

PTC 3450 - Aula 23

5.3 Roteamento intra-AS na Internet: OSPF

5.4 Roteamento entre os ISPs: BGP

(Kurose, p. 280-306)

(Peterson, p. 147-163)

23/06/2017

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

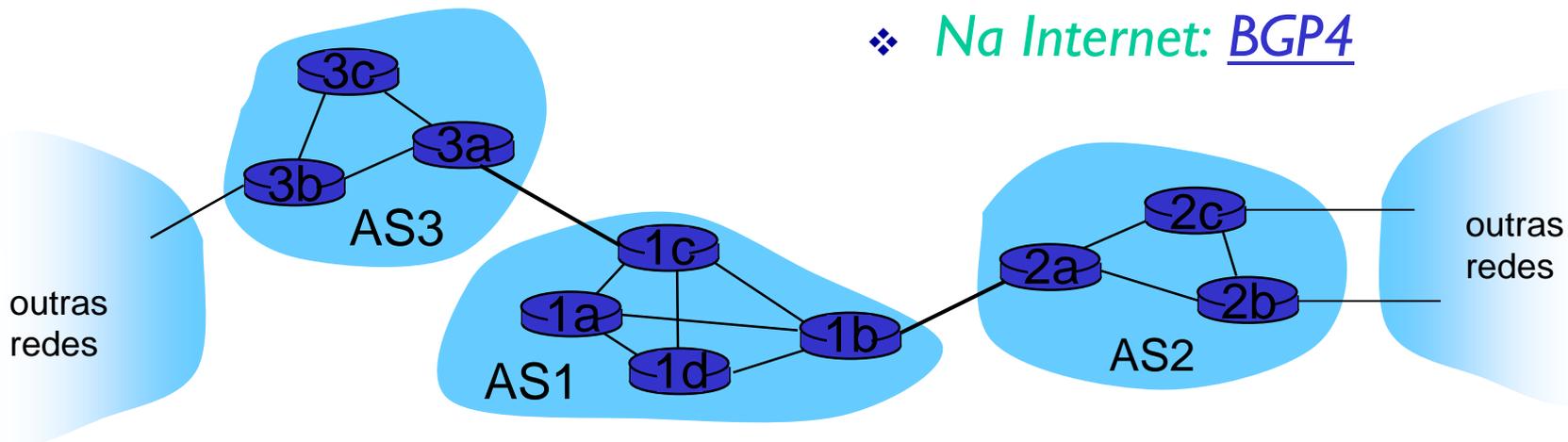
5.5 O plano de controle do SDN

5.6 ICMP: o *Internet Control Message Protocol*

5.7 Administração de rede e o SNMP

Tarefa Inter-AS

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



AS1 precisa:

1. Aprender quais destinos são alcançáveis via AS2, quais via AS3
 2. propagar essa info de alcançabilidade para todos roteadores em AS1
- ❖ *Trabalho do roteamento inter-AS!*
 - ❖ *Todas as ASs precisam rodar mesmo protocolo inter-AS!*
 - ❖ *Na Internet: BGP4*

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 – ISPs de níveis mais baixos e redes corporativas*
 - *OSPF: Open Shortest Path First – ISPs de níveis mais altos*
 - EIGRP: *Enhanced Interior Gateway Routing Protocol* (proprietário - Cisco)

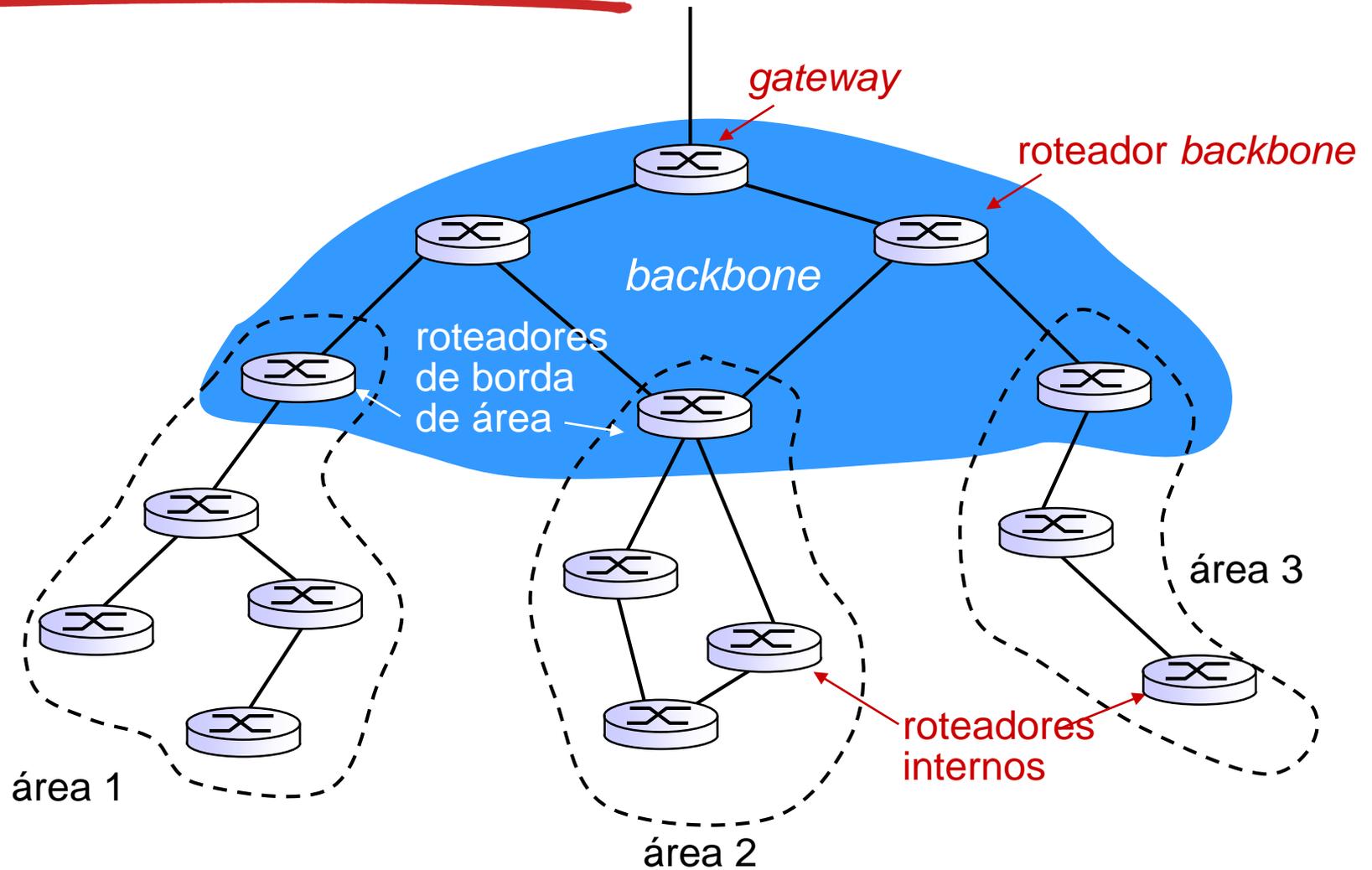
OSPF (*Open Shortest Path First*)

- ❖ “*open*”: disponível publicamente [[RFC 1247 \(1991\)](#) ... [RFC 7474 \(2015\)](#)]
- ❖ “sucessor” do RIP – características mais avançadas
- ❖ usa algoritmo de estado de enlace
 - inundação de pacotes LS
 - 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 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)
- ❖ suporte a **uni e multicast** integrado:
 - OSPF *multicast* (MOSPF) usa mesma base de dados da topologia que o OSPF
- ❖ **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 redes 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 O plano de controle do SDN

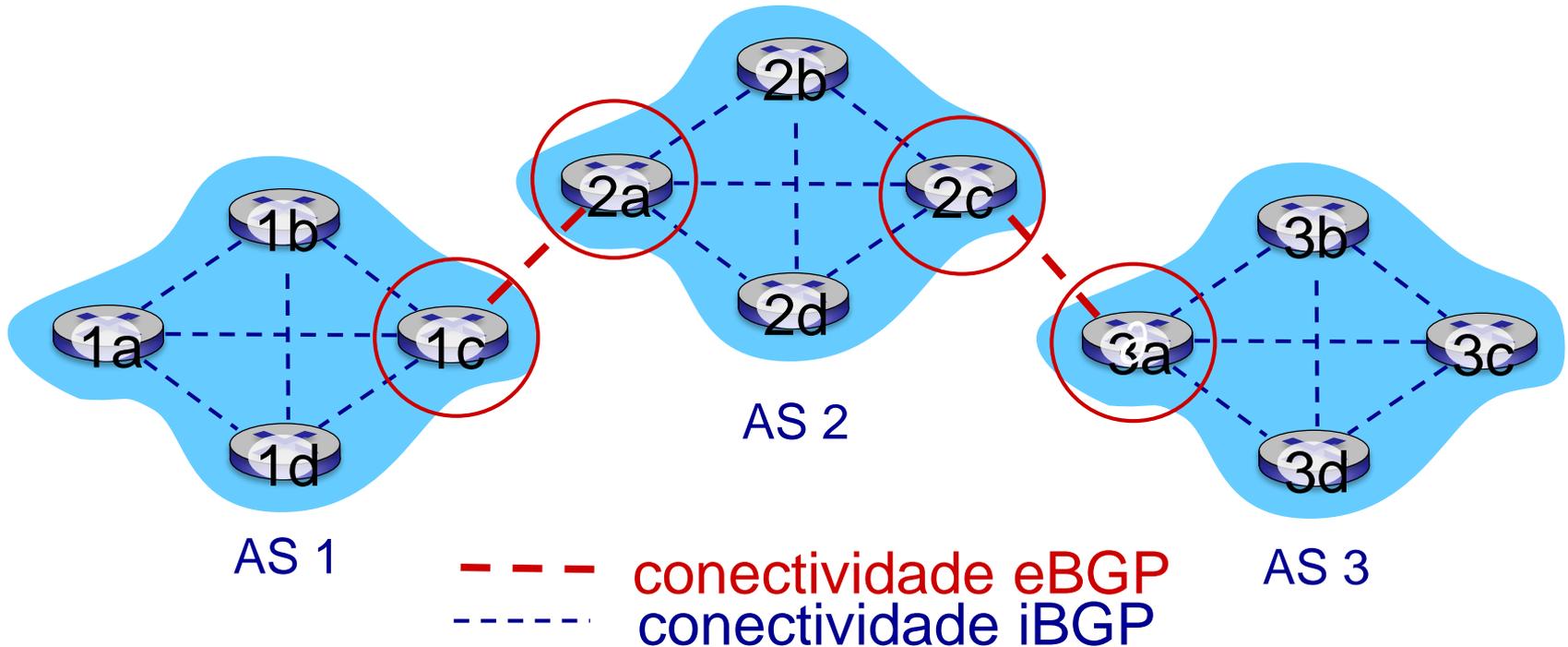
5.6 ICMP: o *Internet Control Message Protocol*

5.7 Administração de rede e o SNMP

Roteamento inter-AS na Internet: BGP

- ❖ **BGP** (*Border Gateway Protocol – versão 4*) [[RFC 1105 \(1989\)](#) ... [7705 \(2015\)](#)]: o protocolo de roteamento inter-domínios *de fato*
 - *“cola que mantém a Internet junta”*
- ❖ BGP provê a cada AS uma forma de:
 - **eBGP**: obter informação de alcançabilidade de sub-rede de ASs vizinhas.
 - **iBGP**: propagar informação de alcançabilidade para todos roteadores internos à AS.
 - 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!”*

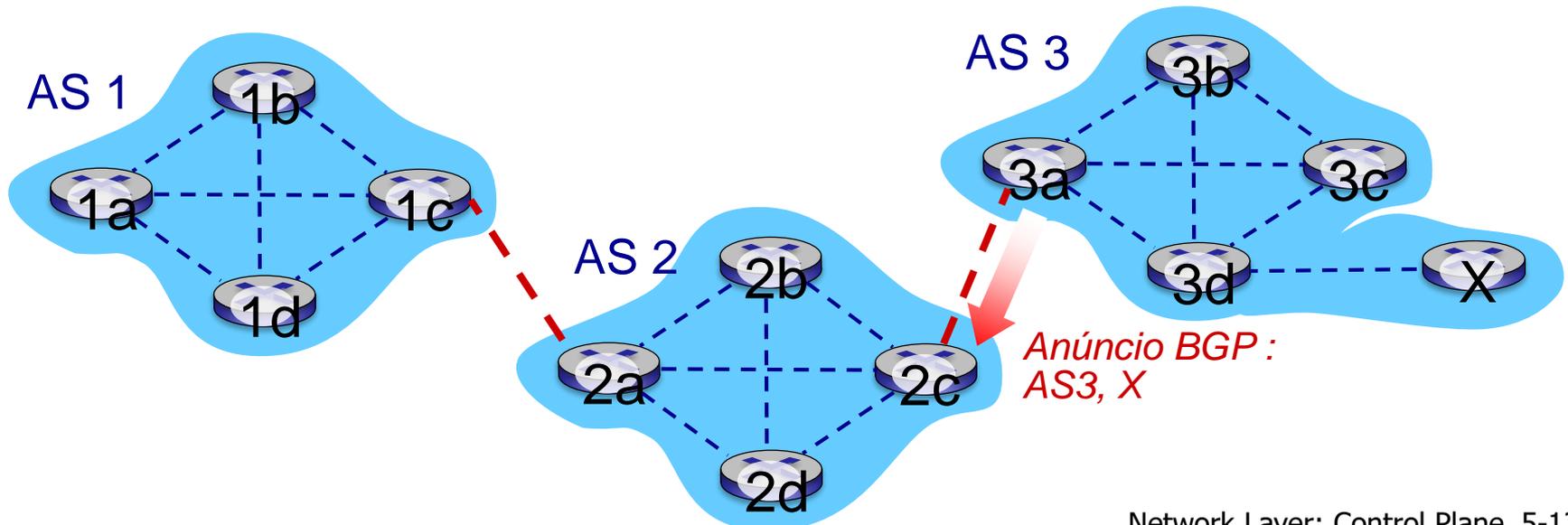
Conexões eBGP, iBGP



roteadores *gateway* rodam ambos protocolos eBGP e iBGP

Básico do BGP

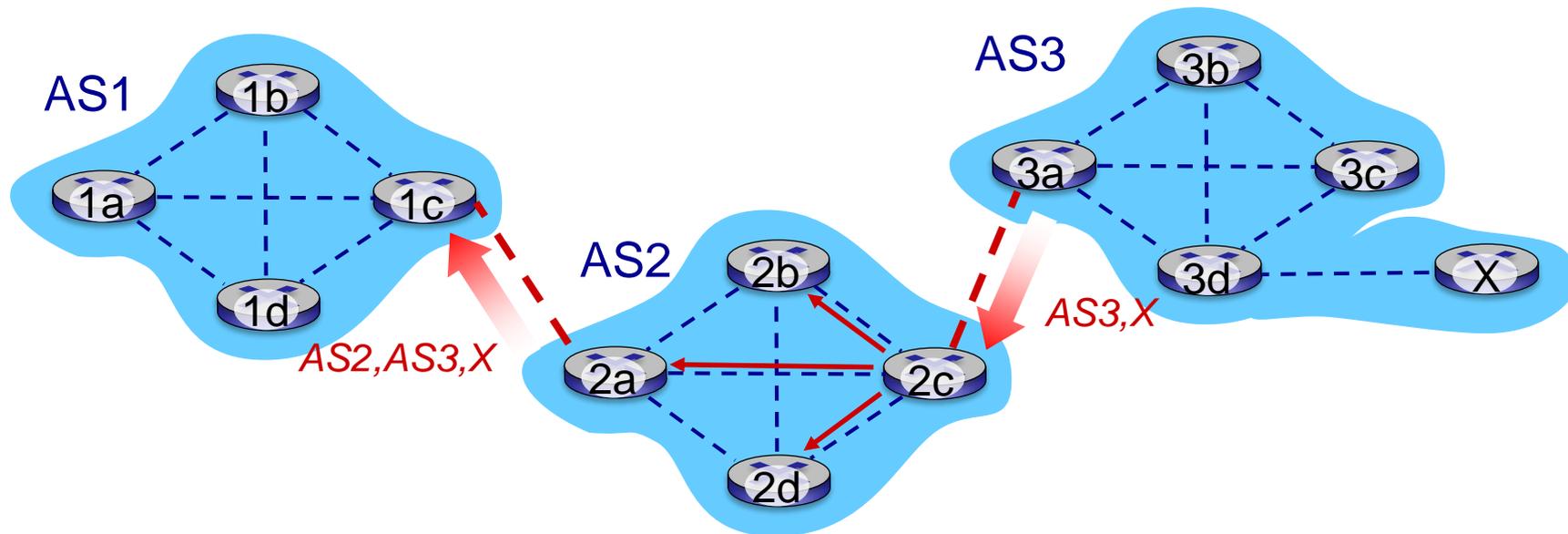
- ❖ Extremamente complicado ☹️
- ❖ **sessão BGP**: 2 roteadores BGP (“peers”) trocam mensagens BGP:
 - anunciando *caminhos* para diferentes *prefixos* de rede destinos (protocolo de “vetor de caminhos”)
 - trocados por conexões TCP (porta 179)
- ❖ quando *gateway 3a* de *AS3* anuncia caminho **AS3,X**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



Atributos de caminhos e rotas BGP

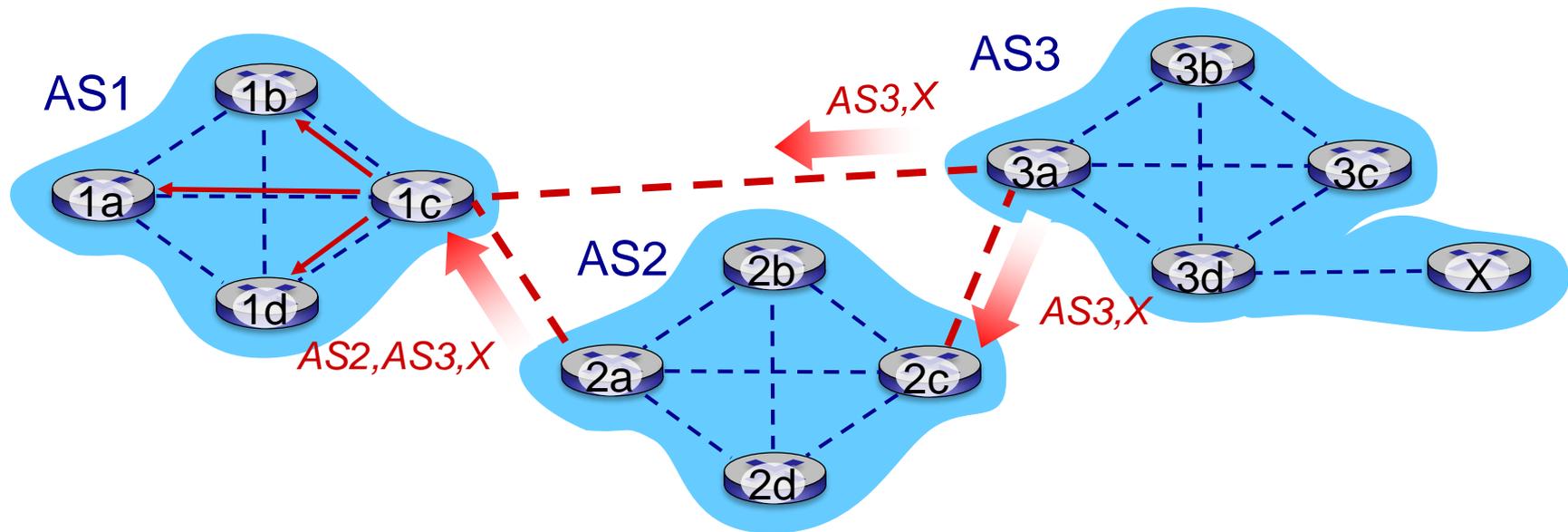
- ❖ prefixo anunciado inclui atributos BGP
 - prefixo + atributos = “rota”
- ❖ 2 atributos importantes:
 - **AS-PATH**: contém ASN das ASs através dos quais anúncio de prefixo foi repassada: e.g., AS 67, AS 17
 - **NEXT-HOP**: indica roteador interno a AS específico na AS de próximo salto. (podem haver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



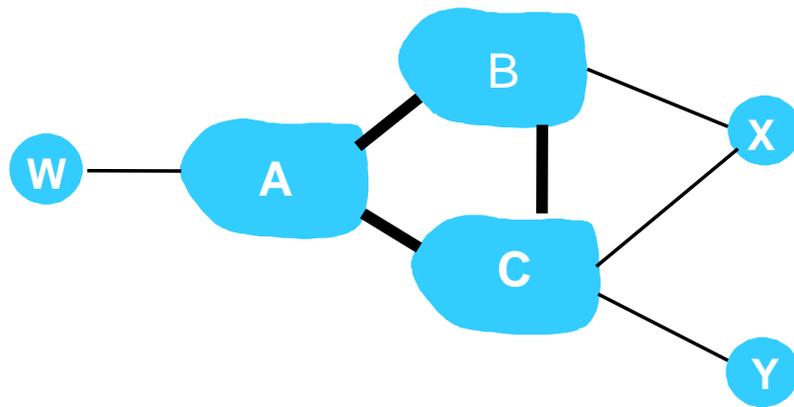
roteador *gateway* pode aprender **múltiplos** caminhos para um destino:

- 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

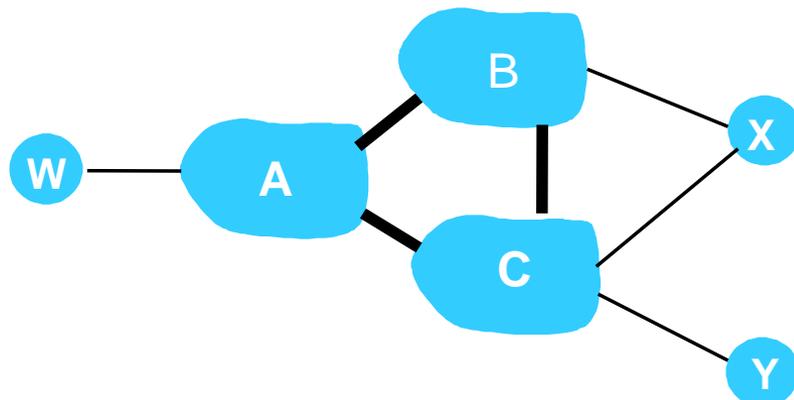
Política de roteamento BGP



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

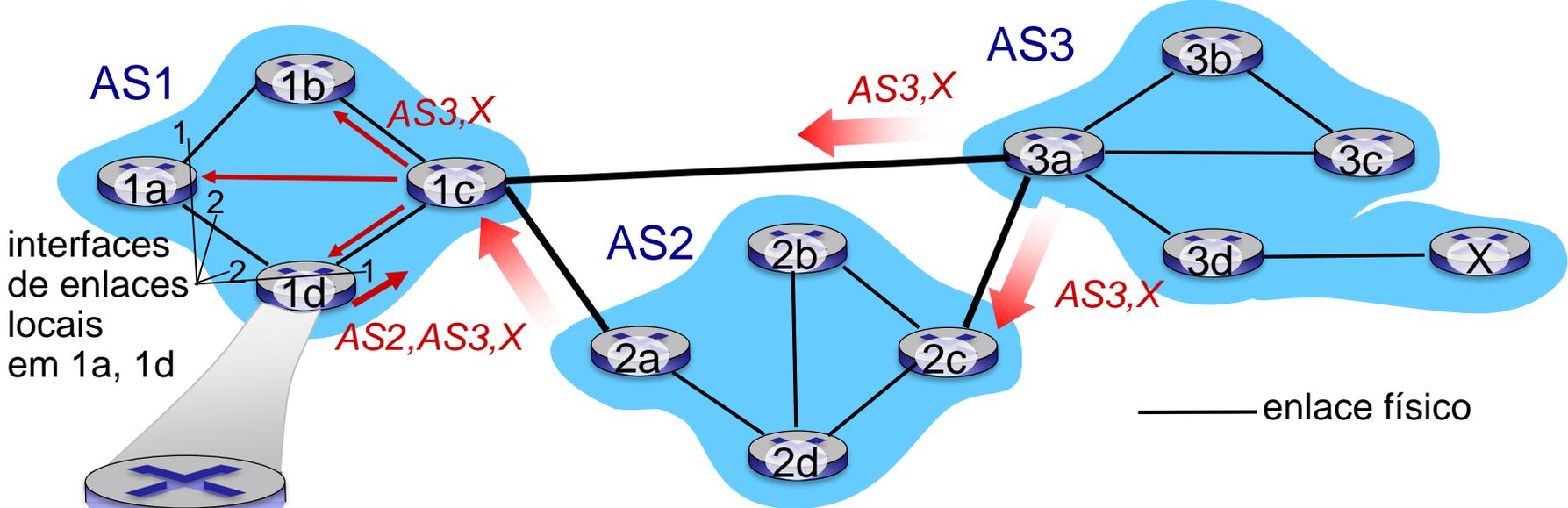
- ❖ A anuncia caminho AW para B
- ❖ B anuncia caminho BAW para X
- ❖ B deve anunciar BAW para C?
 - De forma alguma! B não tem nenhum ganho por rotear CBAW já que nem W nem C são clientes de B
 - B quer forçar C a rotear para w via A
 - B quer rotear *apenas* de/para seus clientes! (regra prática adotada 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?



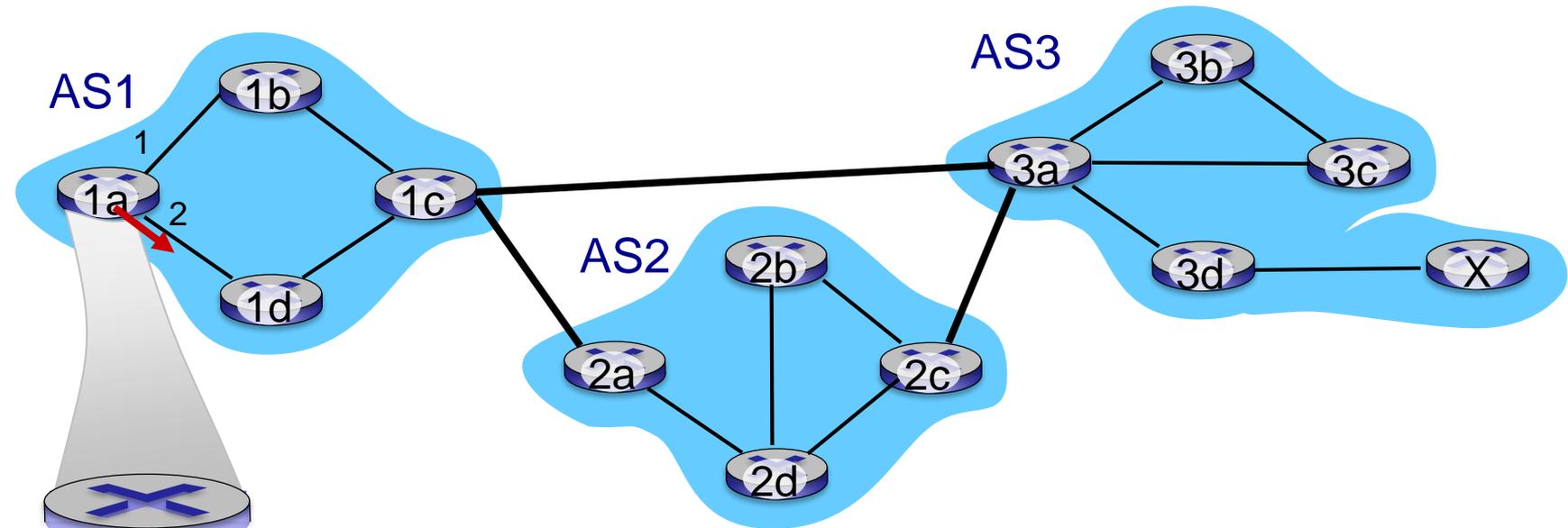
interfaces de enlaces locais em 1a, 1d

dest	interface
...	...
X	1
...	...

- lembrando: 1a, 1b e 1c 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

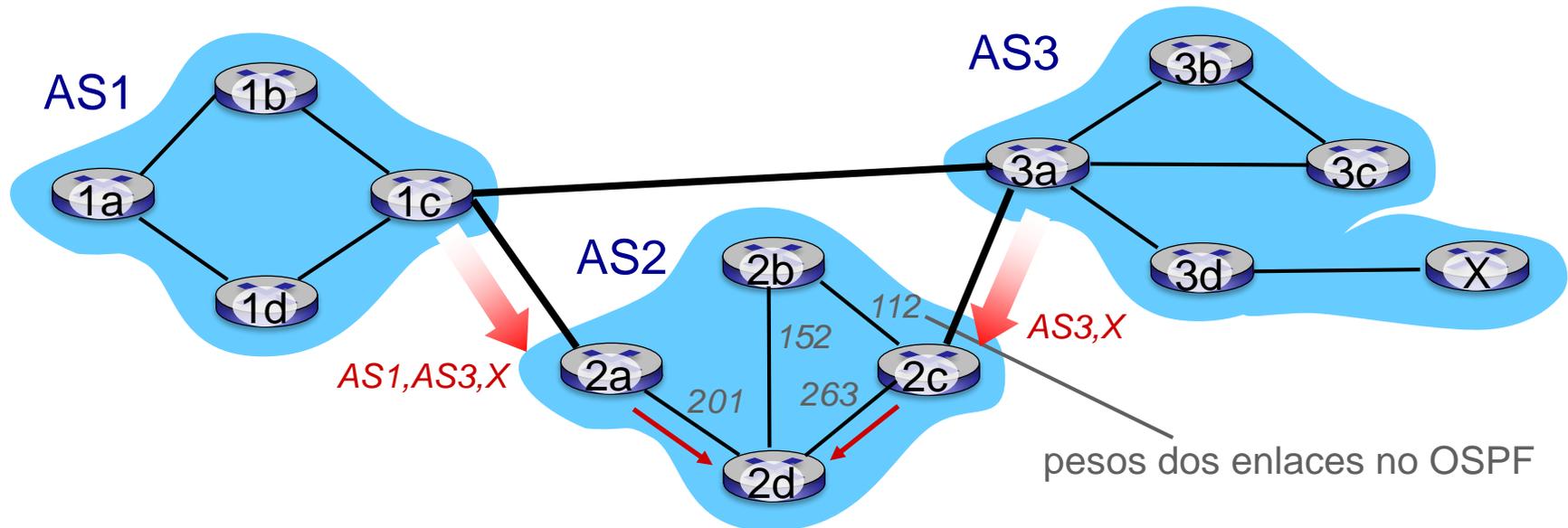
Q: Como roteador define tabela de repasse para prefixo distante?



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...*)

