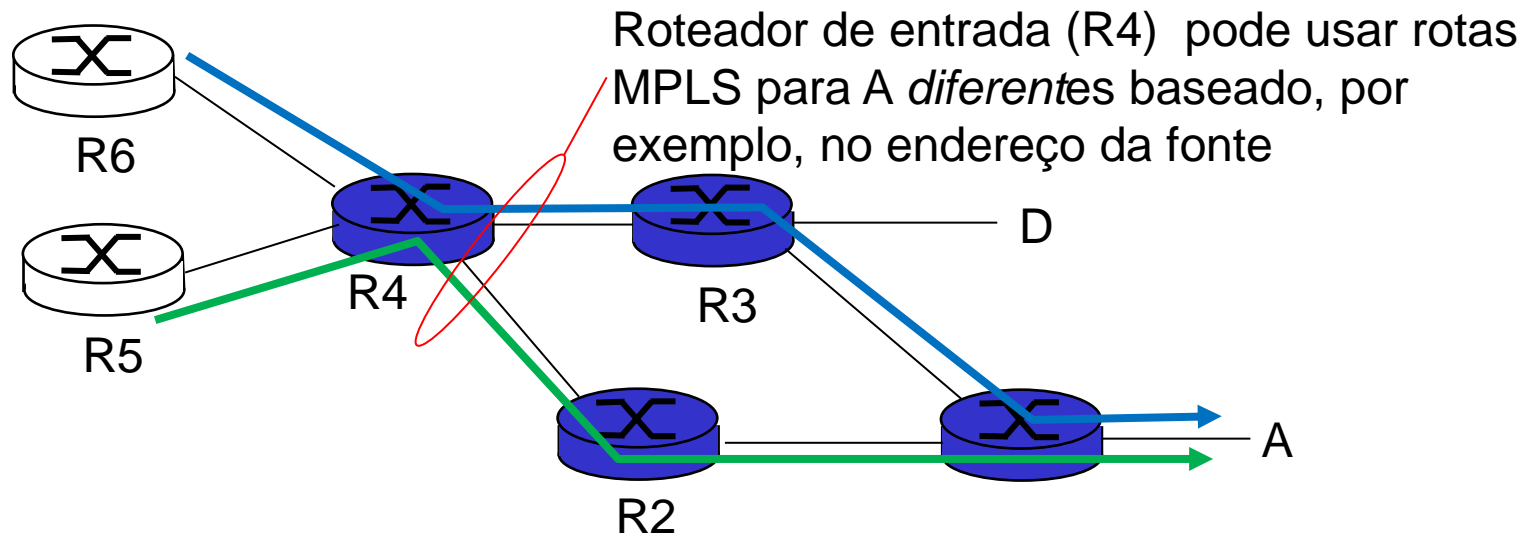
Caminhos MPLS versus IP



- ❖ **Roteamento IP** : caminho para destino determinado apenas pelo endereço de destino



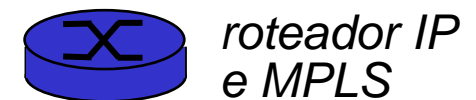
Caminhos MPLS versus IP



❖ **Roteamento IP** : caminho para destino determinado pelo endereço do destino apenas



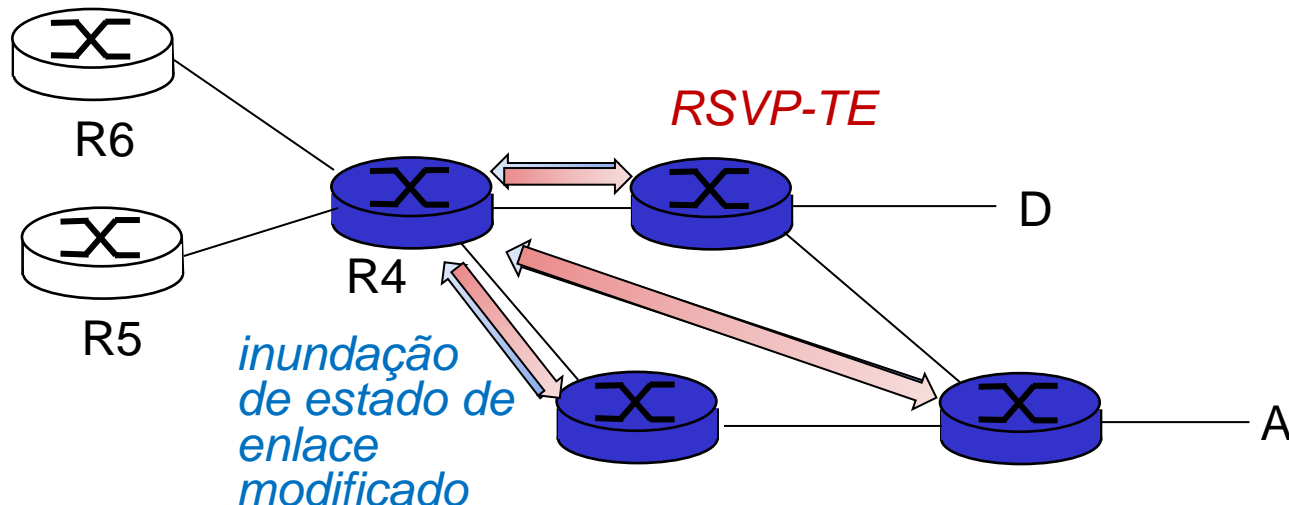
❖ **Roteamento MPLS** : caminho para destino pode ser baseado nos endereços da fonte e do destino



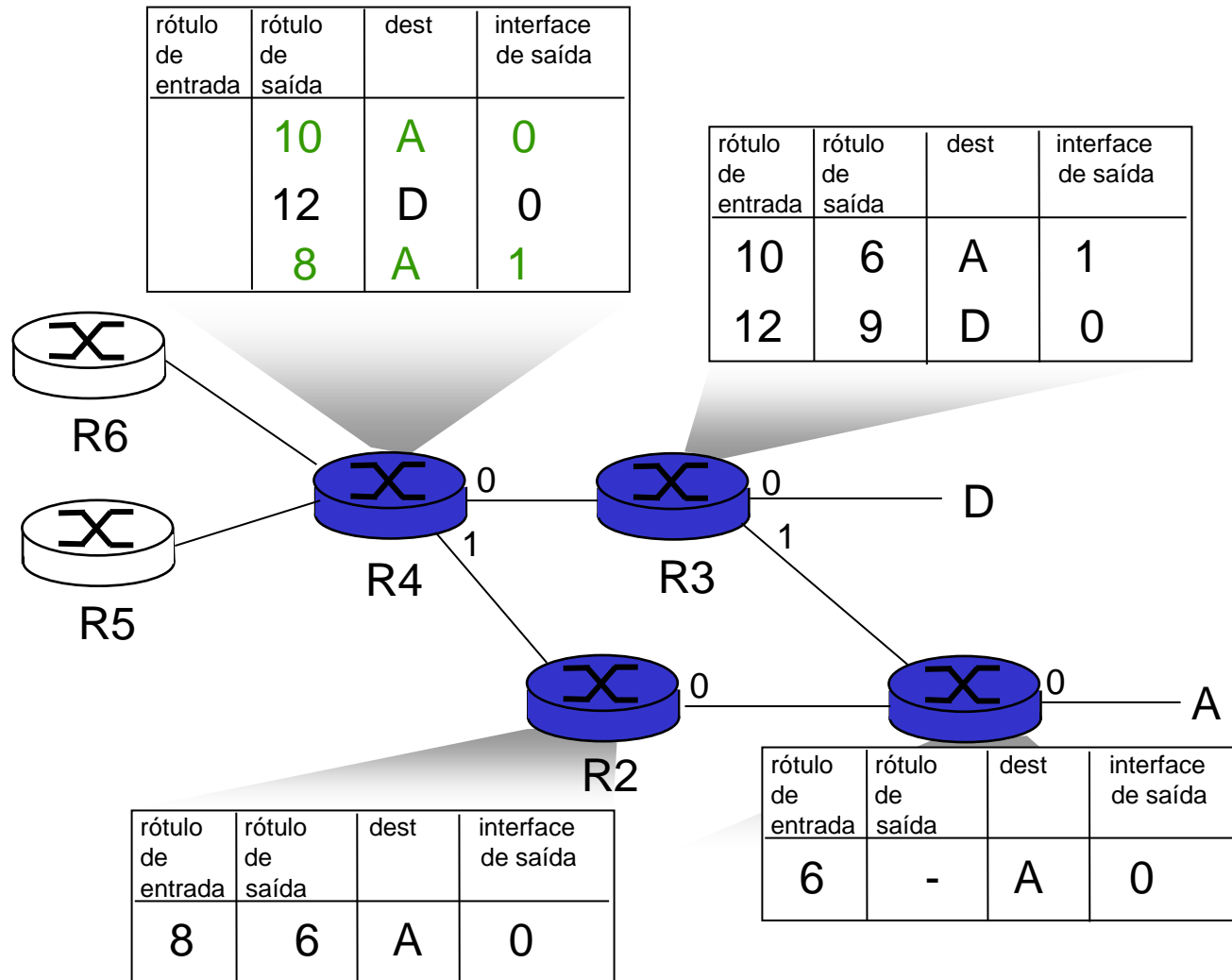
- **Alteração de roteamento rápido** : pré-computa rotas de *backup* em caso de falha de enlace

Sinalização MPLS

- ❖ Modificação dos protocolos de inundação de estado de enlace OSPF para levar informações usadas pelo roteamento MPLS
 - e.g., largura de banda de enlace, quantidade de “largura de banda reservada” da largura de banda do enlace
- ❖ Roteador MPLS de entrada usa protocolo de sinalização RSVP-TE (*Resource Reservation Protocol - Traffic Engineering*) [[RFC 3032 \(2001\)](#)]-... – [7274 \(2014\)](#)] para contruir repasse MPLS nos roteadores a jusante

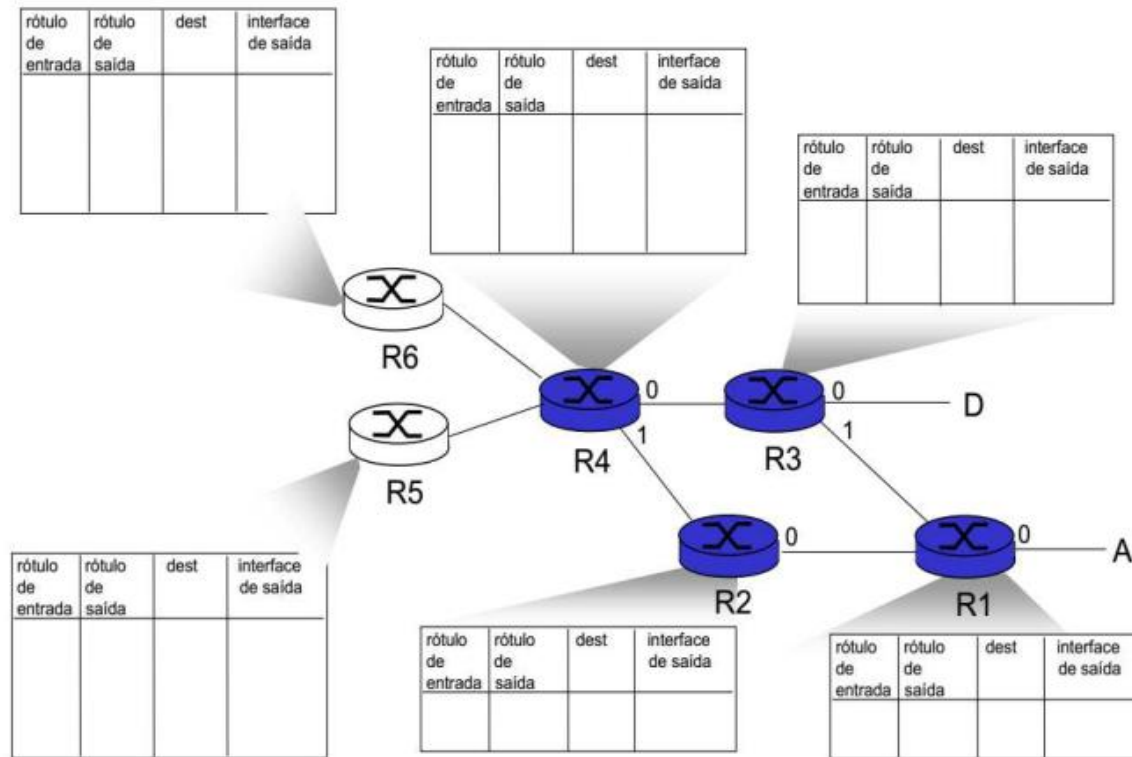


Tabelas de repasse MPLS



Exemplo

2) [Kurose and Ross, 2017, p. 542] (2,0) Considere a rede MPLS ilustrada em aula e mostrada a seguir.



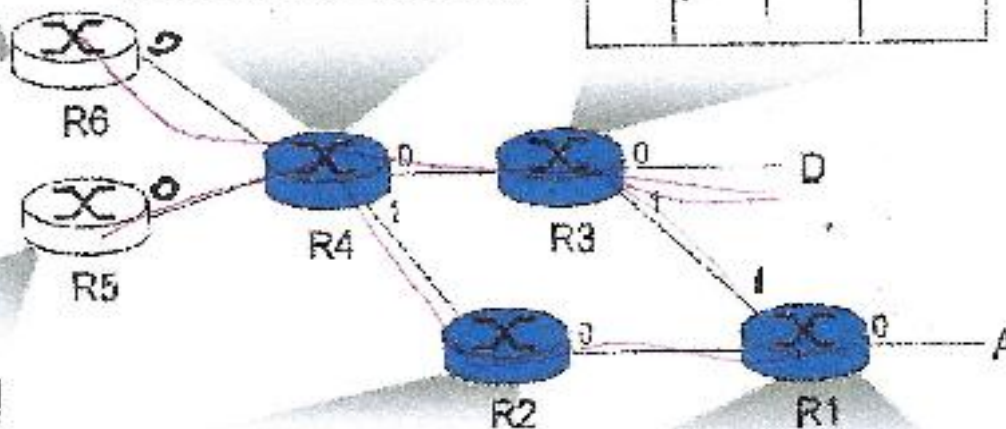
Nesse problema, considere que os roteadores $R5$ e $R6$ também tem capacidade MPLS (*MPLS enabled*). Suponha que os pacotes de $R6$ destinados a D devam ser comutados via $R6 - R4 - R3$ enquanto os pacotes de $R5$ para D devam ser comutados via $R4 - R2 - R1 - R3$. Preencha as tabelas MPLS de todos os roteadores que tornem isso possível. Justifique suas respostas.

Nesse problema, considere que os roteadores *R5* e *R6* também tem capacidade MPLS (*MPLS enabled*). Suponha que os pacotes de *R6* destinados a *D* devam ser comutados via *R6 – R4 – R3* enquanto os pacotes de *R5* para *D* devam ser comutados via *R4 – R2 – R1 – R3*. Preencha as tabelas MPLS de todos os roteadores que tornem isso possível. Justifique suas respostas.

rótulo de entrada	rótulo de saída	dest	interface de saída
1	3	D	0

rótulo de entrada	rótulo de saída	dest	interface de saída
3	12	D	0
2	4	D	1

rótulo de entrada	rótulo de saída	dest	interface de saída
12	1	D	0



rótulo de entrada	rótulo de saída	dest	interface de saída
1	2	D	0

rótulo de entrada	rótulo de saída	dest	interface de saída
4	1	D	0

rótulo de entrada	rótulo de saída	dest	interface de saída
1	12	D	1