

# Chapter 5

## Network Layer: Control Plane

A note on the use of these PowerPoint slides:

We're making these slides freely available to all (faculty, students, readers). They're in PowerPoint form so you see the animations; and can add, modify, and delete slides (including this one) and slide content to suit your needs. They obviously represent a *lot* of work on our part.

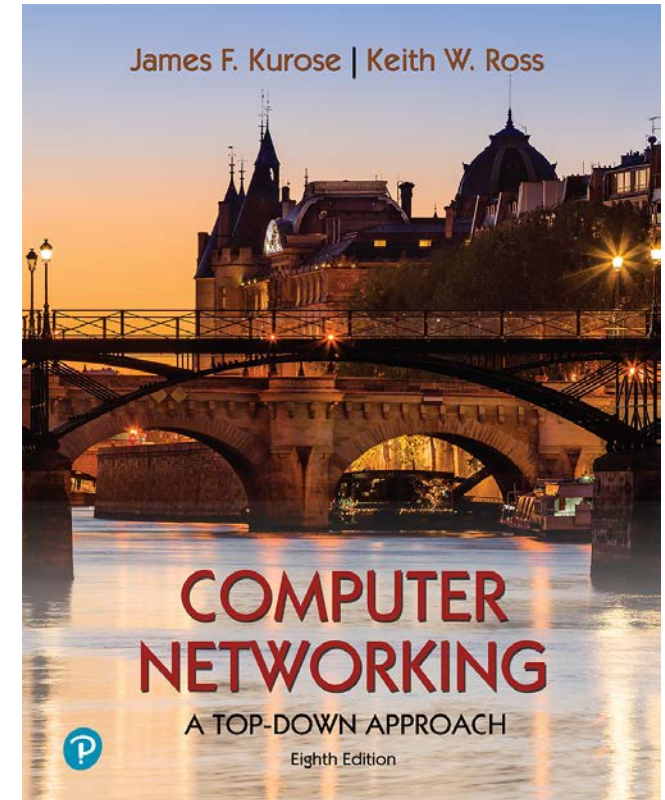
In return for use, we only ask the following:

- If you use these slides (e.g., in a class) that you mention their source (after all, we'd like people to use our book!)
- If you post any slides on a www site, that you note that they are adapted from (or perhaps identical to) our slides, and note our copyright of this material.

For a revision history, see the slide note for this page.

Thanks and enjoy! JFK/KWR

All material copyright 1996-2020  
J.F Kurose and K.W. Ross, All Rights Reserved



*Computer Networking: A  
Top-Down Approach*

8<sup>th</sup> edition

Jim Kurose, Keith Ross  
Pearson, 2020



# Making routing scalable

our routing study thus far - idealized

- all routers identical
- network “flat”

... not true in practice

**scale:** billions of destinations:

- can't store all destinations in routing tables!
- routing table exchange would swamp links!

**administrative autonomy:**

- Internet: a network of networks
- each network admin may want to control routing in its own network

روتینگ پروتکل با روتینگ الگوریتم چه تفاوتی دارد؟

در روتینگ الگوریتم ما دوتا الگوریتم رو گفتیم: این الگوریتم ها اطلاعات شبکه رو در اختیار دارند و الگوریتم می‌گه از این اطلاعات چگونه استفاده بشه برای اطلاعات فورواردینگ در الگوریتم های مبتنی بر لینک استیت هر روتر اطلاعات لینک های در و بر خودش شامل اینکه با چه نودی مرتبطه و هزینه اش برداکست می کنه به همه روترهای دیگر شبکه

سوال : این لینک استیت ها چجوری باید منتقل بشن ؟ محاسبه روتینگ توی یک نود مستلزم این

است که اطلاعات شبکه هم دریافت شده باشه و الگوریتم دایکسترا فقط محاسبه رو می‌گه و بقیش رو دیگه نمی‌گه پس الگوریتم می‌گه چجوری روت رو محاسبه بکنیم ولی پروتکل کل عملیاتی که لازم است که این الگوریتم اجرا بشه و روت ها به دست میاد رو می‌گه

توی شبکه هر کاری براساس پروتکل انجام میشه پس این لینک استیت ها براساس یک پروتکلی باید رد و بدل بشن

پس ما یک پروتکل روتینگ داریم ک همه جوانب کار اونجا پیش بینی میشه و توی پروتکل روتینگ ما از الگوریتم روتینگ استفاده میکنیم که می‌گه چجوری محاسبات رو انجام بشه

چه پروتکل هایی و به چه صورت توی شبکه استفاده میکنیم؟

روشی که گفتیم شبکه رو یک شبکه flat فرض کردیم ینی یک گرافی داریم و یک تعداد نودی داریم و نود ها هم مشابه هم هستند ینی فرقی با هم ندارند و بعد روی اون گراف الگوریتم دایکسترا رو اجرا کردیم ولی در عمل شبکه واقعی این حالت رو ندارند ینی به صورت flat نیستند ینی همه روترها این شرایط رو ندارند بخاطر همین تفاوت هایی که وجود داره ما دوتا مسئله عمده با اون روش های روتینگ که قبلا گفتیم:

1- تعداد destinations ها خیلی زیاد میشه اگر کل اینترنت رو یکپارچه و به عنوان شبکه در نظر بگیریم و در هر روتر، جدول روتینگمون تعداد destinations هاش در حد میلیاردها پس بحث scale رو داریم که واقعا نمیتونه شبکه Flat باشه

بحث دیگر administrative autonomy است این شبکه توسط شرکت های مختلف ایجاد میشه و در قسمت های مختلف توسعه پیدا میکنه و چه شرکت ها و چه دولت ها می خوان روی شبکه خودشون تسلط داشته باشند و اونطوری که دوست دارند مدیریتش بکنن و یک شبکه flat نمی تونه این امکان رو فراهم بکنه



این گرافی هست از شبکه اینترنت  
ارتباطات توی این شبکه باید از هر جایی به هر جای دیگری امکان پذیر باشه

# Internet approach to scalable routing

aggregate routers into regions known as “autonomous systems” (AS) (a.k.a. “domains”)

**intra-AS (aka “intra-domain”):**  
routing among *within same AS*  
(“*network*”)

- all routers in AS must run same intra-domain protocol
- routers in different AS can run different intra-domain routing protocols
- **gateway router:** at “edge” of its own AS, has link(s) to router(s) in other AS'es

**inter-AS (aka “inter-domain”):**  
routing *among* AS'es

- **gateways** perform inter-domain routing (as well as intra-domain routing)

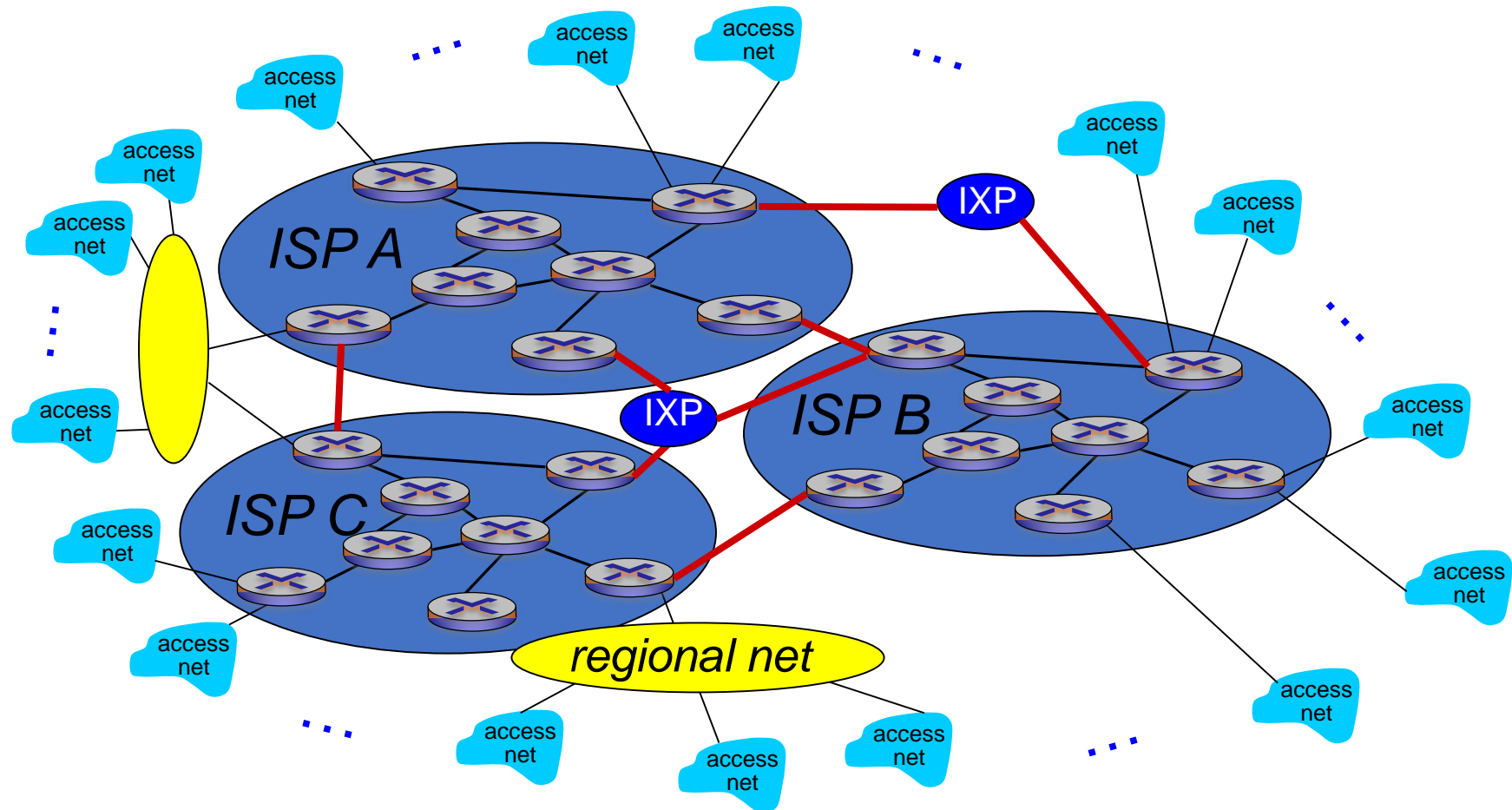
بخاطر همین شبکه اینترنت به این صورت توسعه پیدا کرده که از قسمت های ناحیه های مستقل تشکیل شده که به این ها autonomous systems یا AS میگویم که در واقع همون domains مختلف هستند مثلاً یک domains مال یک شرکت است و یک domains مال فلان کشور پس domains های مختلفی شکل گرفته و فرض بر این است که هر domains مستقل و خود مختار است و به صورت داخلی مدیریت میشه پس گسترش اینترنت براساس اینکه در ناحیه از اون قسمتی که داره گسترش پیدا میکنه به صورت یک ناحیه مستقل میتونه این گسترش انجام بگیره به این نواحی و مجموعه روترها که توی اون ناحیه هستند رو به صورت AS نگاه میکنیم توی کل اینترنت پس بحث روتینگ به دو قسمت می شکنه:

**“intra-AS (aka “intra-domain** : روتینگ داخل یک AS است مثل یک ISP مثلاً توی امریکا شبکه خودش رو توسعه داده و روتینگ خودش رو توی اون شبکه انجام میده - داخل AS ساب نت ها، ساب نت های خودمون هستند و خود این ها هم می تونن AS های مستقل باشند - توسعه شبکه براساس AS ها است و هر AS روتینگ داخل خودش در اختیار خودش است یک پروتکل روتینگ که داخل یک AS استفاده شده همه روترهای اون شبکه باید اون پروتکل رو اجرا بکنند ولی از یک AS به یک AS دیگه می تونه پروتکل های مختلفی استفاده بشه الان این **“intra-AS (aka “intra-domain** به چه دردی می خوره؟ مثلاً توی دانشگاه خودمون یک AS داریم و از یک روتر توی دانشگاه برق وصل می شیم به یک سرور توی مرکز فناوری اطلاعات خب مسیر مناسب برای پکت های این ارتباط چی هست و براساس توپولوژی داخلی شبکه دانشگاه و روتینگ پروتکلی که استفاده میشه این به دست میاد

**inter-AS routing**: روتینگ بین AS ها است و هر AS از طریق gateway به AS های دیگه یا همسایه اش وصل میشه - ارتباط بین AS ها - برای inter-AS باید یک پروتکل استفاده بشه و همه AS ها باید اینو قبول داشته باشند و همه اینو انجام بدن ینی inter-AS یک پروتکل استاندارد بین المللی باید باشه و همه اینو باید رعایت بکنند



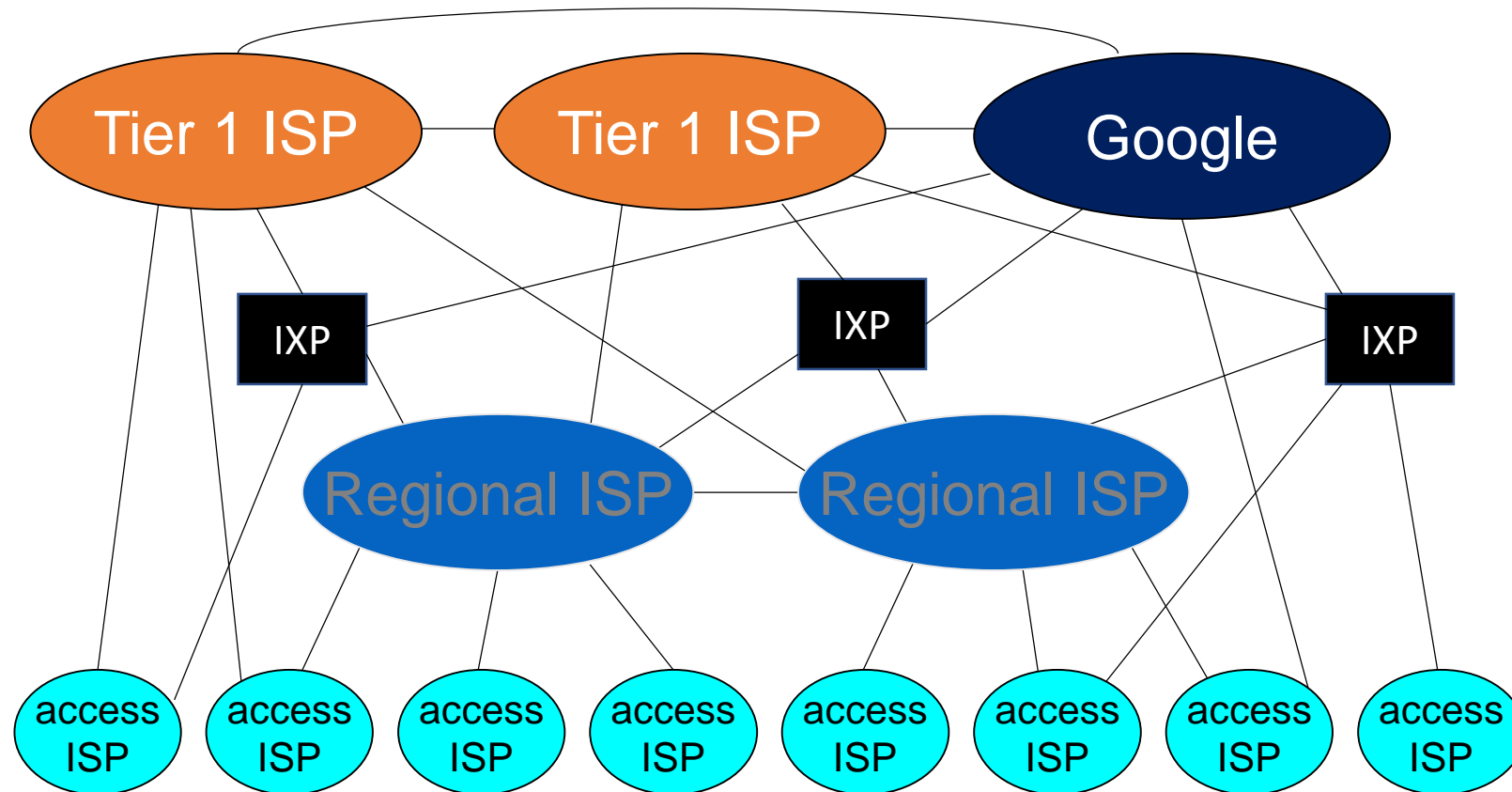
# Internet structure: network of networks



-----

اینترنت به صورت شبکه ای از شبکه ها شکل گرفته که در اون ISP ها در یک ساختار سلسله مراتبی ارتباط نقاط انتهایی رو برقرار میکند

# Internet structure: network of networks

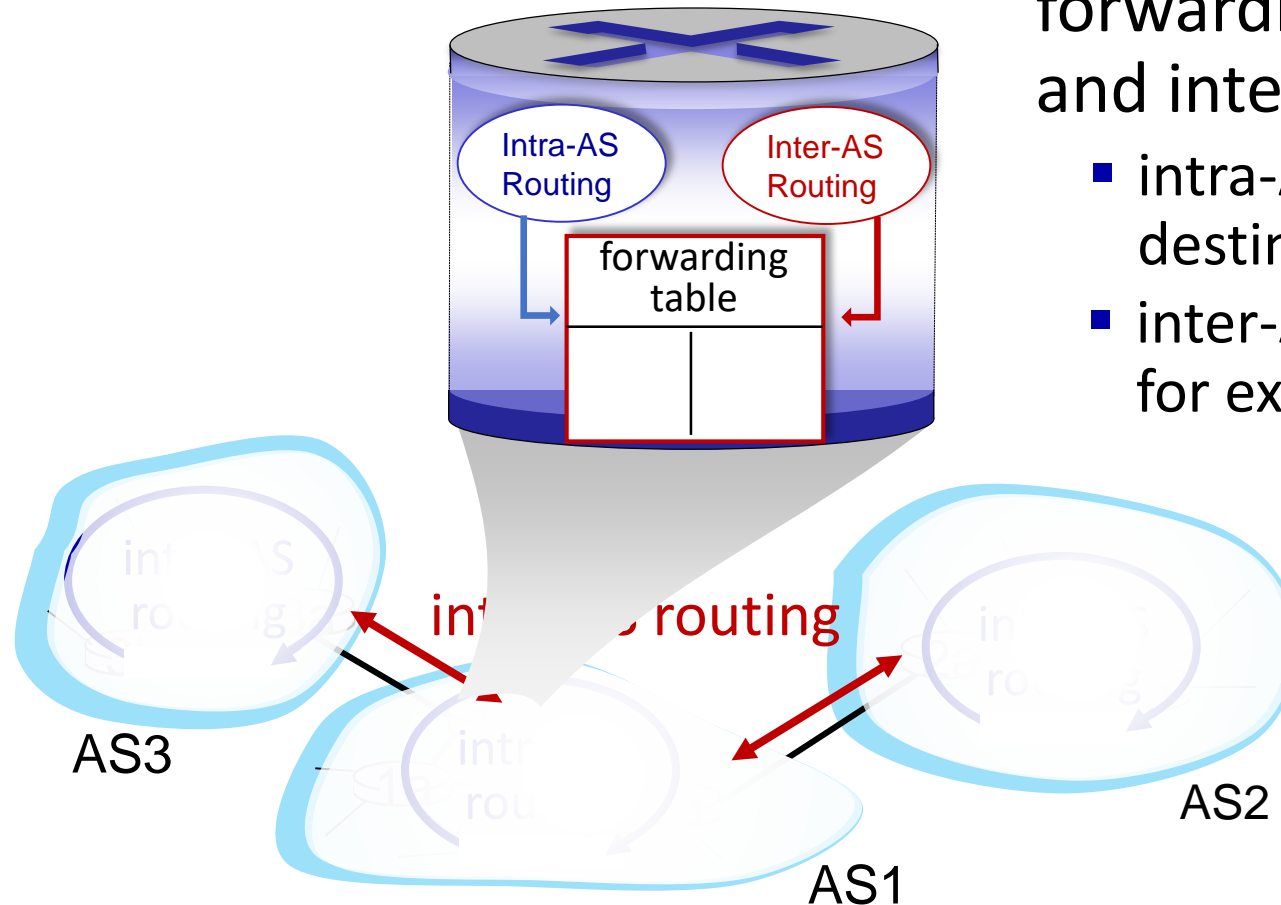


- at center: small # of well-connected large networks
  - “tier-1” commercial ISPs (e.g., Level 3, Sprint, AT&T, NTT), national & international coverage
  - ... and regional networks may arise to connect access nets to ISPs

ISP های مختلفی در سطوح مختلفی در سطح اینترنت به کاربران سرویس می دهند و برای ارتباط بین کاربران خودشان ارتباطاتی بین ISP ها شکل گرفته

انتخاب یک مسیر مناسب برای یک ارتباط بین دو نقطه انتهایی که یکیش از یک ISP داره سرویس میگیره و یکی دیگش از یک ISP دیگه به چه شکلی انجام میگیره؟

# Interconnected ASes



forwarding table configured by intra- and inter-AS routing algorithms

- intra-AS routing determine entries for destinations within AS
- inter-AS & intra-AS determine entries for external destinations

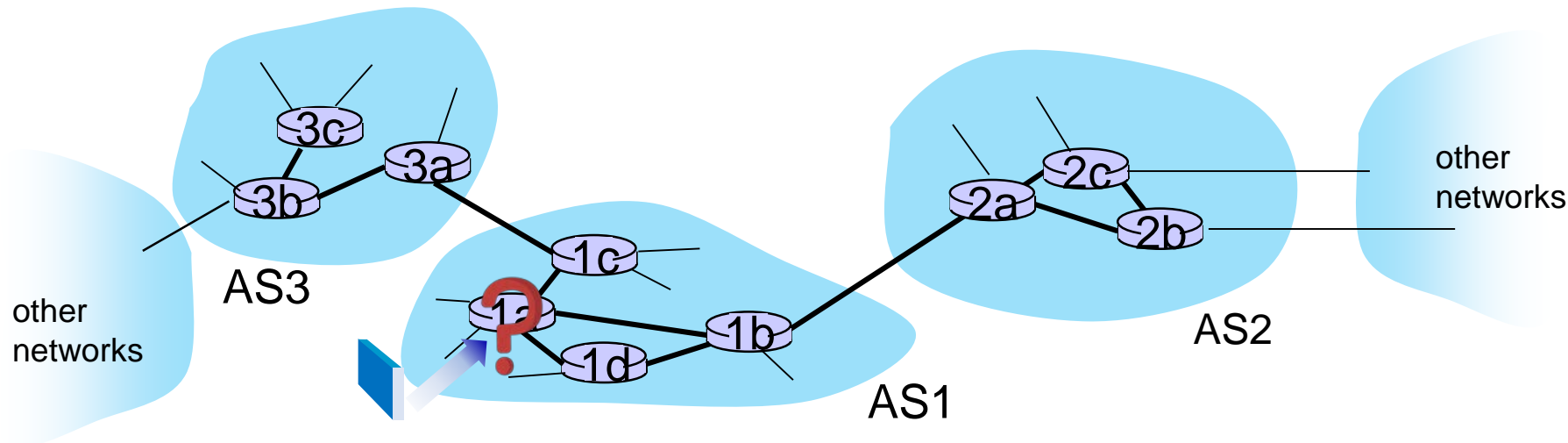
اینترنت از as های مرتبط با هم تشکیل شده و در هر روتری توی هر as ما یک جدول فورواردینگ نیاز داریم که بسته ها که به این روتر می رسن براساس مقصدشون لوکاپ میشه در این lookup در این جدول فورواردینگ و براساس اون مشخص میشه که لوکالی به کدوم پورت خروجی این بسته توی همین روتر باید ارسال بشه ولی این مستلزم اینه که روتر درواقع روتینگ پروتکل ما این پورت رو براساس اینکه مقصد نهایی چه مسیری براش خوبه به دست آورده باشه این مقصد نهایی اگر داخل as است intra\_as اینو تهیه میکنه و مسیر رو براش مشخص میکنه و اگر در یک as دیگه است inter\_as باید این کارو انجام بده پس برای اطلاعات فورواردینگ تبیل بسته به اینکه مقصد داخل همین as است یا خارج as است هم پروتکل intra\_as , inter\_as باید اینجا دخالت داشته باشه

# Inter-AS routing: a role in intradomain forwarding

- suppose router in AS1 receives datagram destined outside of AS1:
- router should forward packet to gateway router in AS1, but which one?

## AS1 inter-domain routing must:

1. learn which destinations reachable through AS2, which through AS3
2. propagate this reachability info to all routers in AS1



در حالتی که بین as ها می خواهیم مسیر رو پیدا بکنیم پروتکل inter\_as باید مشخص بکنه که از چه مسیرهایی باید عبور پیدا بکنه و این مسیر در روتر مورد نظر باید ترجمه بشه در پورت خروجی که جدولش بگه برای اون مقصد این پورت خروجی باید باشه مثلاً می خواهیم از روتر 1d بریم به یک جای دیگه پس باید این مسیر ترجمه بشه در روتر 1d و این پورت خروجی باید پورته باشه که gateway که توی این مسیر باید بسته ازش عبور پیدا بکنه که به as مجاور و از اونجا به مقصد برسه بسته رو برسونه باید این مسیری که توی as دیگه داریم از طریق as های همسایه این مسیر به دست میاد و برای اتصال به همسایه ها هم از طریق gateway شبکه های خودمون

مثلاً ما اینجا دوتا gateway داریم: 1b ما gateway است به as2 و 1c هم gateway است به as3 و همسایه های ما as3 , as2 هستند و بسته ای که از as1 می خواد بره به یک مقصد توی یک as دیگه لاجرم باید از طریق یکی از as های همسایه برود و برای هر کدوم از اون همسایه ها gateway مربوطه باید توی مسیر قرار بگیره

پس inter\_as protocol ما باید به دست بیاره که برای مقصد های مختلف کدوم همسایه اش مناسب است



# Inter-AS routing: routing within an AS

most common intra-AS routing protocols:

- **RIP: Routing Information Protocol** [RFC 1723]
  - classic DV: DVs exchanged every 30 secs
  - no longer widely used
- **EIGRP: Enhanced Interior Gateway Routing Protocol**
  - DV based
  - formerly Cisco-proprietary for decades (became open in 2013 [RFC 7868])
- **OSPF: Open Shortest Path First** [RFC 2328]
  - link-state routing
  - IS-IS protocol (ISO standard, not RFC standard) essentially same as OSPF

-  
interior gateway protocol یا IGP پروتکل هایی هستند که برای روتینگ داخل AS ها استفاده میشن و پروتکل های مشهوری که برای این کار بوده: RIP , OSPF , IGRP که الان OSPF تقریبا هست و همه جا داره استفاده میشه و OSFP یک پروتکل روتینگ برای inter\_as است که مبتنی بر الگوریتم های لینک استیت هم هست ینی توی دلش از دایکسترا استفاده میشه

# OSPF (Open Shortest Path First) routing

- “open”: publicly available
- classic link-state
  - each router floods OSPF link-state advertisements (directly over IP rather than using TCP/UDP) to all other routers in entire AS
  - multiple link costs metrics possible: bandwidth, delay
  - each router has full topology, uses Dijkstra’s algorithm to compute forwarding table
- *security*: all OSPF messages authenticated (to prevent malicious intrusion)

چه ویژگی هایی داره؟

open است ینی open sorce است

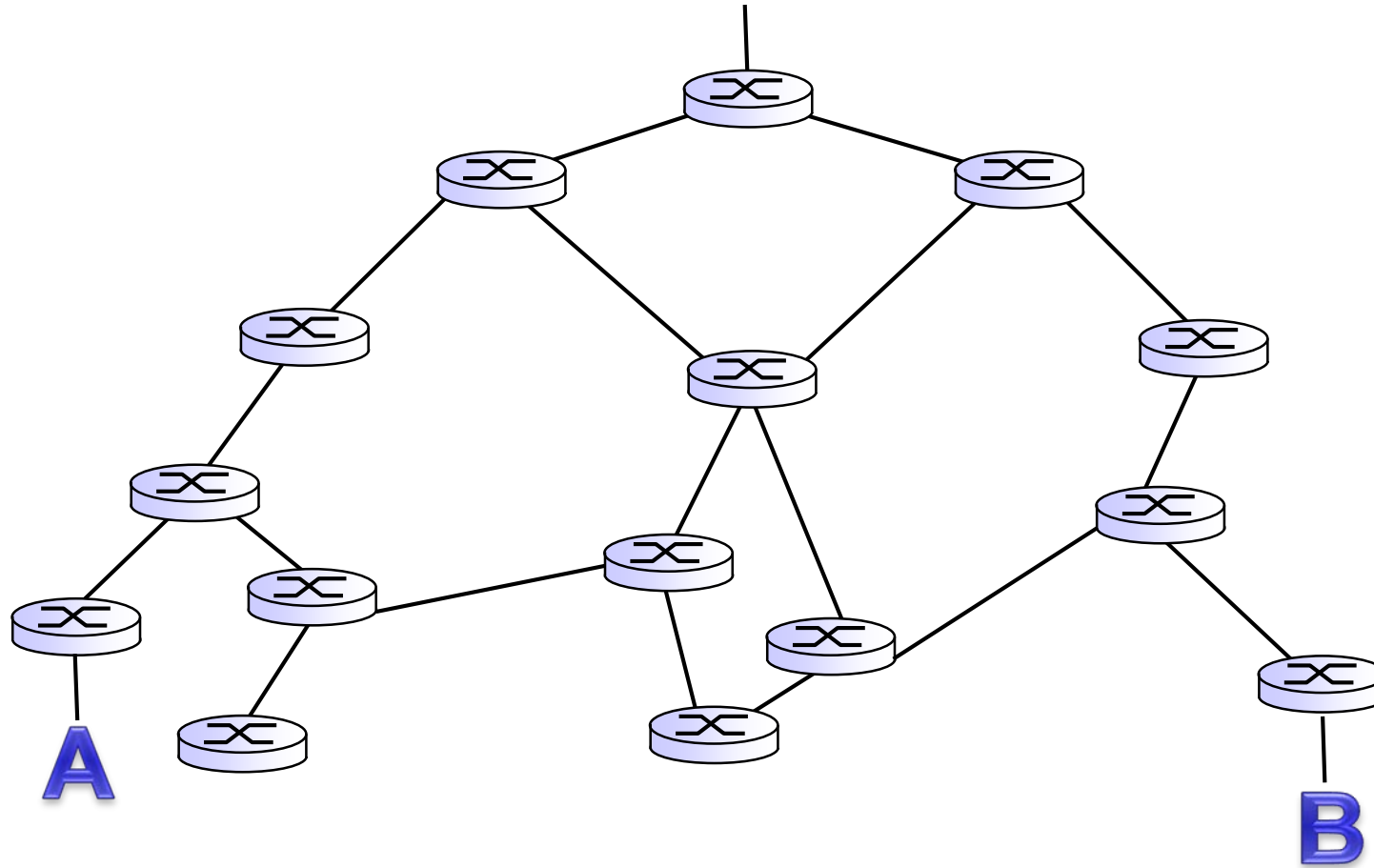
الگوریتمش لینک استتیت است ینی از الگوریتم لینک استتیت استفاده میکنه و از خود دایکسترا استفاده میکنه - پروتکل مشخص کرده که لینک استتیت ها به چه صورت توی شبکه منتشر بشن و براساس اون توپولوژی توی هر نود به دست میاد و بعد دایکسترا اجرا میشه برای به دست آوردن مسیرها

OSPF ترتیبی که برای انتقال لینک استتیت ها در شبکه مشخص کرده این است که این ها باید به صورت فلادینگ در شبکه به همه روترها داده بشه و پروتکل فرمت مسیج های مربوط به لینک استتیت ها هم مشخص کرده و این ها مستقیما روی IP منتقل میشن؟؟

پس OSPF یک پروتکلی هست که از نظر اون لایه بندی پروتکل ها توی خود لایه شبکه اجرا میشه درست مثل TCP , UDP که روی IP قرار دارند، OSPF هم مستقیما روی IP قرار داره توی خود لایه شبکه و بسته های این incapsule میشن توی بسته های IP مستقیما ینی توی لایه اپلیکیشن نیستند توی خود IP هستند

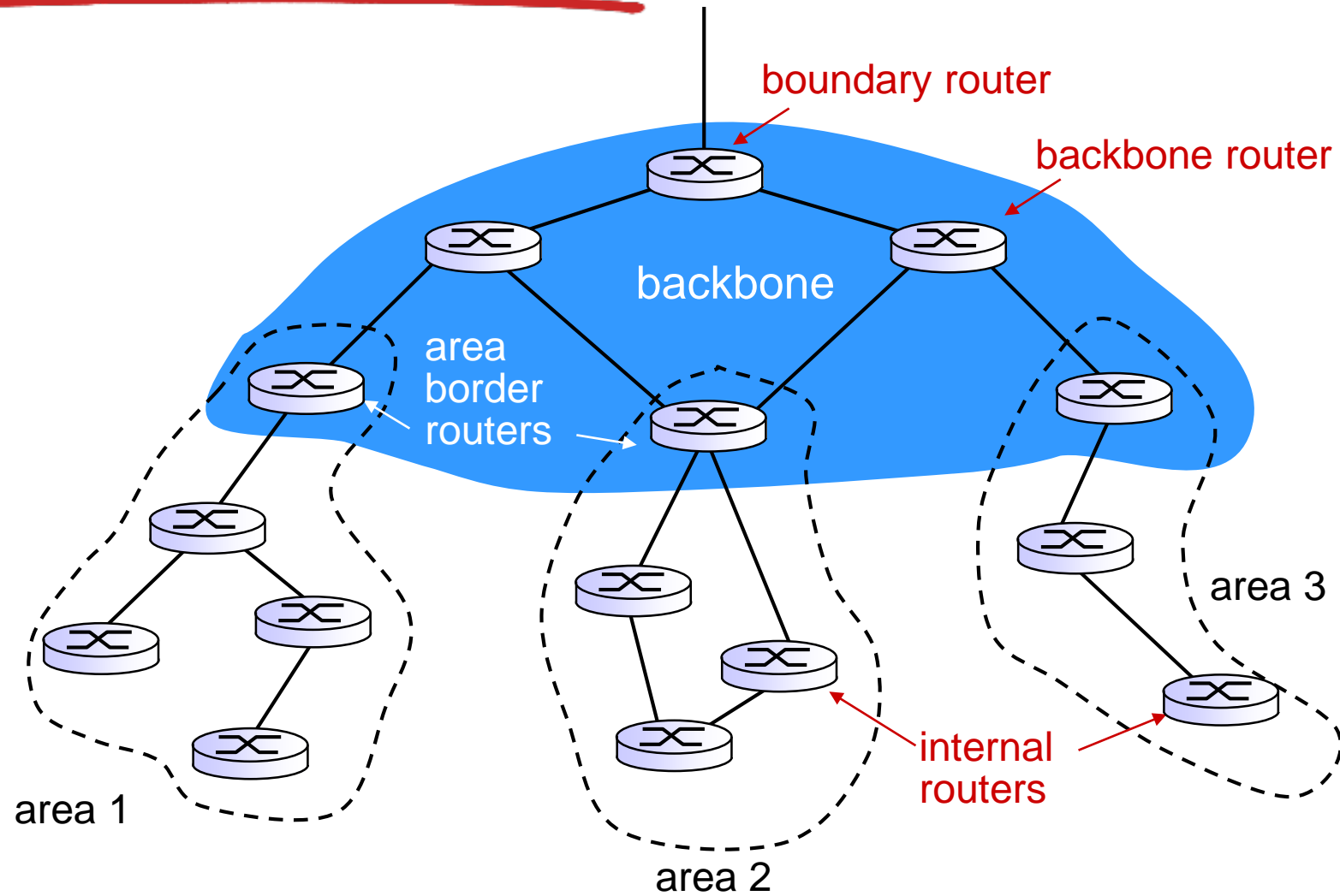
داخل مسیج لینک استتیت OSPF هم برای هر لینک attached برای یک روتر یک سطر است که مشخص میکنه همسایه اش کیه و هزینه اش چیه و یکسری اطلاعات مشابهی مثل این ها در پکت مربوطه گذاشته میشه و ارسال میشه

# Enterprise Network





# Hierarchical OSPF



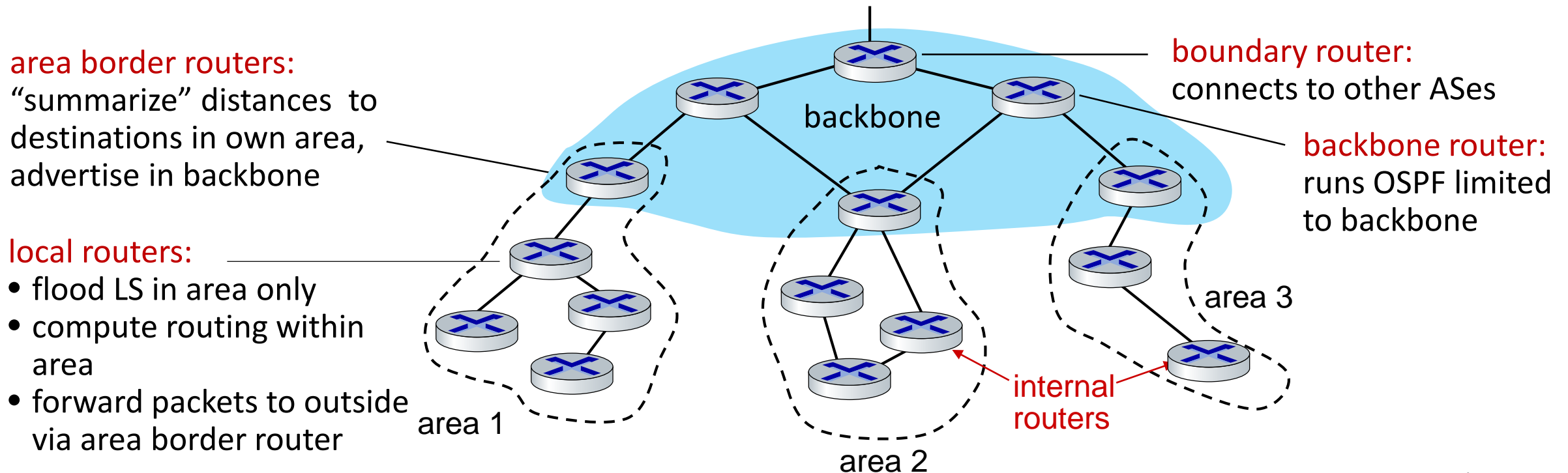
مهمترین ویژگی که OSPF داره Hierarchical OSPF است ینی معماری اون می تونه به صورت Hierarchical انجام بگیره و چرا مهمه؟ چون خیلی از AS ها بزرگ هستند پس اگر شبکه بزرگ باشه و تعداد نودها زیاد باشه داخل AS هم این مطرح میشه ینی ما یک AS داریم و محدود شدیم به AS خودمون ولی حتی این AS هم بزرگه ینی تعداد نودهایش زیاده و OSPF این امکان رو فراهم میکنه که ما به صورت سلسله مراتبی دوباره AS خودمون رو به ناحیه کوچکتر بشکنیم و شبکه ای که اون الگوریتم دایکسترا باید توش اجرا بشه رو کوچکتر کنیم در شبکه های کوچک از این ویژگی استفاده نمیشه ولی توی شبکه های بزرگ خیلی کمک میکنه عملاً OSPF شبکه رو به دو لایه تقسیم میکنه: یک لایه backbone و بقیه شبکه بعد backbone یک area نامیده میشه و area شماره گذاری میشن و backbone ما area صفرمون است و بقیه شبکه در قالب area کوچکتر دسته بندی میشن مثل شکل روبه رو و بعد الگوریتم دایکسترا توی هر area مستقلاً اجرا میشه

پس شبکه رو به 4 شبکه کوچکتر شکوندیم و الگوریتم توی این شبکه کوچکتر اجرا میشه و در نهایت ساده تر میشه



# Hierarchical OSPF

- **two-level hierarchy:** local area, backbone.
  - link-state advertisements flooded only in area, or backbone
  - each node has detailed area topology; only knows direction to reach other destinations



پس link-state advertisements فقط توی هر Area انجام می گیره پس overhead ترافیکی کمی ایجاد میکنه و لینک استیت های area های دیگه توی این ناحیه وارد نمیشن در حالی که اگر این یکپارچه بود لینک استیت تک تک روترها توی این شبکه باید به همه می رسید و هر روتری توی هر area فقط اطلاعات گراف اون area داره و در نتیجه مسیرهای مربوط به مقصدهای توی همین area رو می تونه مشخص بکنه

مسیر بین area های دیگه رو چجوری به دست میاریم و توی روتینگ تیبیلش می نویسیم؟ area ها بهم مرتبط هستند

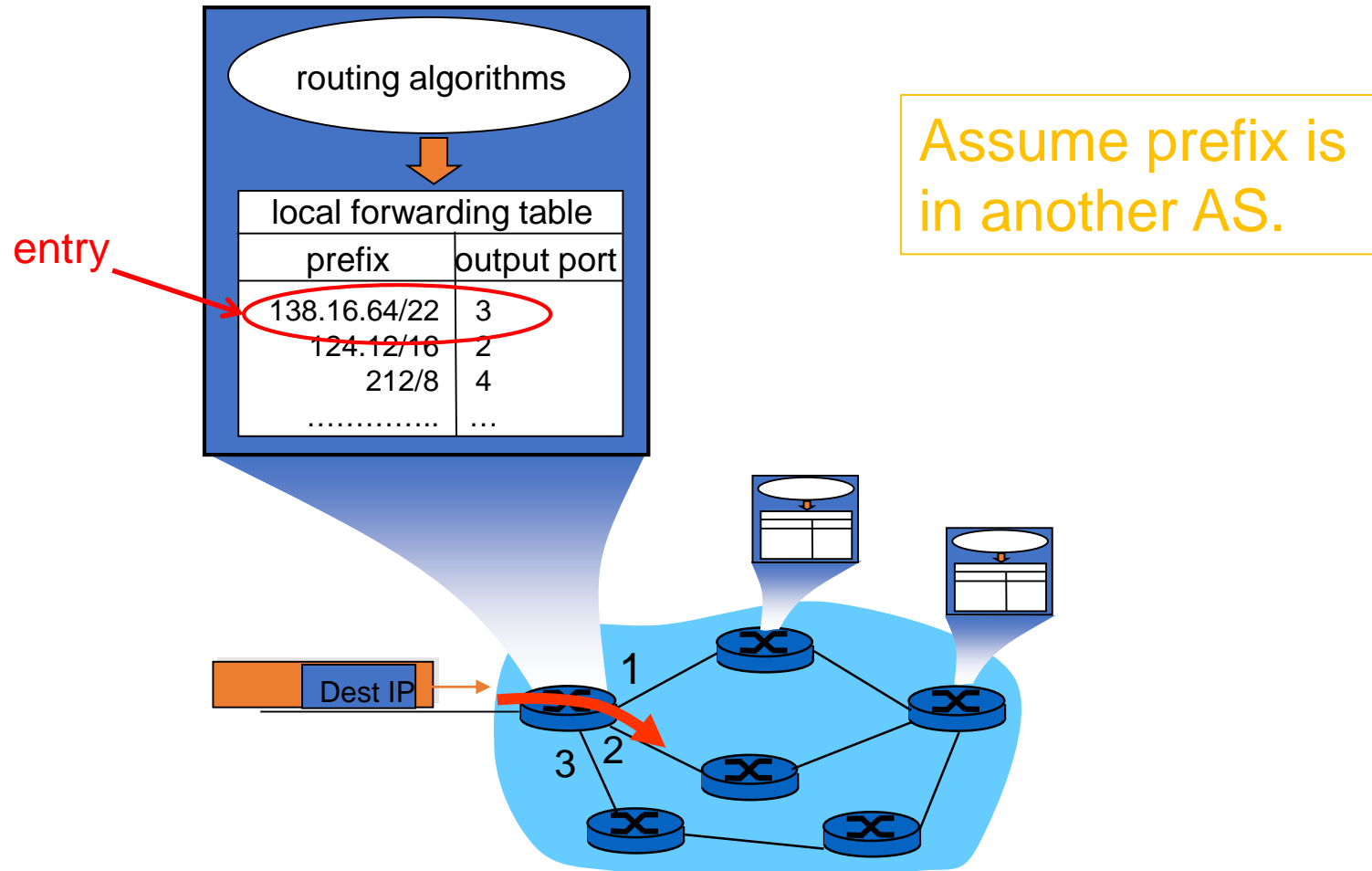
توی backbone روترهایی که داریم backbone router گفته میشن این backbone router ها یکیشون boundary router است که روتری است که کل این as از طریق این as وصل میشه به as های دیگر

backbone router ها دو دسته هستند: این هایی که روترهای داخلی backbone router هستند ینی مال ناحیه صفراند و روترهایی که واسطه هستند با ناحیه های دیگر که این روترها عضو اون ناحیه همسایه هم هستند مثلا روتر سمت چپی در صفحه هم عضو ناحیه صفر است و هم عضو ناحیه یک که به این روترها area border routers میگویم و بقیه روترهای هر ناحیه میشن internal routers

وظیفه پیدا کردن مسیر از ساب نت های توی یک ناحیه به ساب نت های توی یک ناحیه دیگر به عهده این border routers ها هستند این border routers بخشی از این ناحیه است و الگوریتم دایکسترا هم توی این ناحیه اجرا شده و border routers اینو داره و مسیر همه ساب نت های توی این ناحیه رو خودش به دست آورده پس border routers مسیر به همه ساب نت های این ناحیه رو داره و این اطلاعات رو به border routers های ناحیه های دیگر منتقل میکنه

پس OSPF این امکان رو فراهم میکنه که area border routers اطلاعات ساب نت های خودش رو به area border routers های دیگه برسونه و اون ها هم این اطلاعات رو به internal router های خودشون می دن

# How does entry get in forwarding table?



نکته: هدف از پروتکل روتینگ چی هست؟ ینی مسیر رو پیدا کردیم میخوایم باهاش چی کارکنیم؟ در سطح اینترنت ما داریم پروتکل IP رو اجرا میکنیم و طبق پروتکل IP بسته هامون در قالب بسته های IP منتقل میشن و این بسته ها هدر دارن توی هدر این ها یک مقدار فیلد وجود داره برای کارهای مختلف و یکی از این فیلدها، فیلد IP ادرس مقصد است و پروتکل IP به این صورت کار میکنه این بسته به یک روتر می رسه و روتر این فیلد خاص رو توی هدر بسته IP نگاه میکنه ینی فیلد IP مقصد رو و روتر یک جدولی داره به اسم **local forwarding table** که توی این جدول لیست ساب نت های قابل دسترسی رو توی جدول داره و برای هر ساب نت مشخص میکنه کدوم پورت خروجی اون روتر مسیر مناسب برای اون روتر است

به ازای یک ساب نت مقصد ما باید اینجا یک **entry** داشته باشیم که مشخص میکنه برای بسته هایی که برای این مقصد ارسال میشن چه پورته مناسب است این ادرس مقصدی که از هدر بسته ip خونده میشه با این جدول چک میشه و اگر توی این رنج باشه مچ میشه و هدایت میشه به اون پورت خروجی برای ساب نت های توی **AS** روتینگ الگوریتم ما اجرا میشه و این اطلاعات رو به دست میاره و توی جدول می نویسه

اگر ساب نت ما توی یک **AS** دیگه باشد چه اتفاقی می افته؟ اگر ساب نت ما توی یک **AS** دیگه باشد ما باید پروتکل روتینگ مناسبی باید استفاده بکنیم و مسیر مناسب رو برای اون پیدا بکنیم ولی هدف نهایی به هر حال همین است که اون مسیر تبدیل بشه به اطلاعات و توی جدول وارد بشه

# Internet inter-AS routing: BGP

- **BGP (Border Gateway Protocol):** *the* de facto inter-domain routing protocol
  - “glue that holds the Internet together”
- allows subnet to advertise its existence, and the destinations it can reach, to rest of Internet: *“I am here, here is who I can reach, and how”*
- BGP provides each AS a means to:
  - **eBGP:** obtain subnet reachability information from neighboring ASes
  - **iBGP:** propagate reachability information to all AS-internal routers.
  - determine “good” routes to other networks based on reachability information and *policy*

پروتکل روتینگ که بین AS ها اجرا میشه و روتینگ رو انجام میده پروتکل BGP که این پروتکل توسط همه AS ها اجرا میشه

پروتکل روتینگ inter-as در سطح اینترنت باید یکسان باشه و همه as باید همونو اجرا بکنن برای همین نیاز به یک پروتکل استاندارد داریم که همه براساس اون مسیریابی رو انجام بدن که برای اینترنت پروتکل bgp مورد توافق قرار گرفته و همه اینو استفاده میکنند - الان مسیر بسته ها در سطح اینترنت براساس پروتکل bgp است که تعیین میشه

bgp چگونه کار میکنه؟ روشی که bgp برای به دست آوردن مسیر مناسب برای ساب نت ها استفاده میکنه مبتنی بر اطلاع رسانی است در این روش یک as مسیر مناسب برای ساب نت هایی که داخلش هستند رو به همسایه هاش اطلاع میده و به این ترتیب همسایه ها اطلاع دسترسی به این ساب نت هارو به دست میارن و این ها هم می تونن این اطلاعات رو برای به دست آوردن entry در فورواردینگ تیبل های خودشون استفاده بکنند و این اطلاعات رو به همسایه های دیگر منتقل بکنند در نهایت اطلاع دسترسی دست به دست منتقل میشه و به بقیه رسونده میشه

پروتکل bgp این کار رو در دو لایه انجام میده در واقع از دو قسمت تشکیل شده:

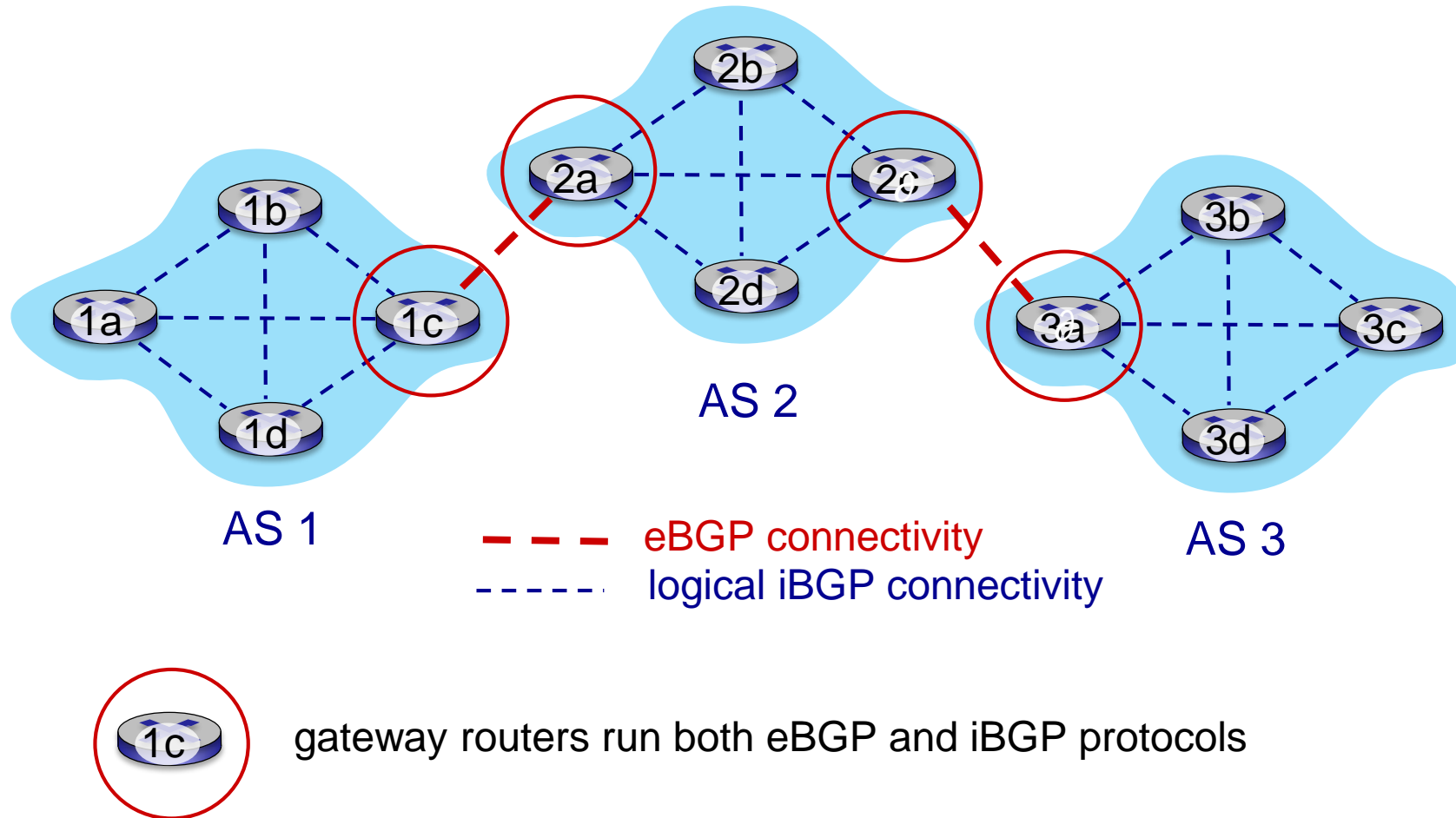
ebgp: همین اطلاعات دسترسی ساب نت هارو از همسایه ها دریافت میکنه و دست به دست به بقیه همسایه ها میده

ibgp: اما وقتی که این اطلاعات توسط روتر مرزی یک as دریافت شد براساس پروتکل ibgp این اطلاعات به روترهای داخلی اون as منتقل میشه

نکته: هر روتری توی هر as باید این اطلاعات رو داشته باشه و فورواردینگ تیبل خودش رو به روز بکنه پس لازمه این اطلاعات به تک تک روترها رسونده بشه

وقتی این اطلاعات به یک روتر رسید روتر entry خودش رو براساس معیارهای داخلی خودش از جمله پالیسی ها و سیاست گذاری هایی که داره تعیین میکنه و فورواردینگ تیبلش رو به روز میکنه پس bgp پروتکلی است که براساس اون هر ساب نتی توی اینترنت می تونه وجود خودش رو و محل خودش رو و مسیر دسترسی به خودش رو اطلاع رسانی بکنه به بقیه و از این طریق بقیه بتونن بسته هاشون رو برای اون بفرستن

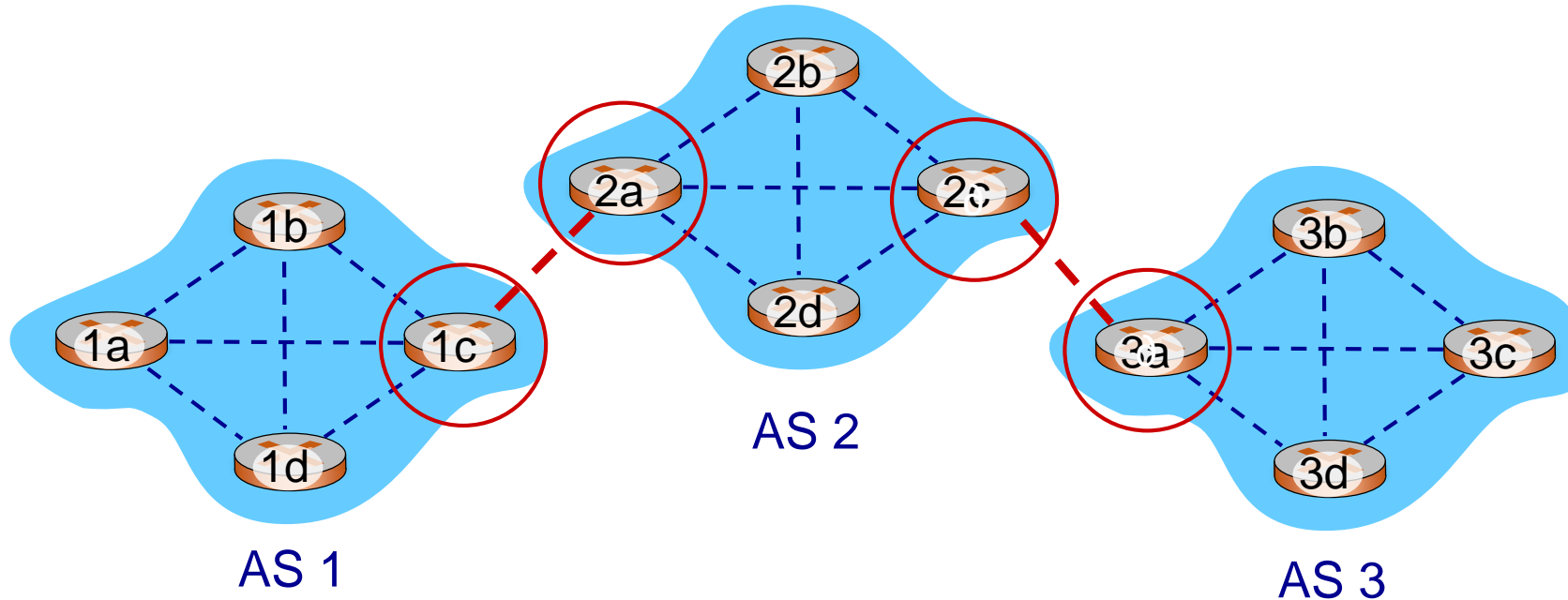
# eBGP, iBGP connections



اطلاع دسترسی ساب نت ها بین As ها براساس پروتکل ebgp منتقل میشه و اطلاعاتی که به هر as رسیده براساس پروتکل ibgp بین روترهای داخلی اون as رد و بدل میشه  
پس روترهای کناری یا gateway ها هم پروتکل ebgp رو اجرا میکنن و هم پروتکل ibgp رو ولی بقیه روترها داخل هر as فقط پروتکل ibgp رو ارسال میکنند



# Route establishment in BGP



- For networks in AS1 and AS2 to communicate:
  - AS1 must announce a route to AS2
  - AS2 must accept the route from AS1
  - AS2 must announce a route to AS1
  - AS1 must accept the route from AS2

حالا برای اینکه یک روت بین دوتا As برقرار بشه و بسته ها در این روت ارسال و دریافت بشن صرفا اطلاع رسانی مسیر دسترسی به یک ساب نت کافی نیست بلکه غیر از اینکه این مسیریابی انجام میگیره مثلا

as1 یک مسیری رو به as2 اطلاع میده و as2 باید این مسیر رو از as1 قبول بکنه و As2 هم همینطور باید مسیر خودش رو به as1 اطلاع بده و as1 این مسیر رو از as2 قبول بکنه به این ترتیب ارتباط بین دوتا as برقرار میشه

مثال: فرض کنید ساب نت n1 در as1 میخواد بسته هایی رو برای مقصد n16 در As16 ارسال بکنه:

ابتدا as16 باید اطلاعات دسترسی n16 رو به عنوان یکی از ساب نت هایی که توی خودش است به همسایه هاش که اینجا as8 است ارسال بکنه و as8 باید این مسیر رو از as16 قبول کرده باشه و بعد که قبول کرد as8 میاد این اطلاع رسانی رو به as34 , as1 منتقل میکنه و این ها هم باید اینو از as8 قبول بکنند و as34 هم اینو به as1 اطلاع رسانی میکنه و as1 بلاخره یکی از مسیرهایی که از as34 یا as8 که بهش اطلاع رسانی شده رو قبول میکنه و براساس اون جدول فورواردینگ روتر n1 رو ستاپ میکنه و بعد از این اگر n1 بسته رو بفرسته این بسته روت میشه و به مقصد رسونده میشه

## Route establishment in BGP



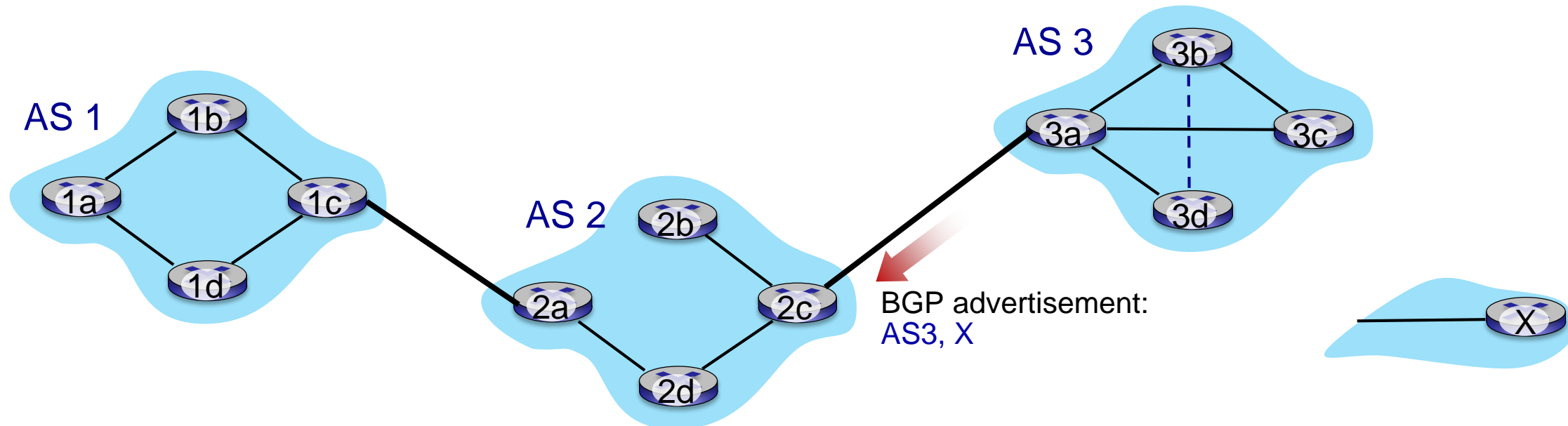
For net N1 in AS1 to send traffic to net N16 in AS16:

- AS16 must originate and announce N16 to AS8.
- AS8 must accept N16 from AS16.
- AS8 must announce N16 to AS1 or AS34.
- AS1 must accept N16 from AS8 or AS34.

For two-way packet flow, similar policies must exist for N1.

# BGP basics

- **BGP session:** two BGP routers (“peers”) exchange BGP messages over semi-permanent TCP connection:
  - advertising *paths* to different destination network prefixes (BGP is a “path vector” protocol)
- when AS3 gateway 3a advertises *path AS3,X* to AS2 gateway 2c:
  - AS3 *promises* to AS2 it will forward datagrams towards X



این اطلاع رسانی ها به چه شکل انجام می گیره؟

اطلاع رسانی این مسیرها بین as ها از طریق روترهای مرزی یا Gateway انجام میگیره و برای این اطلاع رسانی این روترها بین همدیگر یک session برقرار میکنند و این session بین یک tcp connection صورت میگیره که این tcp connection یک tcp connection تقریباً دائمی است و روی این اطلاع رسانی advertising بین این روترها در قالب چیزی که بهش میگیم path vector

مثال: as3 اطلاعات دسترسی ساب نت x به صورت path vector به As2 اطلاع رسانی میکنه و این path vector عبارت از x , as3 به عبارت دیگر می گه برای دسترسی به ساب نت x میتونی از طریق من که همون as3 است بسته هاتو بفرستی و اینو رو به روتر مرزی as2 ارسال میکنه و وقتی as3 همچین اطلاع رسانی انجام میده به as2 بعد as3 تضمین میکنه اگر As2 بسته ای به مقصد x به as3 بفرسته این به X رسونده خواهد شد

# Path attributes and BGP routes

- BGP advertised route: prefix + attributes
  - prefix: destination being advertised
  - two important attributes:
    - **AS-PATH**: list of ASes through which prefix advertisement has passed
    - **NEXT-HOP**: indicates specific internal-AS router to next-hop AS

روت که اطلاع رسانی میشه شامل چه اطلاعاتی است؟

یک روت از دو قسمت تشکیل شده: prefix و attributes

prefix: ادرس اون ساب نتی است که این روت براش داره تبلیغ میشه

attributes: اطلاعاتی است که در کنار prefix قرار میگیرند و این ها کمک میکنن برای به

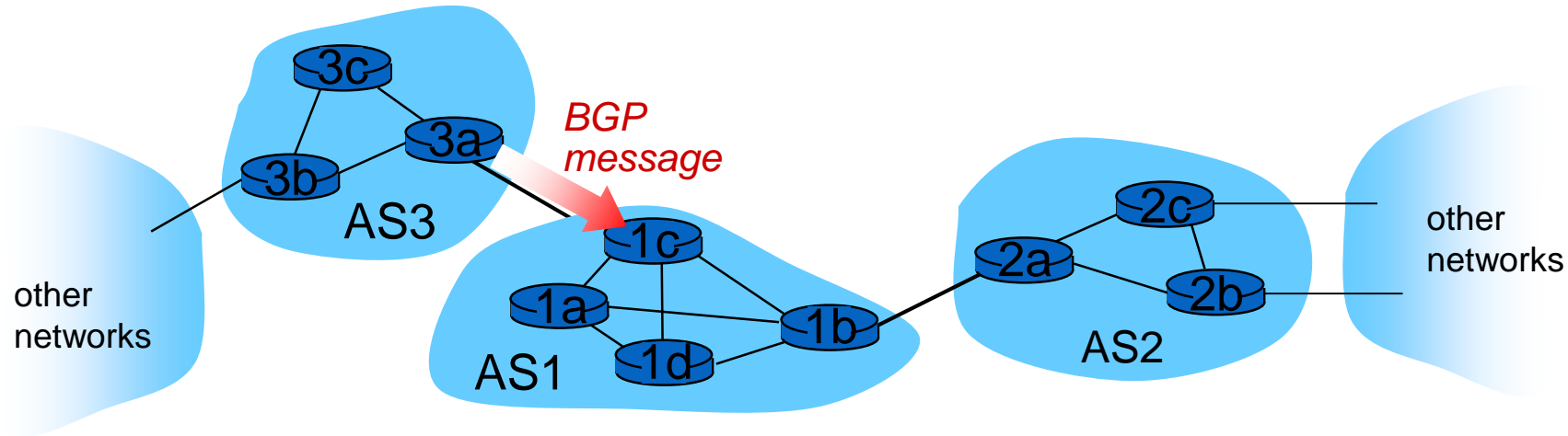
دست آوردن مسیر مناسب

دو قسمت مهم attributes :

AS-PATH: لیست as های مسیر است ینی as های روت از این As تا As مقصد

NEXT-HOP: ادرس ip روتر مرزی اولین as توی این مسیر است

# BGP route advertisement



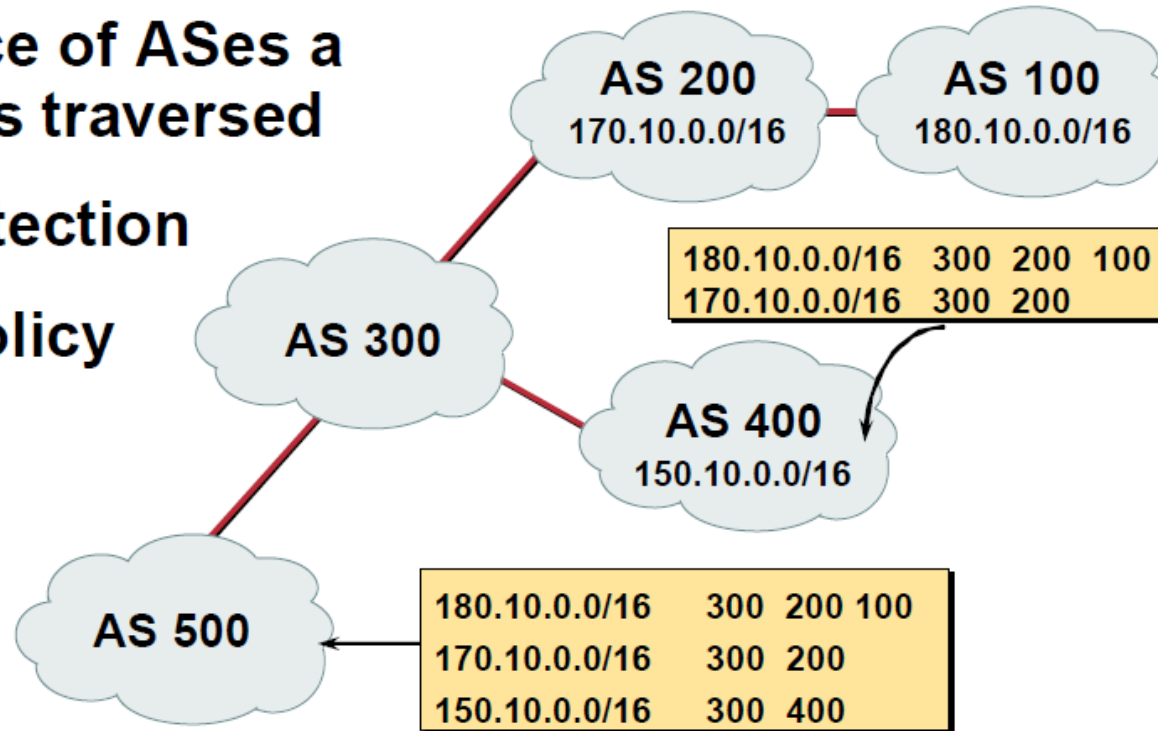
- ❖ BGP message contains “routes”
- ❖ “route” is a prefix and attributes: AS-PATH, NEXT-HOP,...
- ❖ Example: route:
  - ❖ Prefix: 138.16.64/22 ; AS-PATH: AS3 AS131 ; NEXT-HOP: 201.44.13.125

روتر مرزی As3 پیغام bgp به روتر مرزی as1 ارسال می‌کند و این پیغام شامل روت‌هایی است که as3 می‌خواهد برای as1 تبلیغ بکند و این روت از AS-PATH, NEXTHOP در کنار prefix ساب‌نت مقصد تشکیل شده. مثلاً



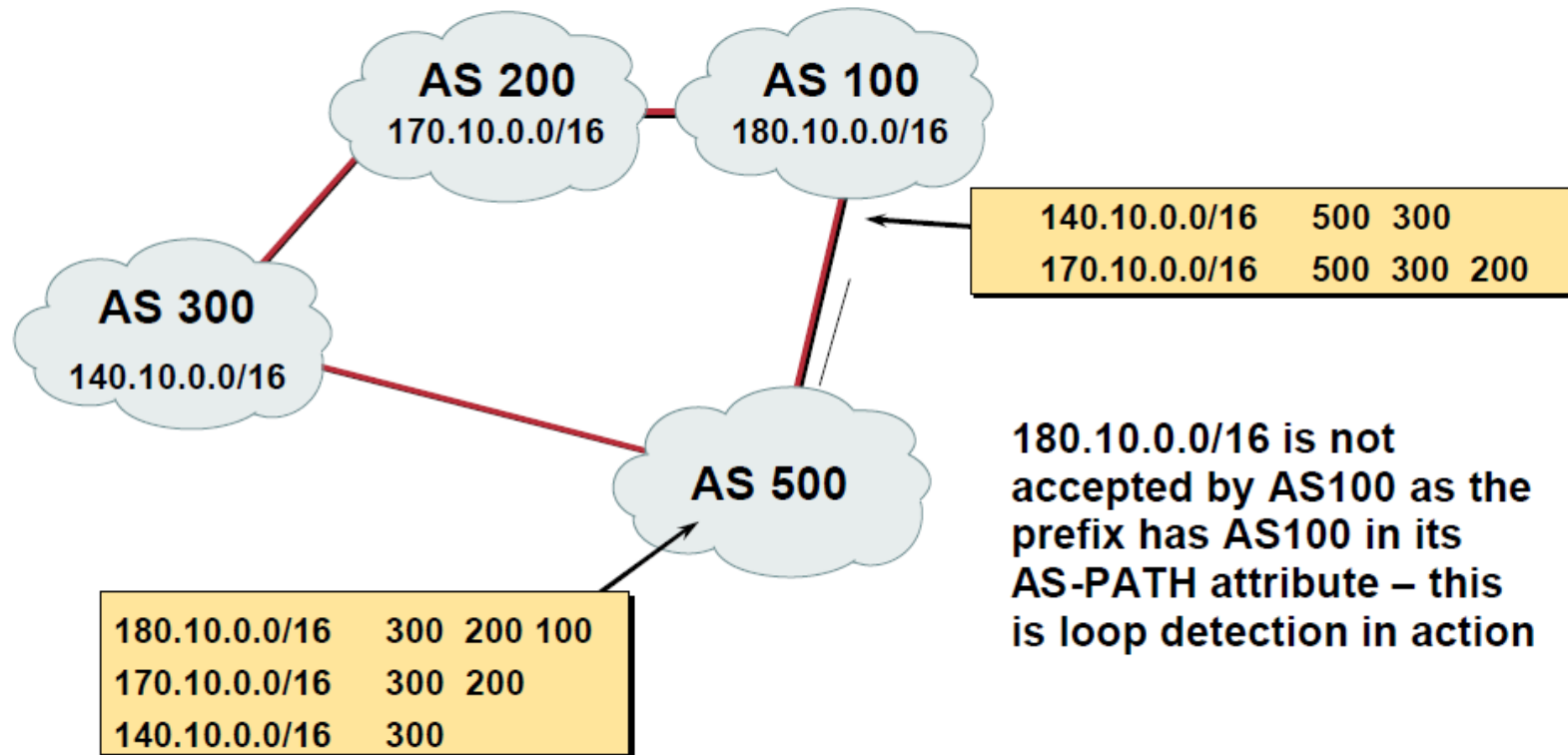
# AS Path

- Sequence of ASes a route has traversed
- Loop detection
- Apply policy



پس مسیر as به عنوان یکی از صفت های path vector ما شامل sequence as هایی هست که این اطلاع رسانی از اون ها عبور کرده و به اینجا رسیده و این چجوری تشکیل میشه؟ هر As که اطلاع رسانی رو دریافت میکنه و قبول میکنه وقتی میخواد این رو به همسایه دیگر خودش بفرسته as خودش رو هم به as path اضافه میکنه چرا این as path به این ترتیب اطلاع رسانی میشه؟ چرا as های توی مسیر هم در روت قرار میگیره؟ برای اینکه وقتی که برای یک as اطلاع رسانی می رسه و میخواد این رو قبول بکنه بدونه بسته هایی که براساس این روت منتقل خواهند شد از چه as هایی عبور می کنند و این برای اعمال سیاست هایی که ممکنه مستلزم این باشه که بسته های ترافیک این شبکه از یک as خاصی عبور نکنند استفاده میشه و کاربر دیگه ای که این As path داره اینه که از لوپ می تونیم براساس اون جلوگیری بکنیم ینی می تونیم براساس مقایسه های لیست as های توی مسیر لوپ رو تشخیص بدیم و مسیرهایی که لوپ ایجاد میکنند رو حذف کنیم

# AS-Path loop detection



مثال:

این لیست روت هایی است که به As500 رسیده و As500 این هارو قبول کرده

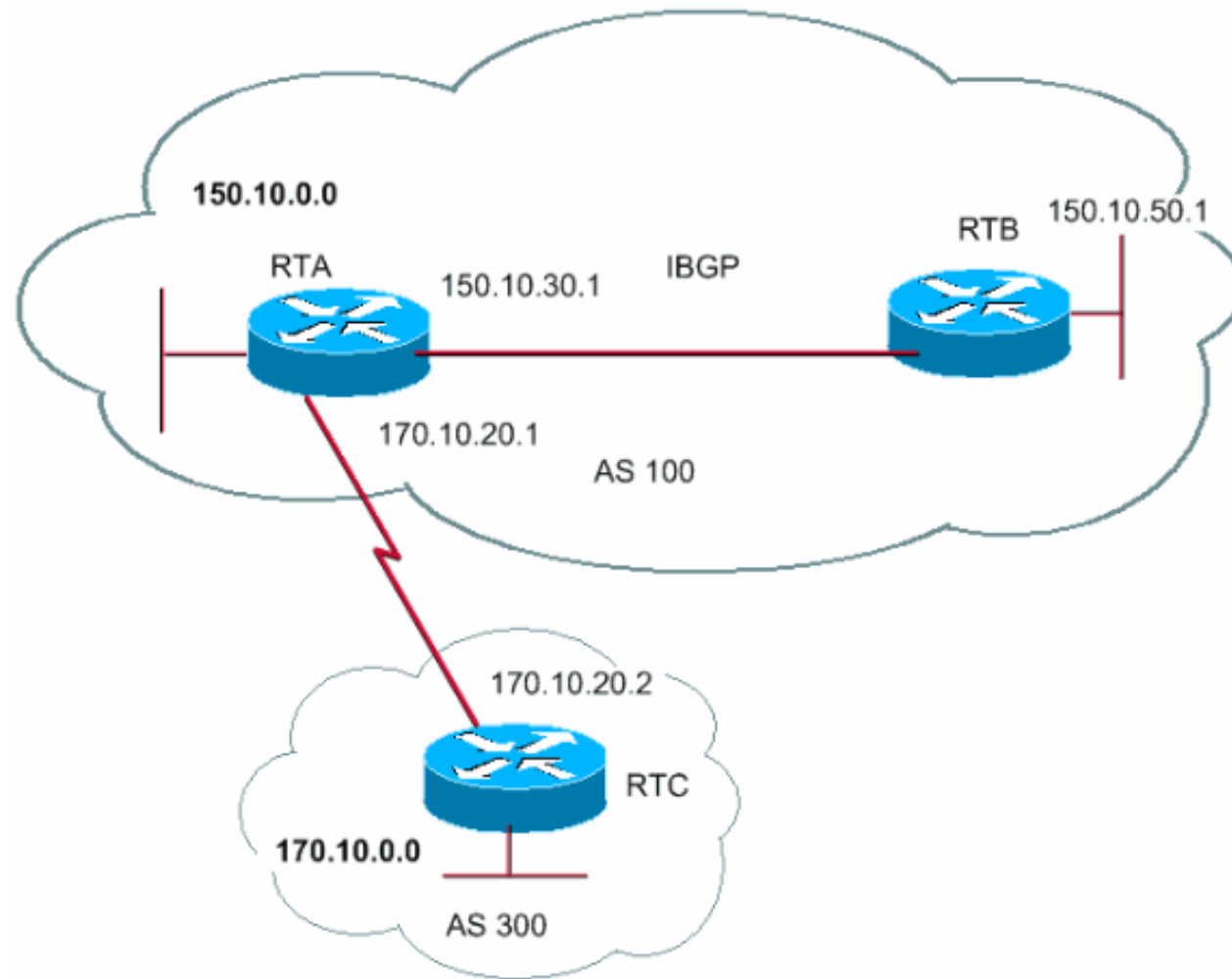
# Next Hop

- **IGP should carry route to next hops**
- **Recursive route look-up**
- **Unlinks BGP from actual physical topology**
- **Allows IGP to make intelligent forwarding decision**

صفت Next Hop در روت چه کاربردی داره؟

وقتی که یک as یک روتی به همسایه اش اطلاع رسانی میکنه با قرار دادن Next Hop در روت به همسایه میگه روترهای اون اگر بخوان بسته ای رو در این روت بفرستن باید اونو به Next Hop بفرستن و همسایه هم وقتی که همین روت رو به همسایه خودش اطلاع رسانی بکنه دوباره Next Hop خودش رو توی روت قرار میده و می فرسته پس هر as با استفاده از Next Hop مسیر این path رو به As قبلی خودش ستاپ میکنه در روترهای خودش پس اگر As path ما as های توی مسیر رو مشخص میکنه در هر قدم As path ما Next Hop مشخص کننده قدم بعدی برای ارسال بسته است و این Next Hop ها هستند که قدم به قدم استفاده میشن برای انتقال بسته توی این مسیر و این باعث میشه که bgp دیگه ضرورتی نداشته باشه که خواد وارد توپولوژی فیزیکی As ها بشه و صرفا براساس لیست as ها مسیر رو مشخص میکنه و مسیر داخلی هر as رو محول میکنه به igp و در هر as ما igp براساس مقصدش که Next Hop است مسیر داخلی خودش رو برای اینکه یک بسته ای رو به این روت بفرسته ستاپ میکنه

# Next Hop



مثال:

یک ساب نت در As300 اطلاع رسانی میشه به as100 و as path که اطلاع رسانی میشه شامل As300 میشه و Next Hop که اطلاع رسانی میشه ادرس روتر اولین روتر توی این مسیر است ینی 170.10.20.2 است و این روت توسط ebgp به روت همسایه میی رسه و روتر مرزی as100 توسط ibgp همین روت رو به روترهای دیگر توی As100 اطلاع رسانی میکنه فرض میکنیم روتر B میخواد توی فورواردینگ تیبل خودش مسیری رو براساس این روت برای ساب نت توی as300 ستاپ بکنه حالا این مسیر رو برای چه مقصدی باید ستاپ بکنه؟ ینی یک هاستی در این ساب نت = 150.10.50.1 میخواد یک بسته ای رو بفرسته ساب نت مورد نظر در as300 <-- این بسته می رسه به روتر B و این روتر باید این بسته رو به چه مسیری بفرسته و این مسیر رو چطوری به دست بیاره؟ براساس Next Hop مسیر مناسب از روترهای داخلی خودش به as همسایه رو پیدا میکنه و از روتر B استفاده میکنه برای ستاپ کردن فورواردینگ تیبل به این ترتیب براساس ادس Next Hop بسته برای اینکه توی این روت بیوفته توی این مسیر به روتر C ارسال میشه و روتر C اینو مسیریابی میکنه و به ساب نت مورد نظر می رسونه

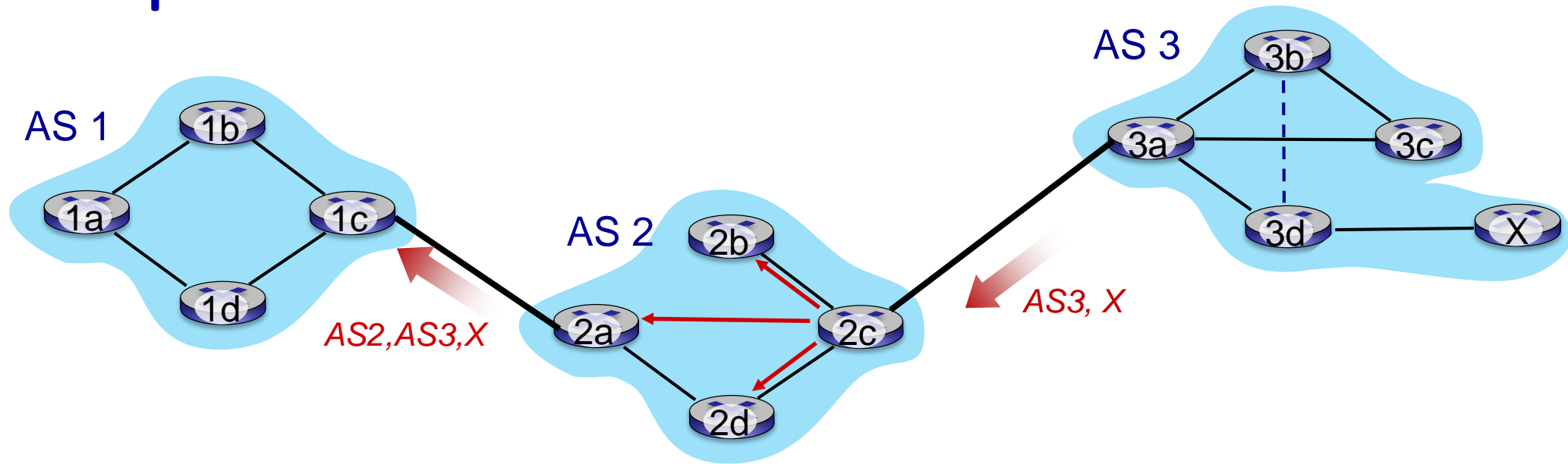


# Policy-based Routing

- gateway receiving route advertisement uses *import policy* to accept/decline path (e.g., never route through AS Y).
- AS policy also determines whether to *advertise* path to other neighboring ASes



# BGP path advertisement

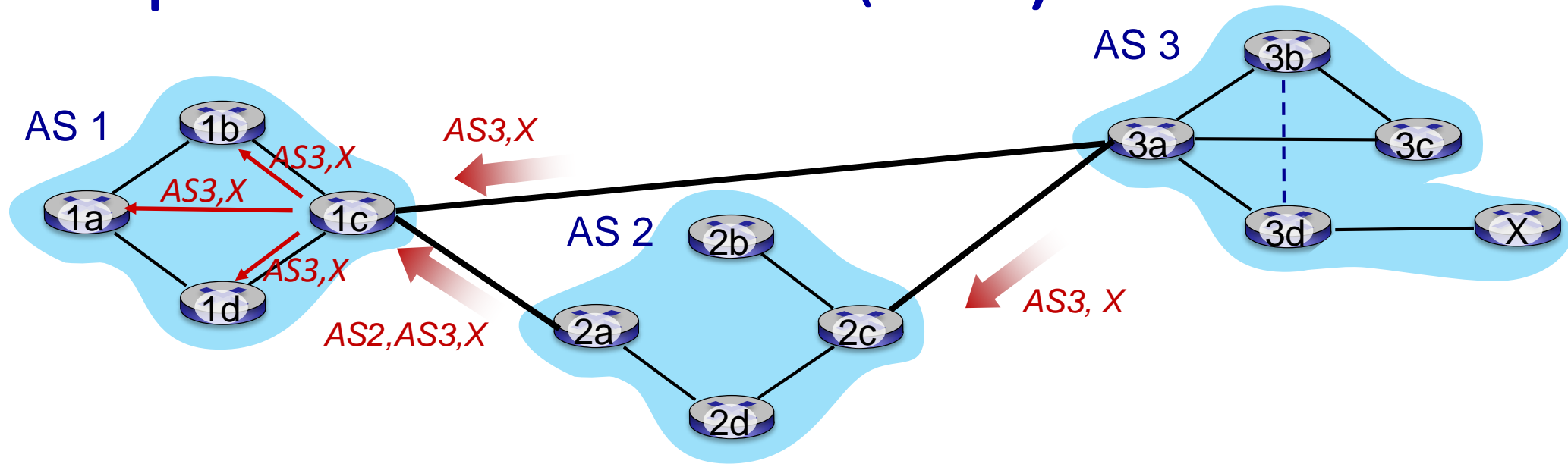


- AS2 router 2c receives path advertisement **AS3,X** (via eBGP) from AS3 router 3a
- based on AS2 policy, AS2 router 2c accepts path AS3,X, propagates (via iBGP) to all AS2 routers
- based on AS2 policy, AS2 router 2a advertises (via eBGP) path **AS2, AS3, X** to AS1 router 1c

فرض کنید  $x$  یکی از ساب نت های  $as3$  است و  $as3$  میخواد اینو اطلاع رسانی بکنه مسیر رسیدن به  $x$

با استفاده از  $ebgp$  روتر مرزی  $3a$  در  $As3$  این روت رو اطلاع رسانی میکنه به  $as2$  و توسط روتر مرزیش ینی  $2c$  دریافت میشه و توسط سیاست های موجود این مسیر قبول میشه در این صورت  $2c$  با استفاده از  $ibgp$  همین روت رو اطلاع رسانی میکنه به روترهای داخلی  $As2$  و با استفاده از  $2a$  می ره به  $As1$  و اگر پالیسی ایجاد بکنه که این روت به  $As1$  اطلاع رسانی بشه و براساس  $ebgp$  ما  $2a$  اینو اطلاع رسانی میکنه به  $as1$  و بعد از طریق  $1c$  دریافت میشه و اگر اینو قبول بکنه با استفاده از  $ibgp$  این اطلاع رسانی میشه به روترهای داخلی

# BGP path advertisement (more)



gateway router may learn about **multiple** paths to destination:

- AS1 gateway router 1c learns path **AS2, AS3, X** from 2a
- AS1 gateway router 1c learns path **AS3, X** from 3a
- based on **policy**, AS1 gateway router 1c chooses path **AS3, X** and advertises path within AS1 via iBGP

یک as ممکنه مسیرهای مختلفی برای یک ساب نت مقصد رو دریافت بکنه

برای مثال اگر ارتباط بینشون خط مشکلی باشه

اطلاع رسانی ساب نت x توسط as3 به As2 و از طریق اون به as1 منتقل میشه و همزمان

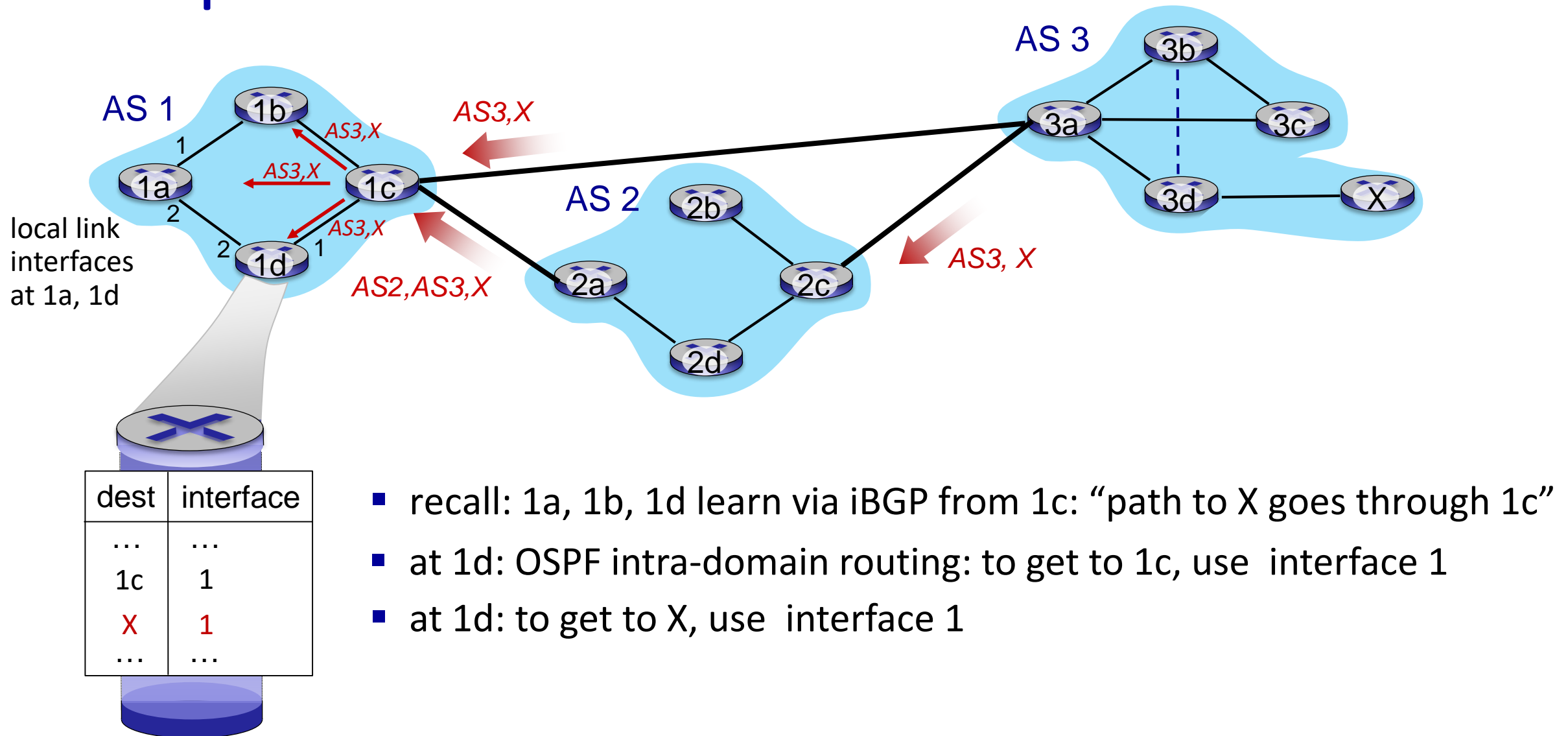
ممکنه همین مسیر از طریق As3 مستقیماً به as1 اطلاع رسانی شده باشه

اطلاع رسانی که از مسیر as3 to as1 است path vector اش as3 رو شامل میشه و اونی که

از طریق As3 to as2 to as1 است path vector اش as2, as3 رو شامل میشه

حالا سوال اینه که روتر 1c باید کدوم یکی از این دو مسیر رو انتخاب بکنه؟

# BGP path advertisement



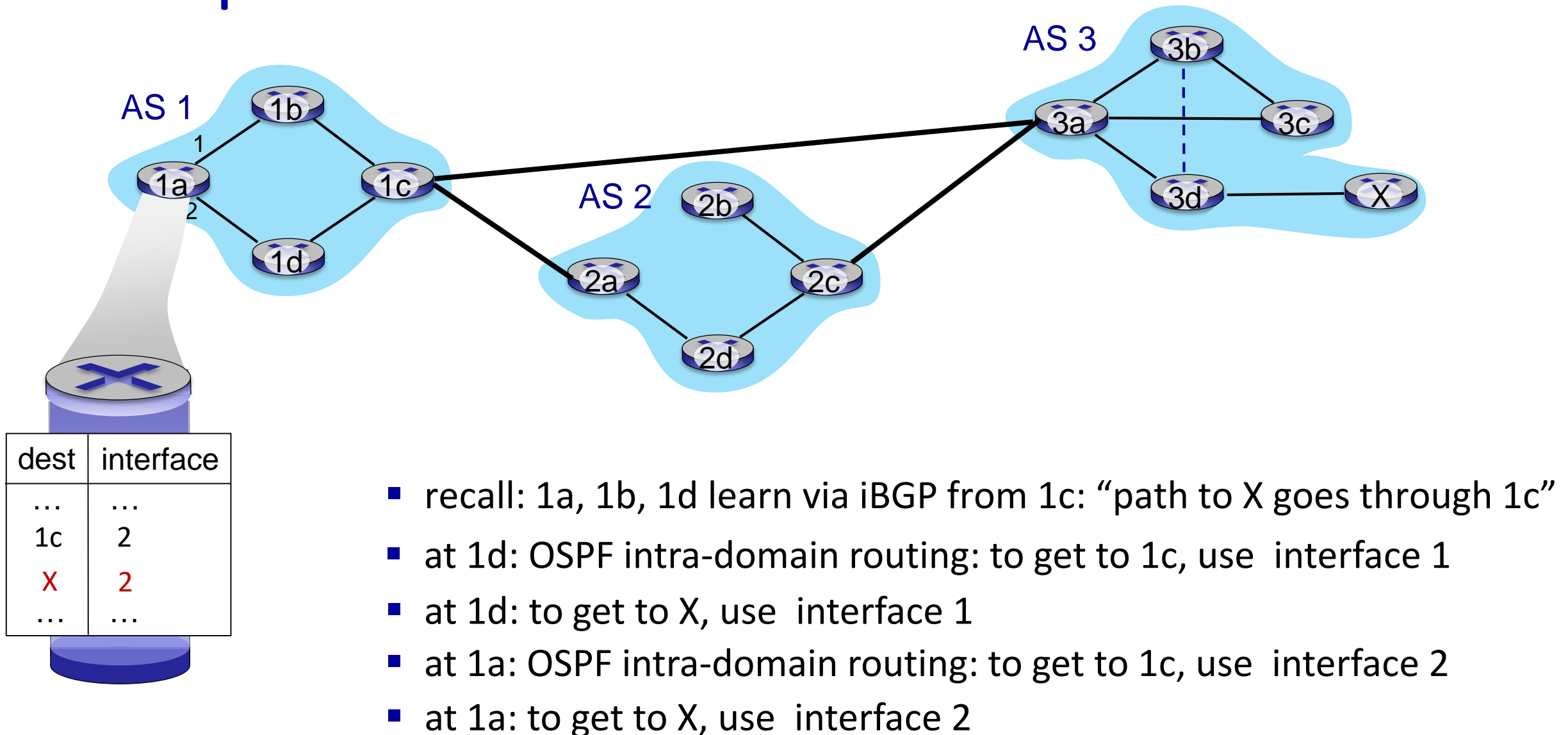
میخوایم ببینیم اطلاعات مربوط به یک  $x$  که در یک ساب نت دیگه قرار داره در روترهای یک  $as$  دیگه مثلاً  $as1$  چجوری ست میشه؟

$1c$  براساس پالیسی که داره مسیری که از  $as3$  to  $as1$  رو قبول میکنه و بعد توسط  $ibgp$  اینو به روترهای داخلی خودش اطلاع رسانی میکنه

حالا روترهای داخلی چجوری  $entry$  مربوط به این ساب نت رو توی فورواردینگ تیبل خودشون ست می کنن؟ براساس اطلاعاتی که دریافت شده روتر  $1d$  میدونه مسیر مربوط ساب نت  $X$  از طریق روتر  $1c$  است و بعد براساس روتینگ داخلی  $as1$  پیدا میکنه کدوم مسیر از  $1d$  به  $1c$  مناسب تر است و براساس پروتکل روتینگ  $intra-domain$  خودش ینی  $OSPF$  این به دست میاره که از طریق اینترفیس یک مسیر بهتری داره پس اون چیزی که توی جدول وارد میشه این اینترفیس یک است برای مقصد  $X$

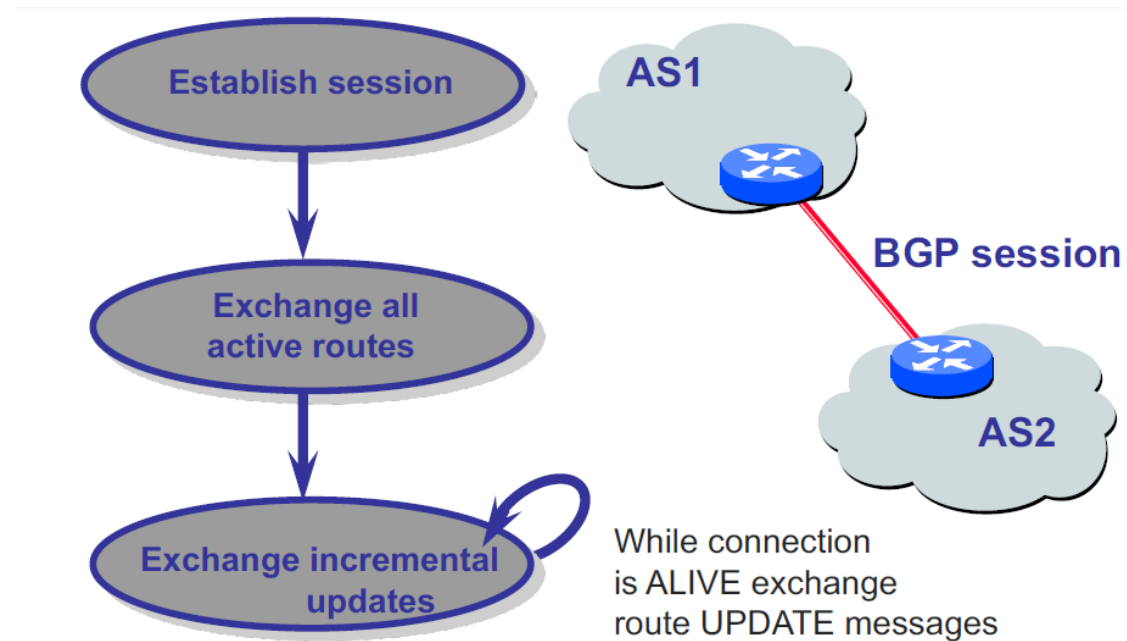


# BGP path advertisement



به همین ترتیب در روتر 1A دوباره براساس OSPF برای مقصد 1C اینترفیس دو بهترین مسیر است و در جدول وارد میشه

# Route establishment and maintenance



اگر پروسه پروتکل bgp رو خلاصه بکنیم:

ابتدا session bgp بین روترهای مرزی As های همسایه برقرار میشه و روی این Session روت ها رد و بدل میشن و براساس این روت ها جداول ستاپ میشه در روترهای هر as و بعد از اون اگر تغییری در این روت ها اتفاق بیوفته این تغییرات به صورت incremental updates بین as ها رد و بدل میشه و براساس اون روت ها اصلاح میشن پس اگر یک روت اطلاع رسانی مثلا از as2 به as1 این روت برقرار است و as1 روی این حساب میکنه تا زمانی که اپدیت اون توسط as2 به as1 ارسال بشه مبنی بر اینکه تغییری در روت است یا روت دیگه معتبر نیست

# BGP messages

- BGP messages exchanged between peers over TCP connection
- BGP messages:
  - **OPEN**: opens TCP connection to remote BGP peer and authenticates sending BGP peer
  - **UPDATE**: advertises new path (or withdraws old)
  - **KEEPALIVE**: keeps connection alive in absence of UPDATES; also ACKs OPEN request
  - **NOTIFICATION**: reports errors in previous msg; also used to close connection

پروتکل bgp با رد و بدل کردن پیغام هایی بین روترها اجرا میشه و اون پیغام ها عبارتند از:  
open: که tcp connection بین روترهای مرزی دوتا as رو برقرار میکنه  
update: برای اطلاع رسانی یک مسیر جدید یا برای حذف یک مسیری که قبلا اعلام شده استفاده  
میشه

keepalive: این tcp connection دائمی است و با ارسال مکرر keepalive این کانکشن  
حفظ میشه  
notification: برای گزارش دادن ارورهای موجود استفاده میشه و همینطور برای بستن کانکشن  
در صورت لزوم

# Why different Intra-, Inter-AS routing ?

## policy:

- inter-AS: admin wants control over how its traffic routed, who routes through its network
- intra-AS: single admin, so policy less of an issue

## scale:

- hierarchical routing saves table size, reduced update traffic

## performance:

- intra-AS: can focus on performance
- inter-AS: policy dominates over performance

چرا بین as ها و داخل As ها پروتکل های روتینگ متفاوتی استفاده میشه؟  
1- پالیسی است :

برای داخل یک As اهمیت چندانی نداره برای اینکه یک as تمام روترهاش ادمینش مشترک است در نتیجه سیاست گذاری های خاصی بین این ها خیلی مطرح نیست ولی بین as ها هر As ادمین خودش رو داره که مستقل است و دوست داره سیاست های خودش رو اعمال بکنه و ما روتینگ پروتکلی رو استفاده میکنیم که این امکان رو فراهم میکنه که سیاست ها استفاده بشن  
2- scale :

پروتکل روتینگ intra-as مون مثلا OSPF برای شبکه های خیلی بزرگ نمیتونه اسکیل بشه حتی در داخل خود یک As هم مجبور شدیم اونو به صورت سلسله مراتبی پیاده سازی بکنیم تا قابل اجرا باشه

3- performance :

در intra-as برامون اهمیت داره و میتونیم روتینگ پروتکلی رو استفاده بکنیم که performance بهتری داشته باشه ینی مسیرهای سریع تری رو در اختیار ما بذاره ولی بین as ها معیارهای بیزنسی و پالیسی اولویت داره نسبت به performance



# Business Relationships

- Neighboring ASes have business contracts
  - How much traffic to carry
  - Which destinations to reach
  - How much money to pay
- Common business relationships
  - Customer-provider
    - E.g., Princeton is a customer of USLEC
    - E.g., MIT is a customer of Level3
  - Peer-peer
    - E.g., UUNET is a peer of Sprint
    - E.g., Harvard is a peer of Harvard Business School

دید بهتری نسبت به ماهیت ارتباط بین as ها به دست بیاریم:  
هر دوتا as همسایه لینک ارتباطی که بین همدیگر برقرار میکنند یک کانال ارتباطی فیزیکی بینشون ایجاد میشه و با استفاده از این لینک ترافیک می تونه بینشون رد و بدل بشه و این ارتباط یک ارتباط تجاری است و براساس یک قرارداد تجاری برقرار میشه و توی این قرار داد مواردی که باید مشخص بشه:

1- چه حجم ترافیکی می تونه از طریق این کانال رد و بدل بشه

2- این ترافیک برای چه مقصد هایی استفاده خواهد شد

3- چه هزینه ای به ازای ترافیک عبوری هر کدوم به دیگری پرداخت خواهد کرد

این موارد تابع این است که جایگاه این دوتا as نسبت به همدیگر به چه صورت است؟

جایگاه as همسایه نسبت به همدیگر می تونه به صورت Customer-provider باشه که در

اینصورت یکی از این ها customer اون یکی است اون یکی هم provider اون هست مثلا

دانشگاه صنعتی اصفهان مثلا عرض باند اینترنت خودش رو از شرکت x دریافت میکنه در این

صورت as دانشگاه صنعتی customer است و x as همیشه provider

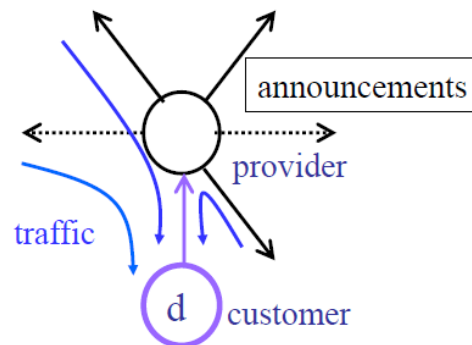
جایگاه بعدی Peer-peer است در این حالت دوتا as نسبت به هم جایگاه یکسانی دارند برای

مثال هر دو provider هستند

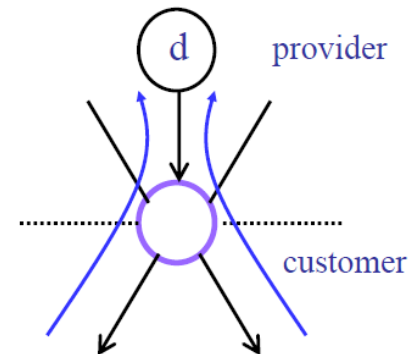
# Customer/Provider

- Customer needs to be reachable from everyone
  - Provider tells all neighbors how to reach the customer
- Customer does not want to provide transit service
  - Customer does not let its providers route through it

Traffic **to** the customer



Traffic **from** the customer



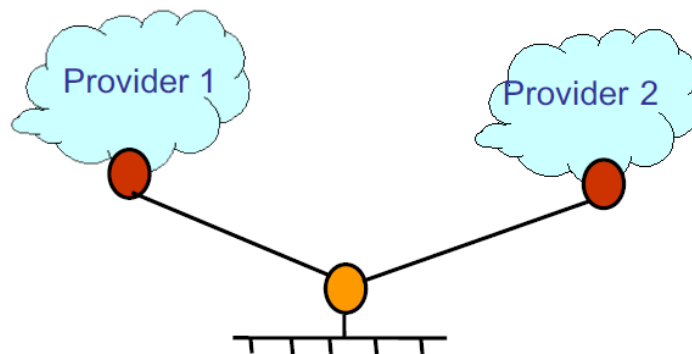
جایگاه ارتباط دوتا as همسایه نسبت به هم Customer/Provider باشد در این صورت ارتباط بین اون ها بر چه اساسی برقرار میشود؟

در ارتباط بین Customer/Provider به هر حال Customer یک شبکه ای است که از طریق provider به اینترنت وصل میشه و طبیعتا انتظار داره تمام ترافیکی که به اینترنت داره یا از اینترنت دریافت میکنه انتظار داره این ترافیک از طریق provider اش رد و بدل بشه و لازمه این کار این است که Provider مسیر دسترسی به این کاستومر رو اطلاع رسانی بکنه هم به کاستومرهای دیگر خودش و هم به peer های خودش و هم به بقیه as ها در اینترنت پس به این ترتیب ترافیکی که as ها برای این مقصد که مقصدش کاستومر d هستند دارند می تونه از طریق این Provider به کاستومر برسه (توی این شکل فلش های سیاه رنگ مسیر اطلاع رسانی رو نشون میدن و فلش های ابی رنگ مسیرهای دیتایی برای این مقصد همون کاستومر از جاهای مختلف ارسال میشه) و انتظار داریم Provider مسیر دسترسی به ساب نت های دیگر در as های دیگر رو اطلاع رسانی بکنه به کاستومر خودش در نهایت کاستومر می تونه ترافیک خودش رو به ساب نت از طریق این Provider ارسال بکنه و همینطور کاستومرهای دیگر این Provider

نکته: کاستومر به هیچ وجه مایل نخواهد بود ترافیک Provider خودش رو عبور بده برای مقصدهای دیگر ینی یک Provider ترافیک یک کاستومر دیگر خودش رو نباید از طریق این کاستومر که اینجا d است به مقصد خودش برسونه البته اگر مقصد خود این کاستومر نباشه و یک جای دیگه باشه

# Multi-Homing

- Customers may have more than one provider
  - Extra reliability, survive single ISP failure
  - Financial leverage through competition
  - Better performance by selecting better path
  - Gaming the 95<sup>th</sup>-percentile billing model





# Export Policies

- Provider to Customer
  - All routes so as to provide transit service
- Customer to Provider
  - Only customer routes
  - Why?
  - Only transit for those that pay
- Peer to Peer
  - Only customer routes

در حالت کلی سیاست هایی که برای Export روت ها اعمال بشه:

انتظار داریم یک provider برای کاستومر خودش همه روت ها رو Export بکنه یا اطلاع رسانی بکنه و provider موظف است که همه ترافیکی کاستومر خودش رو عبور بده و سرویس transit بهش بده

**Customer to Provider** : صرفاً روت هایی که مربوط به خودش است رو Export میکنه چرا؟ برای اینکه روت های مربوط به بقیه رو نمیخواهد که ترافیکش از خودش عبور بکنه و بین دوتا as که peer همدیگر هستند این ها فقط اطلاع رسانی کاستومرهای خودشون رو به peer ما Export میکنند چون ارتباط peer to peer به عنوان transit بقیه نمیخواهد استفاده بشه و فقط برای ترافیک کاستومرهای بین دوتا peer رو برقرار شده



# Import Policies

- Same routes heard from providers, customers, and peers, whom to choose?
  - customer > peer > provider
  - Why?
    - Choose the most economic routes!
    - Customer route: charge \$\$ J
    - Peer route: free
    - Provider route: pay \$\$ L

چه پالیسی هایی برای قبول کردن اطلاع رسانی هایی که دریافت شده انتظار میره که اعمال بشه:  
یک As یک مسیری رو برای مقصد یکسان سه تا اطلاع رسانی بهش میشه یکیش provider اش  
است و یکیش از کاستومرش است و یکش از peer  
کدوم یکی از اینارو باید انتخاب بکنه؟

ترتیب منطقی که as میتونه به کار بگیره برای قبول اطلاع رسانی ها است که ابتدا کاستومر بعد  
peer و بعد provider چرا؟ برای اینکه میخواد مسیری رو انتخاب بکنه که بیشترین سود رو  
براش داشته باشه و اقتصادی تر باشه و مسیر کاستومر اقتصادی تر است چون هزینه لینک رو  
کاستومر داره اینجا می ده و ترافیکی که عبور میکنه رو هزینه اش رو کاستومر میده و هیچ هزینه  
ای برای provider نداره - بعد peer برای اینکه قرارداد بین peer ها تحاطور؟؟ است برای  
انتقال ترافیک و این ها معمولا هزینه ای با هم رد و بدل نمیکند صرفا ترافیک همدیگر رو به  
صورت متقابل عبور میدن و این مسیر هزینه ای برای provider نخواهد داشت - مسیر از طریق  
provider انتخاب میشه چون در ارتباط بین کاستومر و provider ، کاستومر هست که هزینه  
رو میده و برای این مسیر as باید هزینه عبور ترافیک رو پرداخت بکنه

# BGP route selection

- router may learn about more than one route to destination AS, selects route based on:
  1. local preference value attribute: policy decision
  2. shortest AS-PATH
  3. closest NEXT-HOP router: hot potato routing
  4. additional criteria

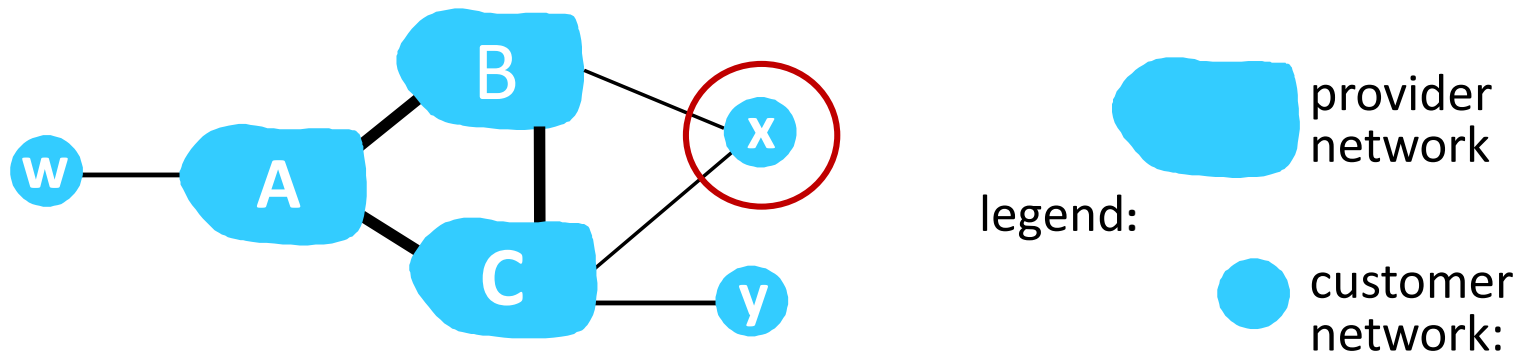
در پروتکل bgp برای یک مقصد مشخص ممکنه مسیرهای مختلفی به یک as اطلاع رسانی بشه و این as باید یکی رو انتخاب بکنه حالا این انتخاب بر چه اساسی صورت میگیره؟  
پالیسی گذاری و پالیسی ها در پروتکل bgp اهمیت به سزایی دارند و این تصمیم گیری تحت شعاع پالیسی ها قرار میگیره

1- در اولین مرحله تصمیم گیری براساس یک مقداری است که به عنوان یکی از صفت های مسیرها در نظر گرفته میشه که بهش local preference میگویم و این local preference براساس پالیسی های تصمیم گیری در as محاسبه میشه و به دست میاد و اولین معیار انتخاب در مقایسه های بین مسیرهای مختلف همین local preference است یا پالیسی های مورد نظر پس مسیری انتخاب میشه که مقدار local preference اون بیشتر باشه که خود این local preference براساس پالیسی ها به دست میاد

اگر تصمیم گیری براساس local preference نتونه انجام بشه یعنی مثلا دوتا مسیر مقدار local preference شون یکسان است در این حالت تصمیم گیری براساس کوتاه ترین AS-PATH انجام می گیره یعنی AS-PATH که از تعداد as کمتری عبور کرده باشه : 2  
و اگر دوتا مسیر AS-PATH شون یکسان باشه در این حالت تصمیم گیری براساس مسیر دسترسی به NEXT-HOP در داخل as انجام می گیره : 3

و اگر از این نظر هم تفاوتی بین دوتا مسیر وجود نداشته باشه معیارهای دیگر رو می بینیم که این معیارها کاملا اختیاری هستند و هر معیاری رو ادمین شبکه می تونه ست کرده باشه برای انتخاب در این مرحله یا کاملا تصادفی یکی رو انتخاب کرده باشه

# BGP: achieving policy via advertisements (more)



ISP only wants to route traffic to/from its customer networks (does not want to carry transit traffic between other ISPs – a typical “real world” policy)

- A,B,C are **provider networks**
- x,w,y are **customer** (of provider networks)
- x is **dual-homed**: attached to two networks
- **policy to enforce**: x does not want to route from B to C via x
  - .. so x will not advertise to B a route to C

چگونه یک As میتونه پالیسی های خودش رو از طریق اطلاع رسانی اعلام بکنه:  
فرض کنید B به عنوان یک provider فقط مایله که به کاستومرهای خودش سرویس بده و مایل نیست که ترافیک کاستومر provider های دیگران رو عبور بده

حالا B چگونه میتونه این پالیسی رو از طریق اطلاع رسانی اعمال بکنه؟

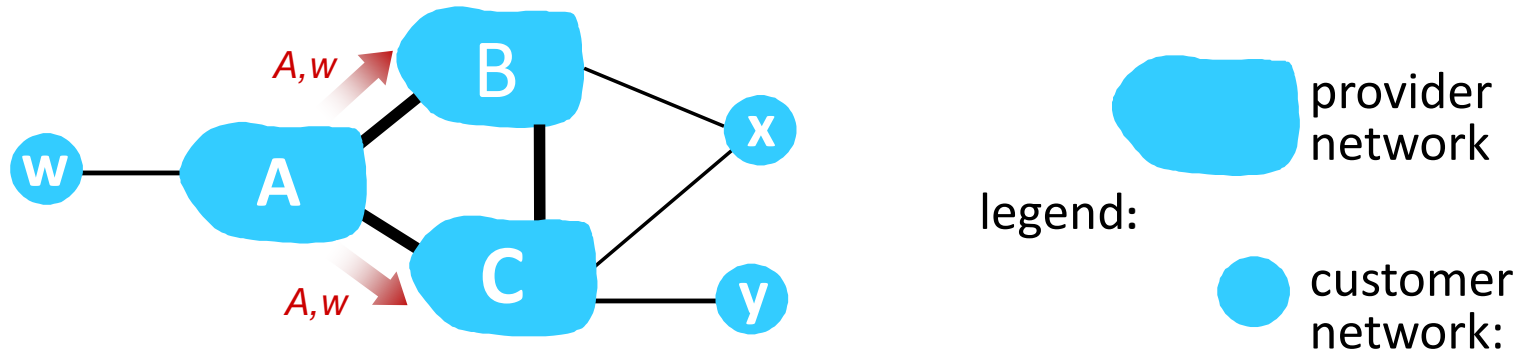
A یک مسیر به w که کاستومر هستش رو به C , B اطلاع رسانی میکنه و B این مسیر که بهش اطلاع رسانی شده رو به C اطلاع رسانی نمیکنه به خاطر این که B از این ترافیک هیچ سودی نخواهد بود چون نه C و نه a و نه w کاستومر B نیستند پس C از وجود مسیر بین خودش و B باخبر نمیشه

C بخاطر اطلاع رسانی Aw که از A دریافت کرده از مسیر CAw باخبر میشه و میتونه از این مسیر برای ترافیک به مقصد w استفاده بکنه --> فک کنم این مال صفحه بعدیه..

مثال: x رو در نظر میگیریم به عنوان یک as که dual-homed است از دو provider A , B سرویس می گیره و x مایل نیست که ترافیک بین C , B رو اطلاع رسانی بکنه  
این رو چجوری x می تونه اعمال بکنه؟

x مسیرهایی که از B یاد میگیره رو به C اطلاع رسانی نمیکنه و از C یاد میگیره رو به B اطلاع رسانی نمیکنه پس نه C از طریق مسیرهایی که از طریق x می تونن بهش دسترسی پیدا بکنند باخبر میشه و نه B از مسیرهایی که از طریق x می تونن دسترسی پیدا بکنند  
البته x خودش رو اطلاع رسانی میکنه به C , B و مسیرهایی که از B بهش اطلاع رسانی میشه رو استفاده می کنه و همینطور مسیرهایی که از C بهش اطلاع رسانی میشه رو استفاده میکنه ولی این مسیر رو به provider های دیگه منتقل نمیکنه که اونا بتونن ازش استفاده بکنند

# BGP: achieving policy via advertisements



ISP only wants to route traffic to/from its customer networks (does not want to carry transit traffic between other ISPs – a typical “real world” policy)

- A advertises path Aw to B and to C
- B *chooses not to advertise* BA<sub>w</sub> to C!
  - B gets no “revenue” for routing CBA<sub>w</sub>, since none of C, A, w are B’s customers
  - C does *not* learn about CBA<sub>w</sub> path
- C will route CA<sub>w</sub> (not using B) to get to w



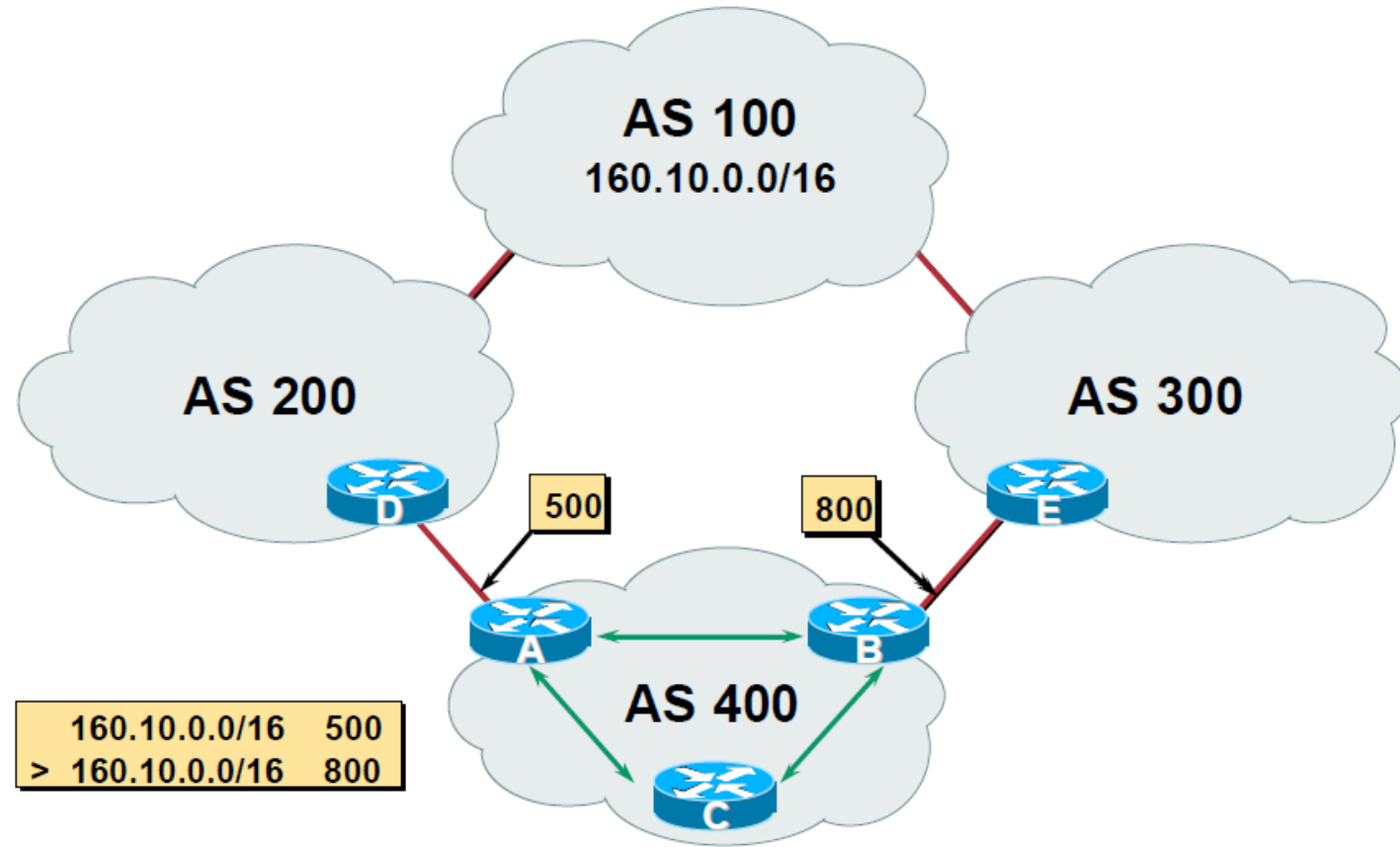


# Local Preferences

- **Local to an AS – non-transitive**  
Default local preference is 100 (IOS)
- **Used to influence BGP path selection**  
determines best path for *outbound* traffic
- **Path with highest local preference wins**

**Local Preferences** یک مقداری است که به یک **as path** نسبت داده میشه و به صورت محلی در **as** این کار انجام میگیره به عبارت دیگر این منتقل نمیشه به **as** های دیگر و مقدار دیفالتش 100 است و کاربردش تاثیرگذاری روی تصمیم گیری روی انتخاب مسیر است ینی مسیرهایی که اطلاع رسانی شده به یک **as** در نتیجه روی مسیری که ترافیک به سمت بیرون عبور خواهد کرد رو تصمیم گیری میکنیم و مسیری رو انتخاب میکنیم که بالاترین **Local Preferences** رو داشته باشه

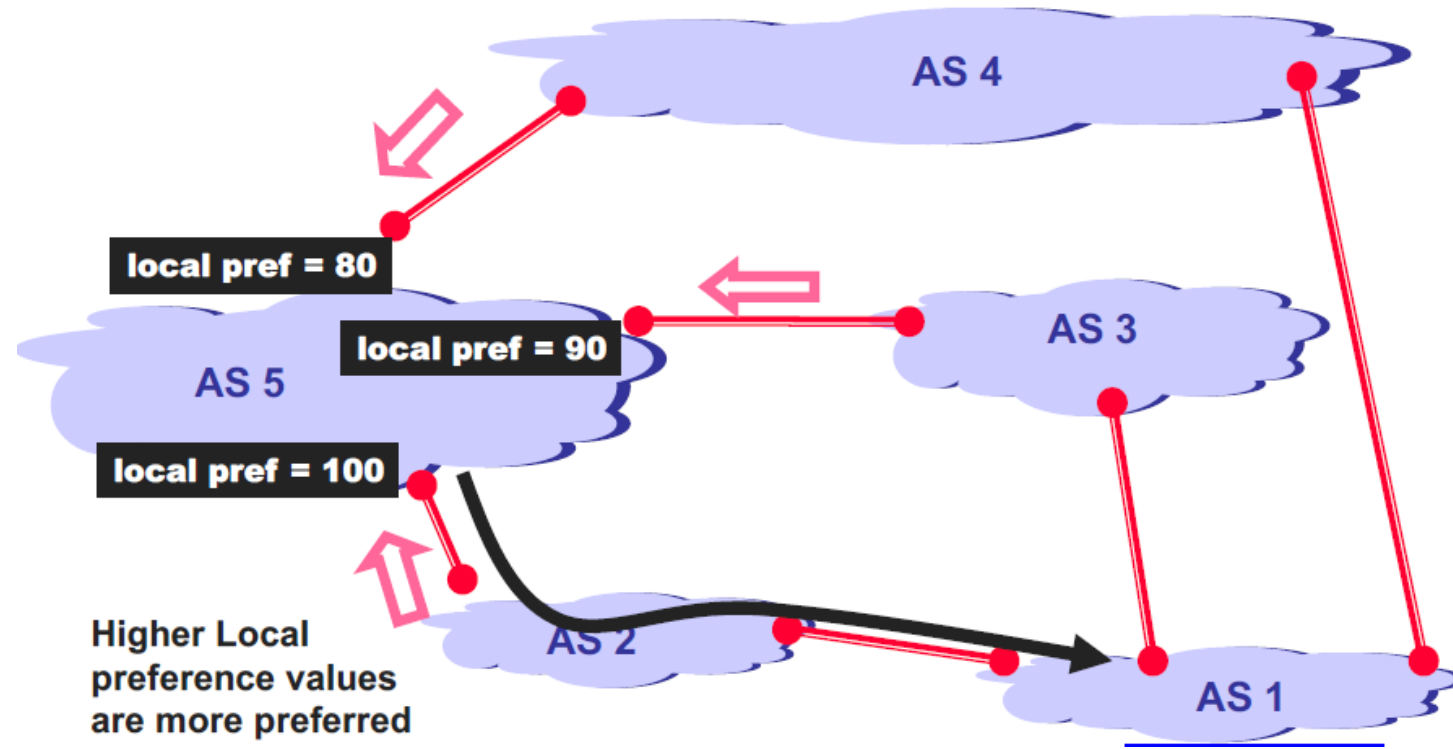
# Local Preferences



فرض میکنیم که as400 یک سرویس گیرنده است و دوتا سرویس provider دارد  
as400 هزینه کمتری برای as300 پرداخت میکند بر این اساس سیاست گذاری کرده که برای  
مقصدهایی که از as300 میتونه به اون ها برسه Local Preferences بالاتری رو اختصاص  
بده و برای as200 چون هزینه بیشتری پرداخت میکنه Local Preferences کمتری قرار  
میده

as400 برای مقصدی اگر مسیر از as300 داشته باشه باید اینو انتخاب بکنه مستقل از Local  
Preferences اش یا اگر برای مقصدی مسیری از as200 داشته باشه باید اونو انتخاب بکنه  
ولی اگر برای یک مقصدی از هر دوی اینا مسیر داشته باشه اونو رو انتخاب میکنه که Local  
Preferences بالاتری دارد

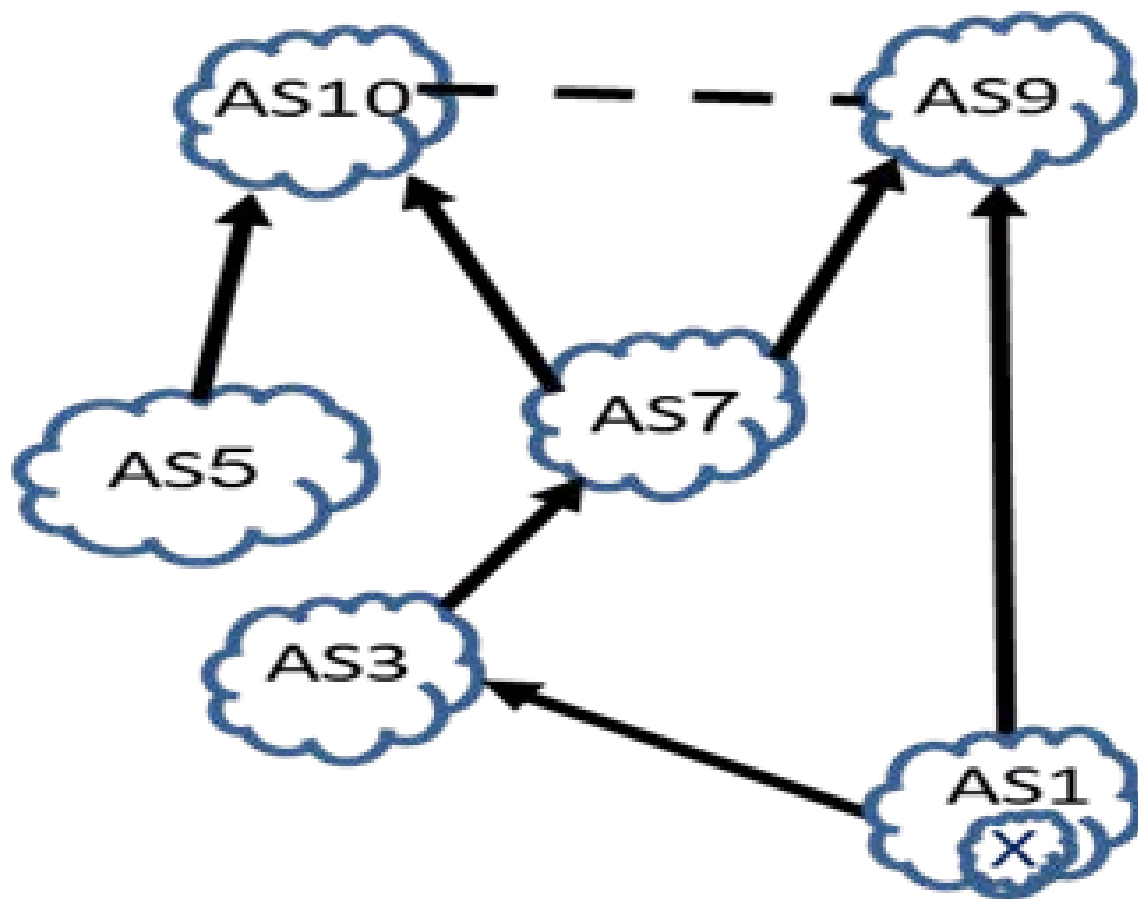
# Local preferences



مثال:

as1 یک as انتهایی است و از سه تا provider سرویس میگیره as2 , as3 , as4  
as2 خودش یک provider محلی است و خودش از as5 سرویس میگیره که یک as ناحیه ای  
است و As3 هم ناحیه ای است و as3 , as5 به عنوان دوتا سرویس provider هم ارز با هم  
قرارداد peer to peer دارند و ترافیک های مشتری هایی همدیگر رو اطلاع رسانی میکنند  
و as4 یک as بالاتر است که As1 مستقیما داره ازش سرویس میگیره در عین حال As5 هم  
خودش دسترسی به اینترنتش رو از As4 میگیره ینی as5 برای as4 یک کاستومر است  
حالا یک ساب نت رو توی as1 در نظر میگیریم و این ساب نت از طریق as1 به همه  
provider هاش اطلاع رسانی میشه حالا این اطلاع رسانی ها میان و به as5 می رسن پس As5  
سه تا مسیر برای این ساب نت می گیره: قبلا گفتیم as5 ترجیح میده اون مسیری رو انتخاب بکنه  
که از کاستومرش ینی as2 میره که هزینه بیوفته روی کاستومر - مسیر peer که میشه as3 رو  
انتخاب نمیکنه هر چند که مجانی است ولی در مقابل ترافیکی که از طریق peer خودش ینی As3  
نواهد فرستاد مجبور است یک ترافیکی از اون هم از خودش عبور بده پس ترجیح میده کمتر از این  
مسیر ترافیک بفرسته - و به هیچ وجه نمی خواهد از طریق provider اش اینو بفرسته چون  
هزینه ی عرض باندی که اینجا استفاده میشه رو As5 باید بده حالا این پالیسی رو چجوری میتونه  
اعمال بکنه؟ as5 ما Local Preferences مسیرهایی که از کاستومر خودش دریافت میکنه رو  
100 می ذاره ینی در واقع بهش دست نمیزنه چون گفتیم دیفالتش 100 هست و Local  
Preferences مسیرهای peer رو 90 می ذاره و Local Preferences مسیرهای  
provide رو روی 80 میذاره و صرفا اینجا اگر براساس Local Preferences اطلاع رسانی  
هایی که بهش شده رو دسته بندی بکنه خود به خود پالیسی اعمال میشه

# Example



در شبکه شکل روبرو خطوط پر نشان دهنده  
ارتباط **Customer-Provider** و خط چین  
نشان دهنده ارتباط **Peer-Peer** است.

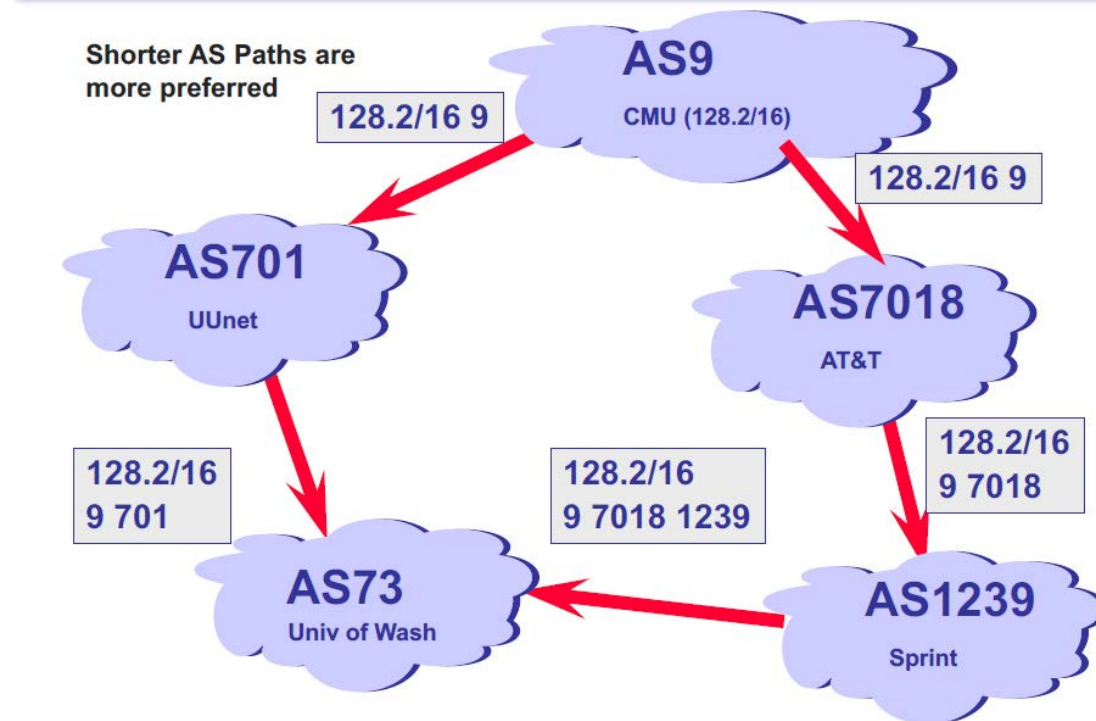
**AS5** چه مسیر (یا مسیرهایی) را برای سابت **x**  
دریافت خواهد کرد؟

کدام مسیر را انتخاب خواهد کرد؟





# Shorter AS path selection



اگر مسیرهایی که برای As اطلاع رسانی میشن از نظر پالیسی ها هیچ ترجیحی بینشون نباشه و Local Preferences یکسانی داشته باشند در این حالت انتخاب بین اون ها براساس طول AS path خواهد بود

مثلا As73 به عنوان یک as انتهایی از دوتا As سرویس میگیره: as701 , as1239 که سرویس provider های as73 هستند

حالا یک ساب نتی از طریق as9 میاد و اطلاع رسانی میشه و تهش از طریق As1239 , as701 به as73 اطلاع رسانی میشه کدوم مسیر؟

سمت راستی از سه تا As رد شده و توی AS path اون لیست سه تا as است در حالی که سمت چپ از دوتا as رد شده و توی لیست اون دوتا As است پس سمت چپی انتخاب میشه

# Select best BGP route to prefix

- ❖ Router selects route based on shortest AS-PATH

- ❖ Example:

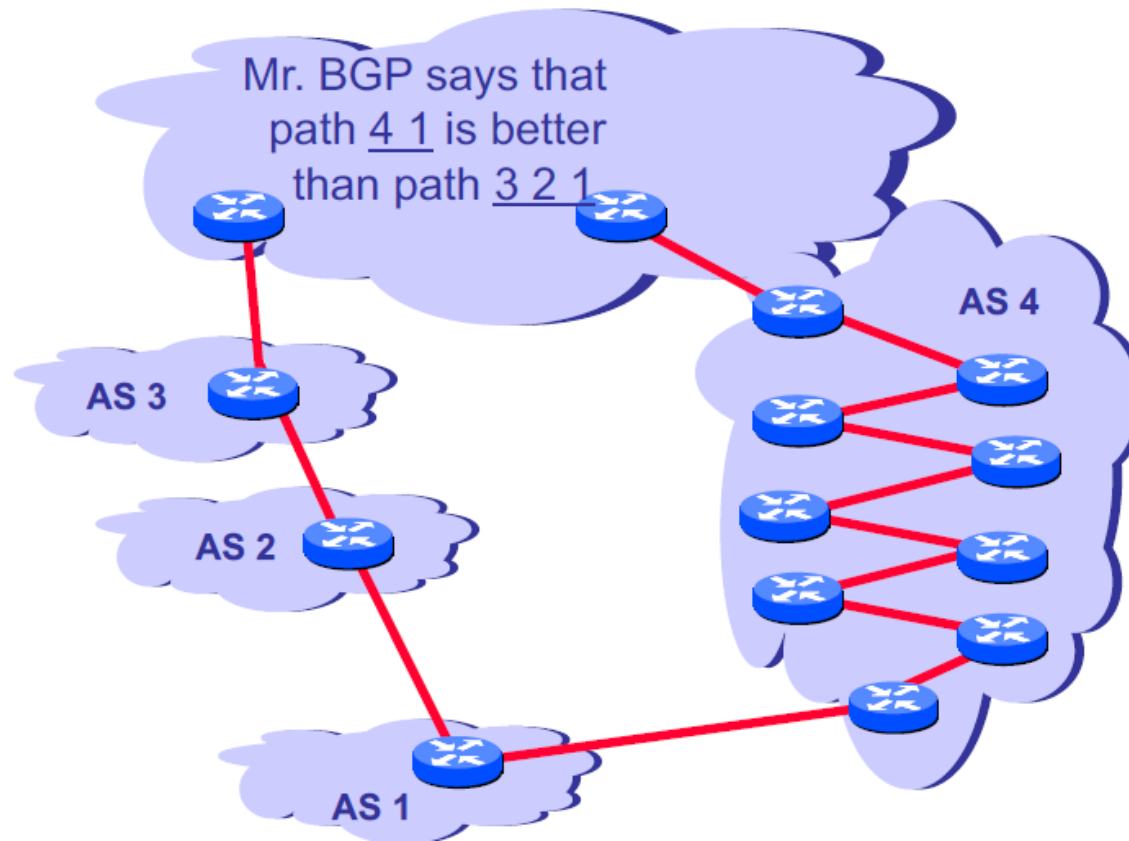
- ❖ AS2 AS17 to 138.16.64/22
- ❖ AS3 AS131 AS201 to 138.16.64/22

select

- ❖ What if there is a tie?

مثال:

# Shorter AS path vs shorter route

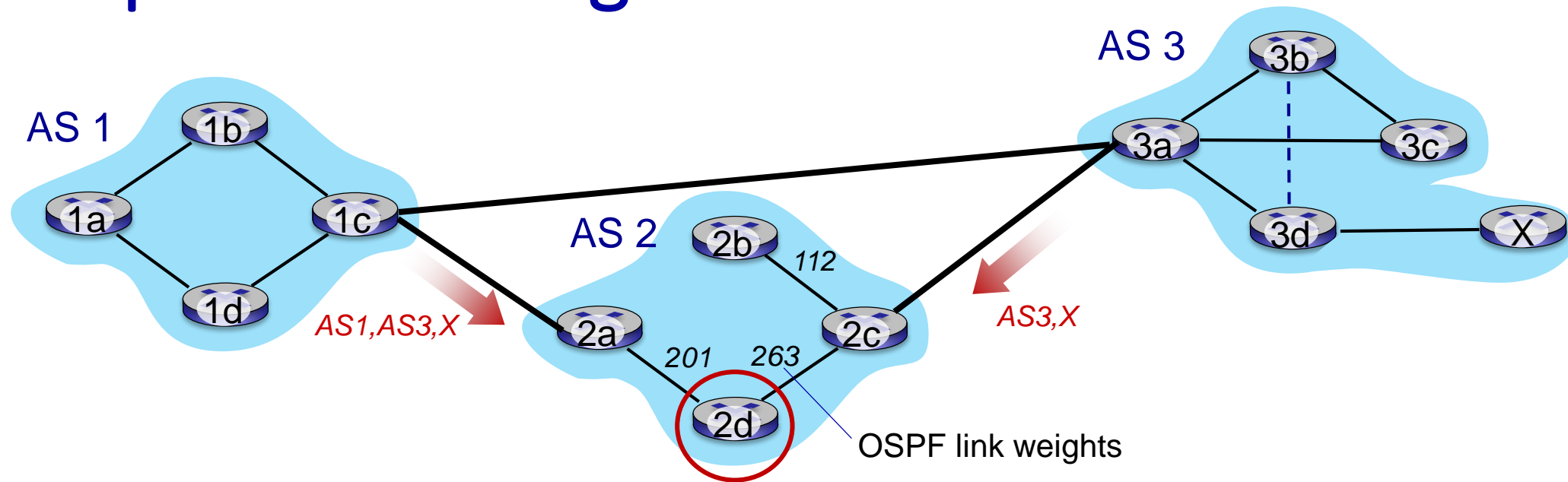


نکته:

کوتاه ترین AS path الزاما به معنی کوتاه ترین مسیر بین روترها نیست چون AS path صرفا as هایی که توی مسیر قرار دارند رو لیست می کنه و از مسیر عبور ترافیک در داخل هر as اطلاعاتی به ما نمیده

مثلا توی شکل سمت راستی ترجیح داده میشه چون AS path کمتری داره نسبت به سمت راستی در حالی که داخل as4 از تعداد روتر بیشتری عبور میکنه و به این معناست که تاخیر این میتونه بیشتر باشه

# Hot potato routing



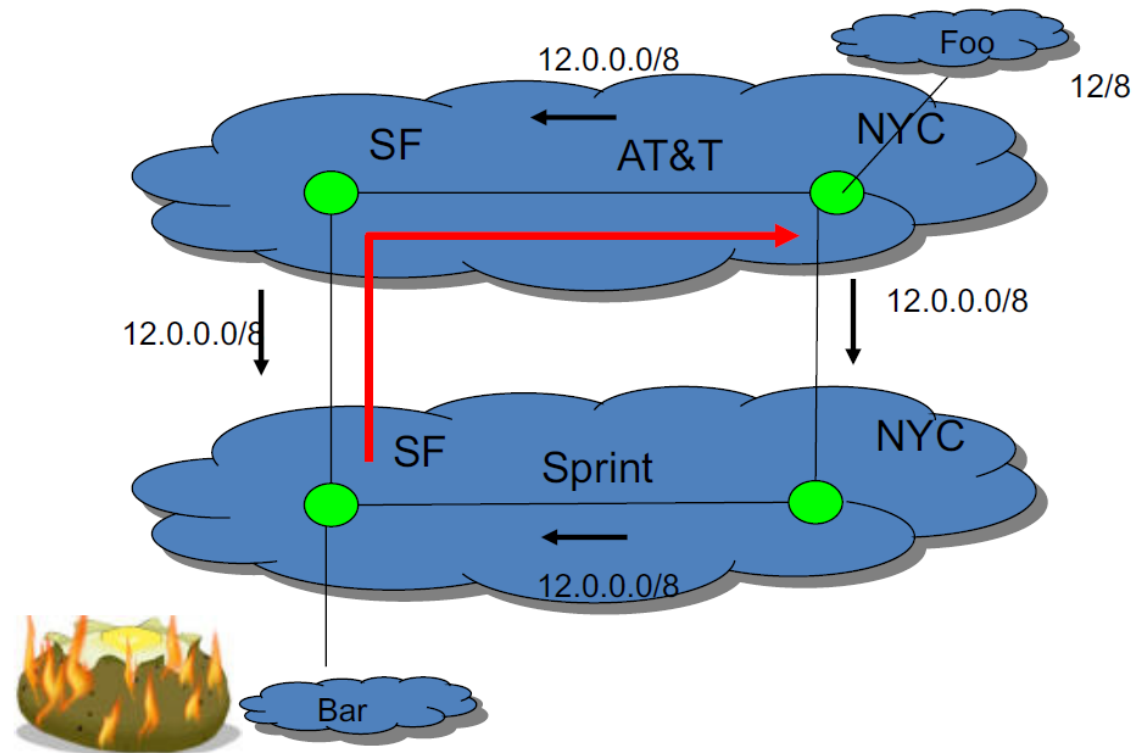
- 2d learns (via iBGP) it can route to X via 2a or 2c
- **hot potato routing**: choose local gateway that has least *intra-domain* cost (e.g., 2d chooses 2a, even though more AS hops to X): don't worry about inter-domain cost!

معیار دیگر: هزینه ای بخشی از مسیر است که در داخل as قرار دارد  
مقصد 2d است و 2d باید انتخاب بکند:

در پروتکل Hot potato routing این انتخاب بر مبنای بخشی از مسیر که در داخل as2 است انجام می‌شود یعنی از 2d تا gateway مربوطه انجام می‌شود  
پس 2d to 2a همیشه 201 که کمتر از 2d to 2c است پس 2a انتخاب می‌شود هر چند که این مسیر as path طولانی‌تری دارد ولی بر اساس این پروتکل این مسیر انتخاب می‌شود



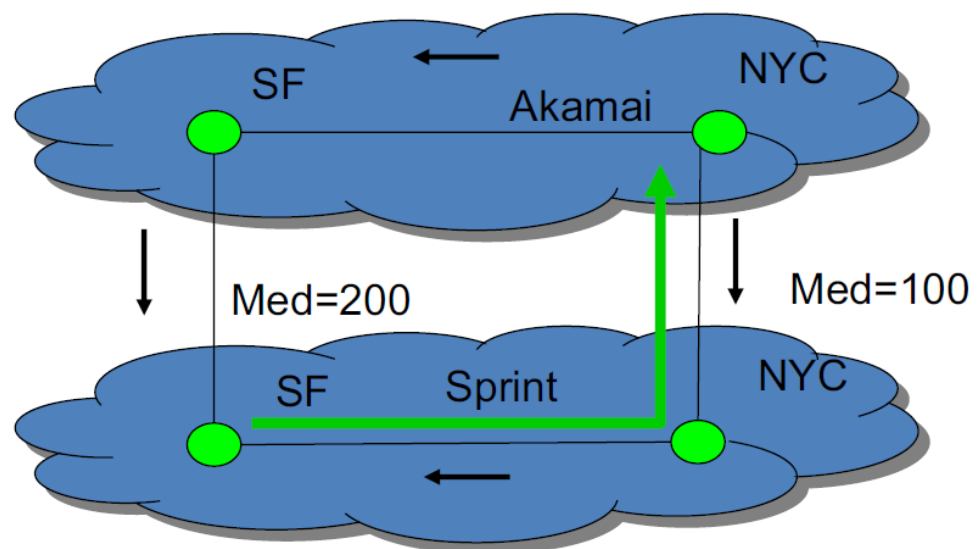
# Hot potato routing



شبکه ها چه انگیزه ای برای استفاده از پروتکل Hot potato routing ممکنه داشته باشند؟  
سناریو:

کاستومر bar میخواد ترافیک رو بفرسته به کاستومر foo

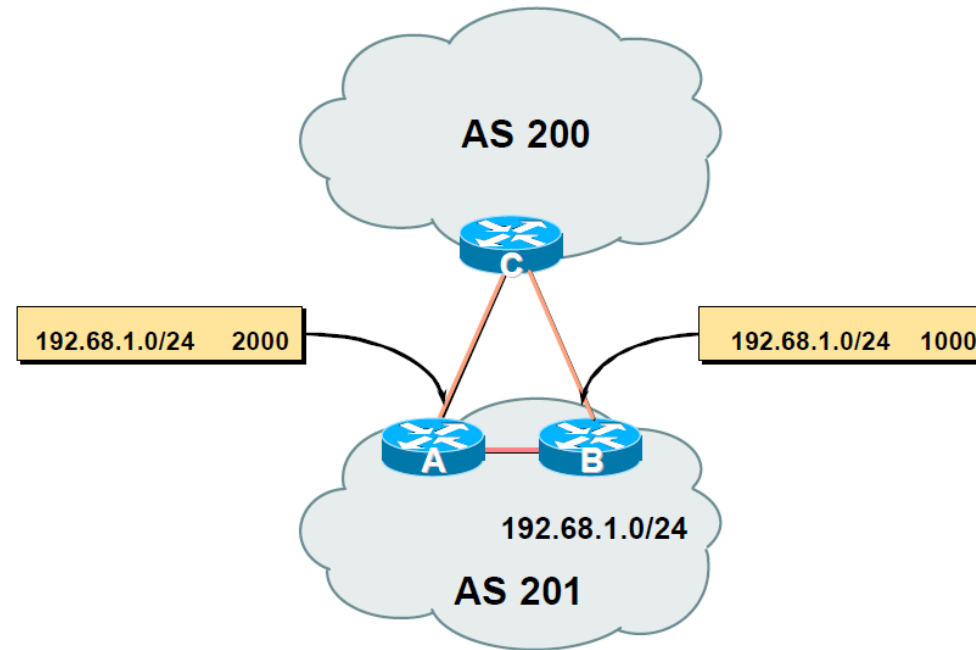
# Cold potato routing



MED: Multi Exit Discriminator

اگر از سبز بره پروتکل ما میشه Cold potato routing سناریو:  
اونی که Med کمتری داشته باشه انتخاب میشه اینجا

# MED: Multi-Exit discriminator



حالا MED: Multi-Exit discriminator چی هست؟

در سناریویی که یک شبکه بیش از یک درگاه با یک شبکه دیگه داره ینی as201 از طریق دو تا درگاه مختلف می تونه به as200 مرتبط بشه و as200 ممکنه ترجیح بده این ترافیک رو از روتر مربوط به B دریافت بکنه در مقایسه با روتر A پس med کمتری رو برای B اختصاص میده و بیشتری رو برای A و این باعث میشه که as201 ما med کمتری رو انتخاب بکنه

# MED

- Inter-AS – non-transitive
- Used to convey the relative preference of entry points
  - determines best path for *inbound* traffic
- Comparable if paths are from same AS
- IGP metric can be conveyed as MED
  - set metric-type internal* in route-map

MED یک پارامتر inter-as است و بین دو تا As مجاور مفهوم داره و منتقل نمیشه  
و مشخص میکنه که As ترجیح میده این ترافیک رو روی کدوم یکی از این درگاه ها دریافت بکنه  
و این میتونه منعکس کننده هزینه اون بخش داخلیه مسیر در اون شبکه باشه



# Routing Protocols

- IGP:
  - Intra-AS routing protocols
    - OSPF
      - Dijkstra
- EGP:
  - Inter-AS routing protocol
    - BGP-4
      - eBGP
      - iBGP

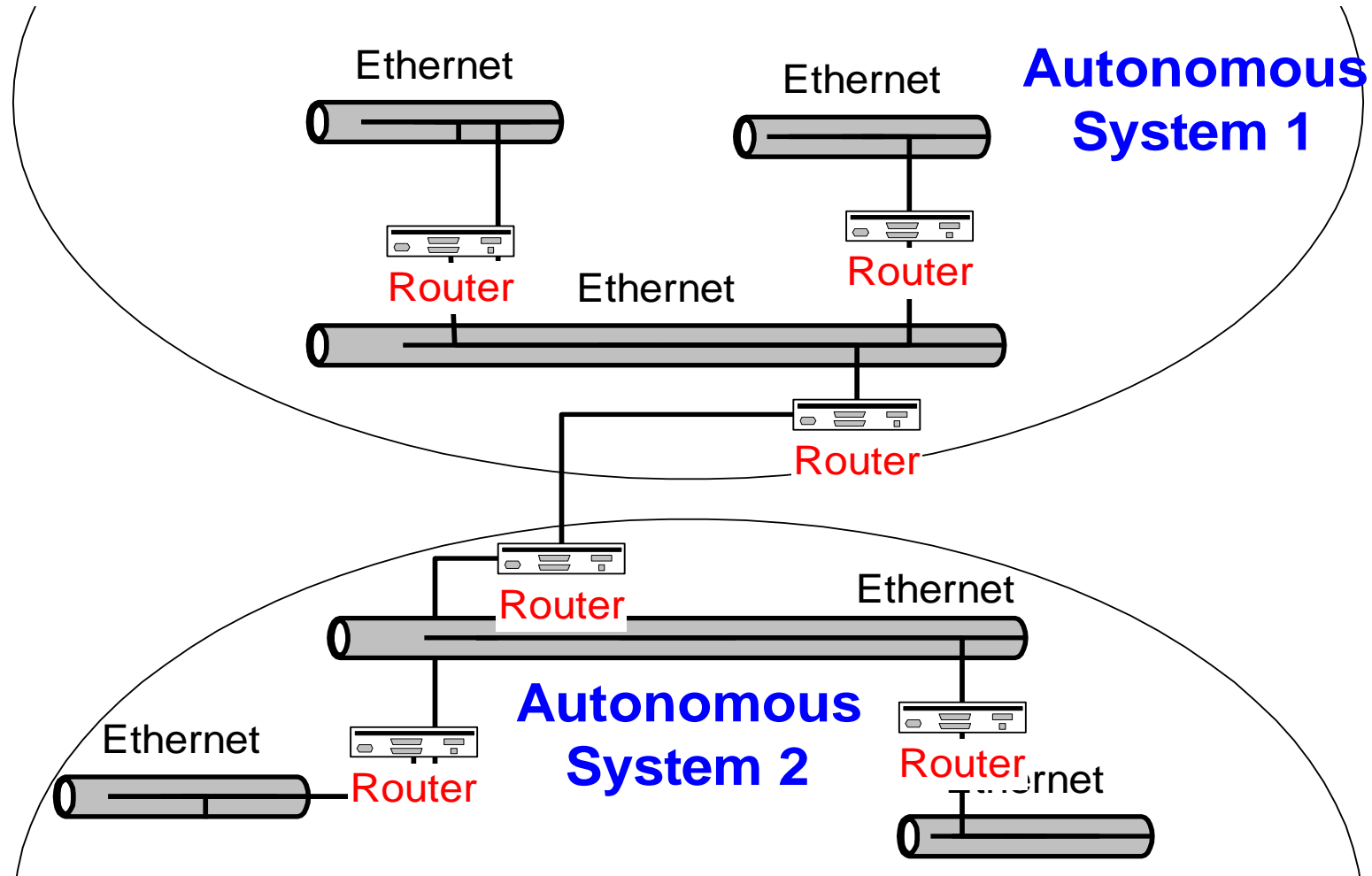


# Autonomous Systems

- An **autonomous system** is a region of the Internet that is administered by a single authority.
- Examples of autonomous regions are:
  - UVA's campus network
  - MCI's backbone network
  - Regional Internet Service Provider
- Types of autonomous system (AS):
  - Stub AS: has connection to only one AS, only carry local traffic
  - Multihomed AS: has connection to >1 AS, but does not carry transit traffic
  - Transit AS: has connection to >1 AS and carries transit traffic
- Routing is done differently within an autonomous system (**intradomain routing**) and between autonomous system (**interdomain routing**).



# Autonomous Systems (AS)





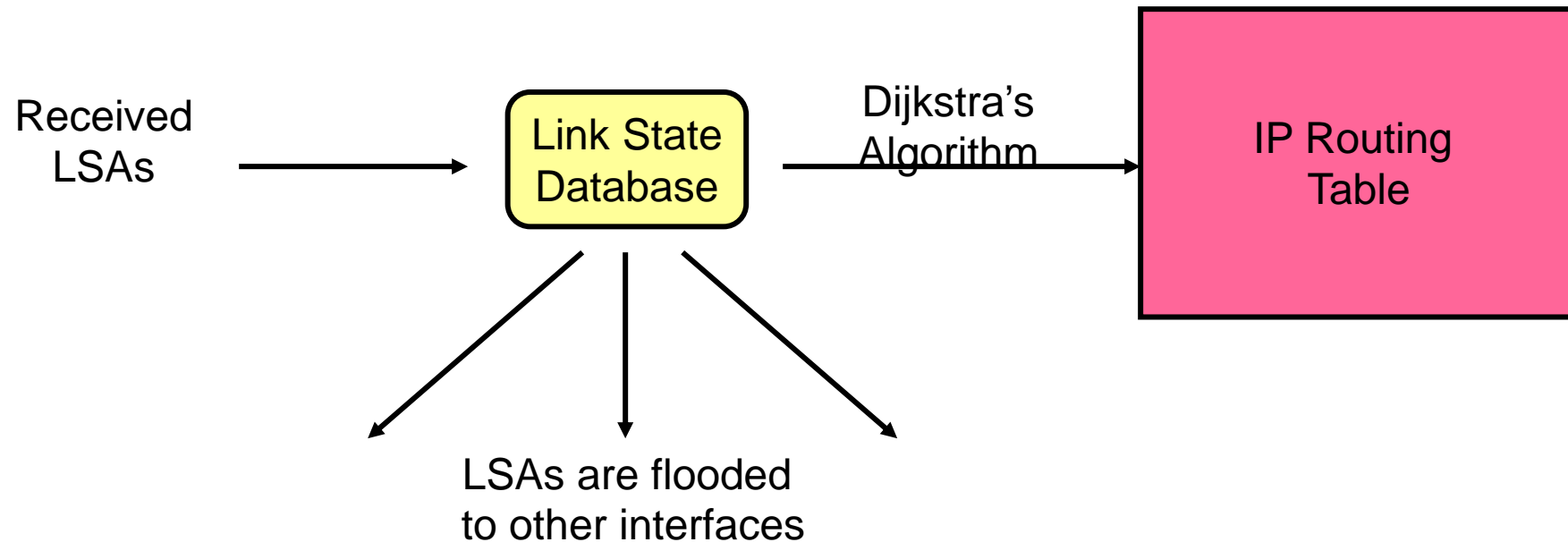
# Link State Routing: Basic principles

1. Each router establishes a relationship (*“adjacency”*) with its neighbors
2. Each router generates *link state advertisements (LSAs)* which are distributed to all routers  
LSA = (link id, state of the link, cost, neighbors of the link)
3. Each router maintains a database of all received LSAs (*topological database* or *link state database*), which describes the network as a graph with weighted edges
4. Each router uses its link state database to run a shortest path algorithm (Dijkstra's algorithm) to produce the shortest path to each network





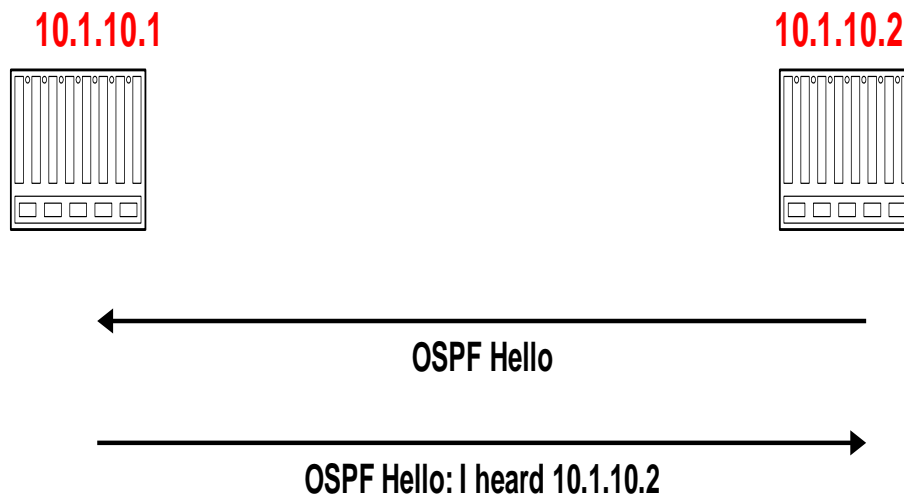
# Operation of a Link State Routing protocol





# Discovery of Neighbors

- Routers multicasts **OSPF Hello packets** on all OSPF-enabled interfaces.
- If two routers share a link, they can become neighbors, and establish an adjacency



Scenario:  
Router 10.1.10.2 restarts

- After becoming a neighbor, routers exchange their link state databases



# Synchronizing OSPF Databases

- While the Hello packet was used to establish neighbor adjacencies, the other four types of OSPF packets are used during the process of exchanging and synchronizing link state databases.

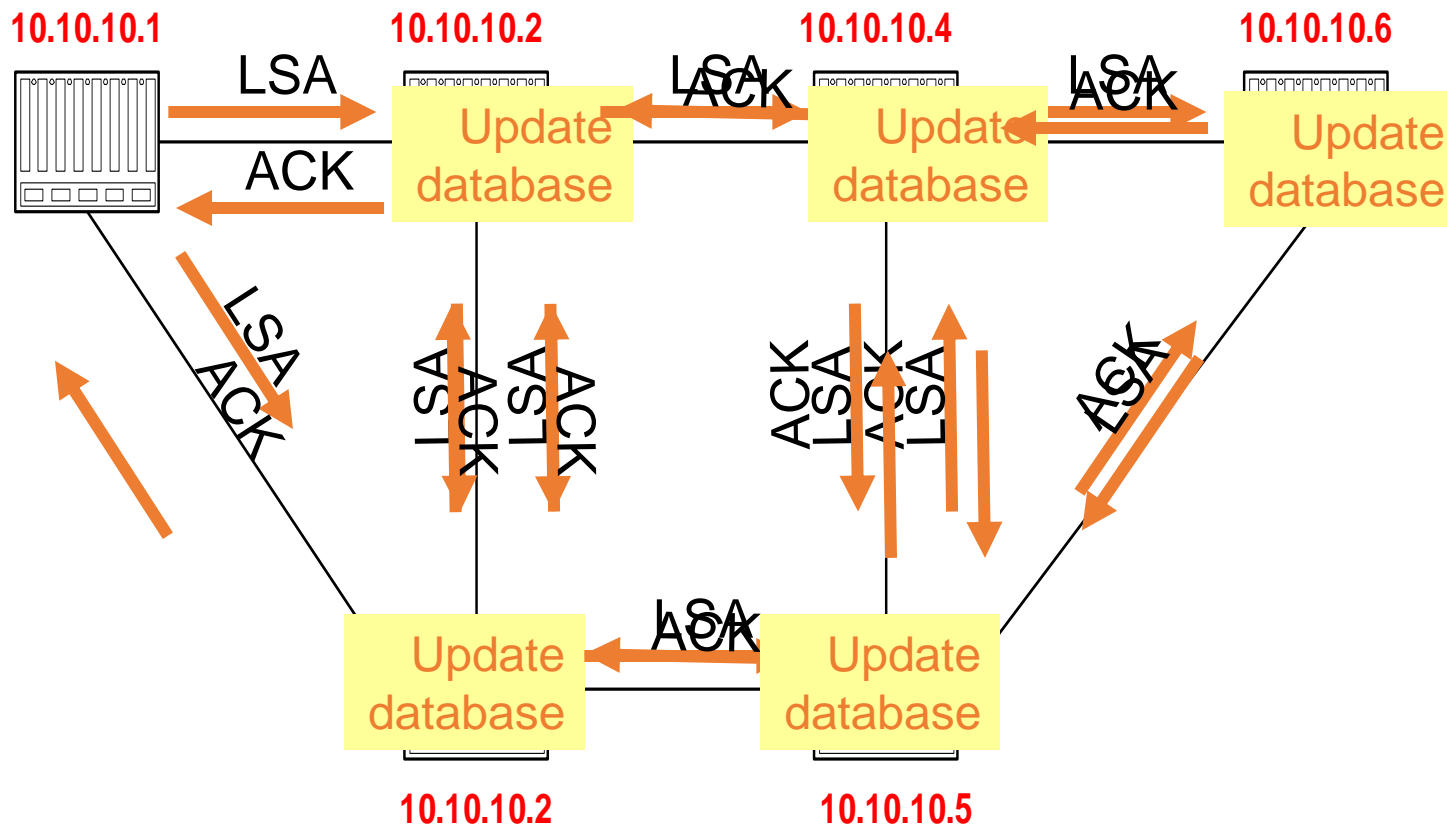
OSPF Packet Descriptions

Type	Packet Name	Description
1	Hello	Discovers neighbors and builds adjacencies between them
2	Database Description (DBD)	Checks for database synchronization between routers
3	Link-State Request (LSR)	Requests specific link-state records from router to router
4	Link-State Update (LSU)	Sends specifically requested link-state records
5	Link-State Acknowledgment (LSAck)	Acknowledges the other packet types



# Routing Data Distribution

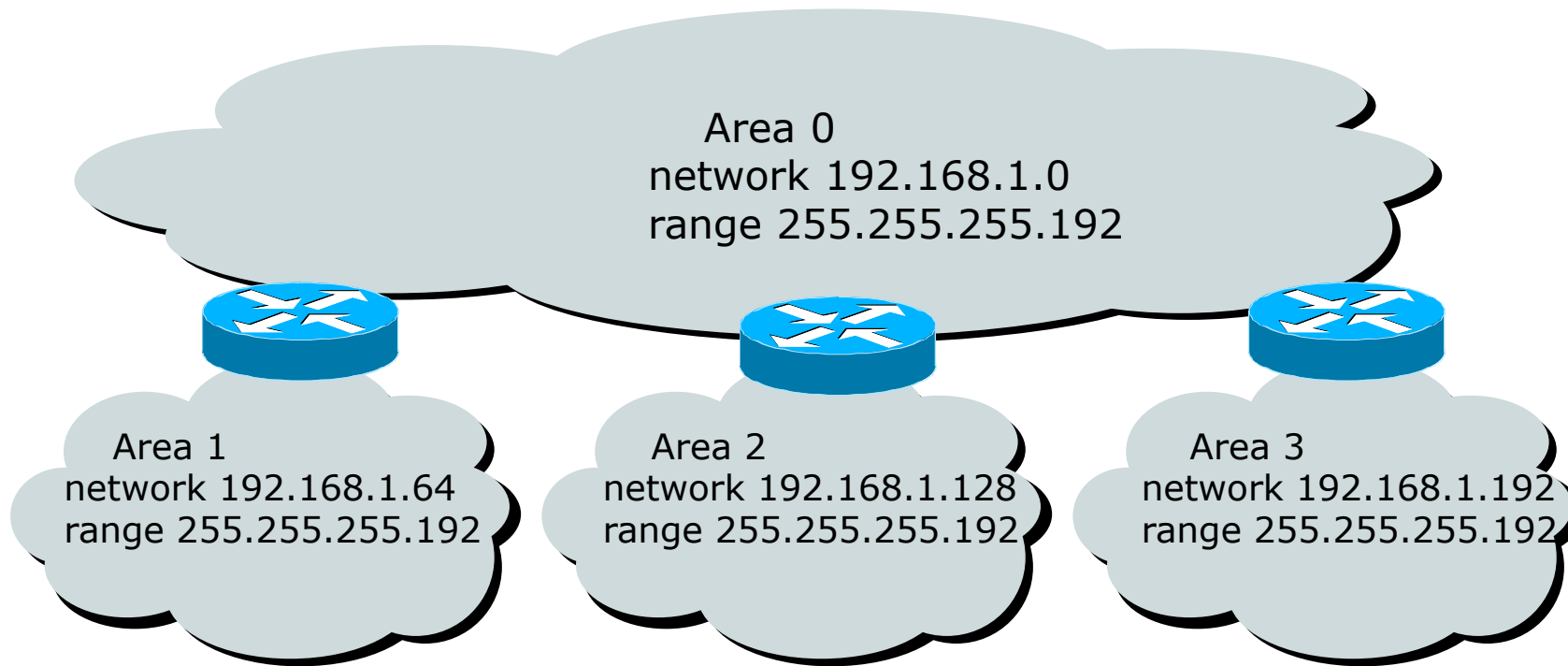
- LSA-Updates are distributed to all other routers via **Reliable Flooding**
- **Example:** Flooding of LSA from 10.10.10.1



-  
اگر تکراری بود حذف میشه و اگر نرسیده بود دوباره می فرسته



# Addressing for Areas

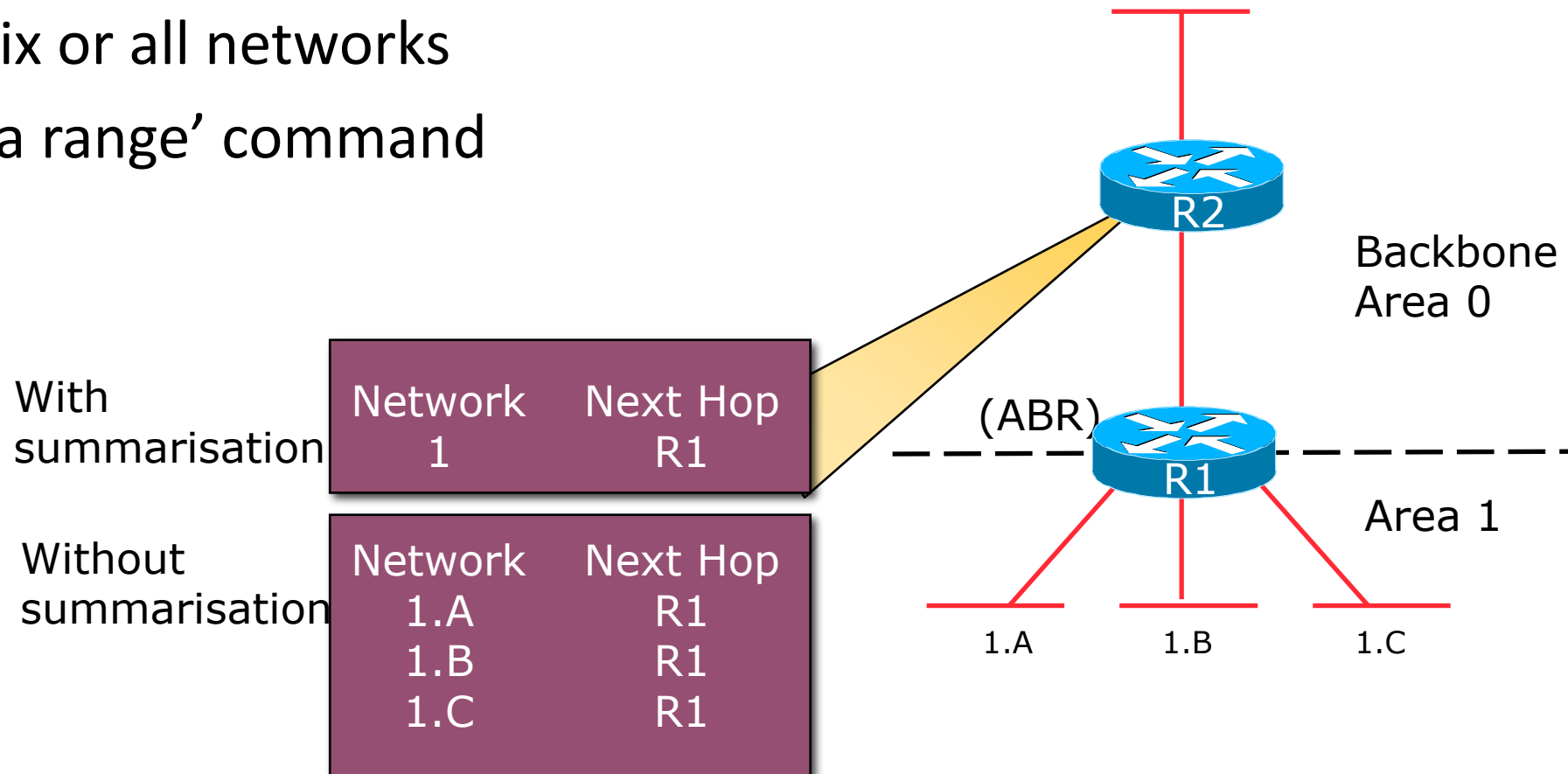


Assign contiguous ranges of subnets per area to facilitate summarisation



# Inter-Area Route Summarisation

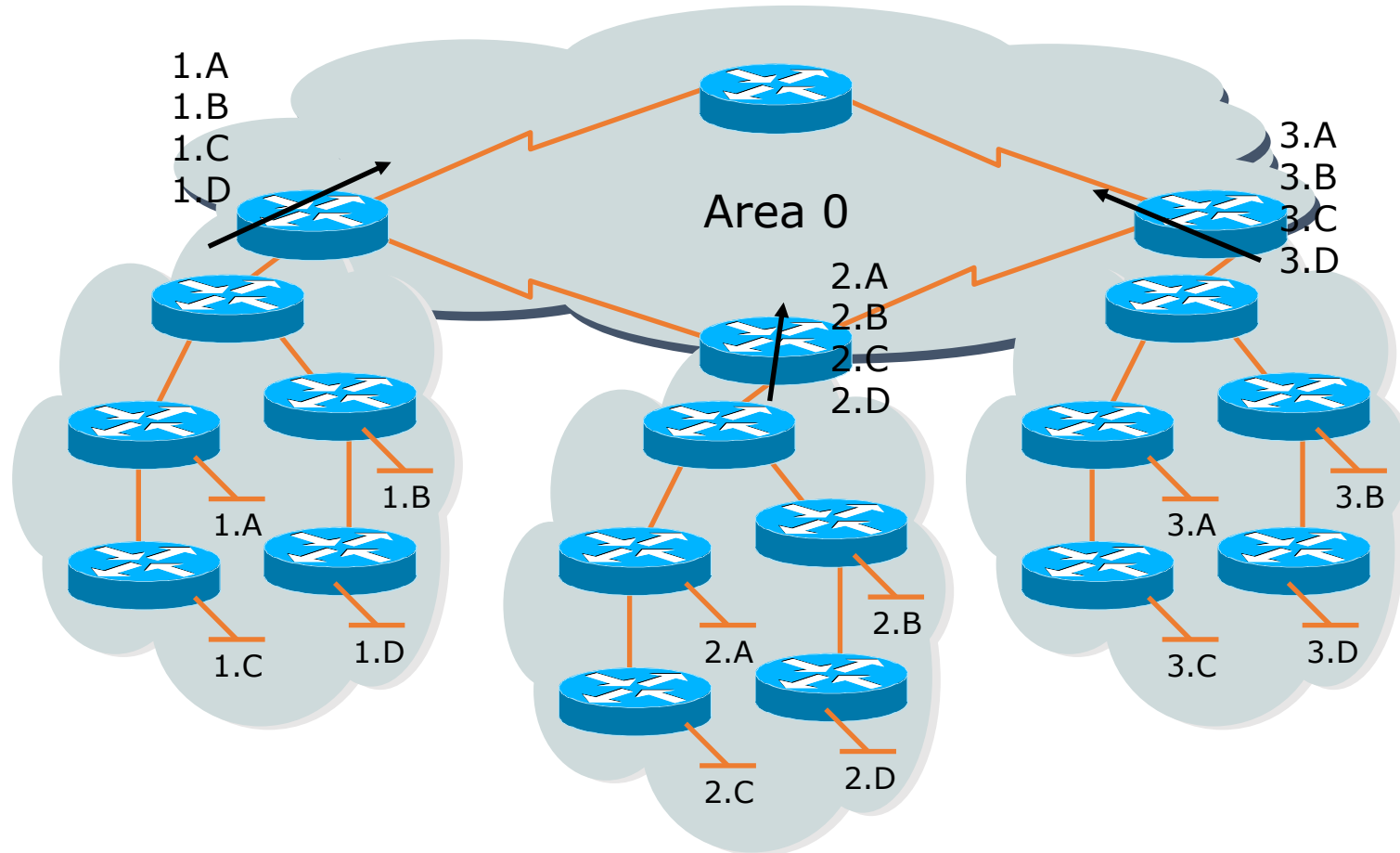
- Prefix or all subnets
- Prefix or all networks
- 'Area range' command





# No Summarisation

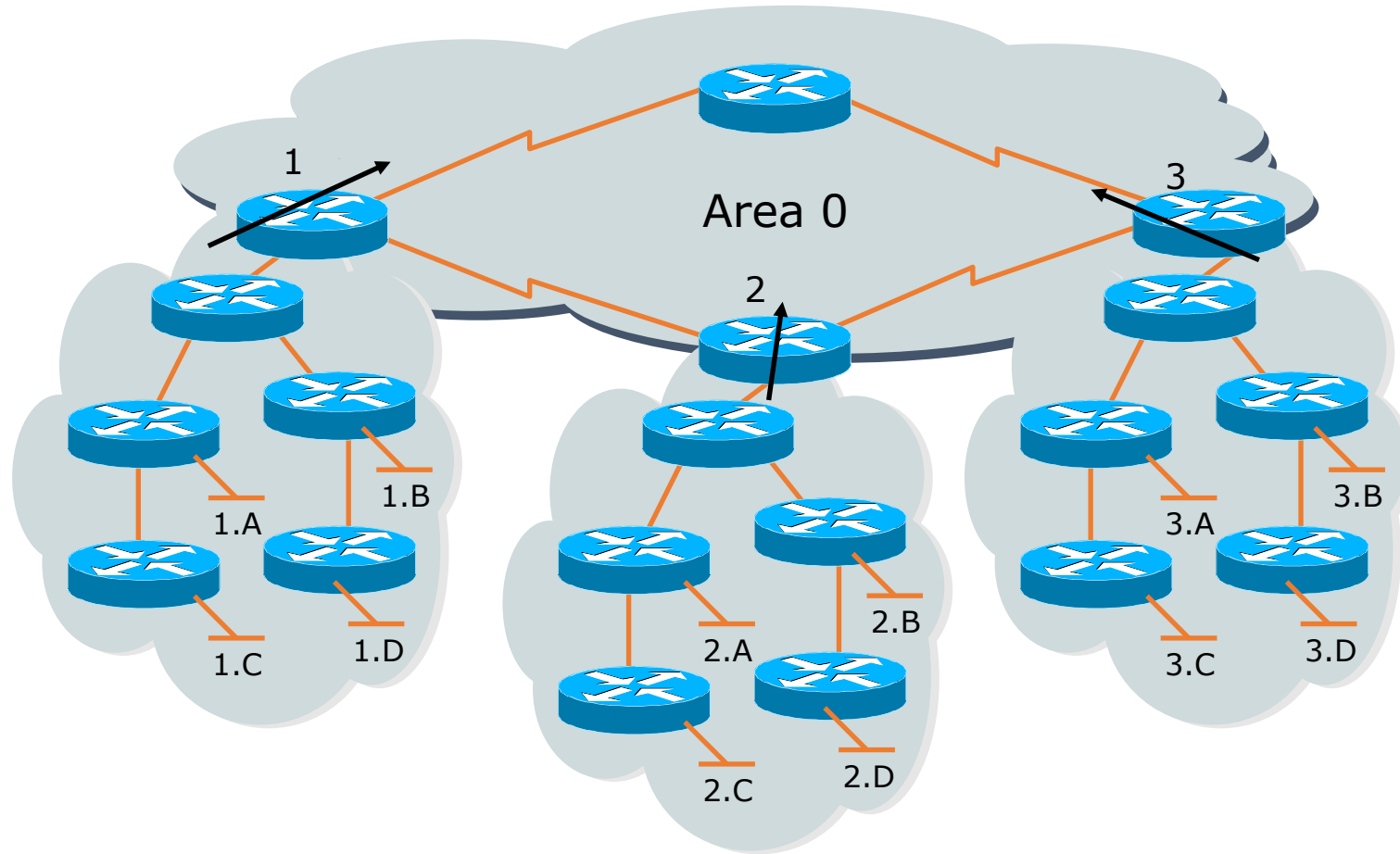
- Specific Link LSA advertised out of each area
- Link state changes propagated out of each area





# With Summarisation

- Only summary LSA advertised out of each area
- Link state changes do not propagate out of the area

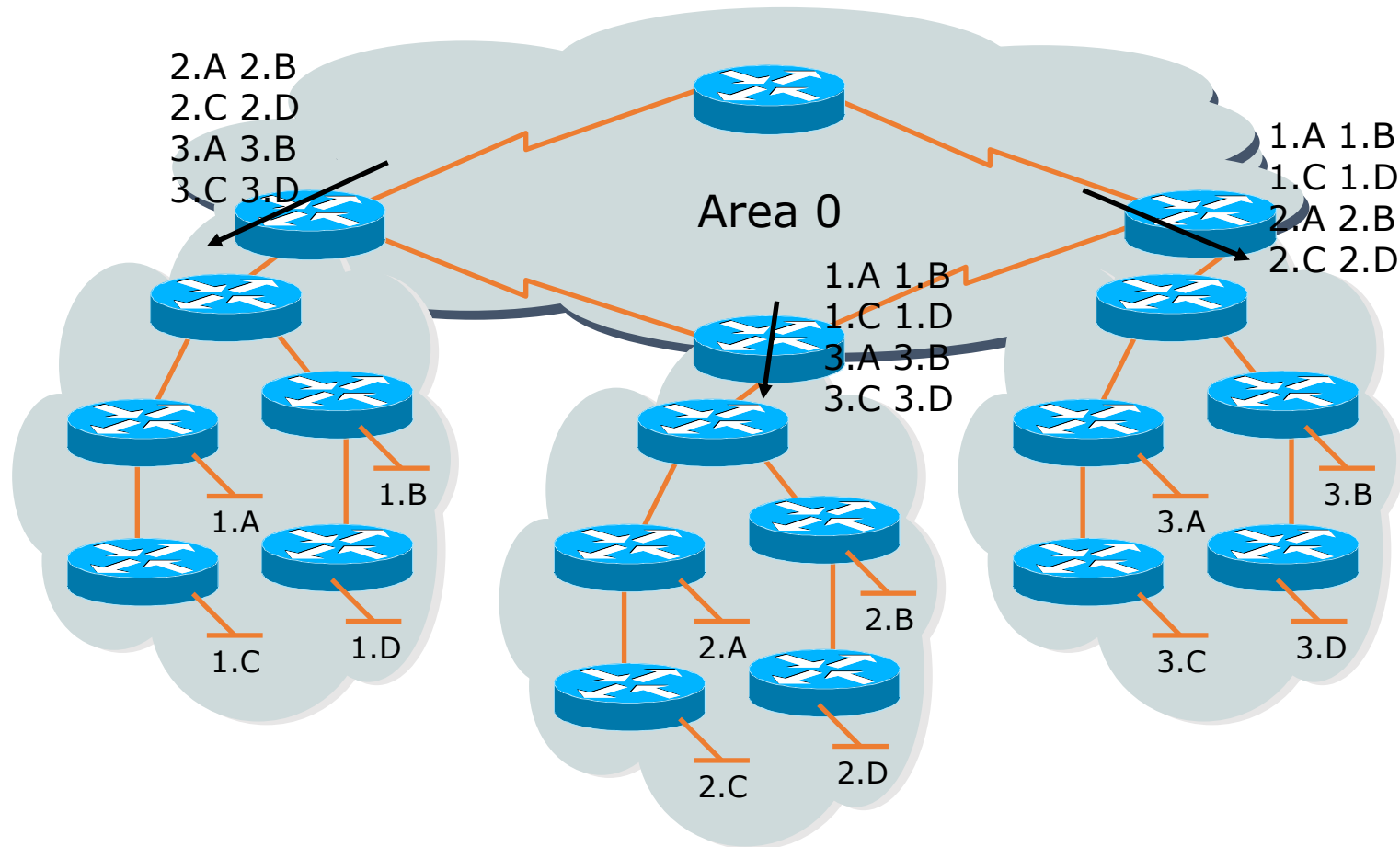


-  
این حالت تجميع شده است



# No Summarisation

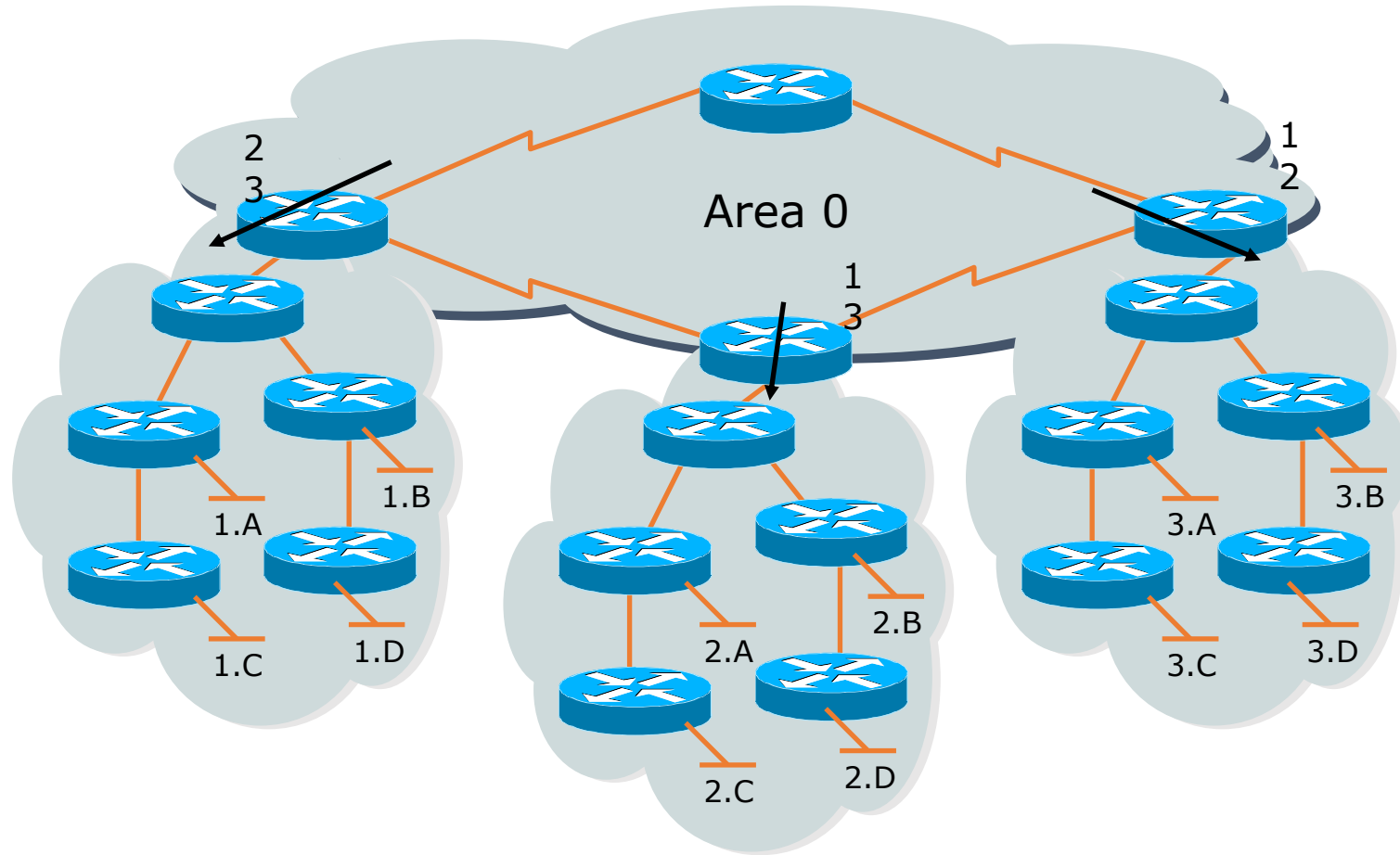
- Specific Link LSA advertised in to each area
- Link state changes propagated in to each area





# With Summarisation

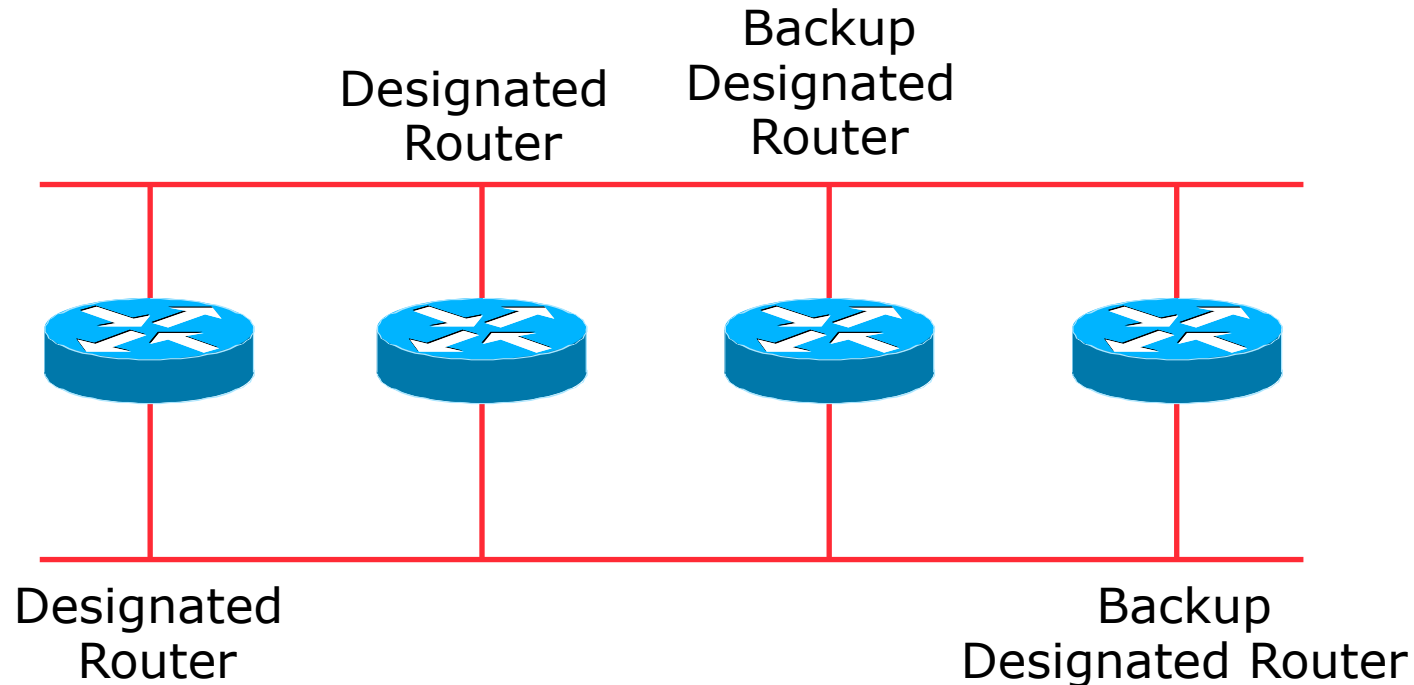
- Only summary link LSA advertised in to each area
- Link state changes do not propagate in to each area





# Designated Router

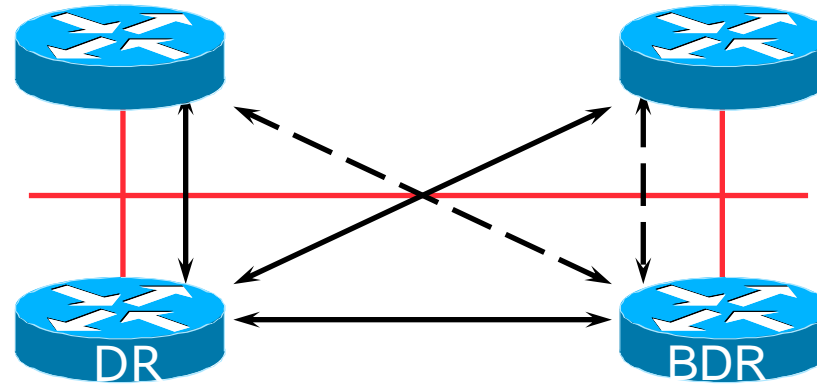
- There is ONE designated router per multi-access network
  - Generates network link advertisements
  - Assists in database synchronization



-

این روترها توی یک lan هستن  
اینجوری کوچکتر هم میشه باز

# LSAs Propagate Along Adjacencies



- LSAs acknowledged along adjacencies

- به روترهای دیگر توی اون lan منتقل بکنه



# Routing Protocol Packets

- Share a common protocol header
- Routing protocol packets are sent with type of service (TOS) of 0
- Five types of OSPF routing protocol packets
  - Hello – packet type 1
  - Database description – packet type 2
  - Link-state request – packet type 3
  - Link-state update – packet type 4
  - Link-state acknowledgement – packet type 5



# Different Types of LSAs

- Six distinct type of LSAs

- Type 1 : Router LSA
- Type 2 : Network LSA
- Type 3 & 4: Summary LSA
- Type 5 & 7: External LSA (Type 7 is for NSSA)
- Type 6: Group membership LSA
- Type 9, 10 & 11: Opaque LSA (9: Link-Local, 10: Area)



# Router LSA (Type 1)

- Describes the state and cost of the router's links to the area
- All of the router's links in an area must be described in a single LSA
- Flooded throughout the particular area and no more
- Router indicates whether it is an ASBR, ABR, or end point of virtual link



# Network LSA (Type 2)

- Generated for every transit broadcast and NBMA network
- Describes all the routers attached to the network
- Only the designated router originates this LSA
- Flooded throughout the area and no more





# Summary LSA (Type 3 and 4)

- Describes the destination outside the area but still in the AS
- Flooded throughout a single area
- Originated by an ABR
- Only inter-area routes are advertised into the backbone
- Type 4 is the information about the ASBR



# External LSA (Type 5 and 7)

- Defines routes to destination external to the AS
- Default route is also sent as external
- Two types of external LSA:
  - E1: Consider the total cost up to the external destination
  - E2: Considers only the cost of the outgoing interface to the external destination
- (Type 7 LSAs used to describe external LSA for one specific OSPF area type)

