

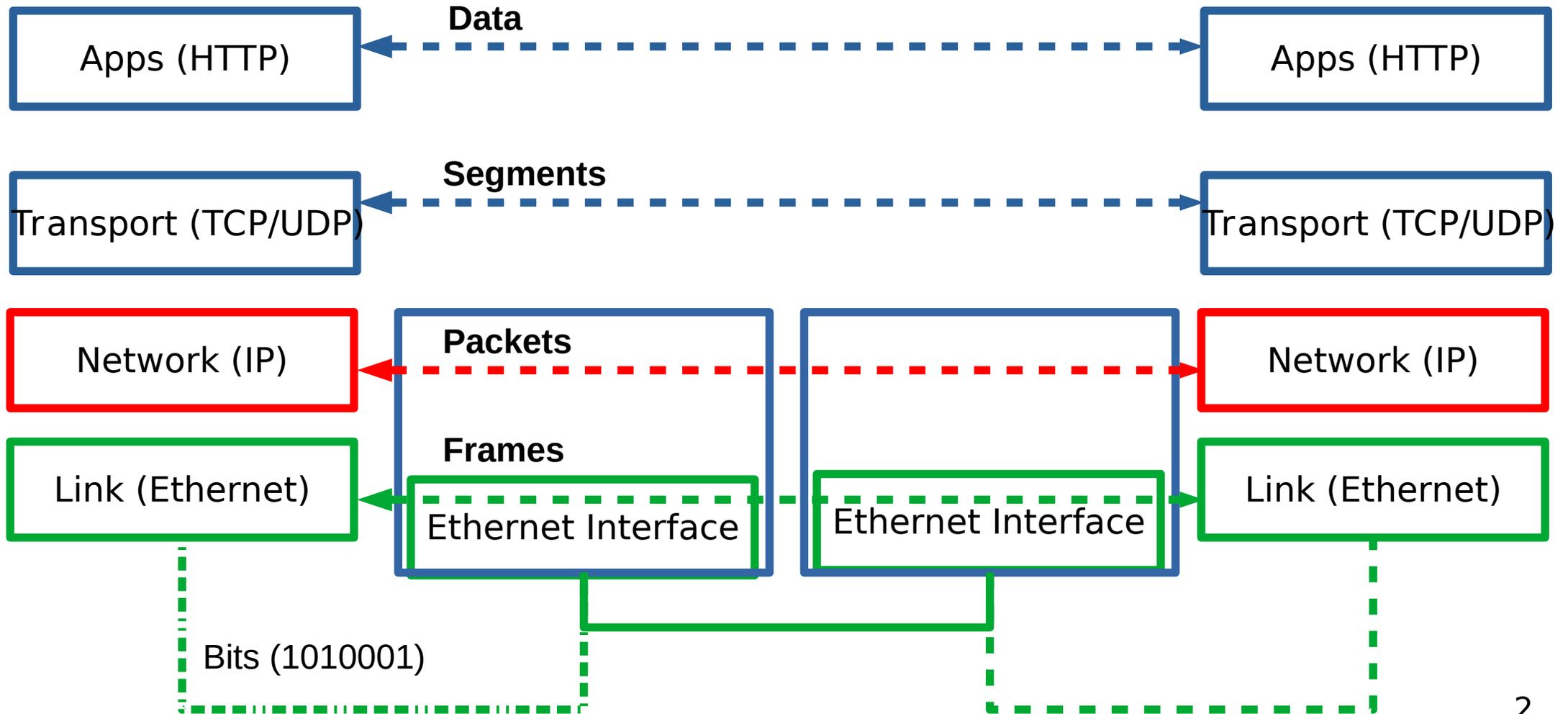
CSC4200/5200 – COMPUTER NETWORKING

Instructor: Susmit Shannigrahi

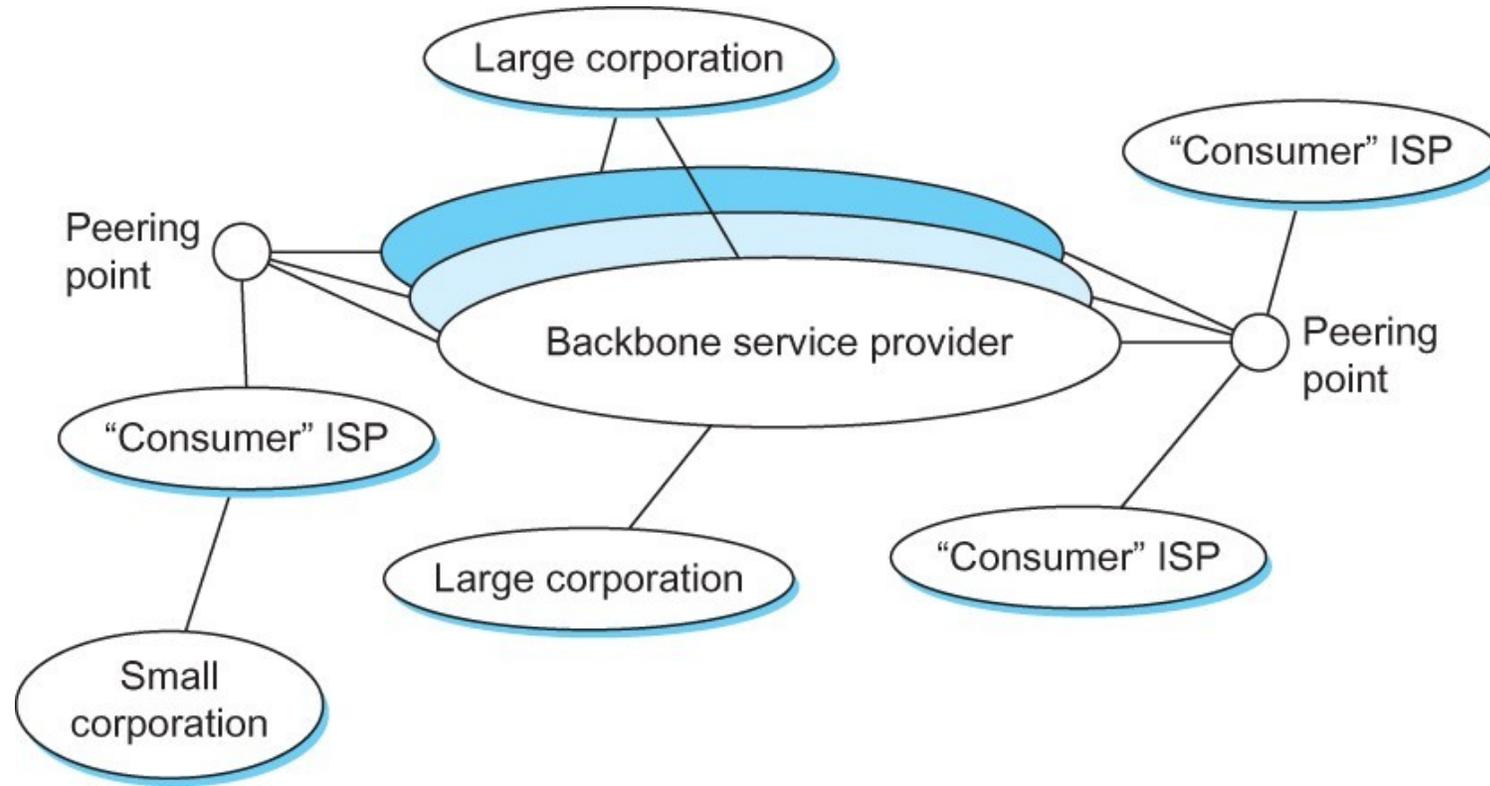
BGP - CONTINUED

sshannigrahi@tntech.edu



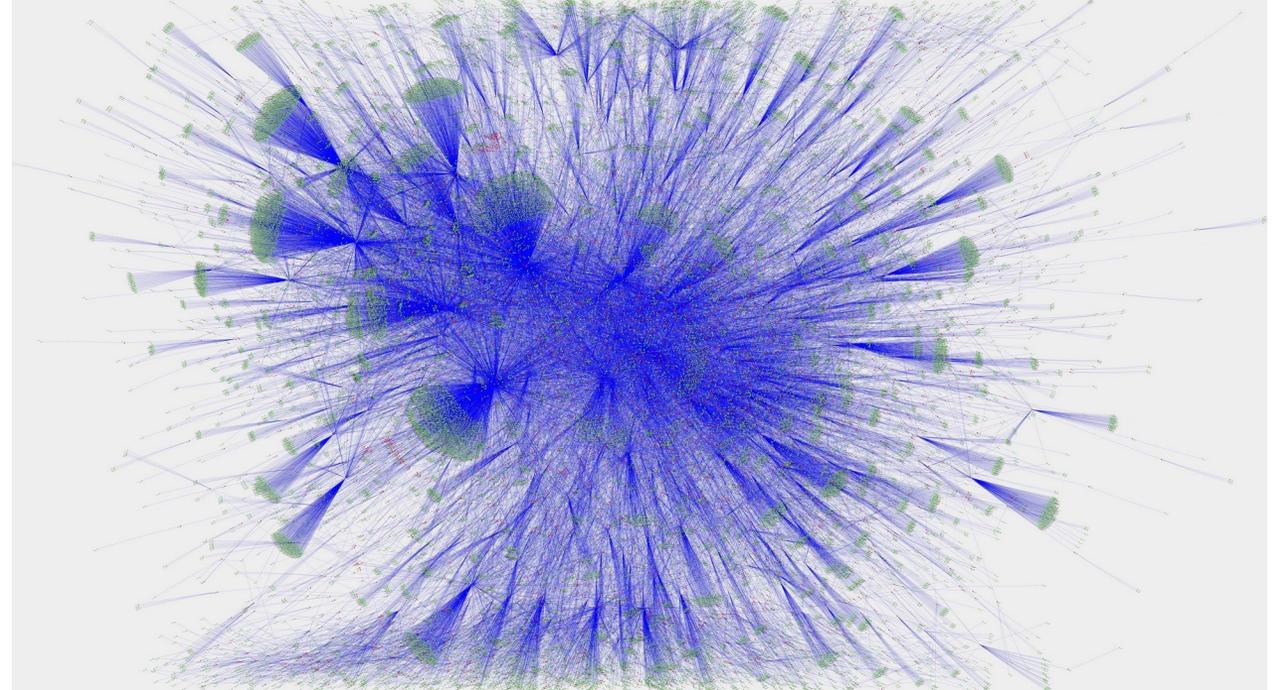


Internet now



So far...

- How do we scale routing?
 - BGP
 - Only connectivity, not optimality



Hierarchical routing - Policy

scale: with 600 million destinations:

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

administrative autonomy

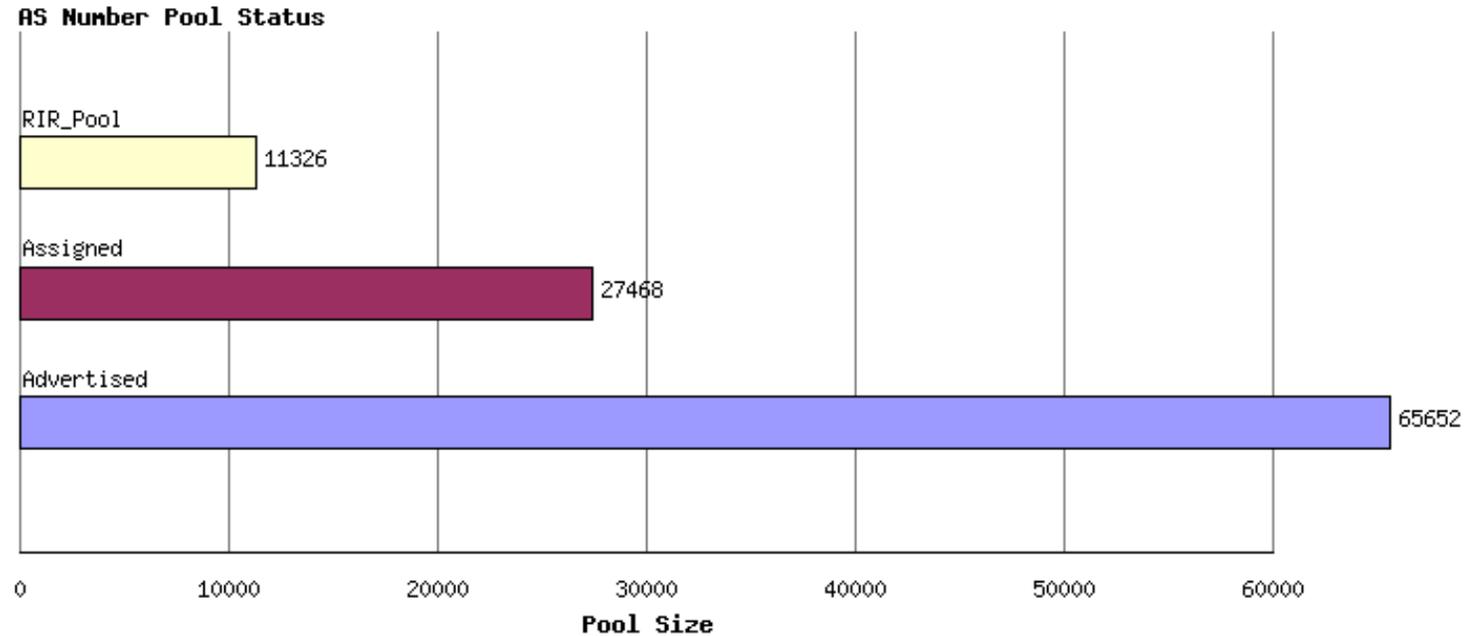
- internet = network of networks
- each network admin may want to control routing in its own network

Autonomous systems (ASes)

- AS
 - A set of routers under a single technical administration
 - Uses IGP within the AS to route packets
 - Uses BGP between Ases to route packets
- What happens inside an AS stays within that AS!
 - That is, AS decides routing metrics internally

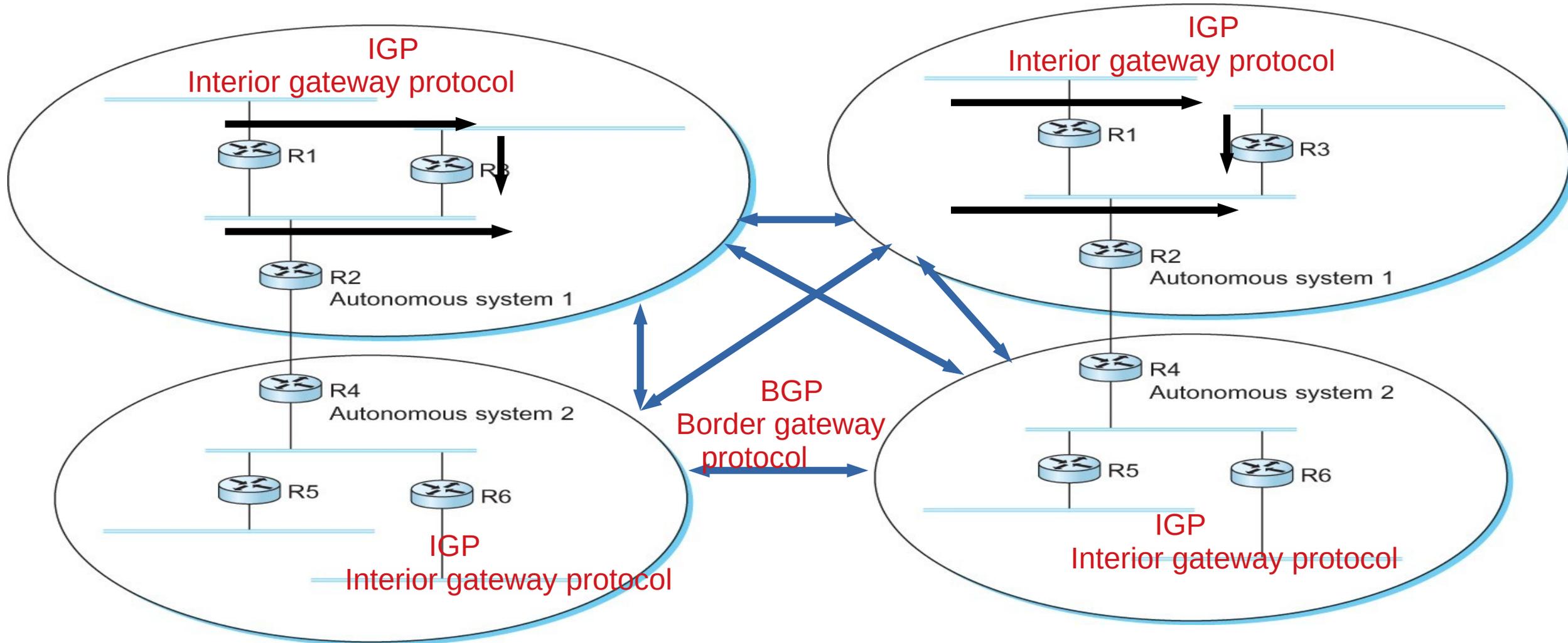
Status of ASNs

Status	AS Pool	16-bit	32-bit
IETF Reserved	95033874	1042	95032832
IANA Unallocated Pool	4199828976	0	4199828976
Allocated	104446	64494	39952
RIR Data			
AFRINIC	2302	1278	1024
APNIC	19093	8539	10554
ARIN	31567	25522	6045
RIPE NCC	39453	25729	13724
LACNIC	12031	3426	8605



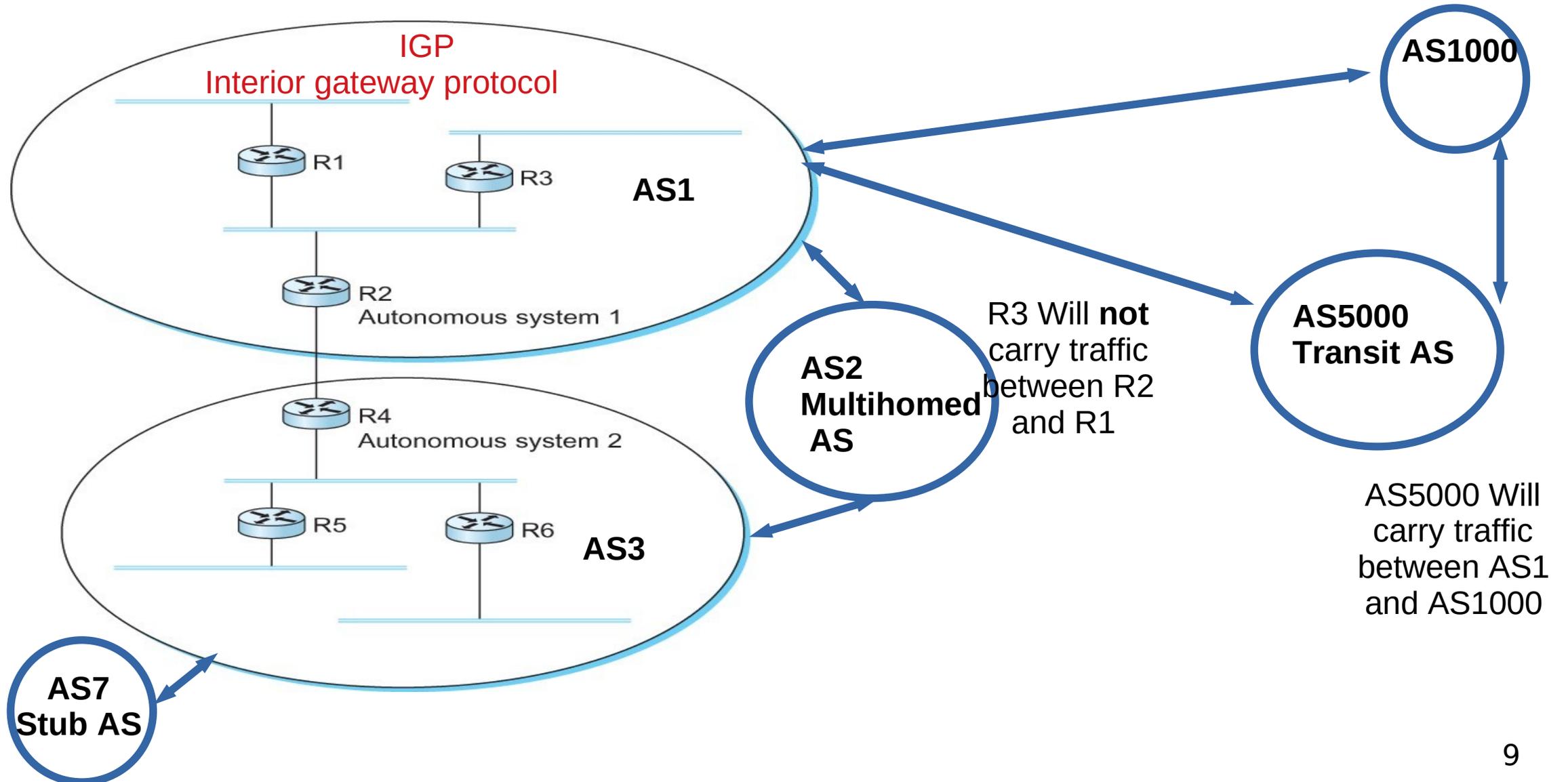
<http://www.potaroo.net/tools/asn32/>

Interdomain Routing



A network with four autonomous systems

BGP-4: Border Gateway Protocol



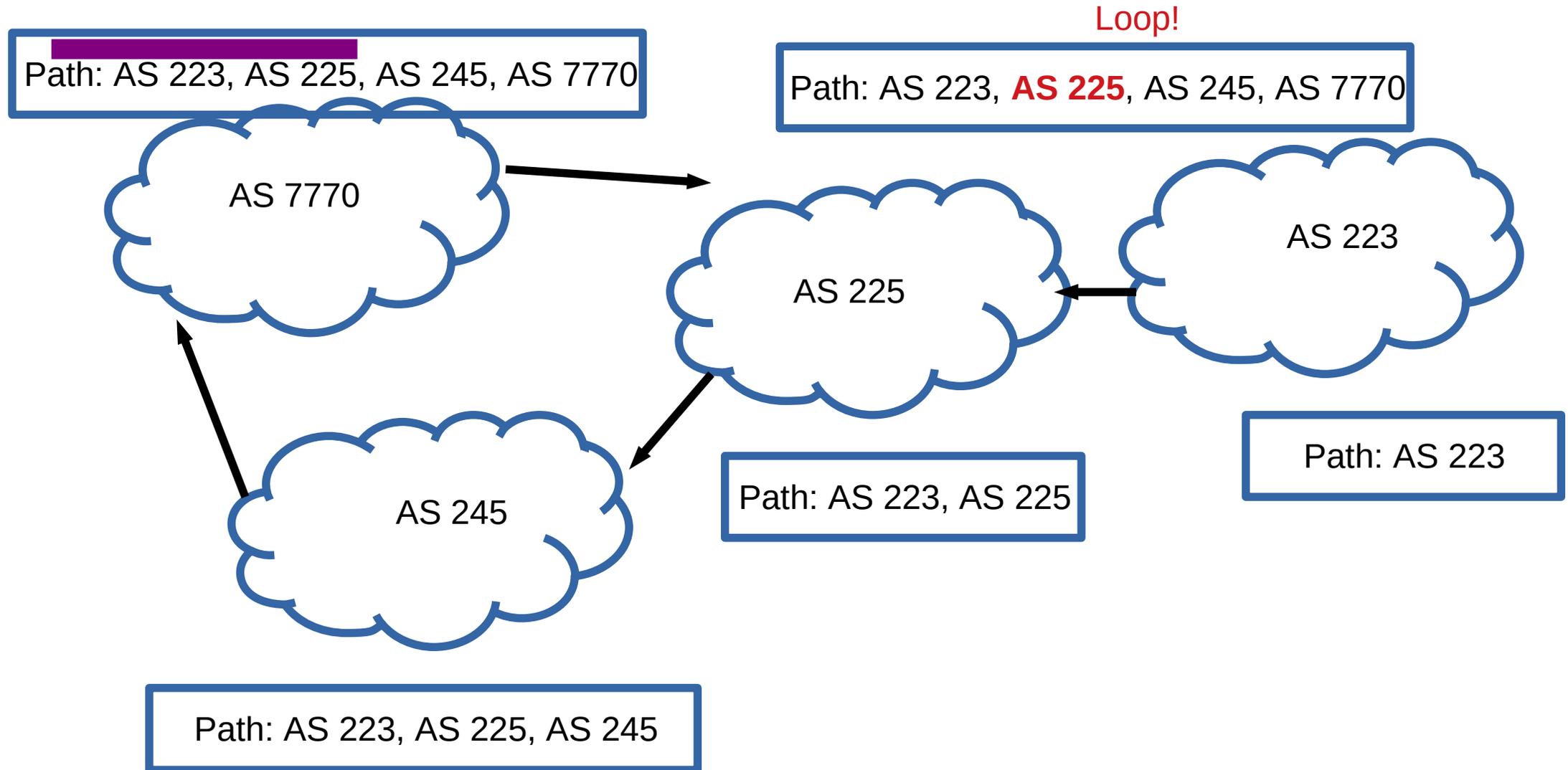
BGP - goals

- The goal of Inter-domain routing is to find **any path** to the intended destination that is **loop free**
 - **We are concerned with reachability than optimality**
 - Finding path anywhere close to optimal is considered to be a great achievement
- Why?

BGP: Path vector protocol

- Send the whole path with the routing update
- Loops are detected if an AS finds itself in the path
 - Reject if so
 - Accept otherwise
- Add self to the path and advertise to the neighbors
- Advantage: No loops, Local decision before advertising

BGP: Path vector protocol



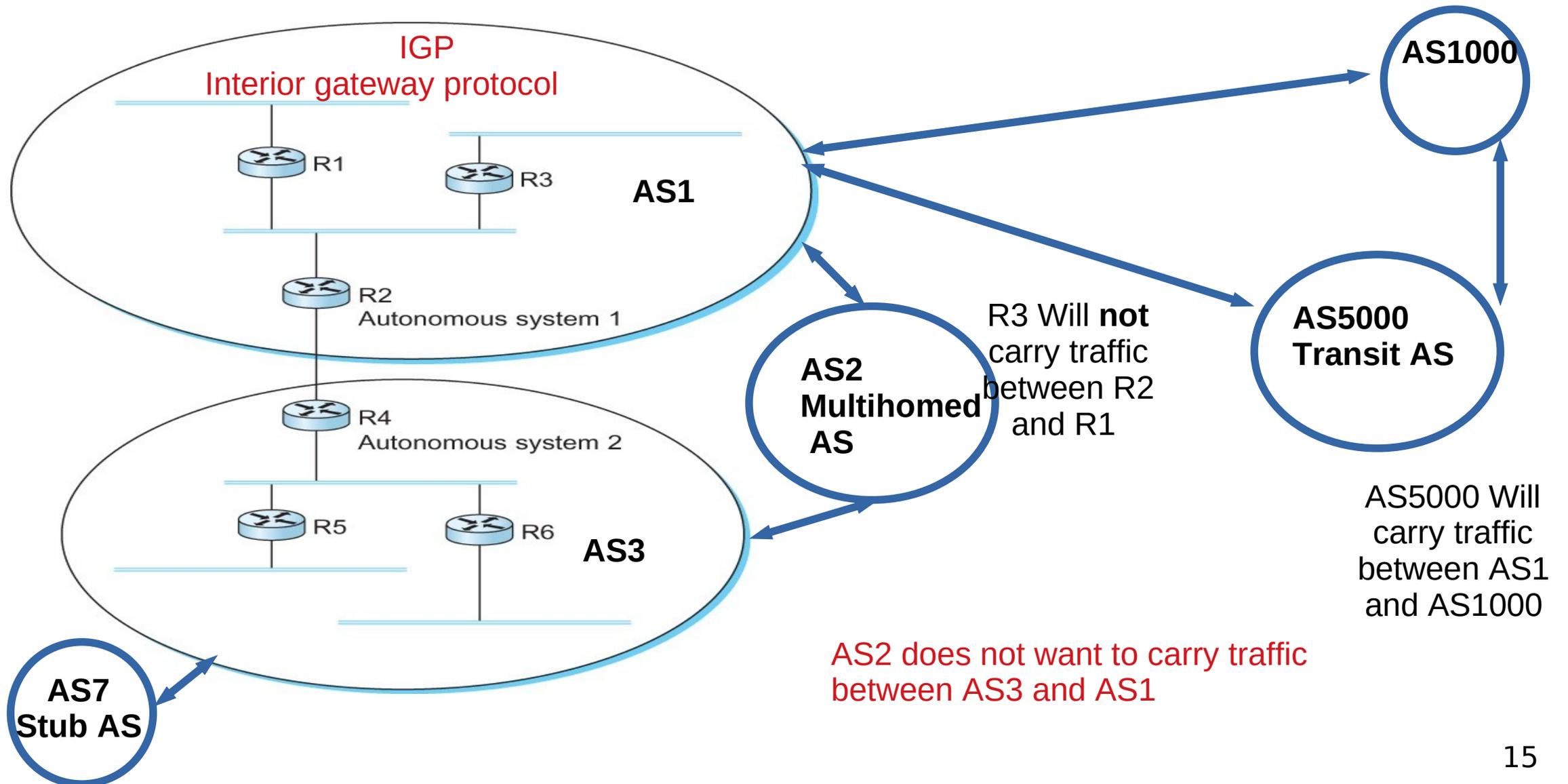
BGP: Allows for policy

- Capable of enforcing various policies
 - AS2 → Don't use AS1 to get to AS3
- Not part of BGP – configuration information that controls propagation of paths

BGP: Hop by Hop model and control what you tell your neighbors

- You can only tell others what you are using
 - But you control what you say
- BGP advertises only to peers
 - Tell them what you are using
 - Hop-by-hop model

What should AS2 (multihomed) tell AS3?



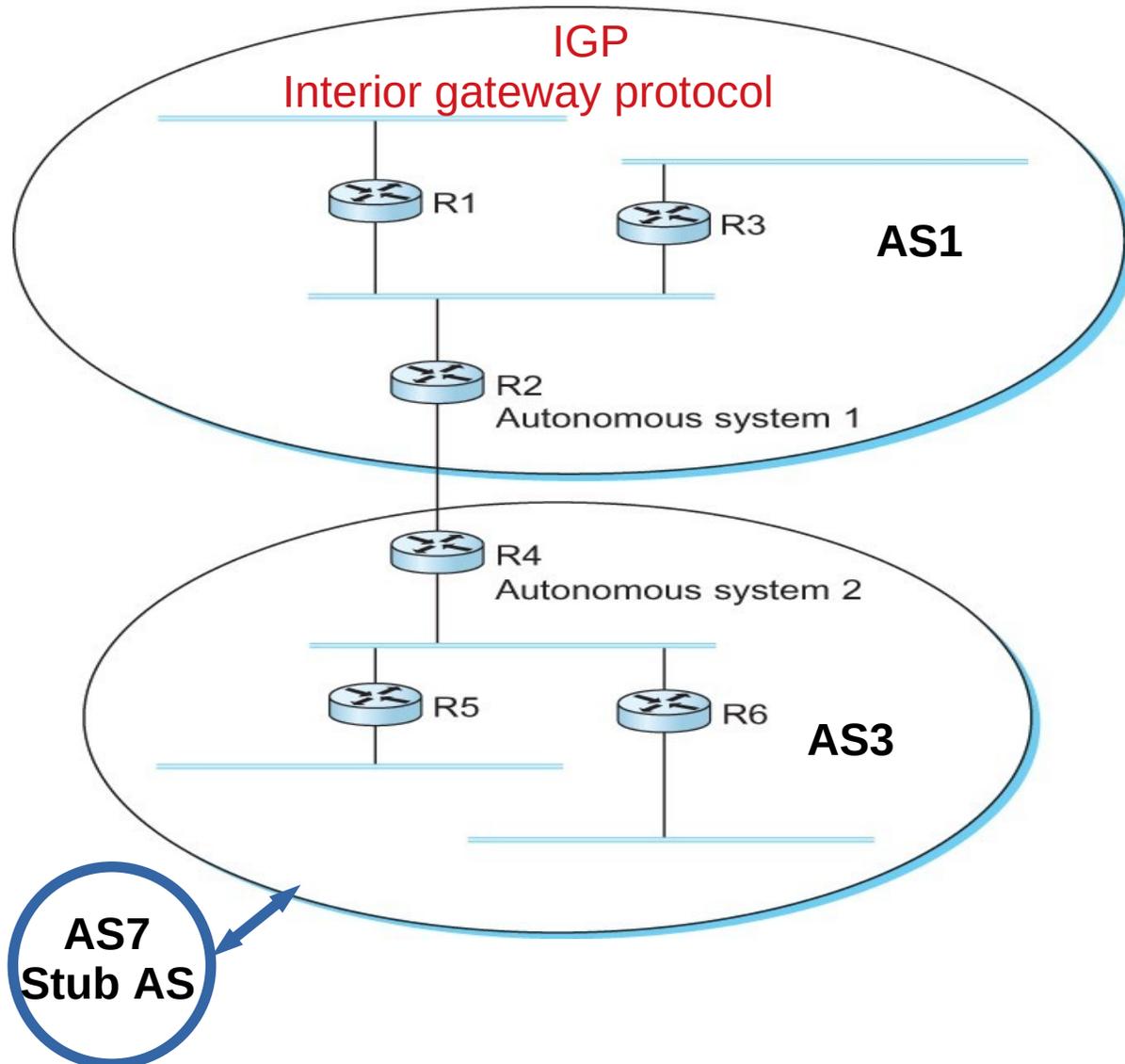
Examples BGP Policies

- Multihomed AS100 does not want to act as a transit
 - Limit advertisement
- If someone pays AS100 – yes
 - Advertise only to those who are paying
- Prefer one path over the other
 - Play with the cost, artificially increase path length and so on ← more on this later

Examples BGP Policies

- Multihomed AS100 does not want to act as a transit
 - Limit advertisement
- If someone pays AS100 – yes
 - Advertise only to those who are paying
- Prefer one path over the other
 - Play with the cost, artificially increase path length and so on ← more on this later

You don't need BGP for Stub ASes

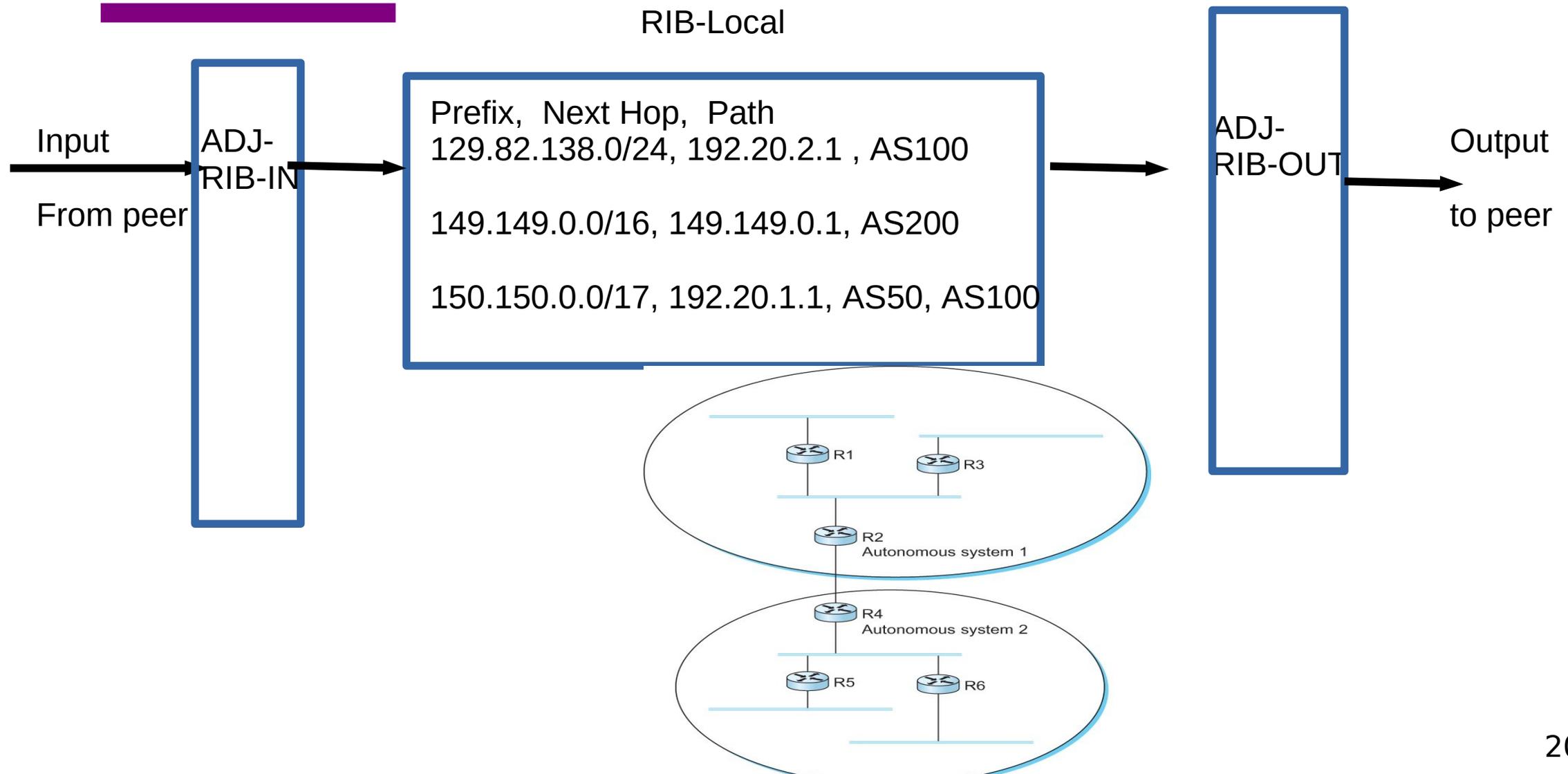


Default IP route should be sufficient

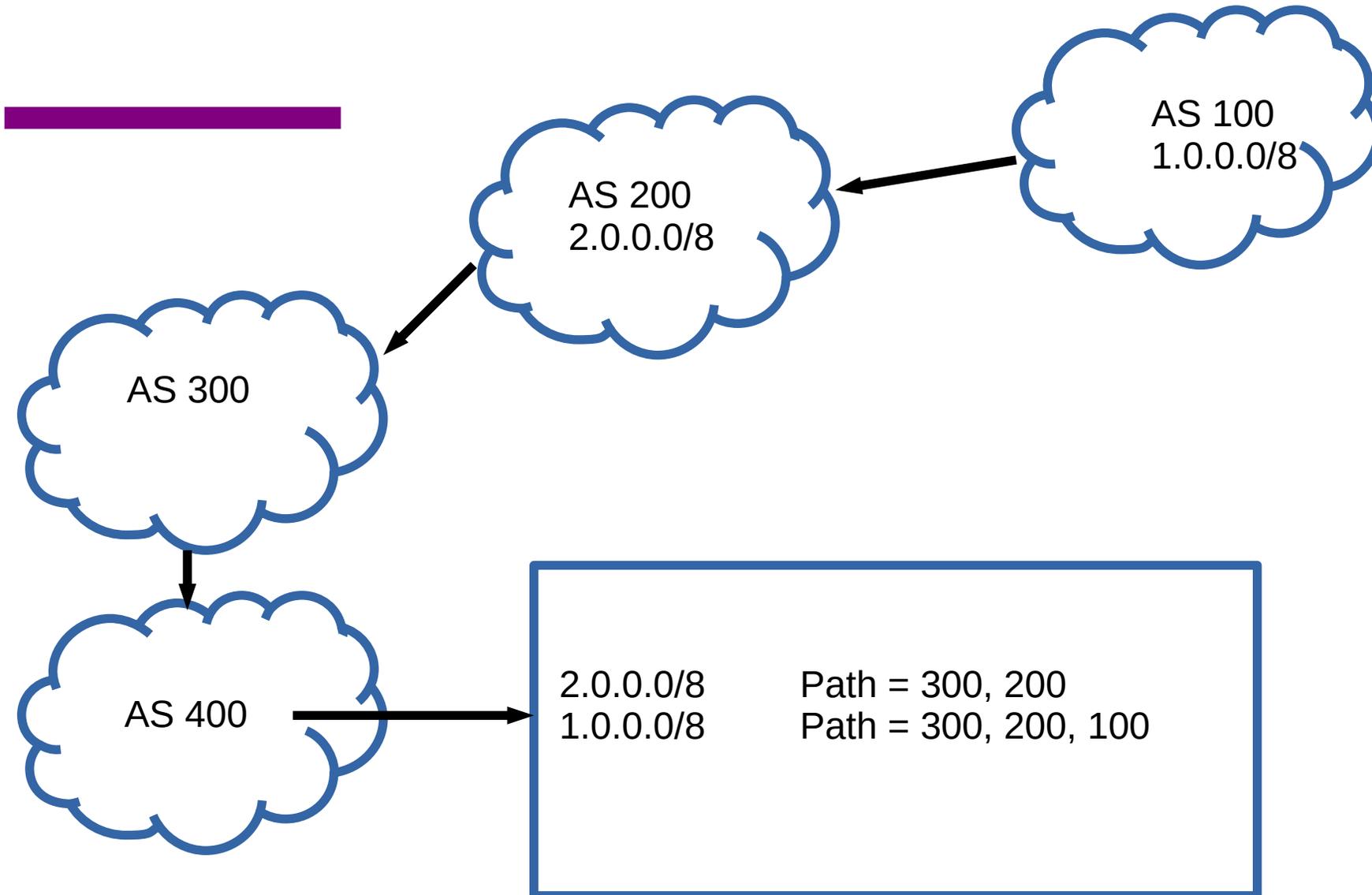
BGP Messages

- Open – Open a TCP connection to a peer
- Update – Update route attributes or withdraw
- Notification – Error notification, close connection
- Keep alive – Periodic update to peers

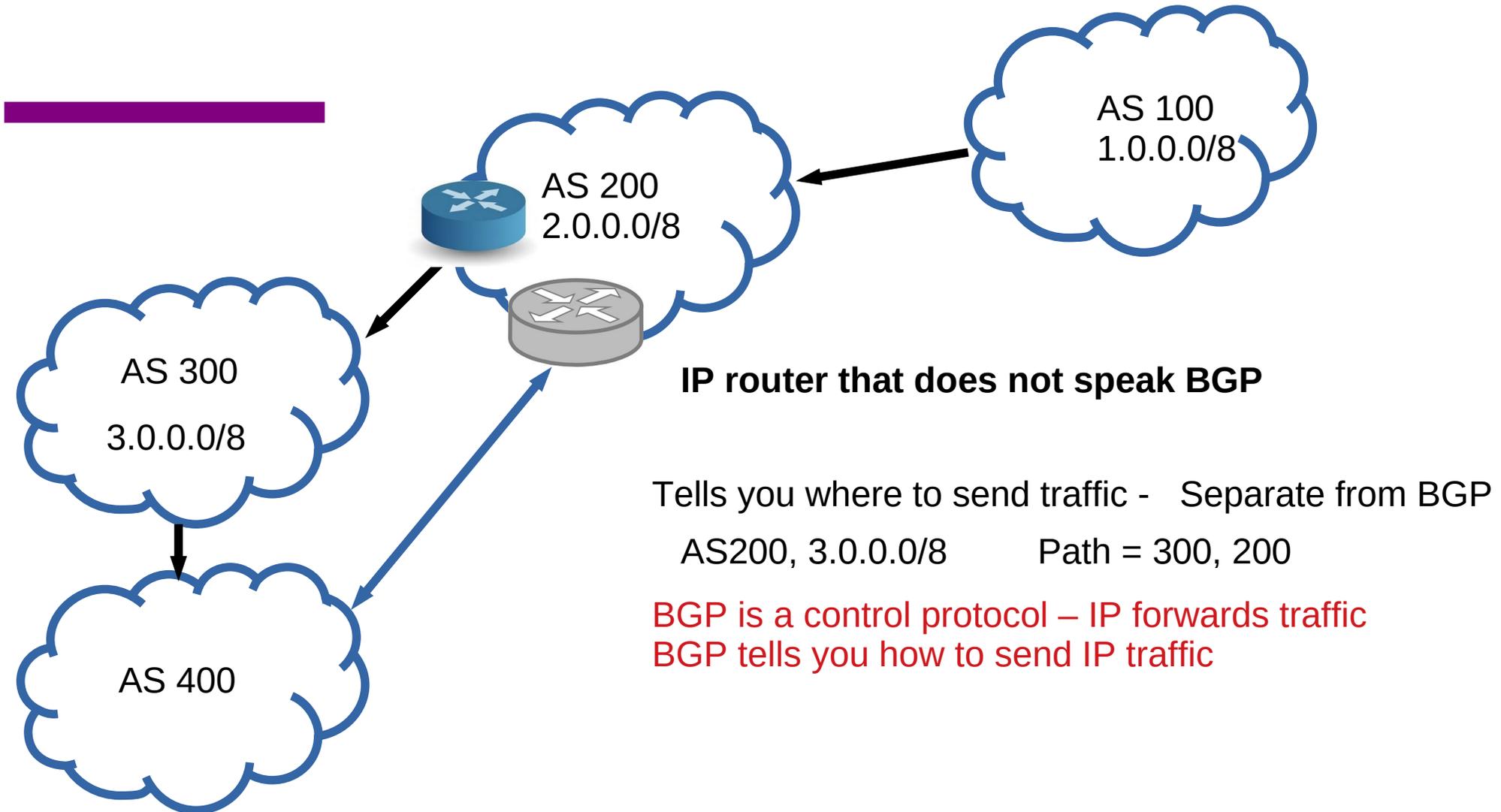
Routing Information Bases (RIB)



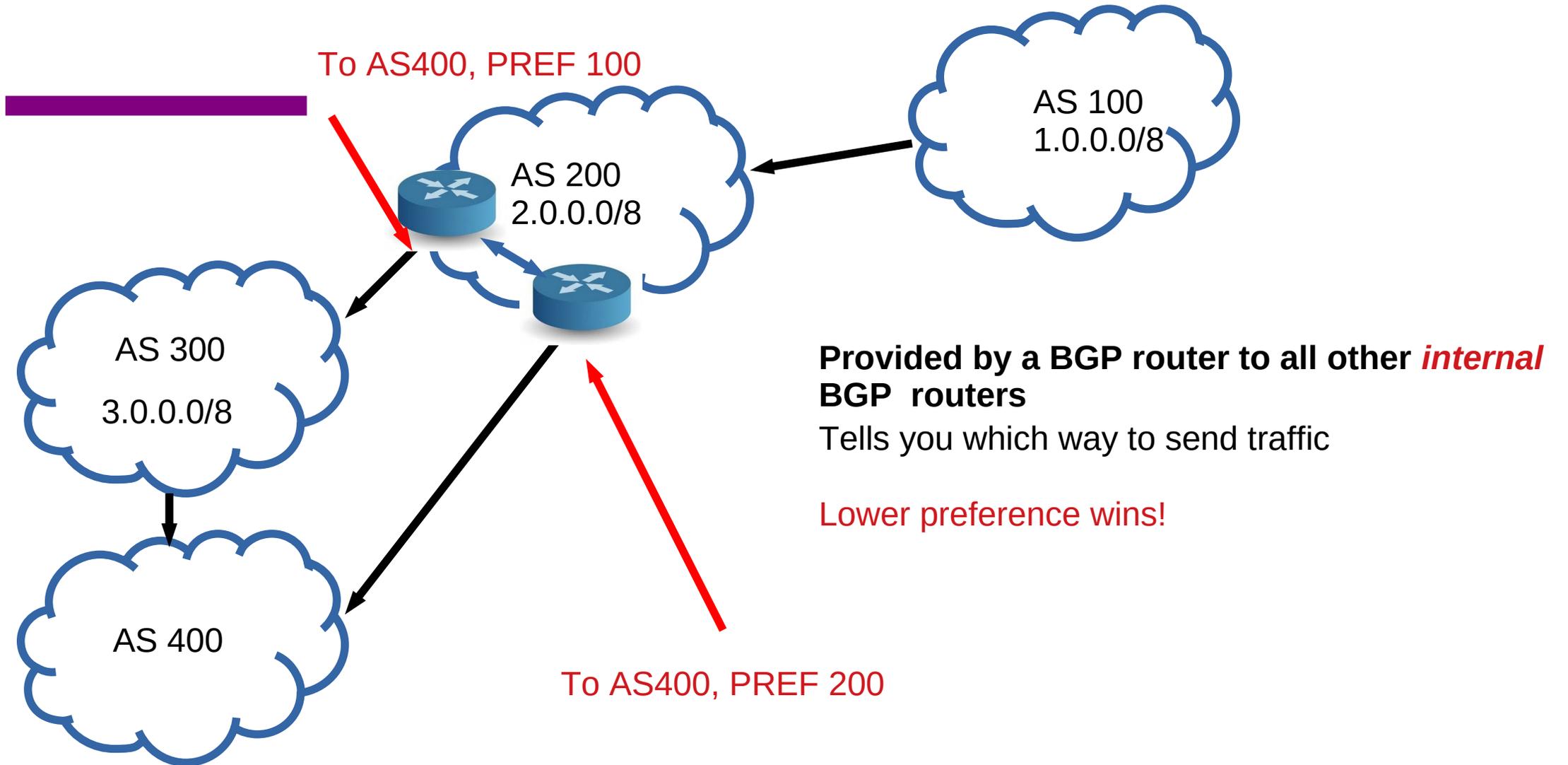
BGP Attributes - AS_PATH



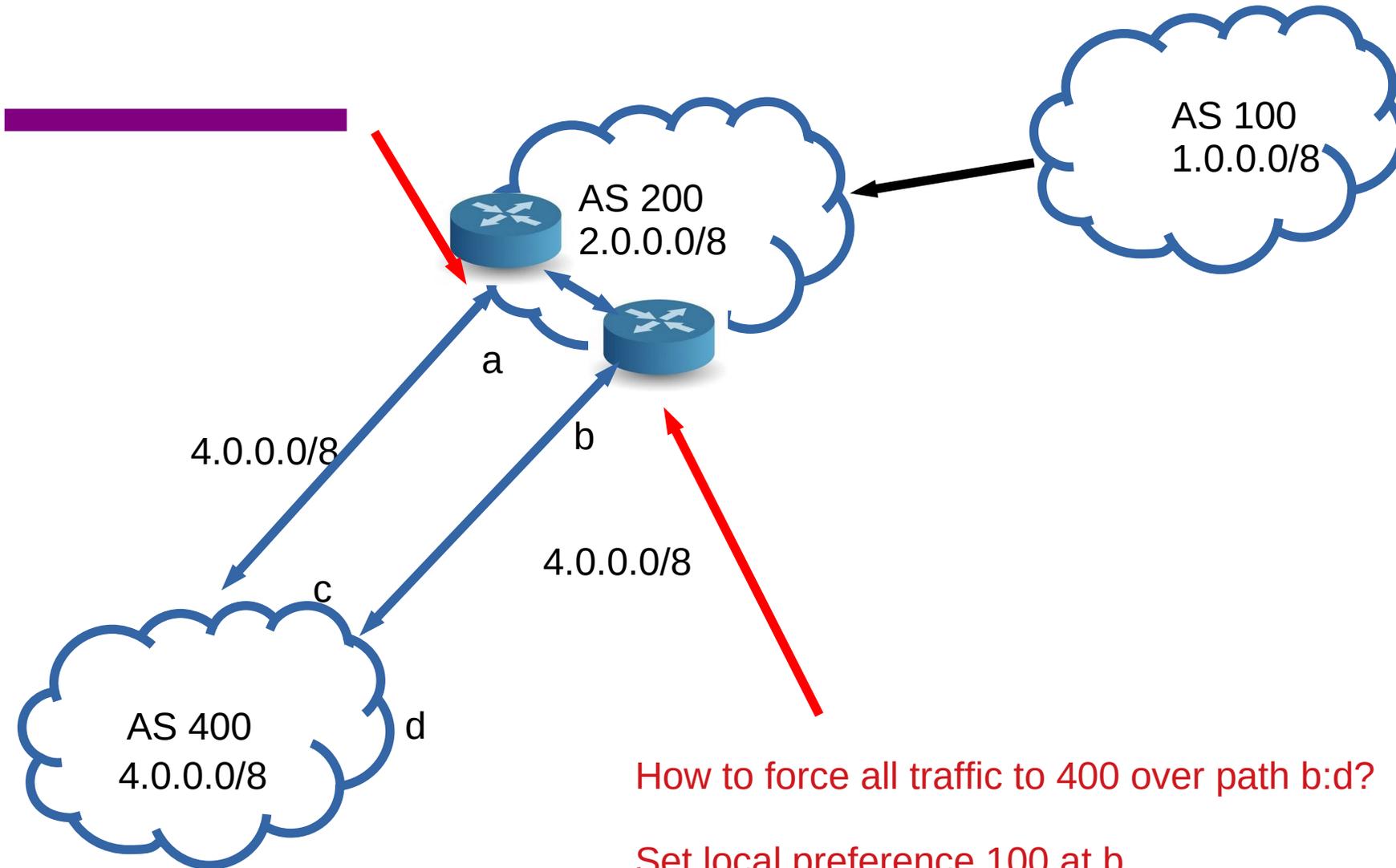
BGP Attributes - Next hop?



BGP Attributes - LOCAL-PREF



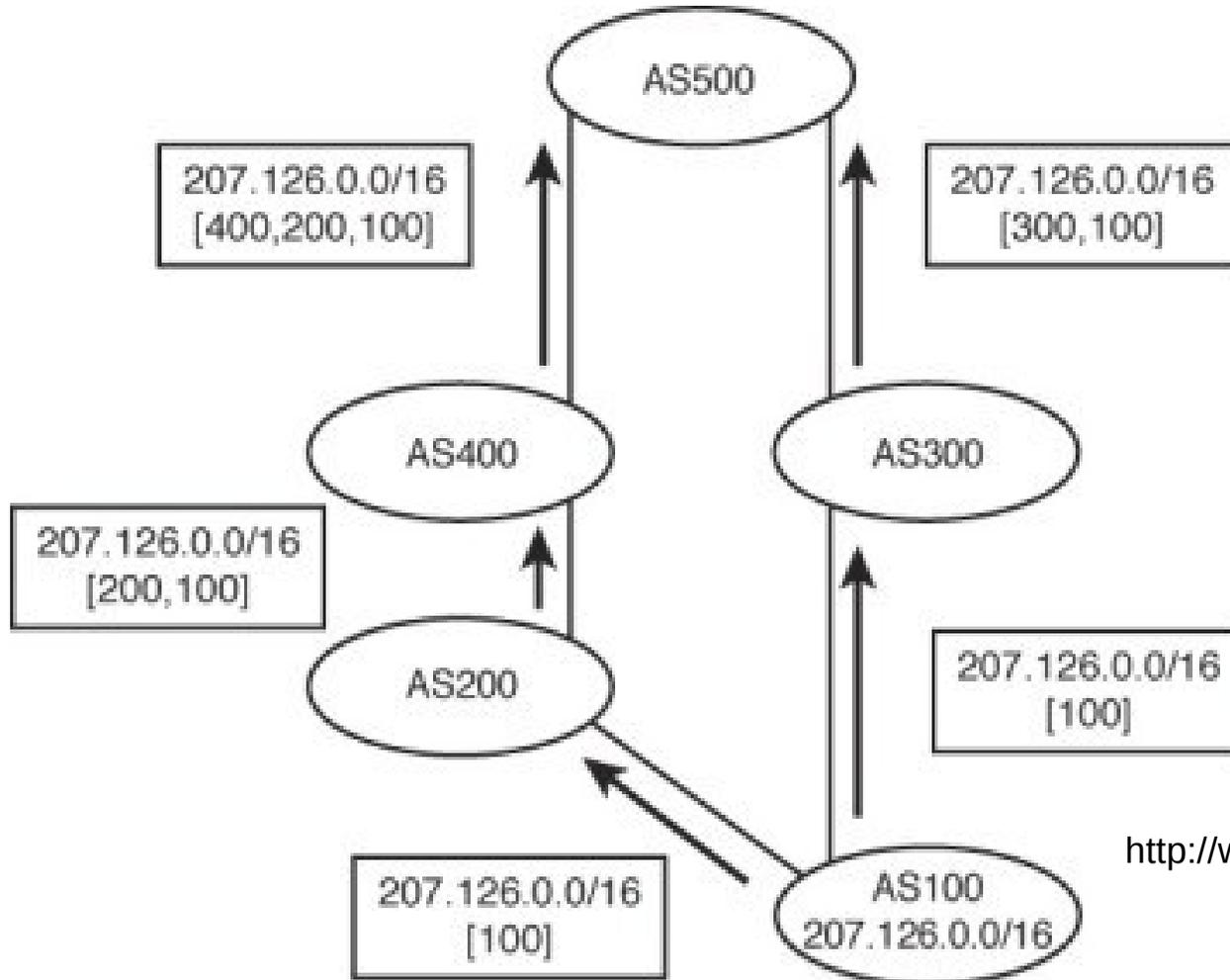
BGP Attributes - LOCAL-PREF Example



How to force all traffic to 400 over path b:d?

Set local preference 100 at b
Set local preference 200 at a

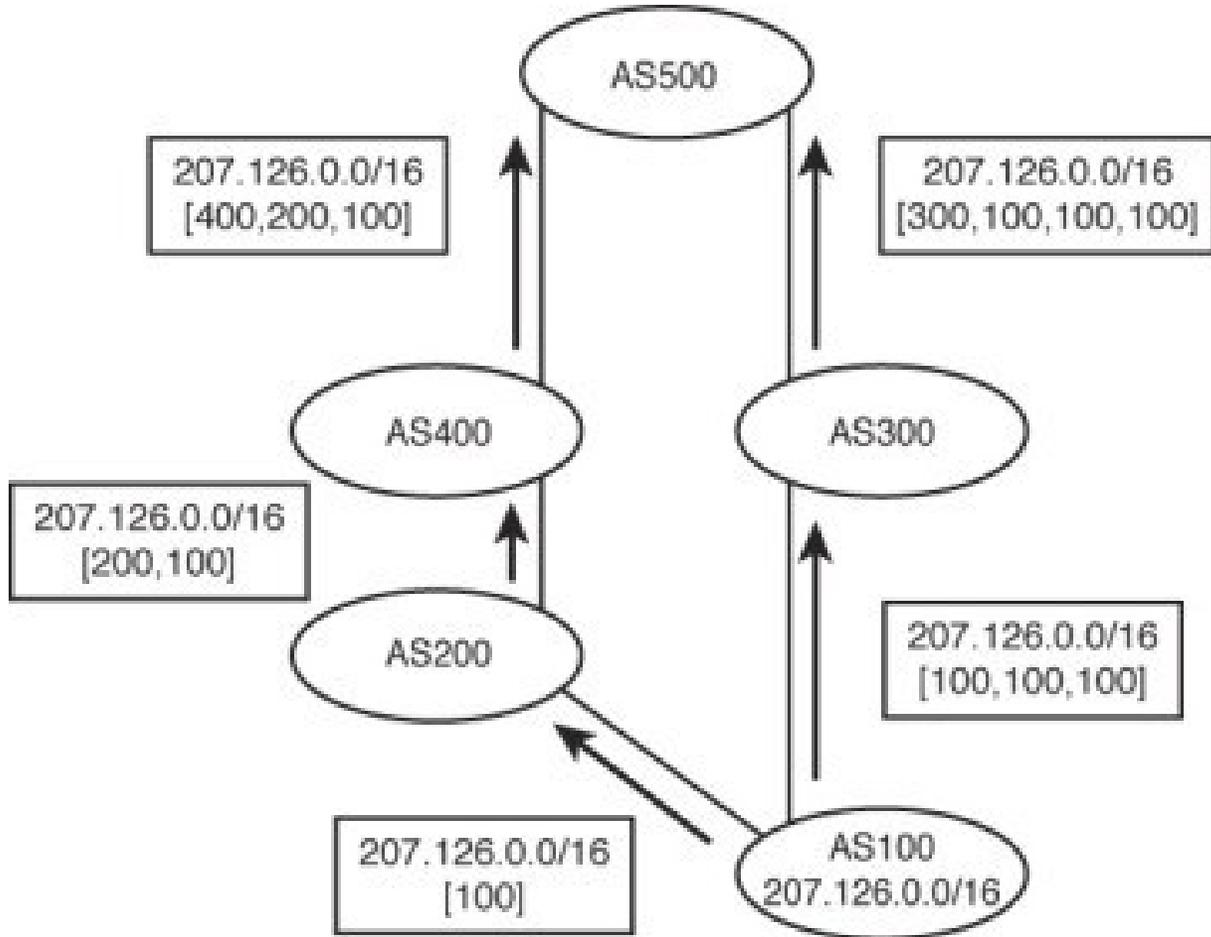
BGP Attribute - AS PATH



Each hop adds ASN to the path
-Only externally

<http://www.ciscopress.com/articles/article.asp?p=2738462&seqNum=2>

BGP Attribute - AS PATH



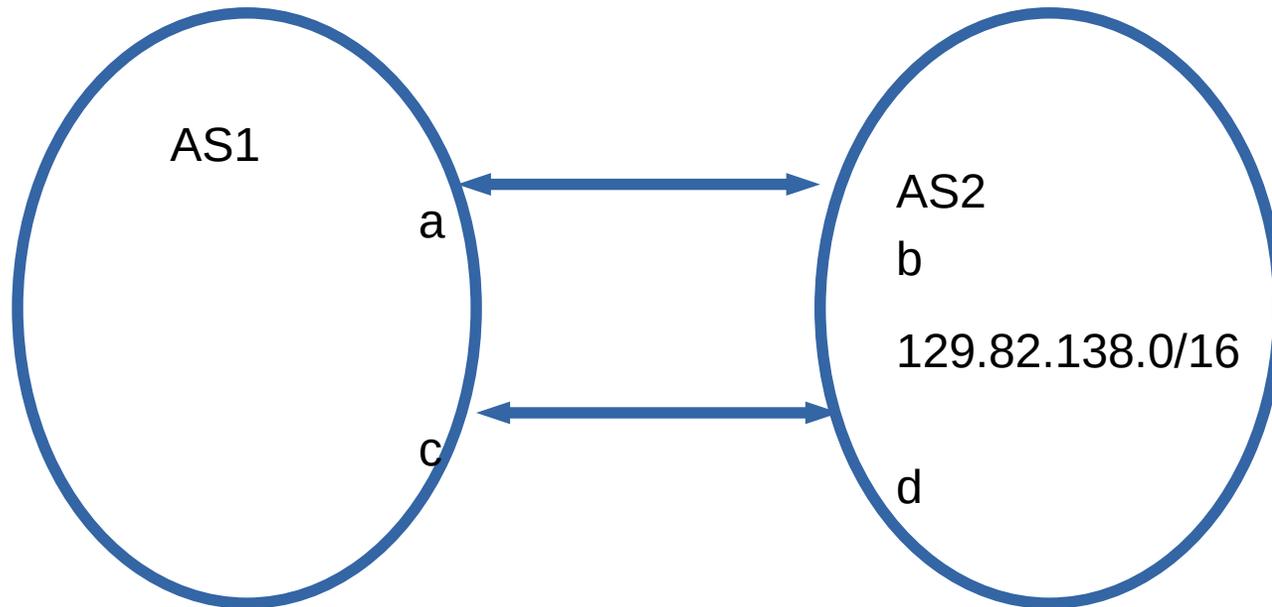
AS100 trying to influence path selection at AS500

- Append multiple path

<http://www.ciscopress.com/articles/article.asp?p=2738462&seqNum=2>

BGP Attribute - Local Preference

How do you load balance between two links using BGP?



At A:

129.82.138.0/17 → 10

129.82.138.128/17 → 5

At C:

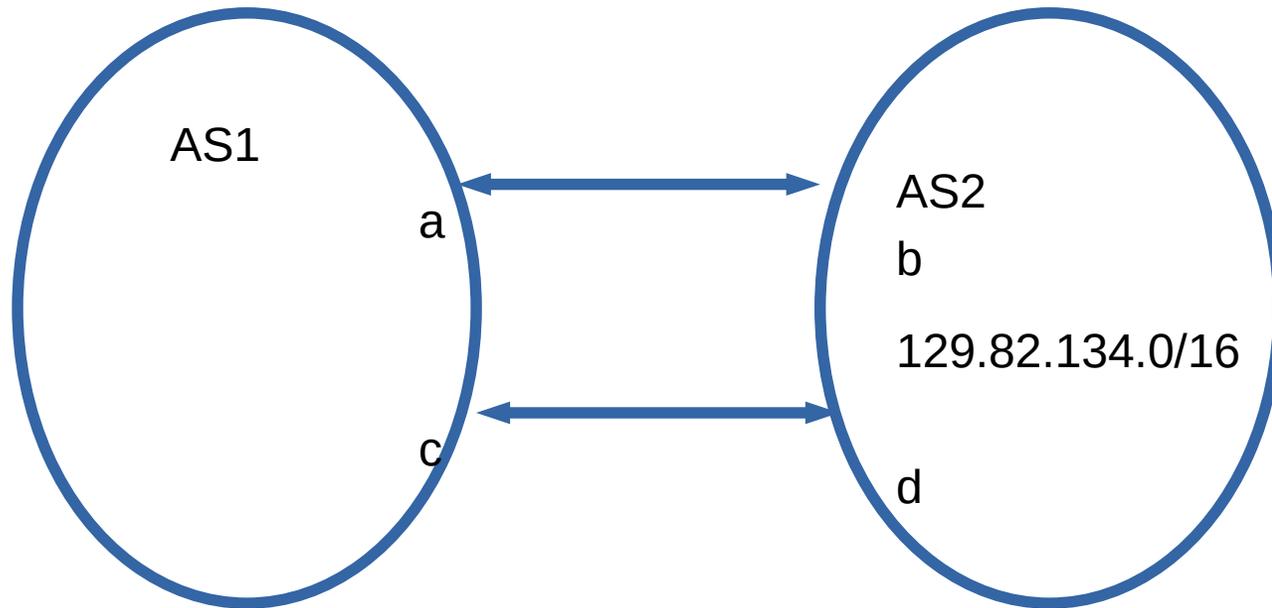
129.82.138.0/17 → 5

129.82.138.128/17 → 10

<http://www.ciscopress.com/articles/article.asp?p=2738462&seqNum=2>

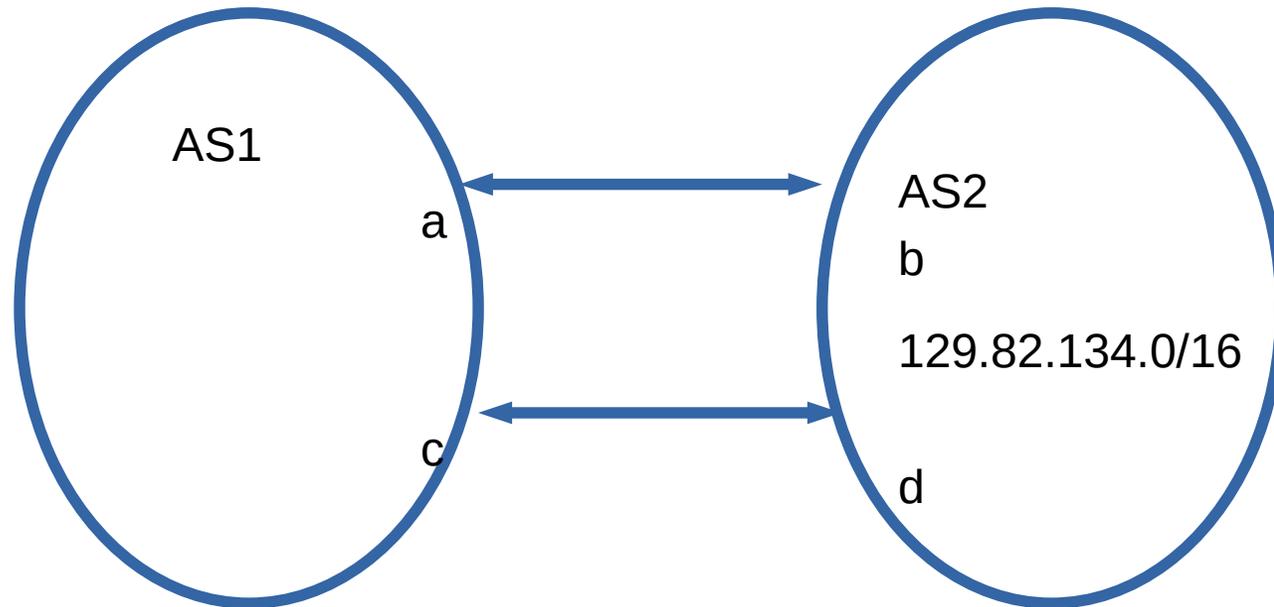
BGP Attribute - Local Preference

How does AS1 prefer a-b over c-d?
Higher preference wins!



<http://www.ciscopress.com/articles/article.asp?p=2738462&seqNum=2>

BGP Attribute – MED (Multi exit discriminator)

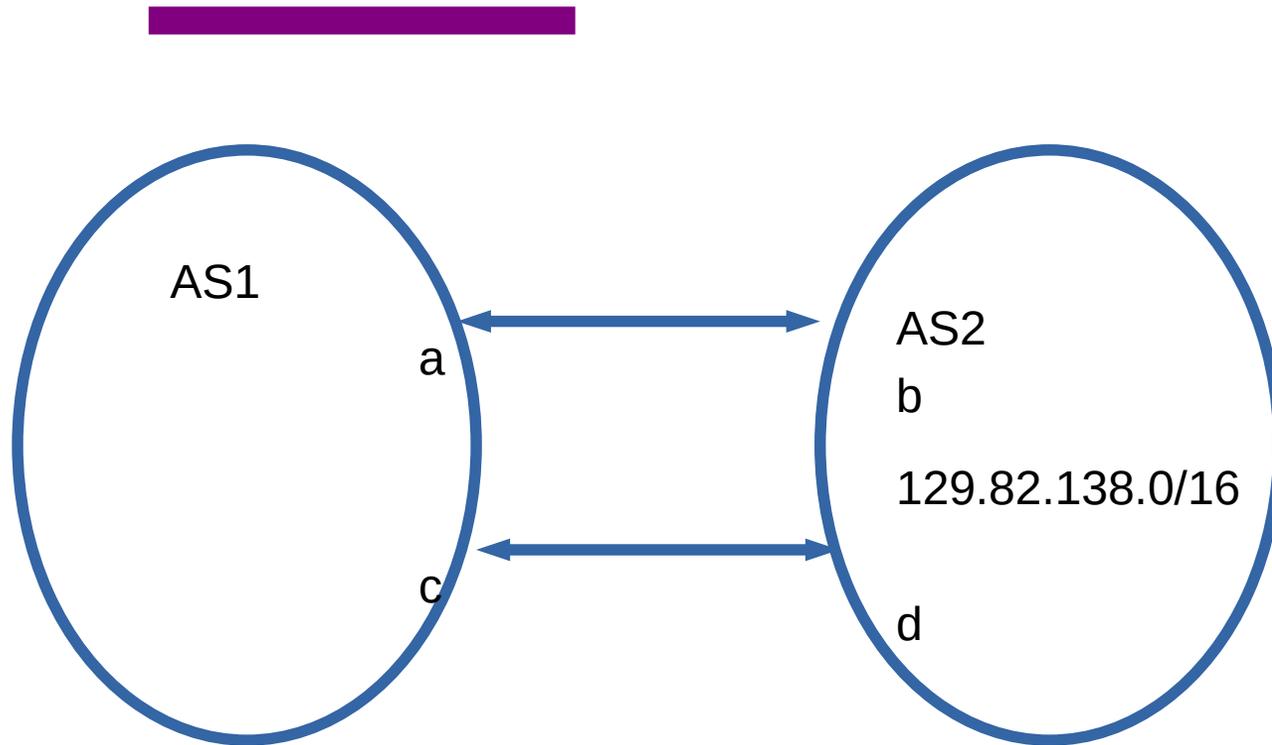


AS1 and AS2 has two paths between them

AS1 tells AS2 it's MED for influencing AS2's path selection

Lower cost wins

BGP Attribute - MED



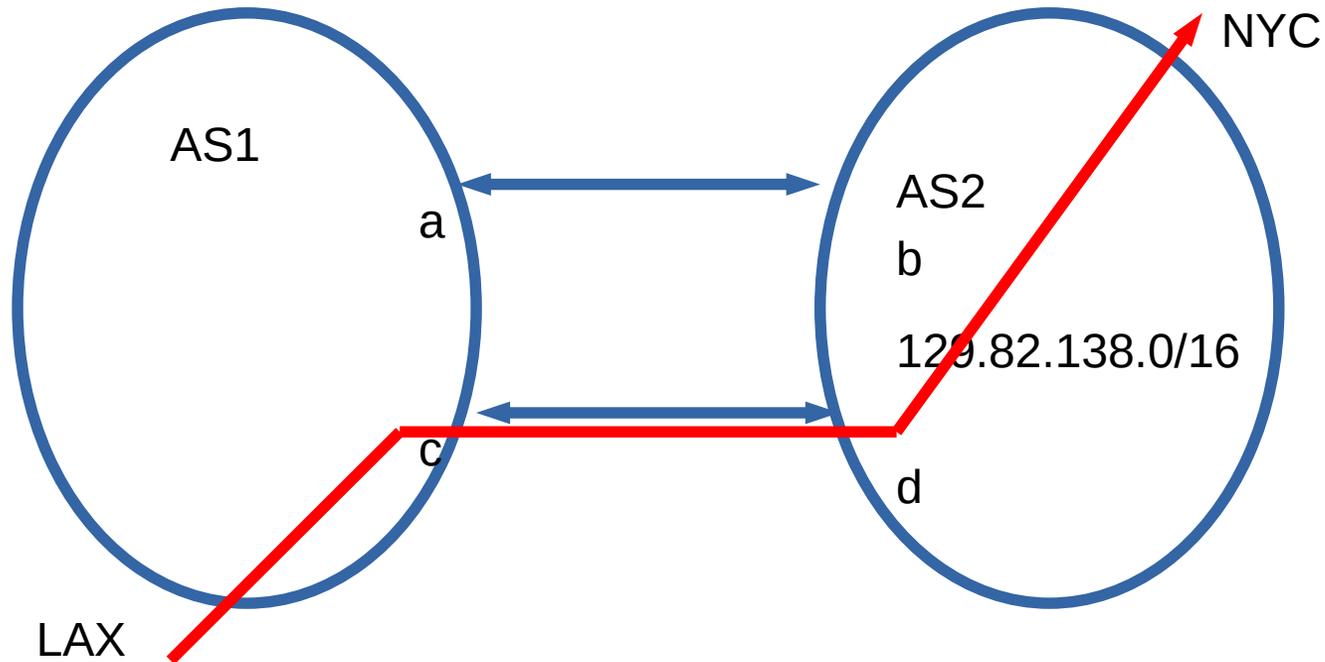
How would AS1 make AS2 send
129.82.138.0/17 over a-b
and
129.82.138.128/17 over c-d?

AS1 tells AS2

129.82.138.0/17 MED 5 via a
129.82.138.128/17 MED 10 via a

129.82.138.0/17 MED 10 via c
129.82.138.128/17 MED 5 via c

BGP Attribute - MED



Typically used in provider/subscriber
Not between peers – why?

One AS may force the other to carry traffic
for it

Local Pref vs MED

LOC_PREF → Internal – you tell your routers which route to use

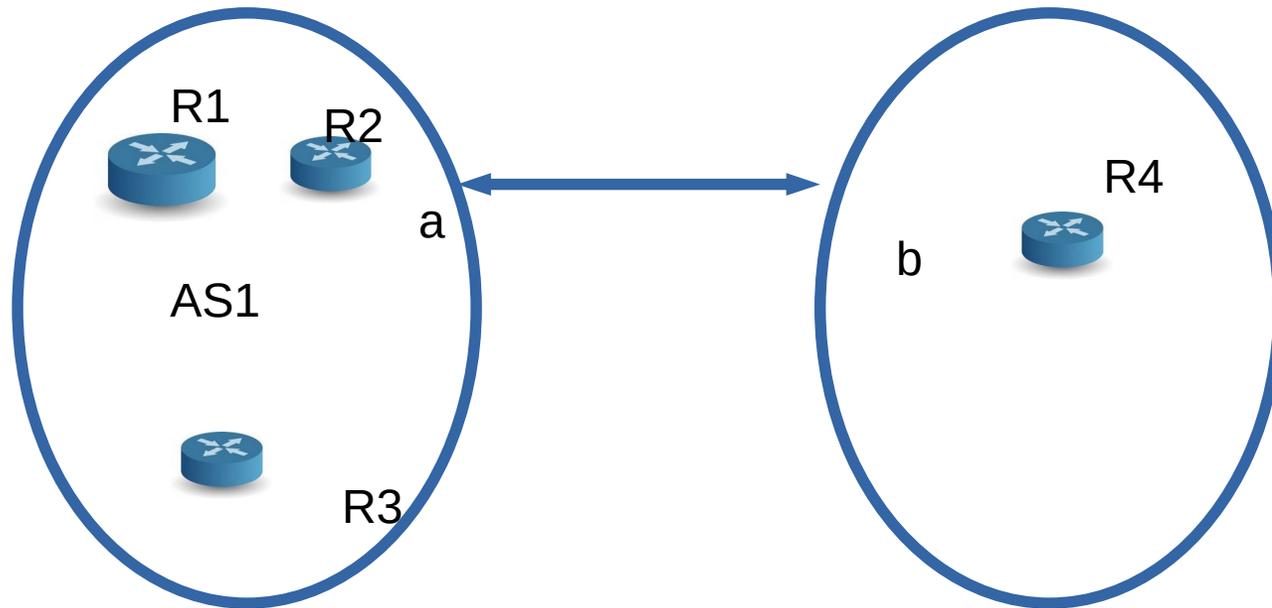
MED → External – you tell you neighbors which route you prefer
Neighbor is an autonomous system, so it can ignore you

BGP Attribute - Community

Put anything you want – between Ases, not known publicly

COMMUNITY: 17:210 17:13 4195:10 416:13 45:1103

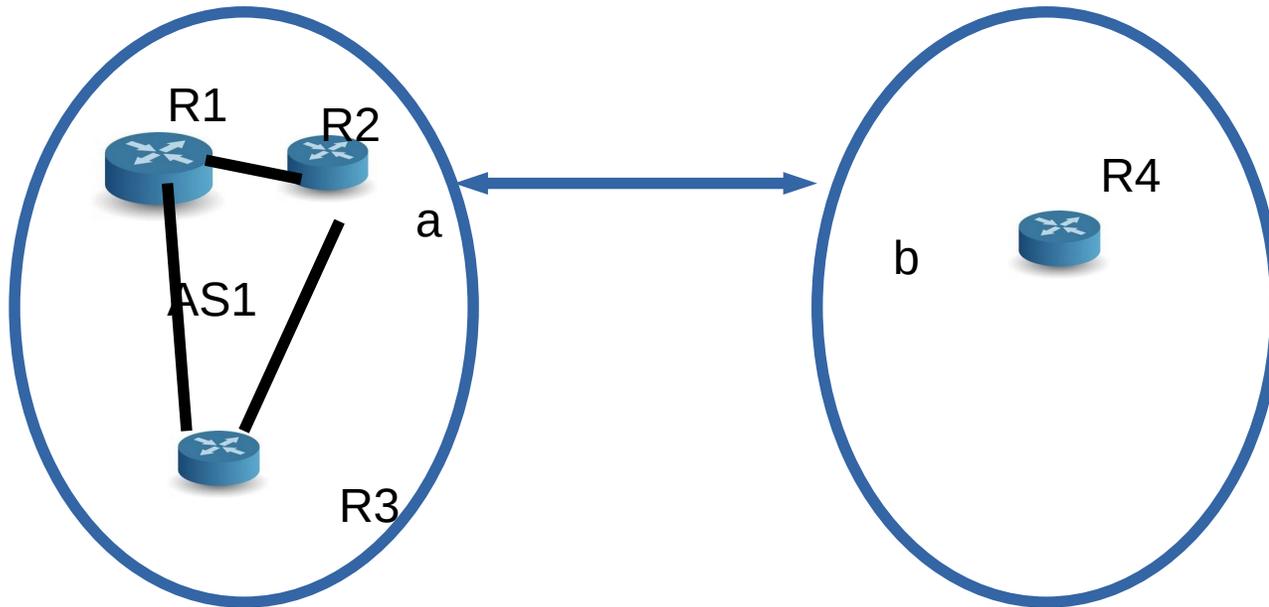
Internal vs External BGP



BGP between R2 and R4

What is between R1, R2, and R3?

Internal vs External BGP



BGP between R2 and R4

What is between R1, R2, and R3?

IBGP (Internal)

Different rules:

If you learn from outside, advertise
If you learn from inside, don't

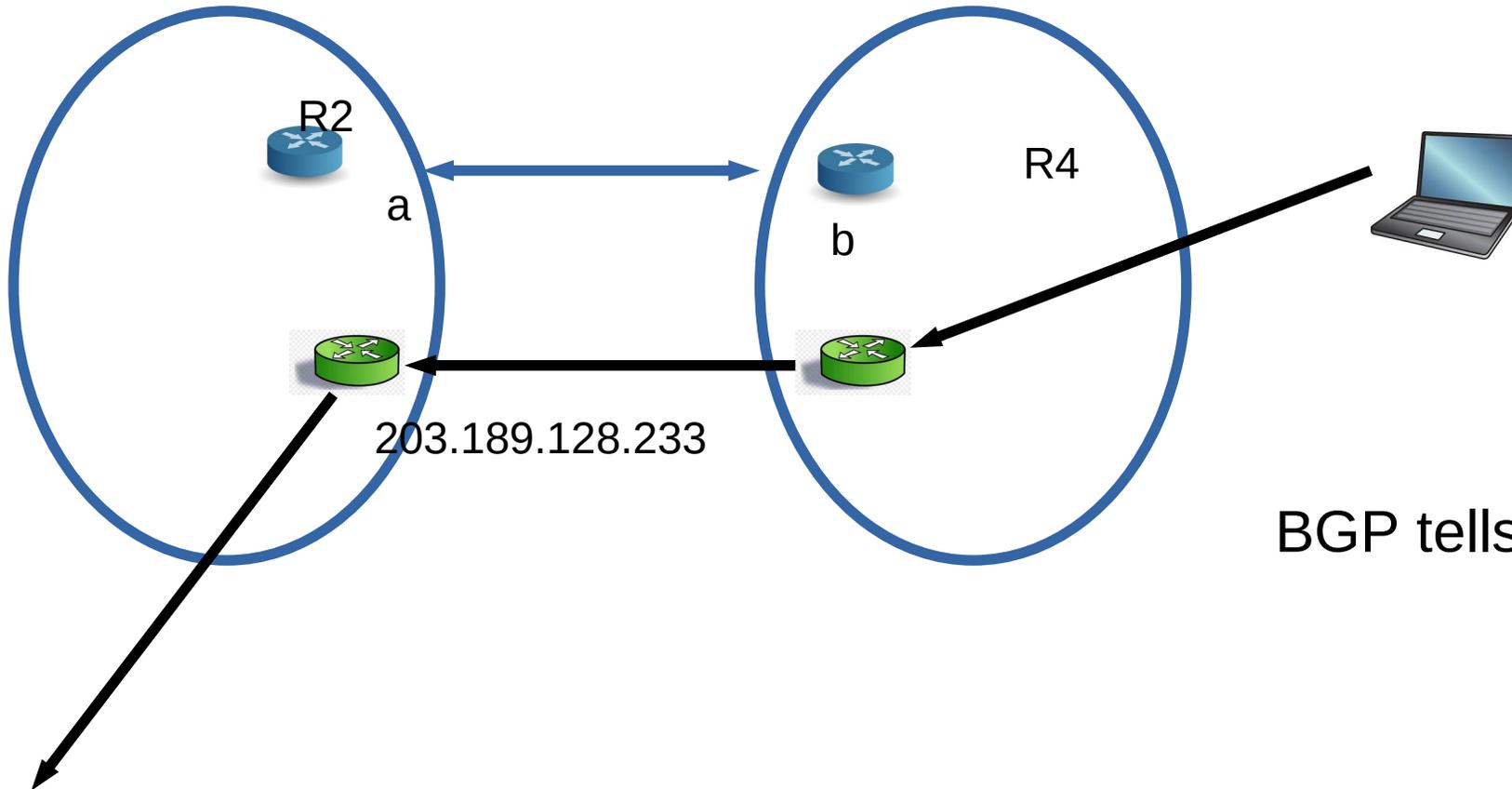
R2 can tell R3 and R1 about R4
R2 can not tell R1 about prefixes from R2 -loop!

IBGP must be a mesh!

BGP vs IP routers

Next hop | Announcing AS| Target Prefix| Path

203.189.128.233 | 23673 | 149.149.0.0/16 | 23673 1299



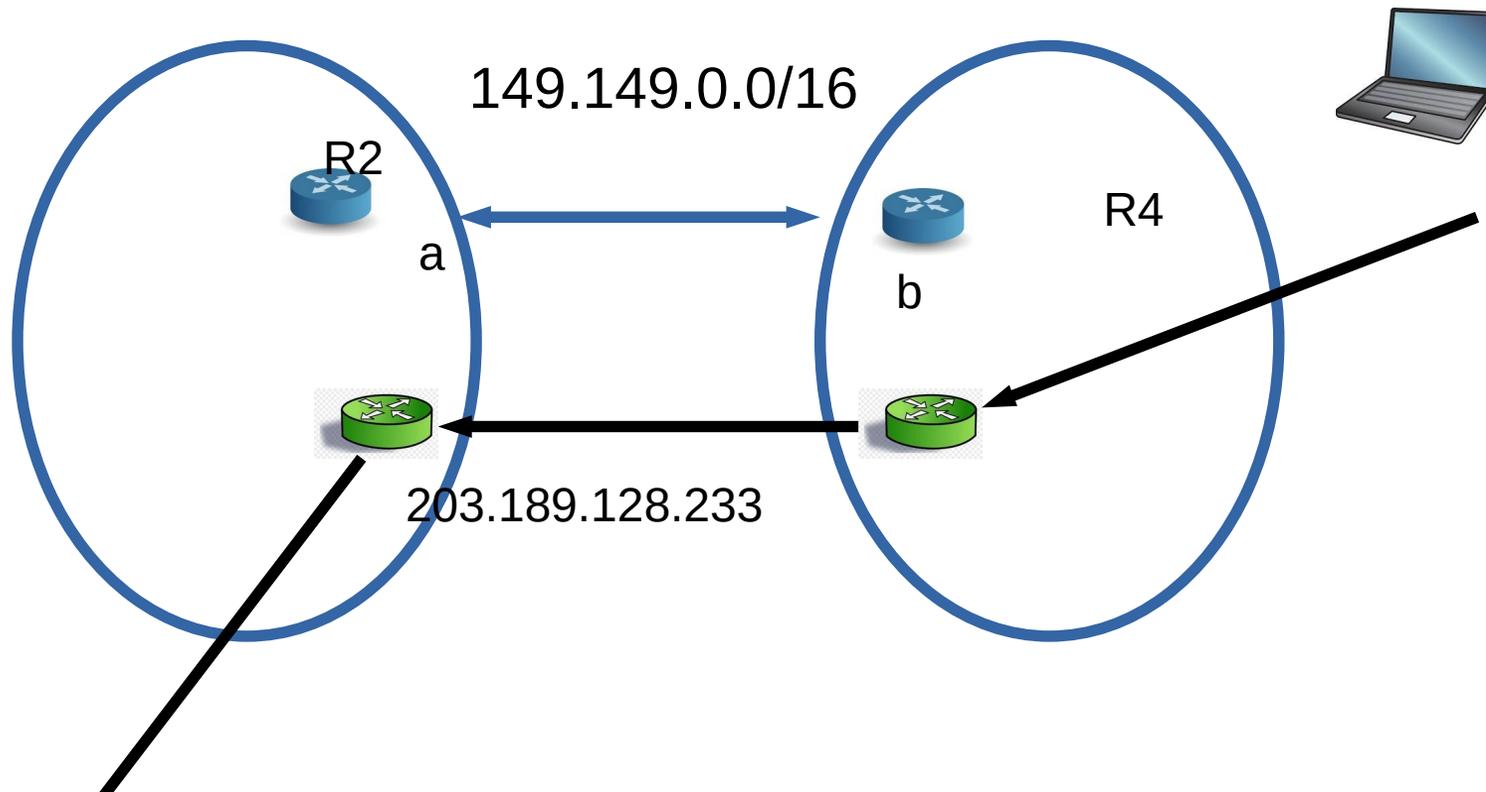
BGP tells you which IP router to use

BGP Decision process

Next hop | Announcing AS | Target Prefix | Path | LOCAL_PREF | MED | Next Hop Cost

203.189.128.233 | 23673 | 149.149.0.0/16 | 23673 1299 | 10 | 5 | 100

203.189.128.233 | 23673 | 149.149.0.0/16 | 23673 1299 | 100 | 50 | 10

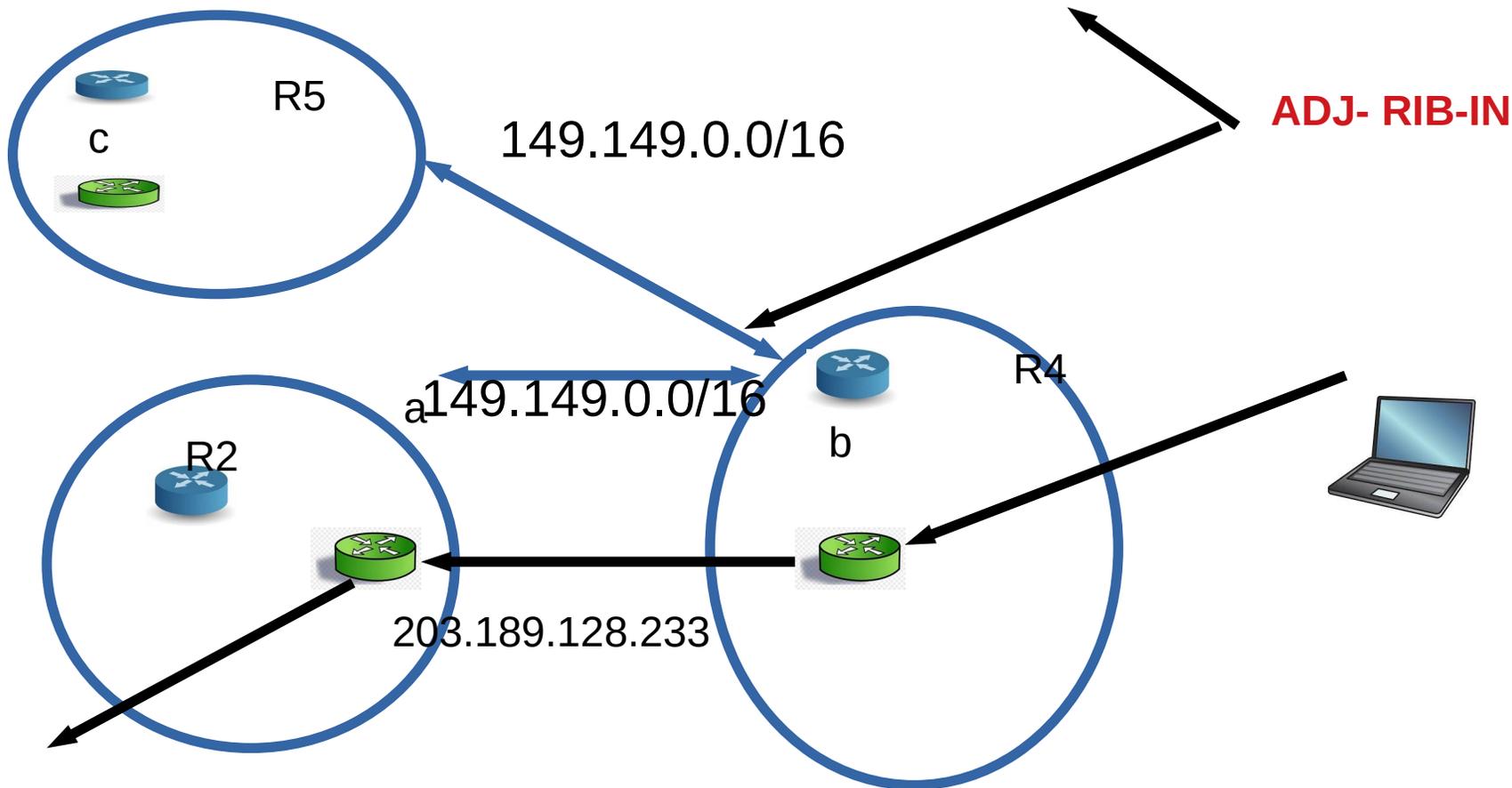


BGP Decision process

Next hop | Announcing AS | Target Prefix | Path | LOCAL_PREF | MED | Next Hop Cost

203.189.128.233 | 23673 | 149.149.0.0/16 | 23673 1299 | 10 | 5 | 100

203.189.128.233 | 23673 | 149.149.0.0/16 | 23673 1299 | 100 | 50 | 10

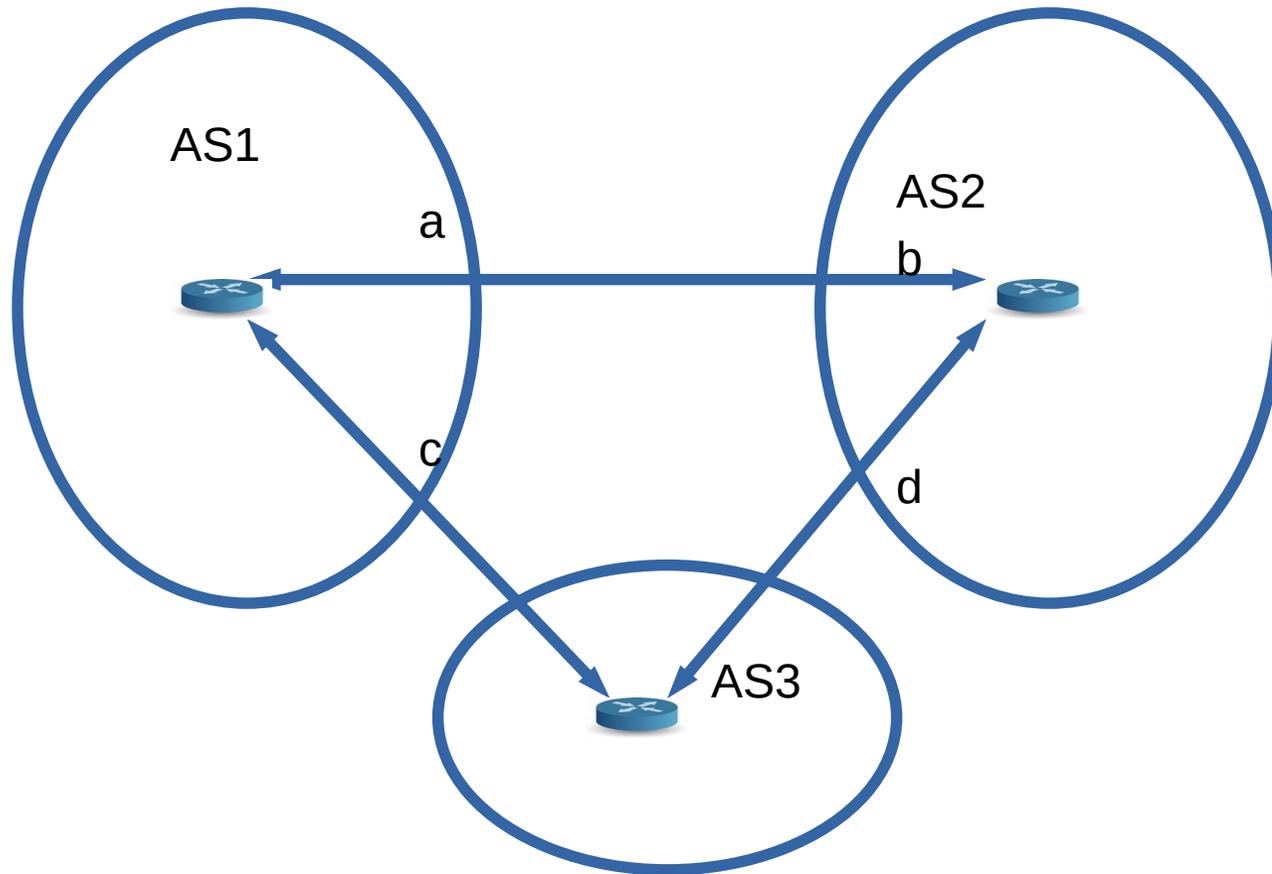


BGP Decision process

At ADJ-RIB-IN calculate degree of preference until **one route for each destination remains!!**

- select route with highest LOCAL-PREF
 - Select route with shortest AS-PATH
 - Select route with lowest MED
 - Select route with smallest NEXT-HOP cost
 - Select route learned from E-BGP peer with lowest ID
 - Select route learned from I-BGP peer with lowest ID
-
- Install selected route in LOC-RIB
-
- Update ADJ-RIB-OUT, notify peers
 - You can only send what is in LOC-RIB (or a subset of it)

BGP



- 1 will prefer 2 over 3
- 1 will not accept traffic from 3
- 2 will prefer path to 3 via 1
- 3 will utilize both paths

Reading Assignment

- BGP
 - <https://book.systemsapproach.org/scaling/global.html#interdomain-routing-bgp>
 - 30-40 minutes read