Lecture 6: Inter-Domain Routing

Chapter 4.1 of the Peterson and Davie book

(these slides borrow some content from that chapter)

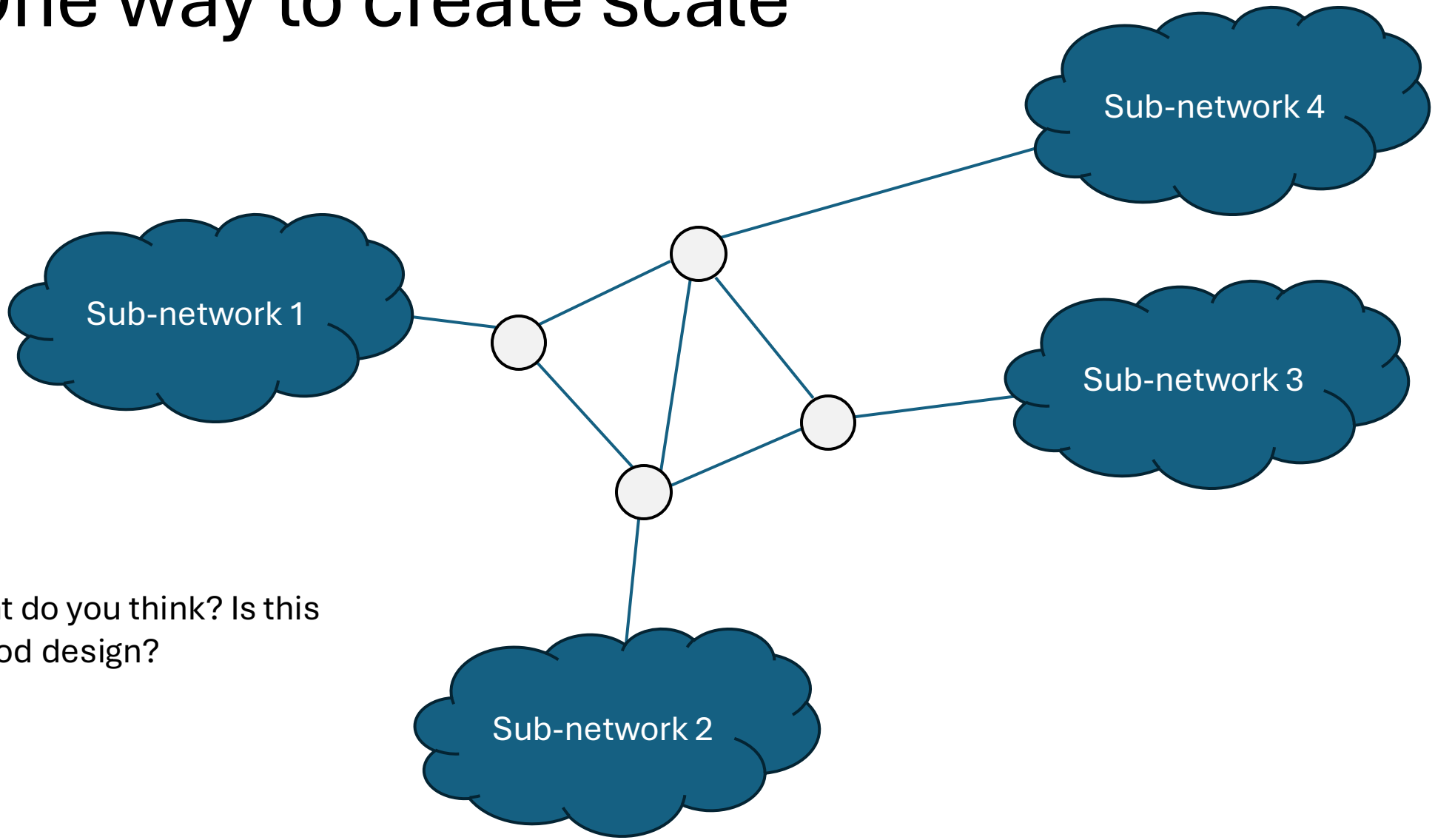
Venkat Arun

We will address two problems today

- Scale
  - The internet has doubled in size every year for 30 years!
  - Scale can only be achieved by having a hierarchical structure and hiding detail

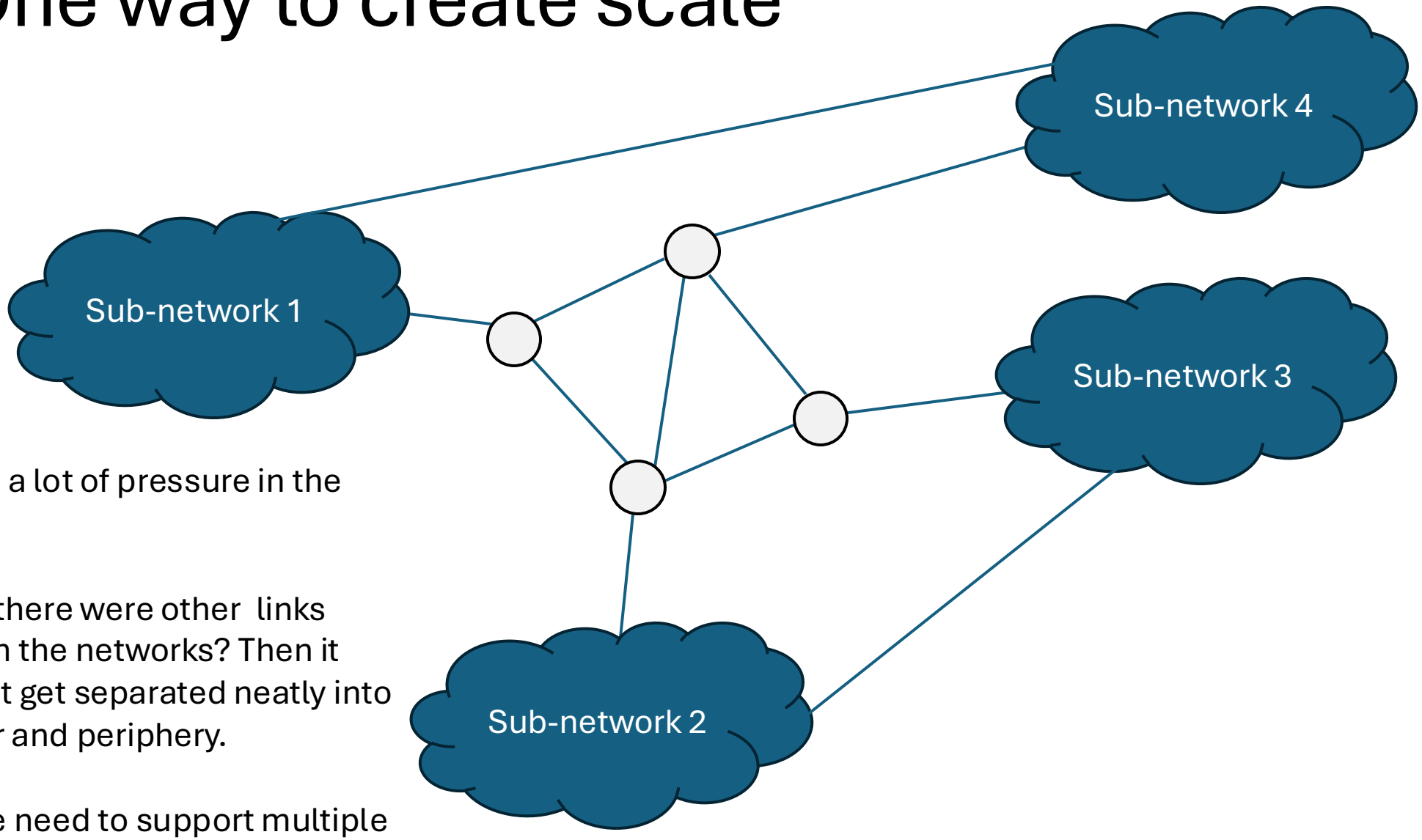
- Coordination between different entities
  - Everyone needs to speak the same protocol (BGP – Border Gateway Protocol)
  - This protocol must allow everyone to express their own priorities and policies, which can be quite complex
    - Ideally, it should keep these policies secret, which is important for commercial negotiations
  - The protocol must be distributed, so that the network can be self-repairing and to avoid having a global internet dictator
    - Q: Have we discussed any internet dictators yet?

One way to create scale



What do you think? Is this a good design?

One way to create scale

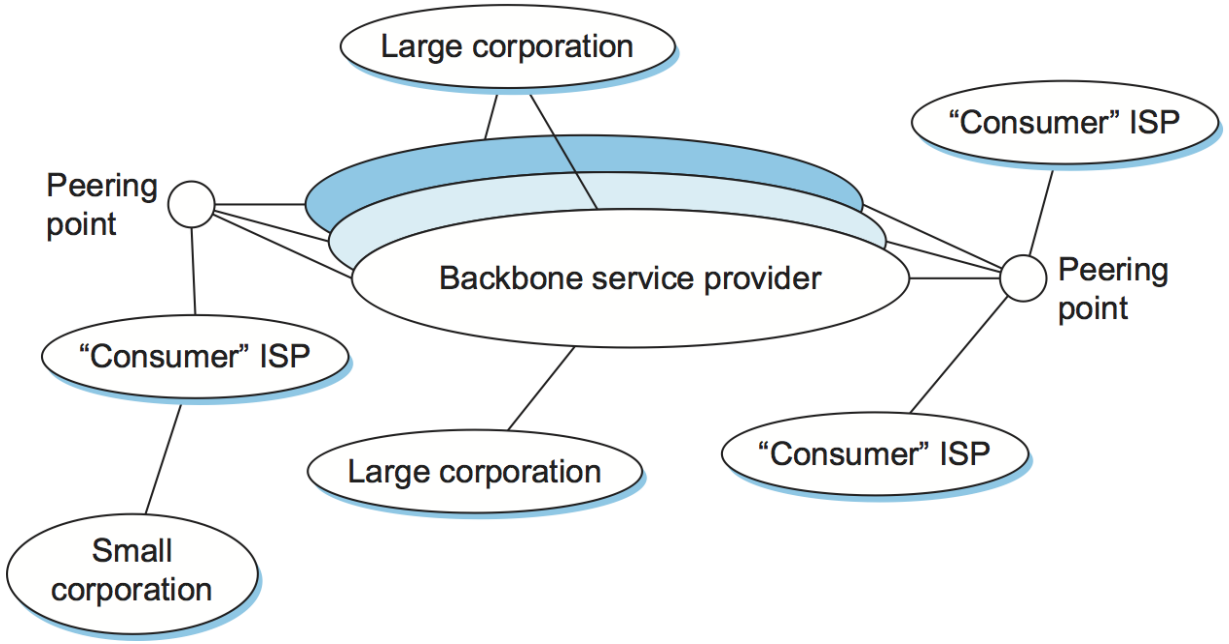


Creates a lot of pressure in the center.

What if there were other links between the networks? Then it does not get separated neatly into a center and periphery.

Plus, we need to support multiple border routers

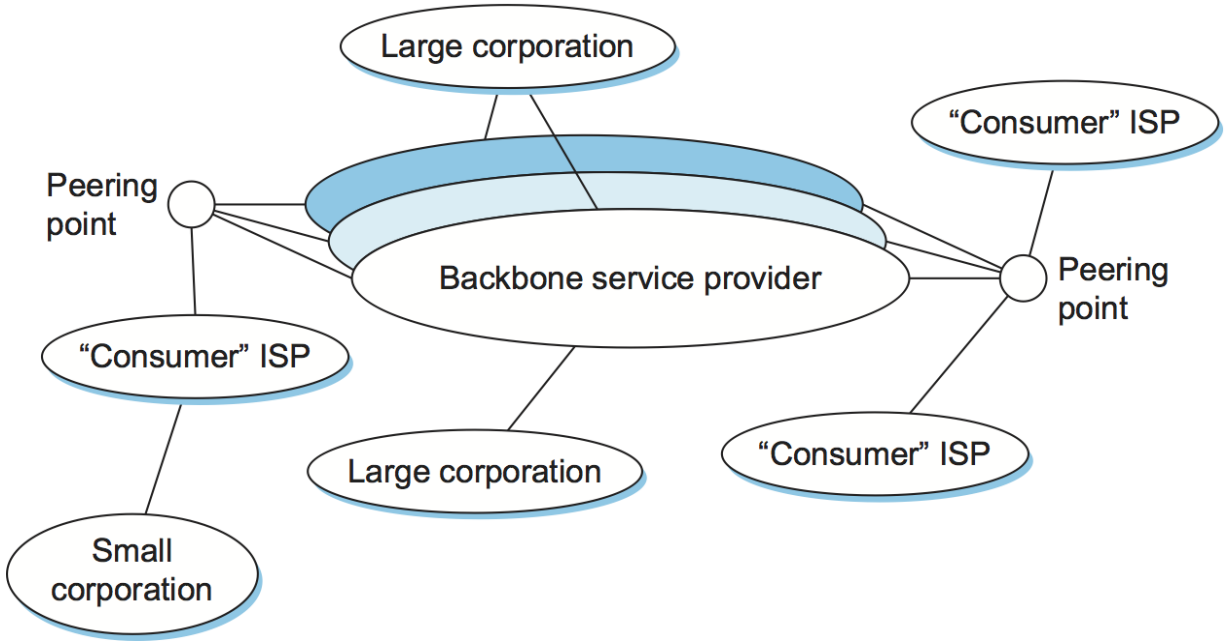
Inter-Domain routing



Each "Autonomous System (AS)", represented by ellipses here, can do whatever they want internally. They can be connected in complex topologies*. As long as they all speak BGP, the internet works*

*terms and conditions apply

Inter-Domain routing



Each AS is given a globally unique number. There are about 115,000 ASes today globally.

A partial list is available here: https://ftp.ripe.net/ripe/asnames/asn.txt
UT's AS number is 18

Example of a policy an AS might want to enforce

- Whenever possible, I prefer to send traffic via AS X than via AS Y. I'll use AS Y if it is the only path
- I never want to carry traffic from AS X to AS Y or *vice versa*

- *Scenario:* I am paying X and Y, but I prefer X

Note: This policy flexibility has a cost; BGP is not guaranteed to converge, as we shall see later

BGP scales to the entire internet and deals with the myriad commercial agreements between different entities

Its specification is relatively simple, but the behavior is anything but simple

If there is a large internet outage, chances are, BGP was involved

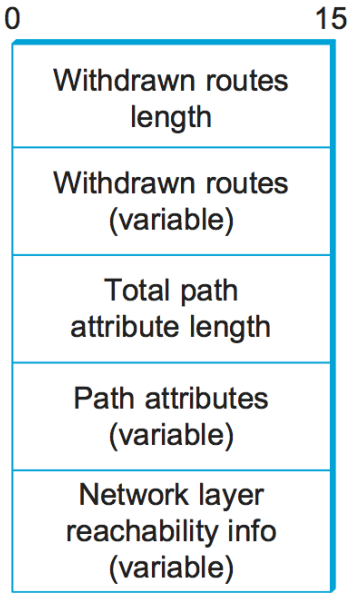
"It is always BGP!"

BGP Principle: Speakers speak for ASes

- Each AS has at least one "speaker" that advertises BGP routes
- Border routers are also speakers, but they do not have to be
- Speakers communicate via TCP, which ensures reliable delivery

BGP Principle: Path vector routing

- Advertisements are of the form "If you send packets to me, I can send them to 82.128.0.0/9 through the path AS4123, AS1045, AS9"
- Advertisements can be withdrawn explicitly
- Since the entire path is included, it is easy to avoid loops
- Advertisements include preferences on which entry-point is prefered for a given destination



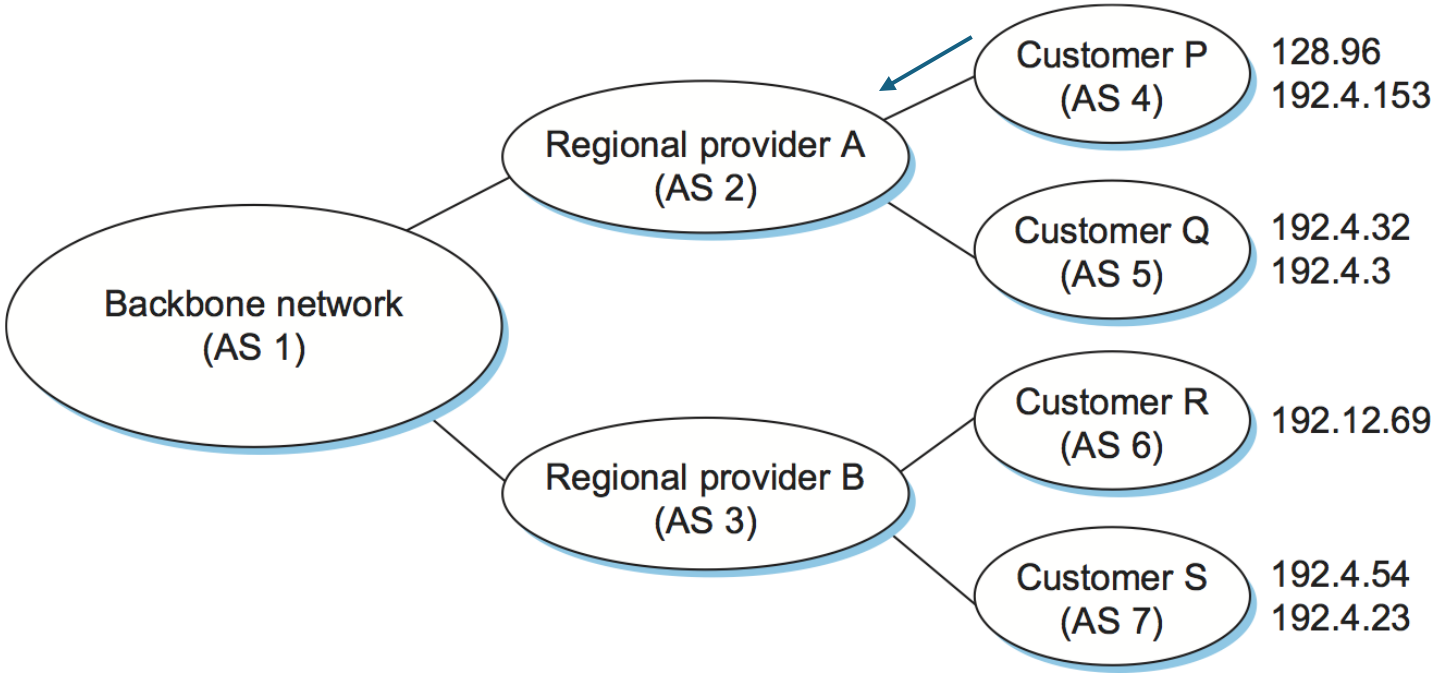
BGP update packet format

BGP principle: Everybody lies

- ASes can choose which routes they advertise to whom
  - For example, if I am paying both ATT and Comcast to be my ISP, I do not want to let them use me to send packets between each other. Thus, I will not send ATT an advertisement for routes to Comcast and vice versa
- ASes are free to disbelieve certain routes from some other ASes
  - Misconfigurations happen. If an AS says they can route to everybody in the world, and they are not a tier 1 ISP, I will not believe them
  - Real example: an (unnamed) country that wanted to block YouTube advertised routes to YouTube with the intention of dropping all of those packets. Unfortunately, the advertisement went everywhere in the world, and caused a global YouTube outage. They only wanted to cause an outage within the country.
- If someone sends me packets to a destination I did not advertise, I am free to drop the packets.
  - Note: the internet operates on "best effort", so I can always drop whatever packets I want. It is just "not nice" to advertise a route and refuse to honor it
- Unlike intra-domain routing protocols, BGP does not try to be "plug-and-play". Configuring policies is complicated

Let us work through an example

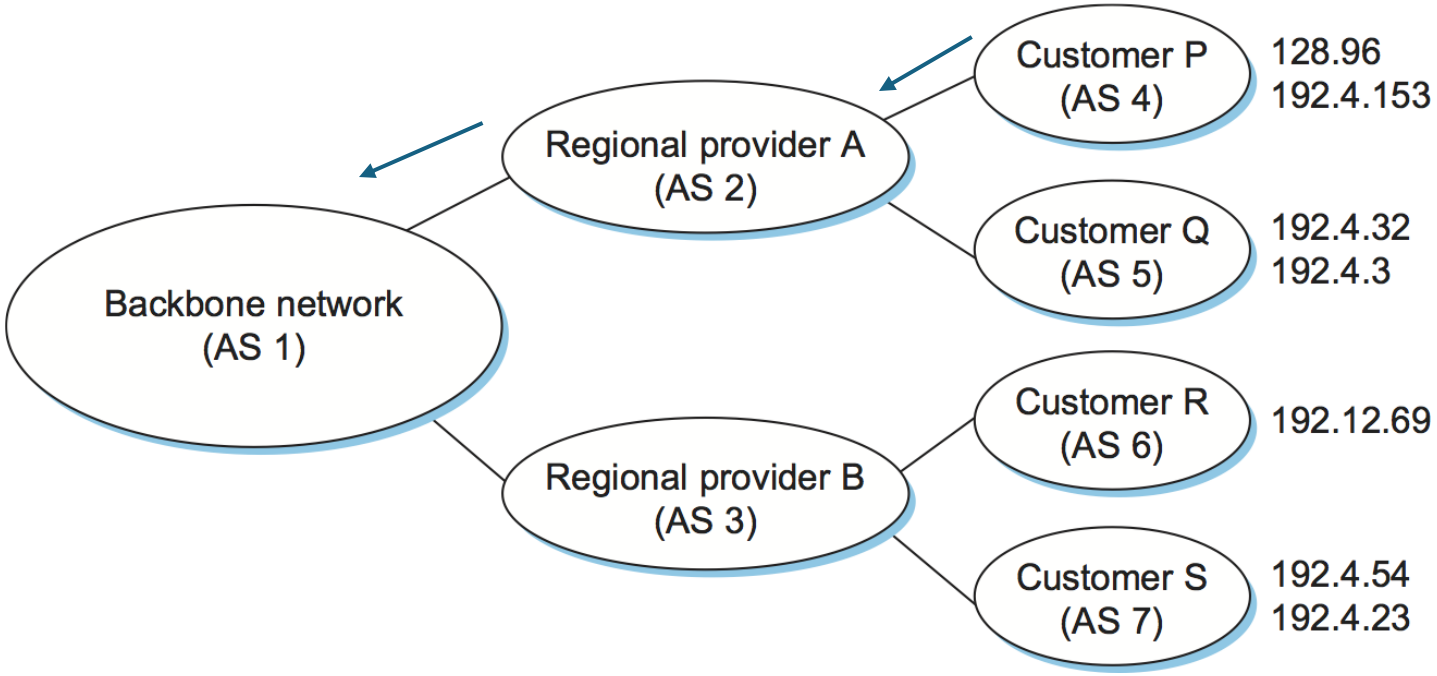
"I can reach 128.96.*.* through (AS4)"
"I can reach 192.4.153 through (AS4)"



Let us work through an example

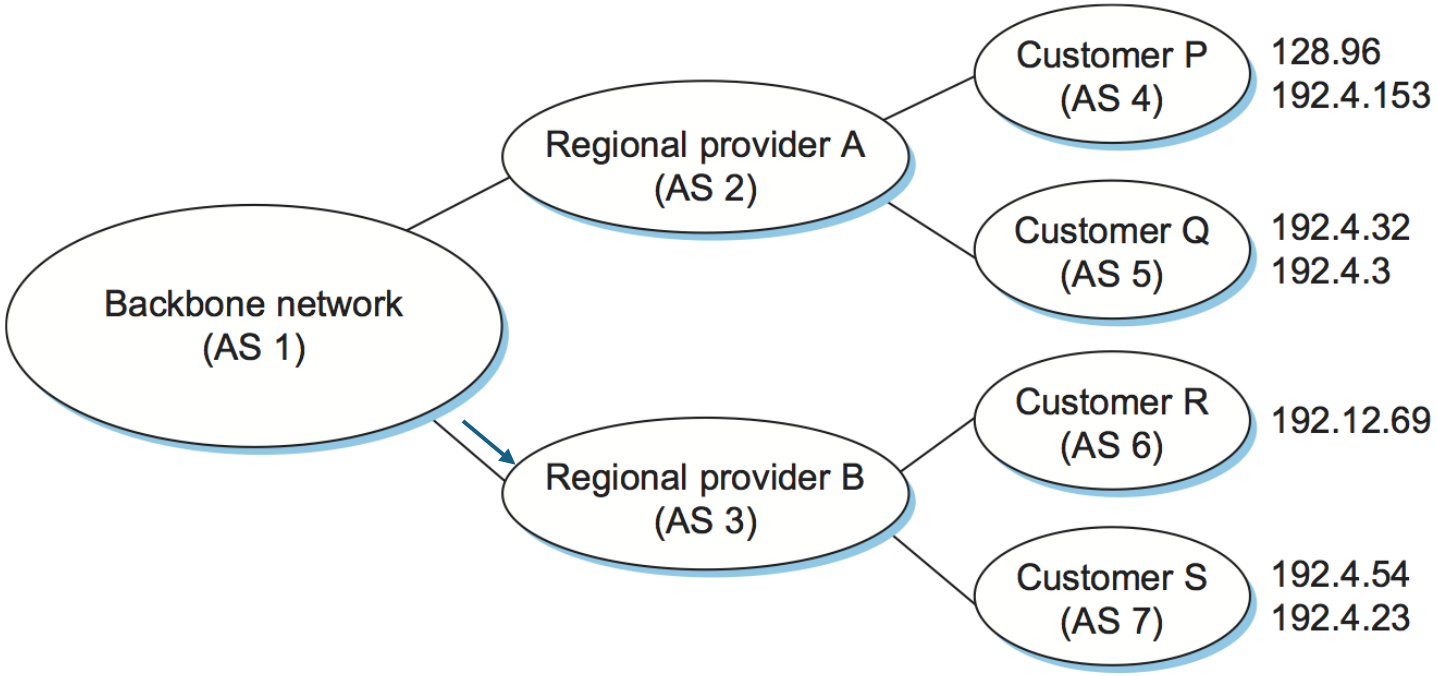
"I can reach 128.96.*.* through (AS2, AS4)"
"I can reach 192.4.153 through (AS2, AS4)"
"I can reach 192.4.32.* through (AS2, AS5)"
"I can reach 182.4.3 through (AS2, AS5)

"I can reach 128.96.*.* through (AS4)"
"I can reach 192.4.153 through (AS4)"



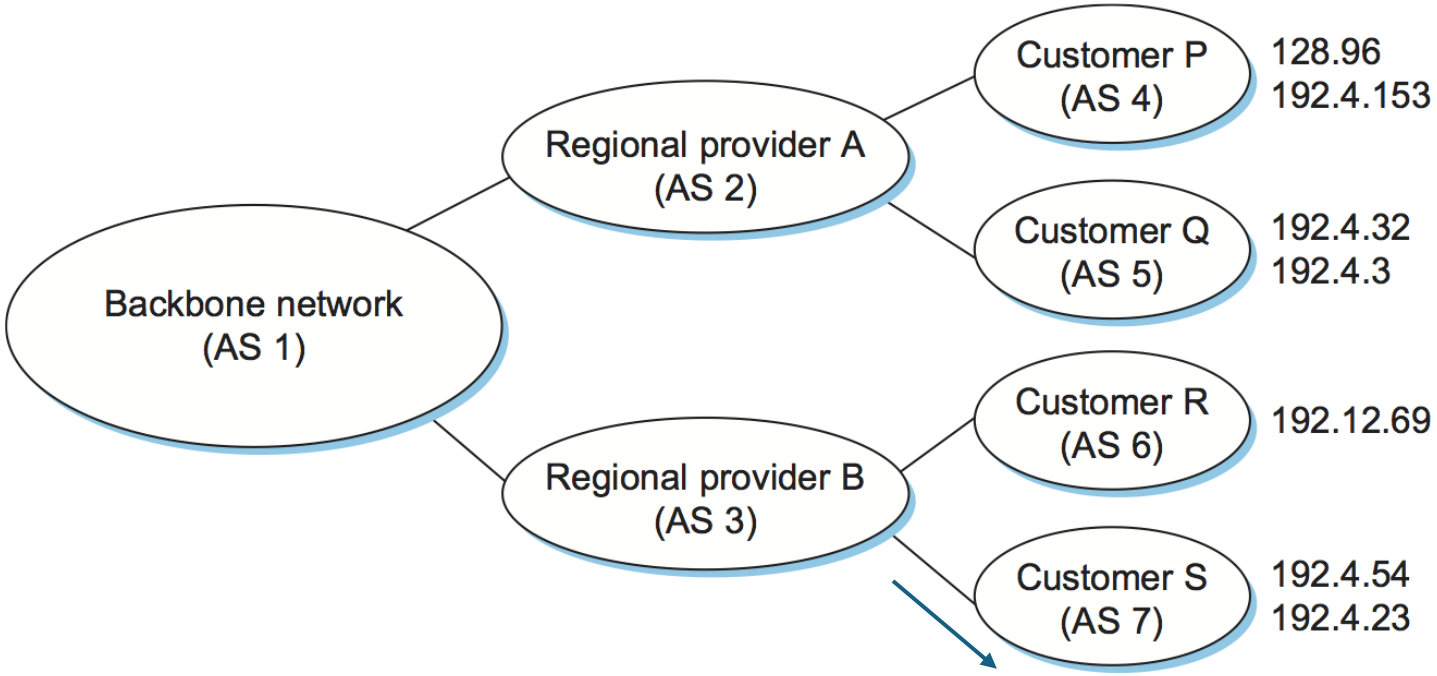
Let us work through an example

"I can reach 128.96.*.* through (AS1, AS2, AS4)"
"I can reach 192.4.153 through (AS1, AS2, AS4)"
"I can reach 192.4.32.* through (AS1, AS2, AS5)"
"I can reach 182.4.3 through (AS1, AS2, AS5)



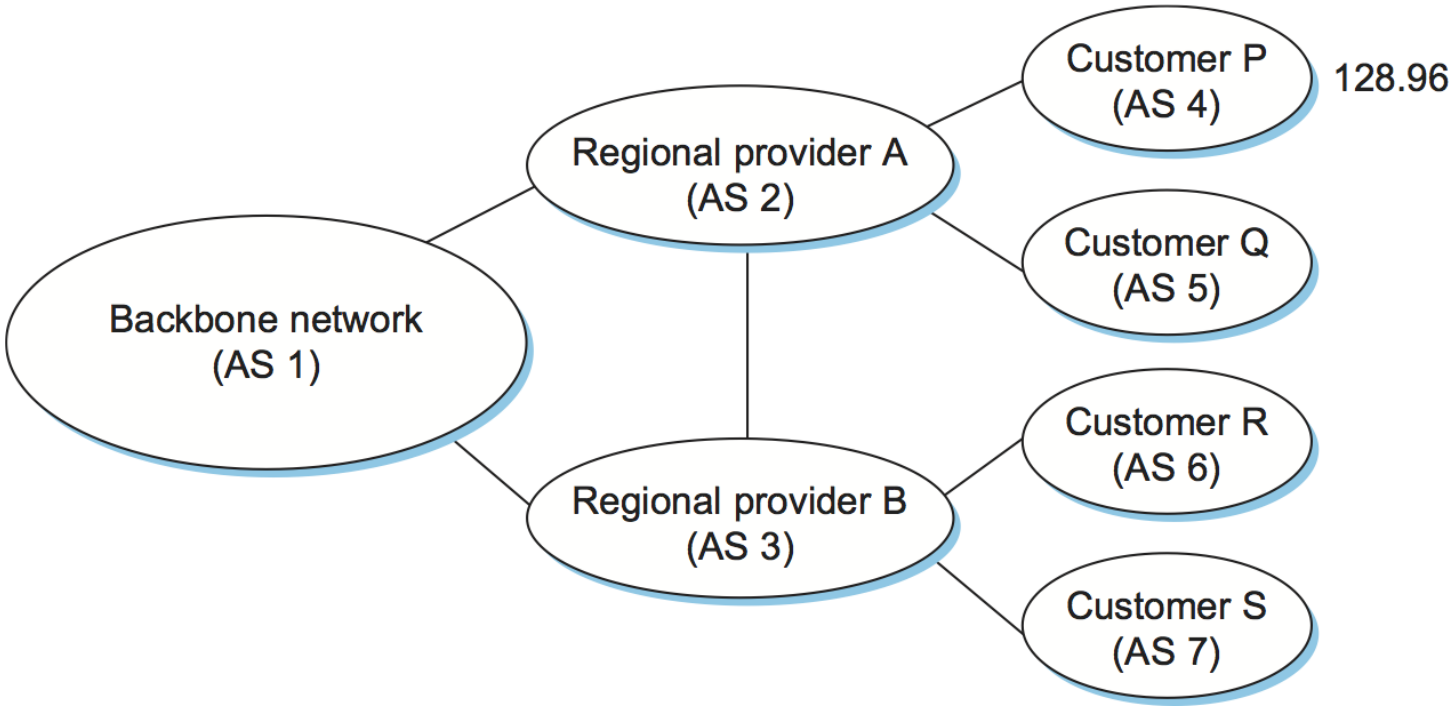
Let us work through an example

"I can reach 128.96.*.* through (AS3, AS1, AS2, AS4)"
"I can reach 192.4.153 through (AS3, AS1, AS2, AS4)"
"I can reach 192.4.32.* through (AS3, AS1, AS2, AS5)"
"I can reach 182.4.3 through (AS3, AS1, AS2, AS5)"
"I can reach 192.12.69.* through (AS3, AS6)"

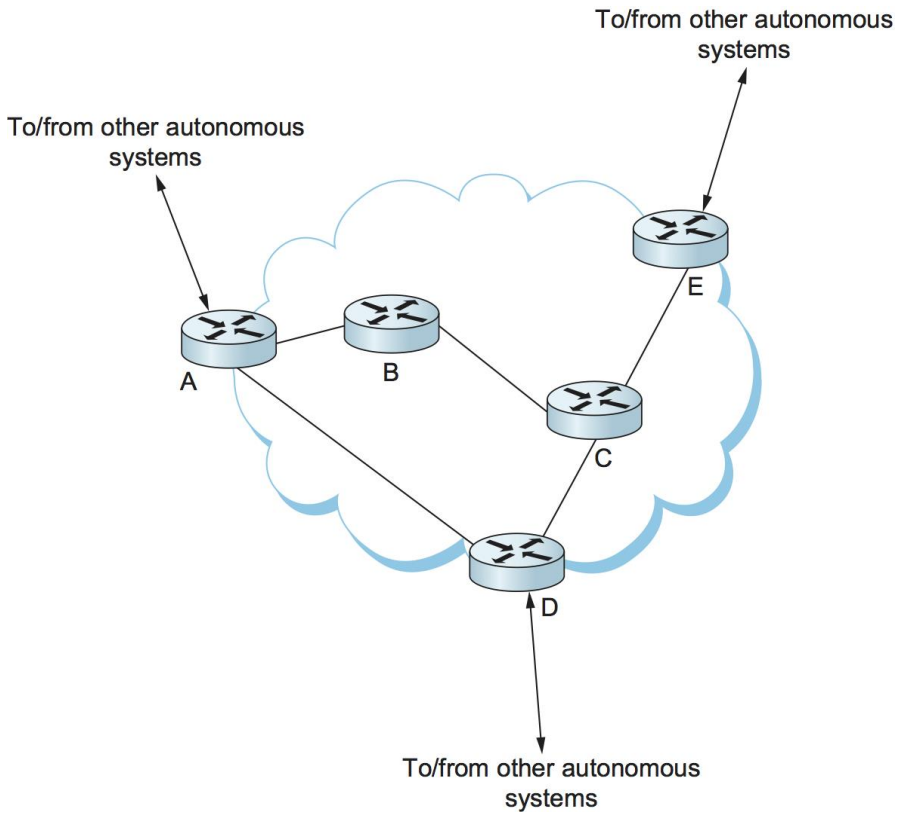


This way, AS7 can reach everybody in the network!

Let us work through an example: loop avoidance



Connecting intra- and inter- domain routing



Prefix	BGP Next Hop
18.0/16	E
12.5.5/24	A
128.34/16	D
128.69./16	A

BGP table for the AS

Router	IGP Path
A	A
C	C
D	C
E	C

IGP table for router B

Prefix	IGP Path
18.0/16	C
12.5.5/24	A
128.34/16	C
128.69./16	A

Combined table for router B

Next lecture: Advanced BGP topics

- Stability and route-flapping
- Load balancing across different ingress points
- Examples of BGP failure