CSC4200/5200 – COMPUTER NETWORKING

Instructor: Susmit Shannigrahi

GLOBAL INTERNET

sshannigrahi@tntech.edu

GTA: dereddick42@students.tntech.edu

Tennessee TECH

Apps (HTTP)

Data

Apps (HTTP)

Transport (TCP/UDP)

Segments

Transport (TCP/UDP)

Network (IP)

Packets

Network (IP)

Link (Ethernet)

Frames

Ethernet Interface

Ethernet Interface

Link (Ethernet)

Bits (1010001)

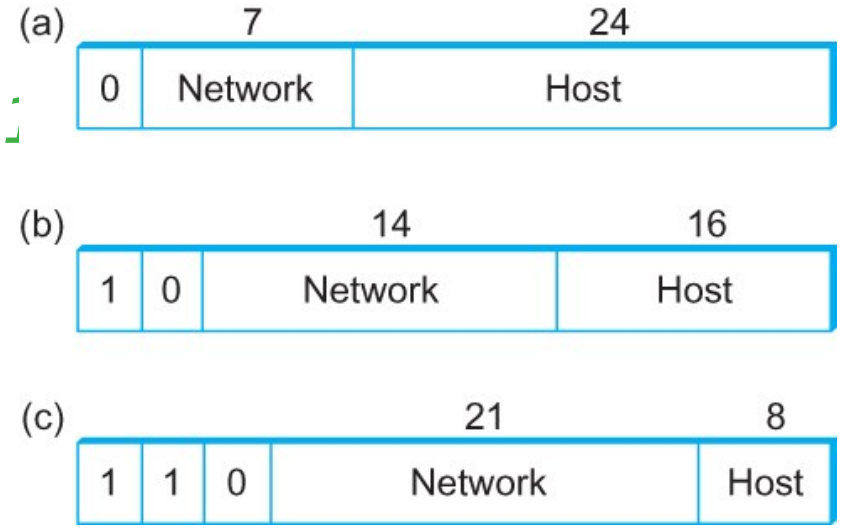
Logistics

- PA2 will post tomorrow – due March 31st.
- Those who asked for PA1 extension – email me the code.
- Homework2 will post next week.
- Second exam – end of March.

Back to Addressing

- A 32 bit number in quad-dot notation
- Identifies an ***Interface***
  - ***A host might have several interfaces!!!***
- *129.82.138.254*

  10000001.01010010.10001010.11

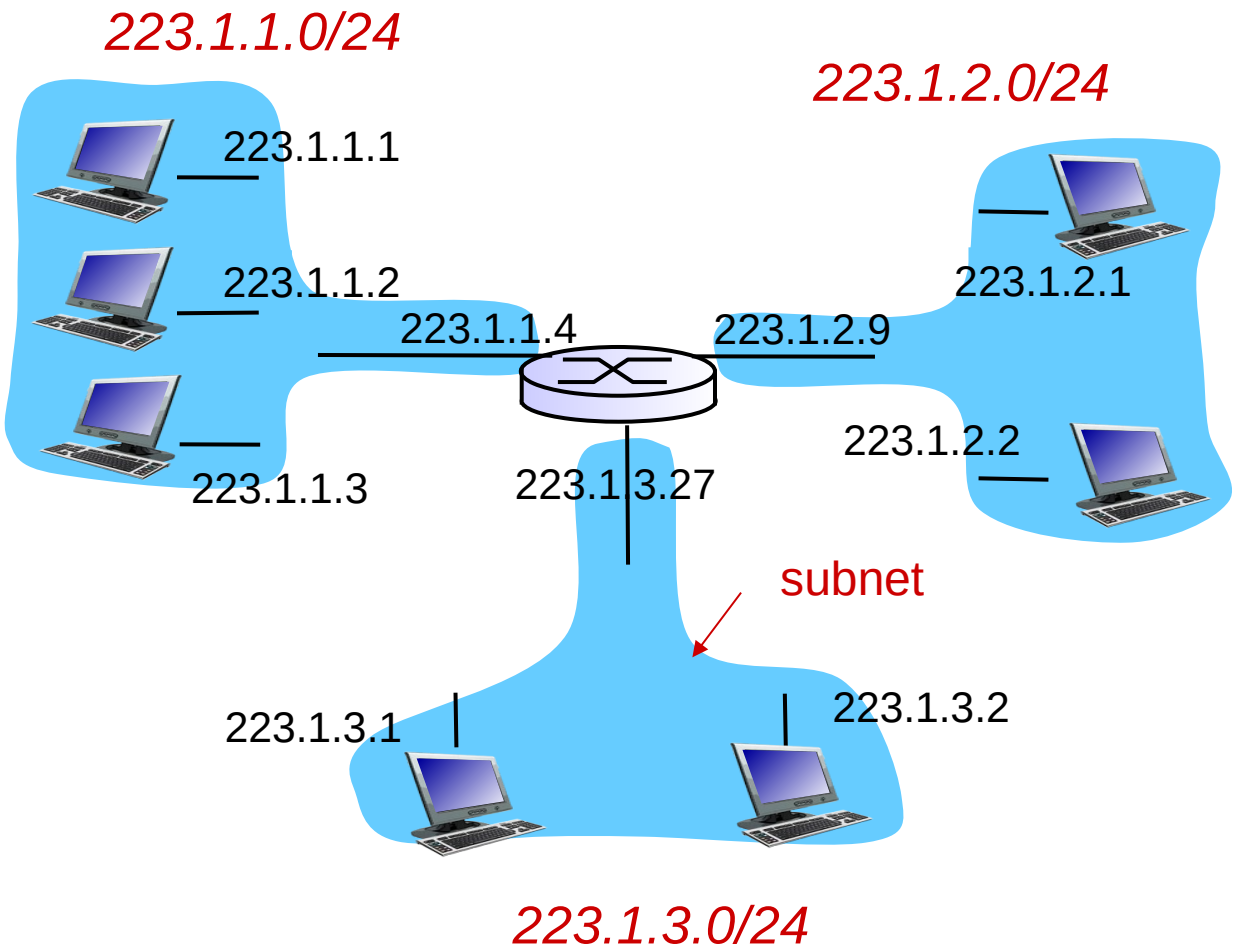


Subnets Revisited

Recipe:

- Create isolated networks – *subnets*

- No longer need to know individual Ips – knowing the subnet is enough
  - *223.1.1.0/14 → Interface 2*



subnet mask: /24

Subnets – Class based

- Originally – only rigid boundaries
  - Class A – 5.0.0.0/8 - 0*
  - Class B – 149.149.0.0/16 - 10*
  - Class C – 129.82.138.0/14 - 110*
  - Class D – 224.0.0.0 – 1110*
  - Class E - reserved

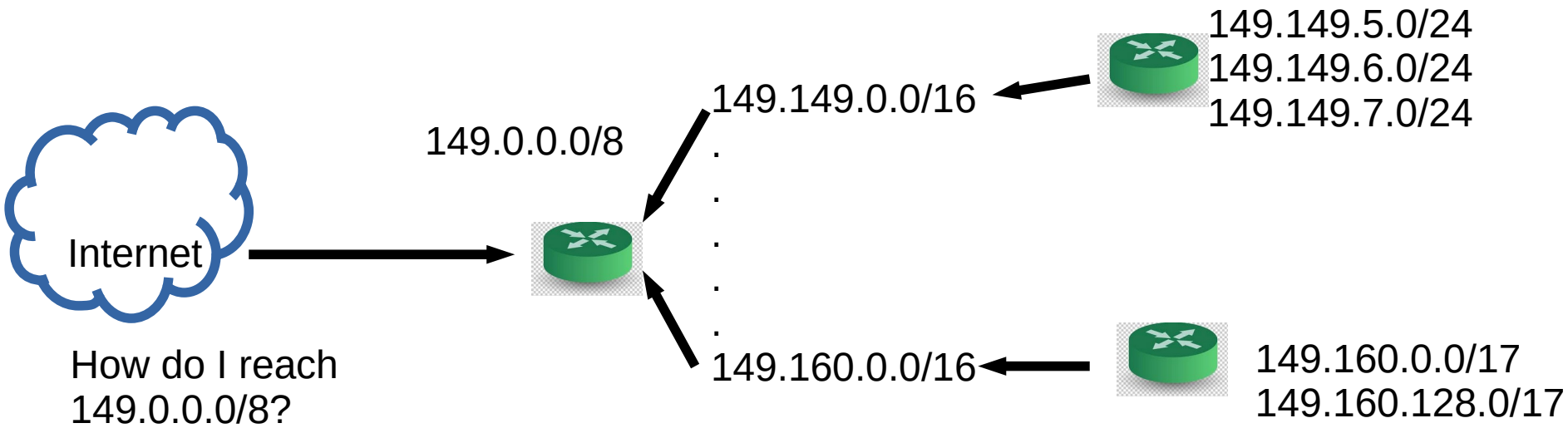
Subnets – Classless CIDR

- No rigid boundaries
  - *129.82.138.0/25*

    10000001.01010010.10001010.10000000

Subnets (Prefixes) scales the Internet

- Addresses are allocated in contiguous prefixes (tntech 149.149.0.0/16)
- Routing protocols operate based on prefixes (how do I reach 149.149.0.0/16)?

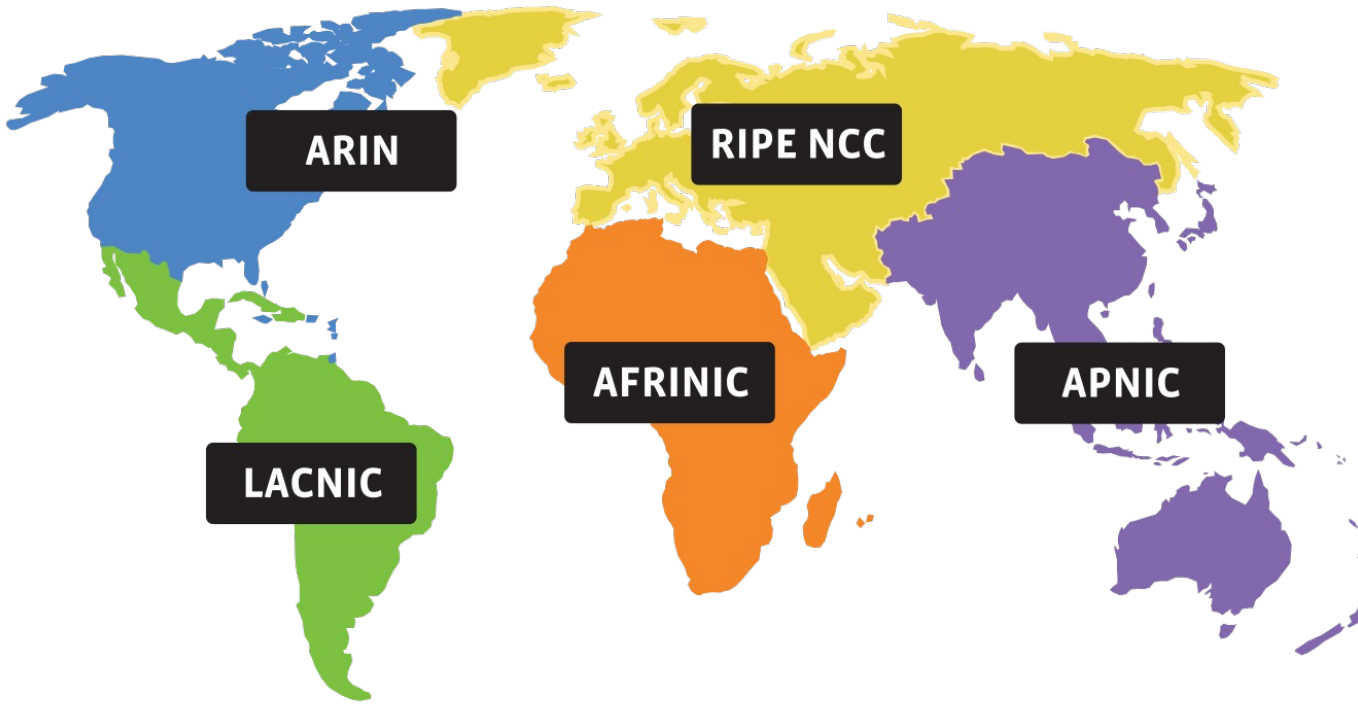


Not

How do I reach 149.149.5.0/24

How do I reach 149.149.6.0/24

Who gets what prefix?



0. Internet Corporation for Assigned Names and Numbers (ICANN) – Decides which RIRs get what address

1. Regional Internet Registries (RIRs) – Which orgs get what address

2. ISPs – Which customers get which address

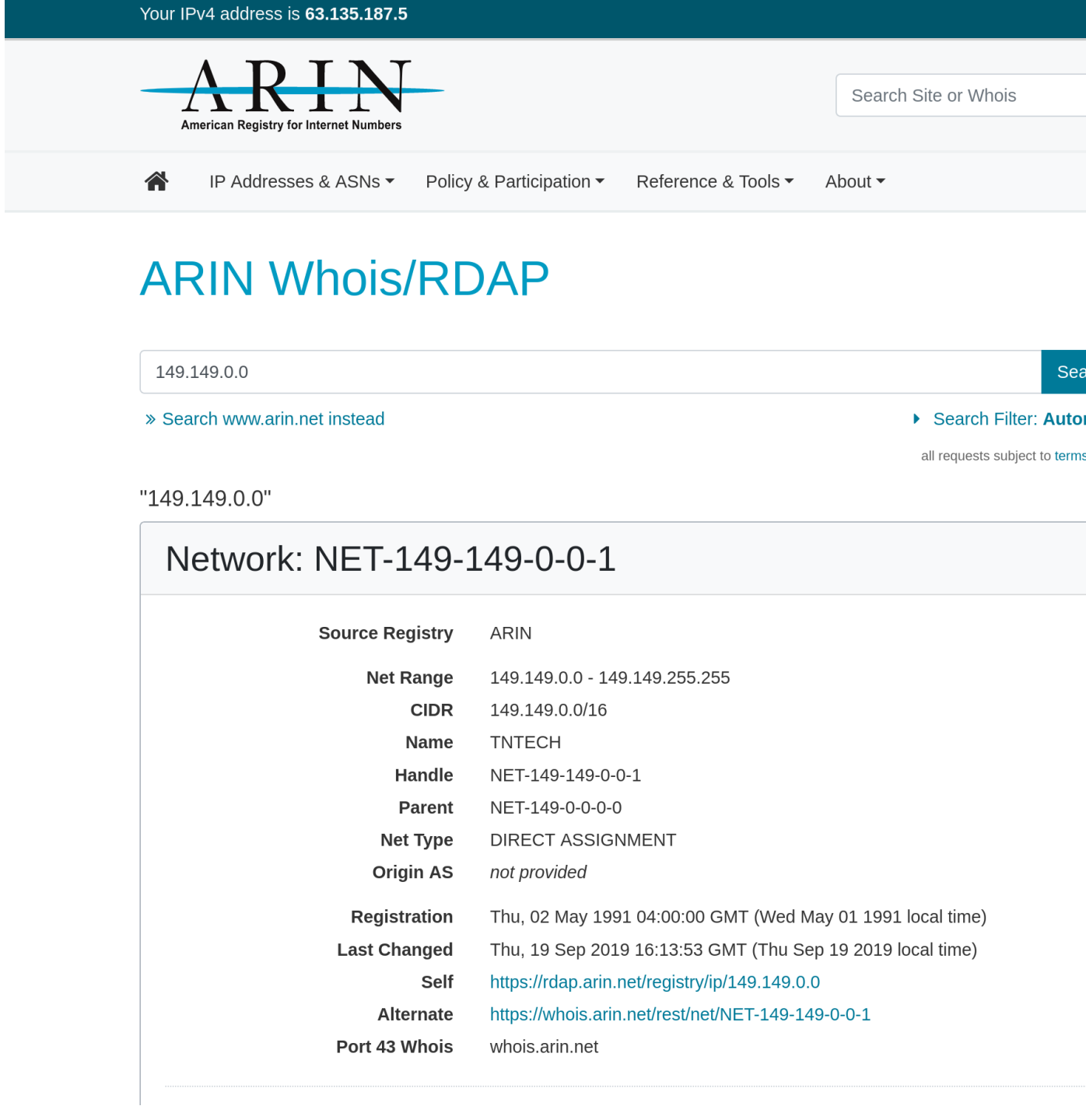
How do you know who has a prefix? “whois”

$ whois tntech.edu

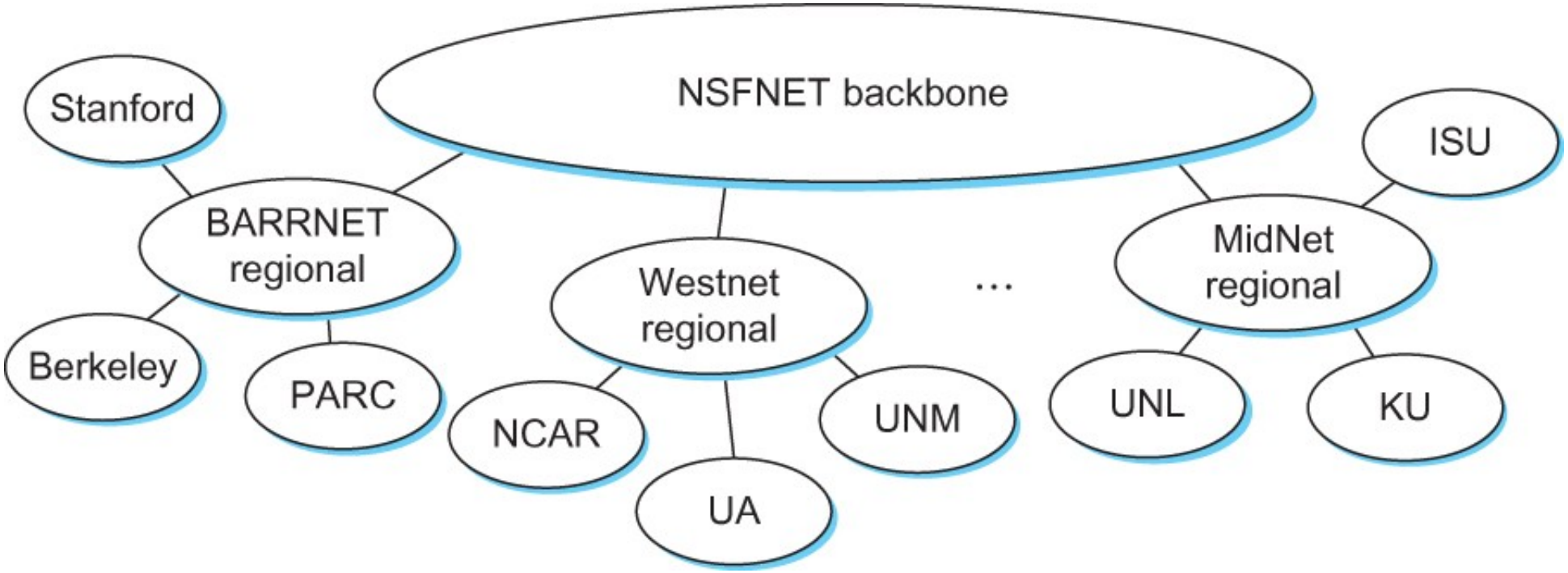
Domain Name: TNTECH.EDU

Registrant:
Tennessee Technological University
Information Technology Service
1010 N. Peachtree Street
Cookeville, TN 38505
USA

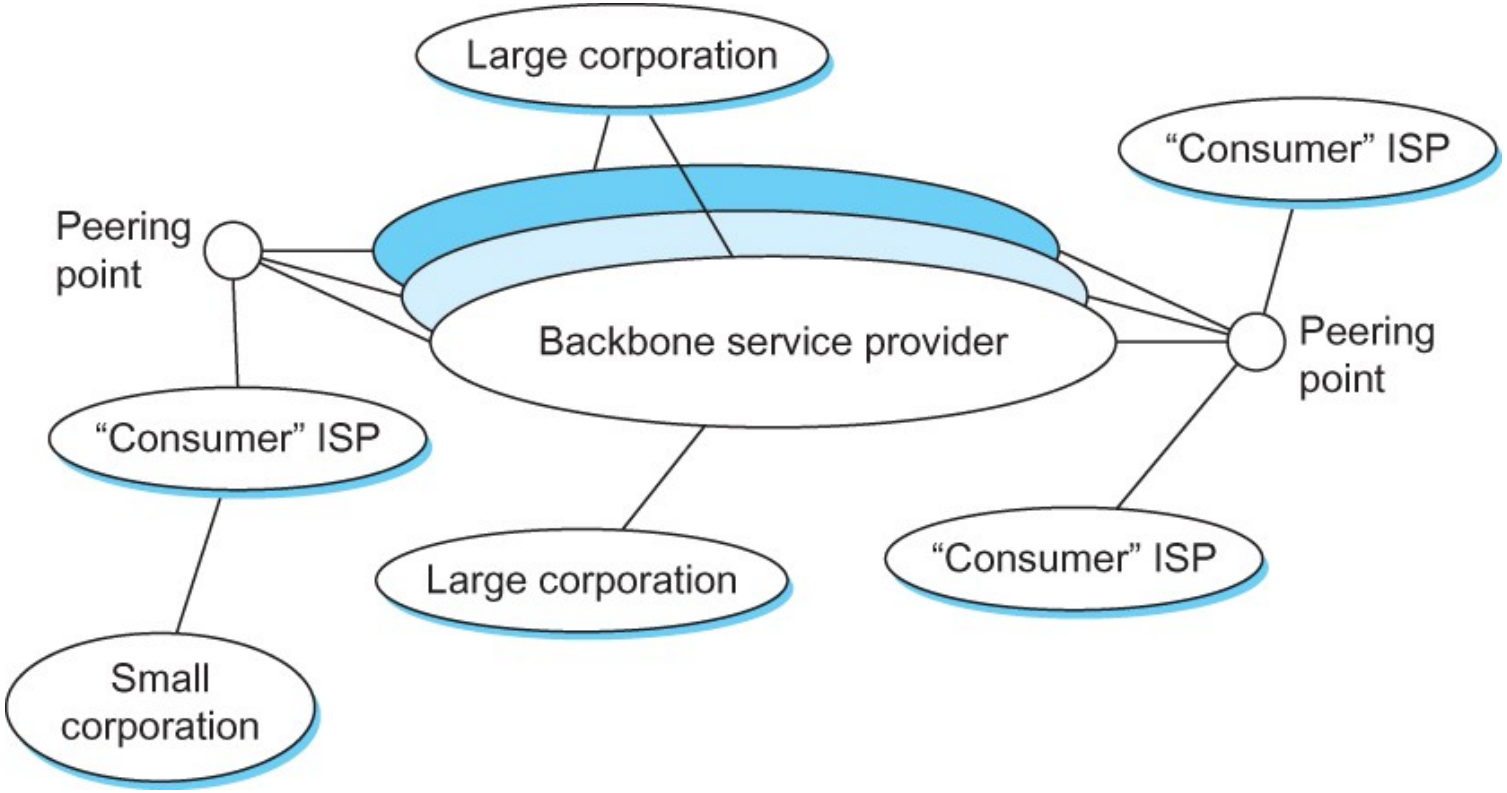
Domain record activated: 09-Sep-1992
Domain record last updated: 26-Sep-2019
Domain expires: 31-Jul-2020



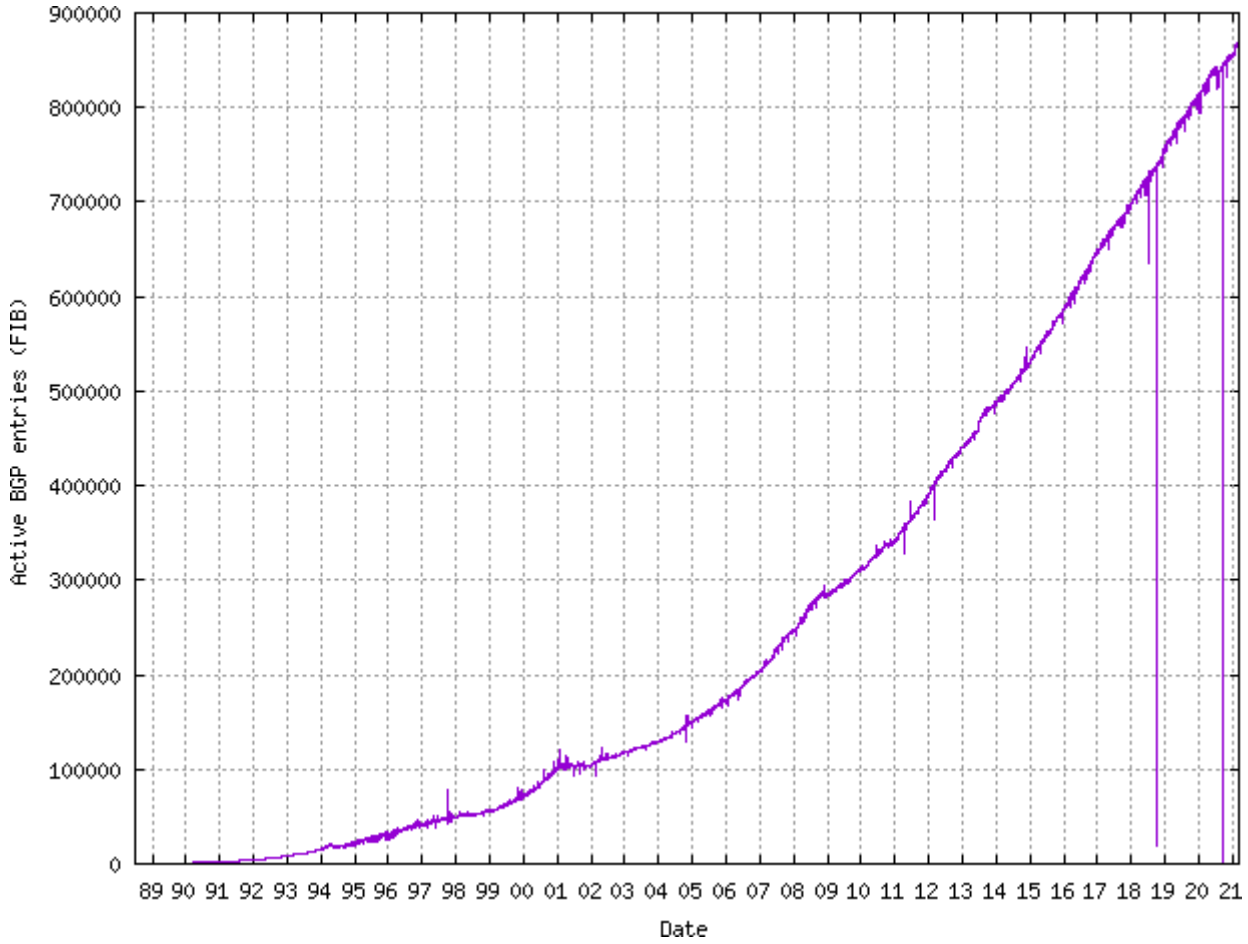
Internet in the 1990s



Internet now

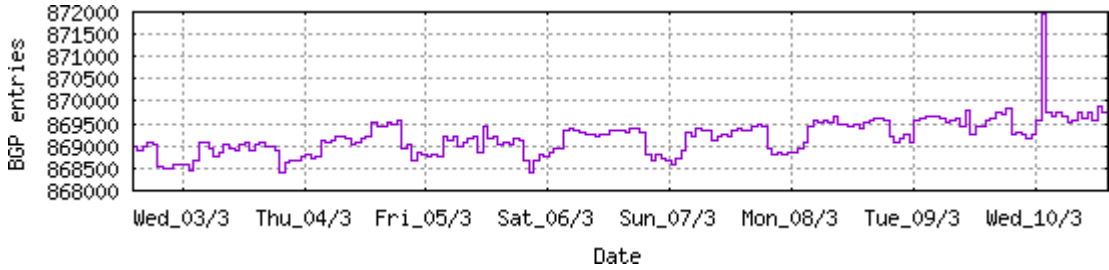


How many prefixes are there?

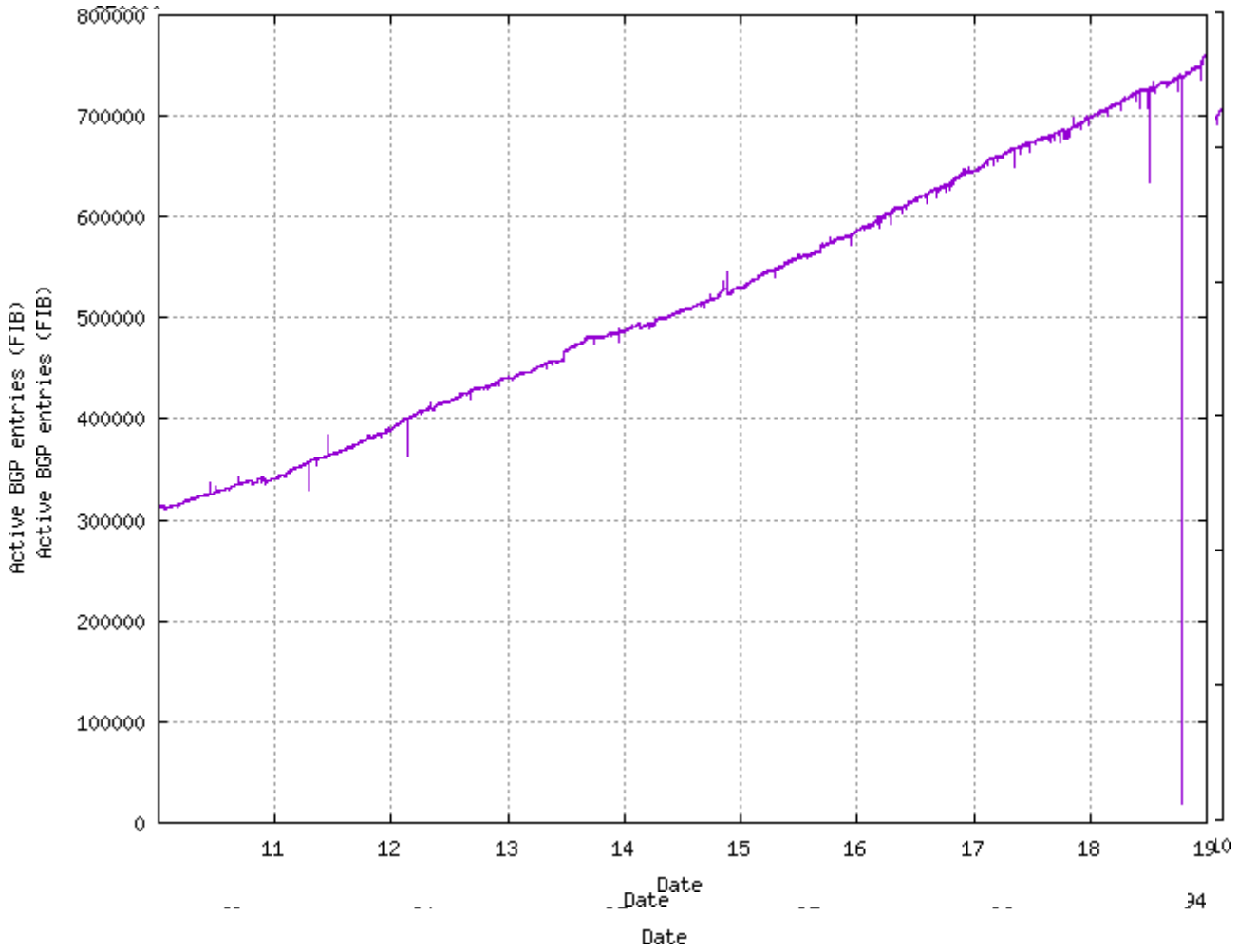


https://www.cidr-report.org/

100K in 2001 → 800K in 2019



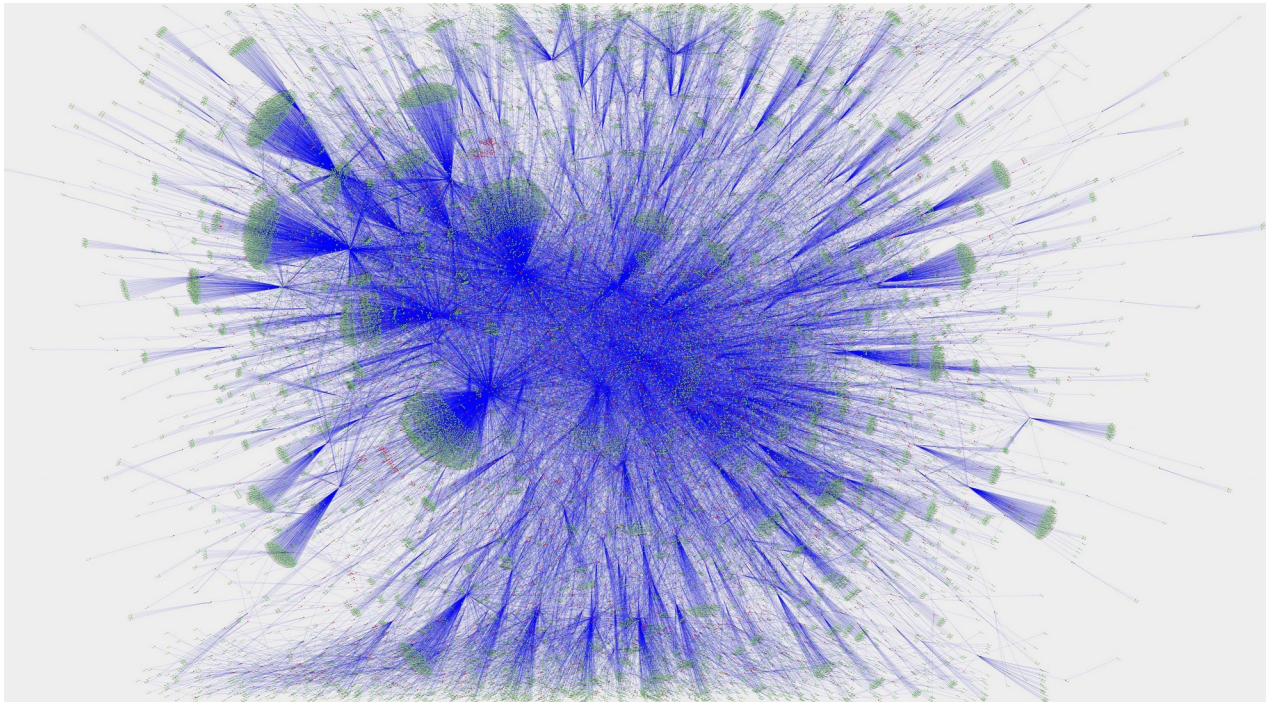
Bit of history – how the Internet evolved



‘88-’94 - 0 → 14000
‘94-’00 – 90000 – Linear growth
’00-’10 – up to 300,000
’10-’19 – up to 800,000

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

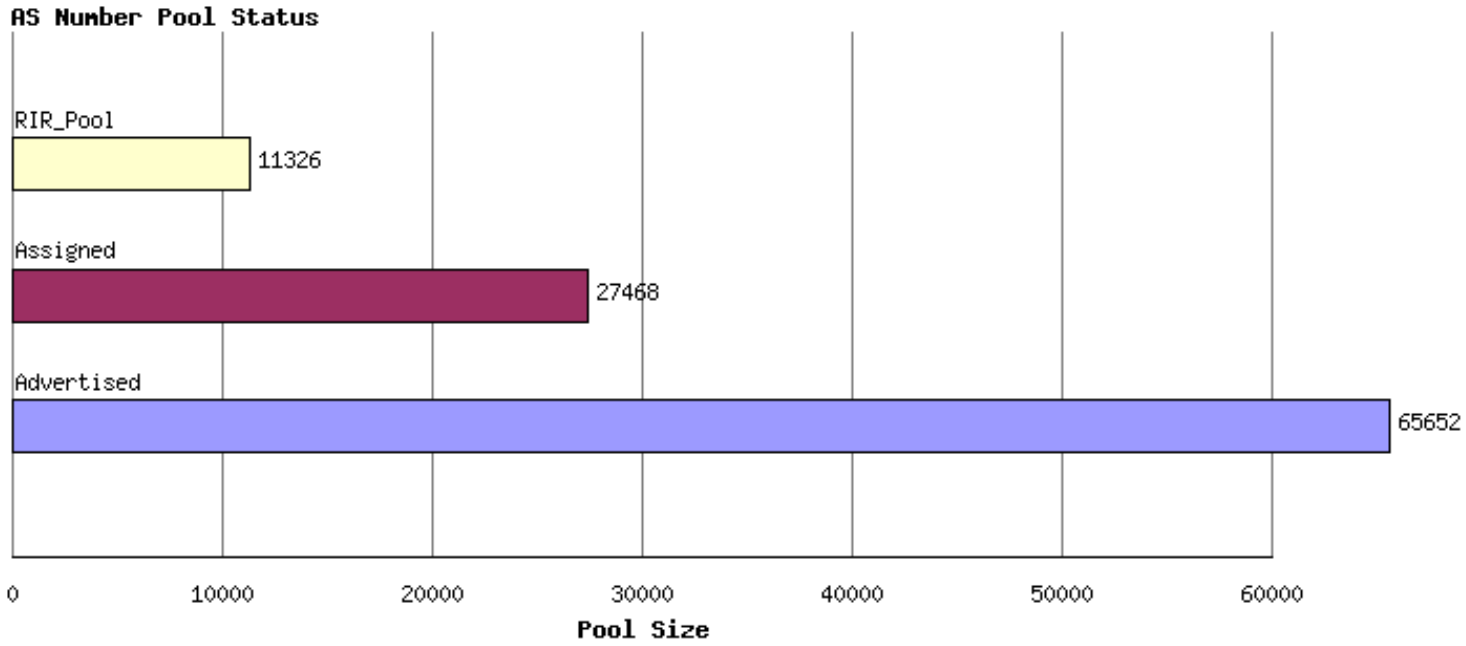
Autonomous systems (ASes)

- 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

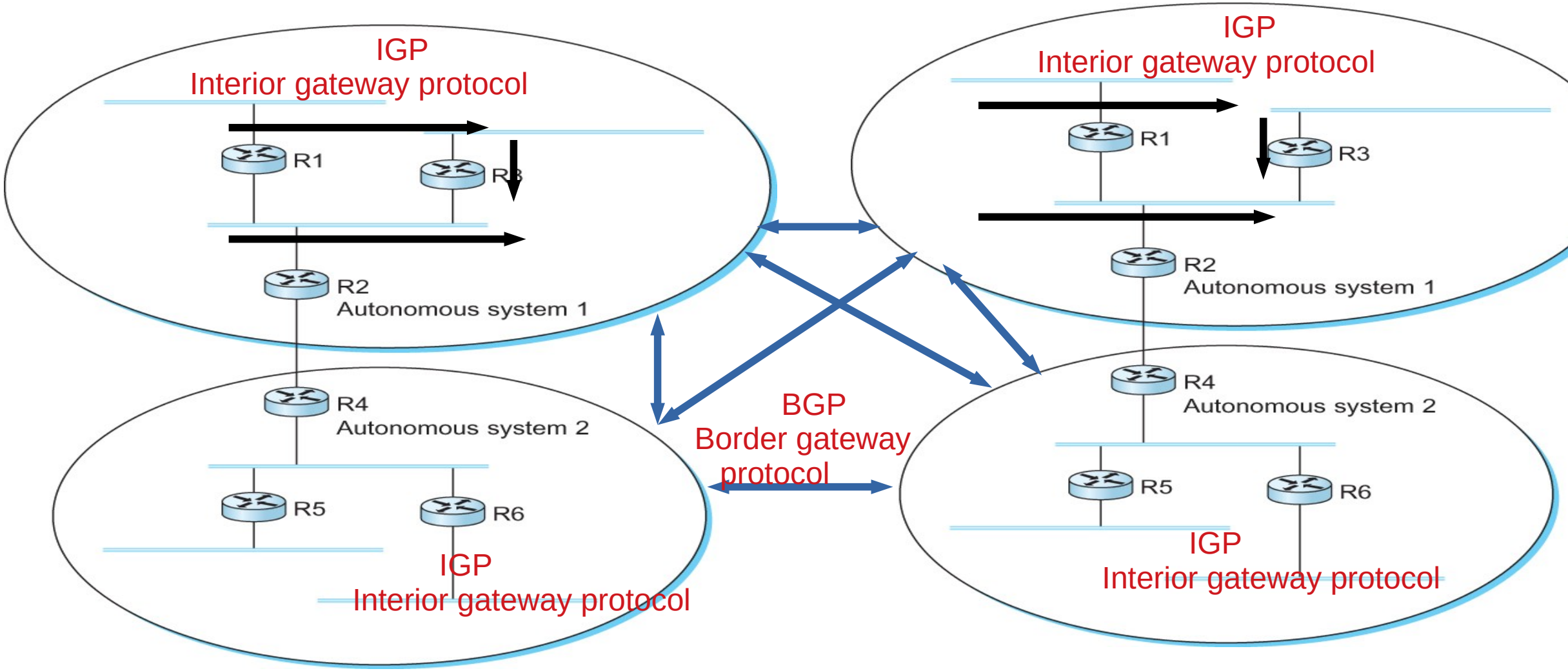
Status of ASNs

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



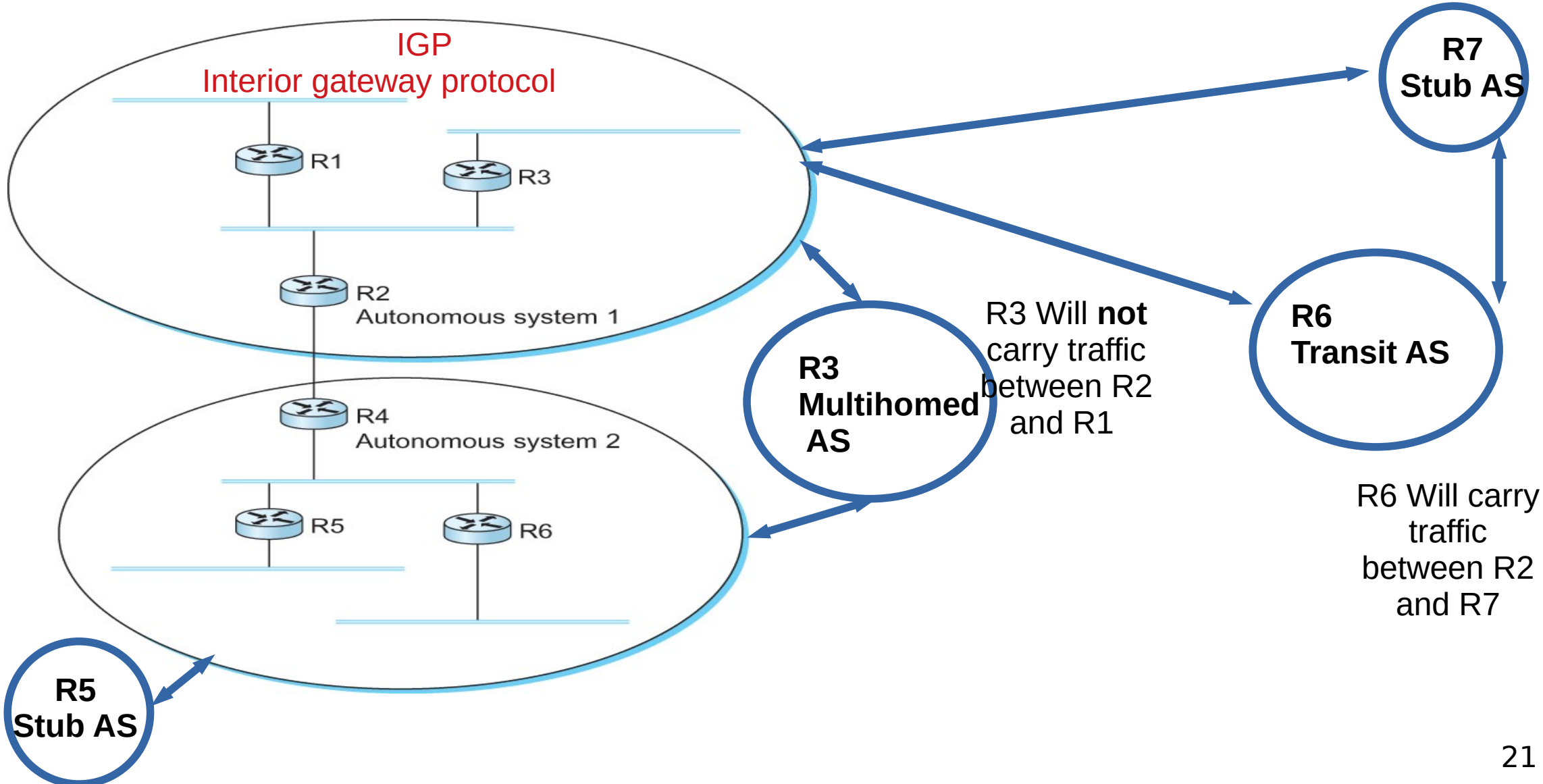
A network with four autonomous systems

BGP-4: Border Gateway Protocol

- Assumes the Internet is an arbitrarily interconnected set of AS's.

- Local traffic – within the AS
- Transit traffic – from AS1 to AS3 via AS2
- Three types of AS’s
  - *Stub AS*
  - *Multihomed AS*
  - *Transit AS*

BGP-4: Border Gateway Protocol



BGP: Which routing protocol?

Link state?

- Does not scale
- you can have loops
- exposes routing costs to others

Distance vector?

- Slow to converge, count-to-infinity
- No universal metrics

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 - Goals

- Scalability: Forward any packet destined anywhere in the Internet
  - Having a routing table that will provide a match for any valid IP address
- Autonomous nature of the domains
  - impossible to calculate meaningful costs for a path crossing multiple ASs
  - A cost of 1000 is great at provider 1, terrible at provider 2
- Issues of trust
  - Provider A might be unwilling to believe certain advertisements from provider B

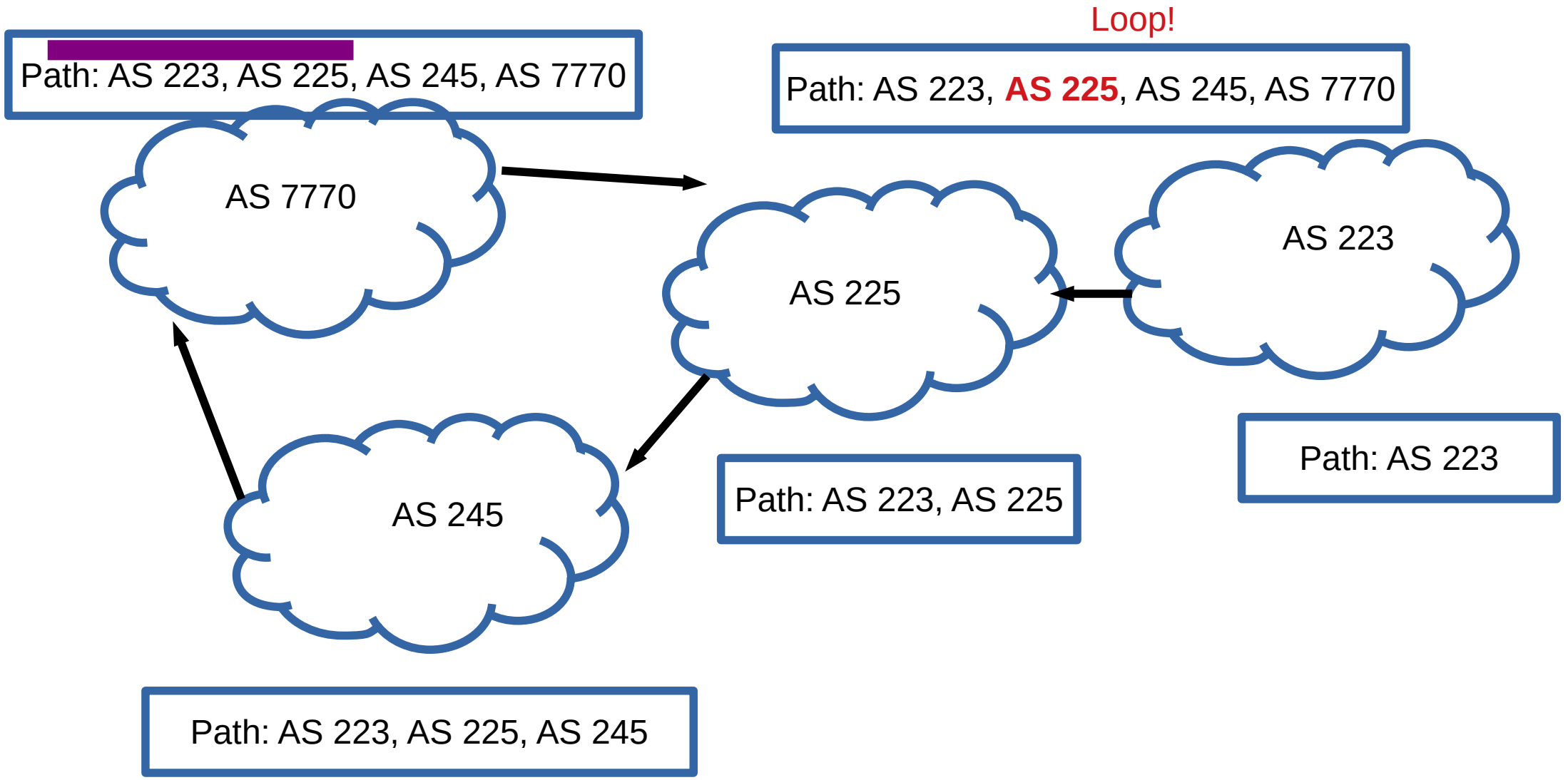
BGP: Path vector protocol

- 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 advertizing

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: Interconnections

- Uses TCP port 179 to connect to **peers**
- Arbitrary connections between AS's
- Advantages:
  - Much simpler, no periodic update
  - Valid as long as TCP connection is valid (or withdrawn)
  - Incremental update (only a portion of the routing table)
- Disadvantages:
  - No security
  - Congestion control on routing messages

BGP: Security problems

Anyone can advertise anything!!!

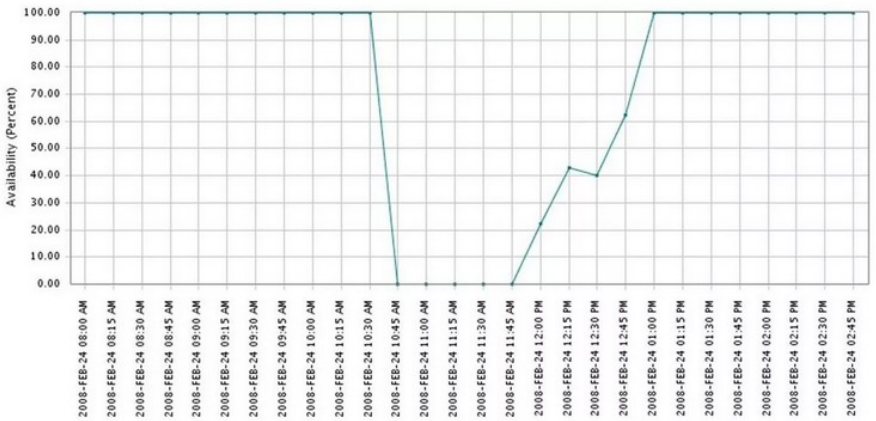
c|net BEST PRODUCTS REVIEWS NEWS VIDEO HOW TO SMART HOME CARS DEALS DOWNLOAD JOIN / SIGN IN

CULTURE

How Pakistan knocked YouTube offline (and how to make sure it never happens again)

YouTube becoming unreachable isn't the first time that Internet addresses were hijacked. But if it spurs interest in better security, it may be the last.

BY DECLAN MCCULLAGH | FEBRUARY 25, 2008 4:28 PM PST



This graph that network-monitoring firm Keynote Systems provided to us shows the worldwide availability of YouTube.com dropping dramatically from 100 percent to 0 percent for over an hour. It

BGP: Hop by Hop model

- 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

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: Allows for policy

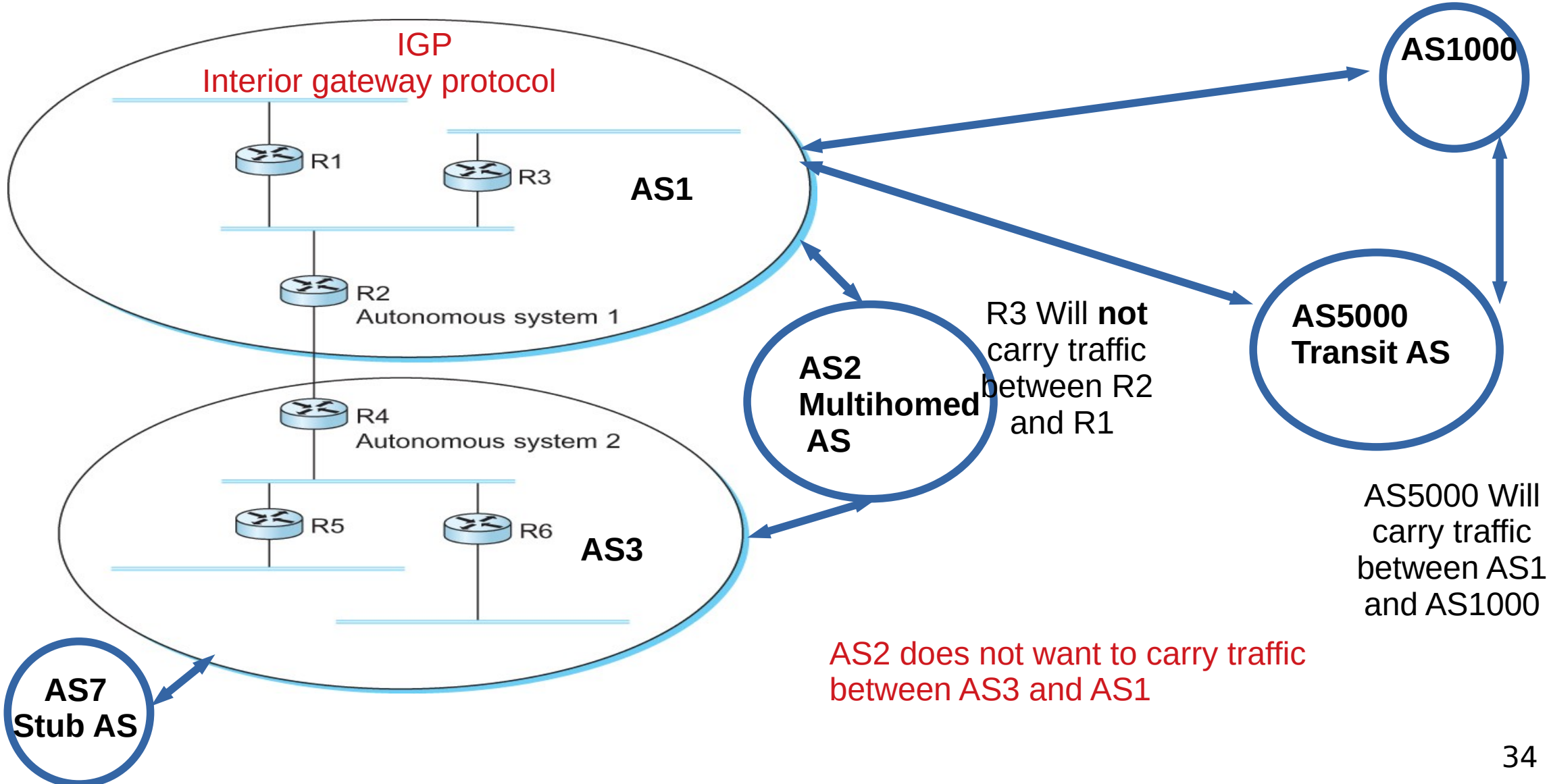
- Capable of enforcing various policies
  - AS2 → Don’t use AS1 to get to AS3
- Not part of BGP – configuration information that controls propagation of paths

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?



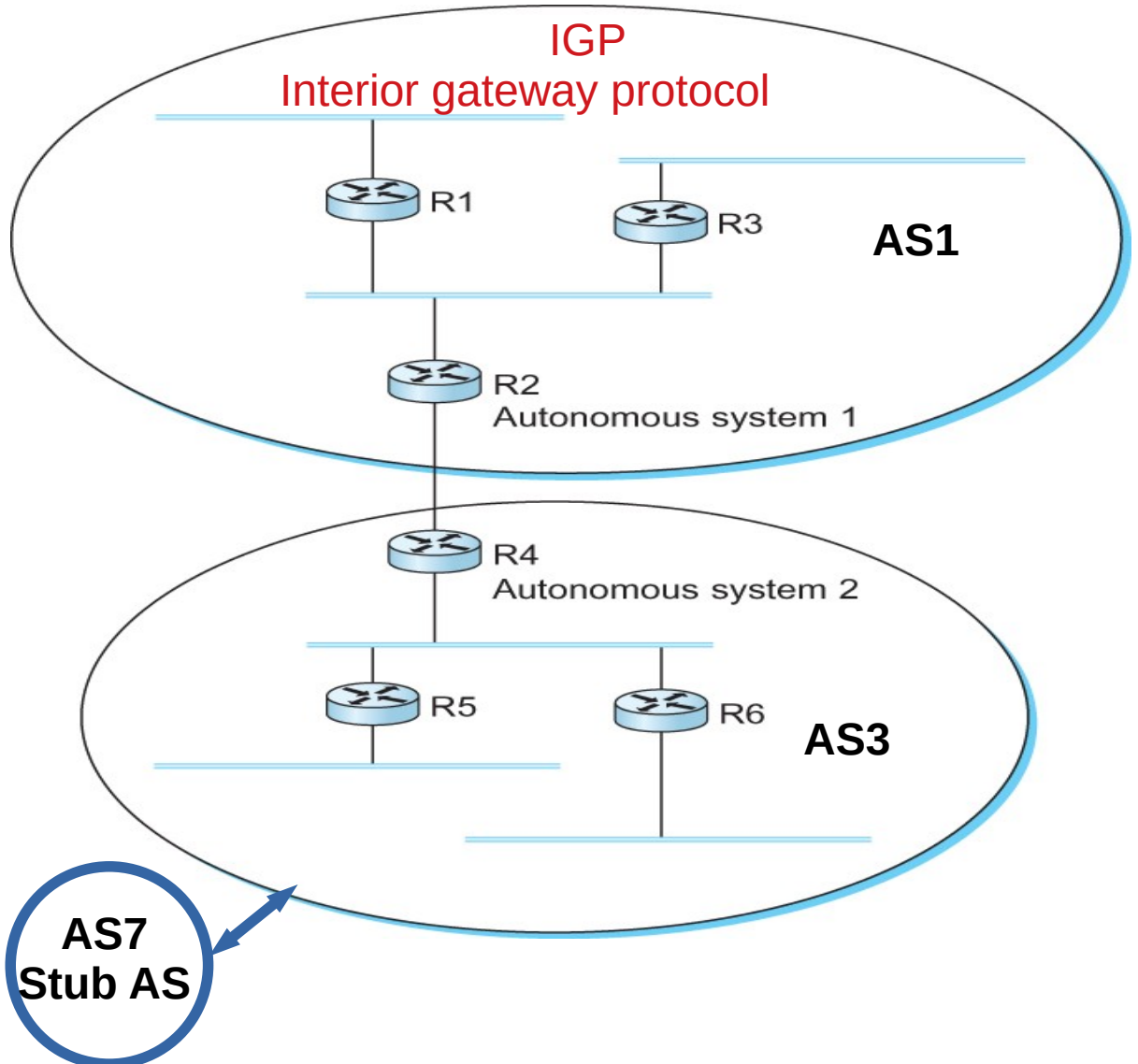
Examples BGP Policies

- 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 late

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 late

You don’t need BGP for Stub ASes

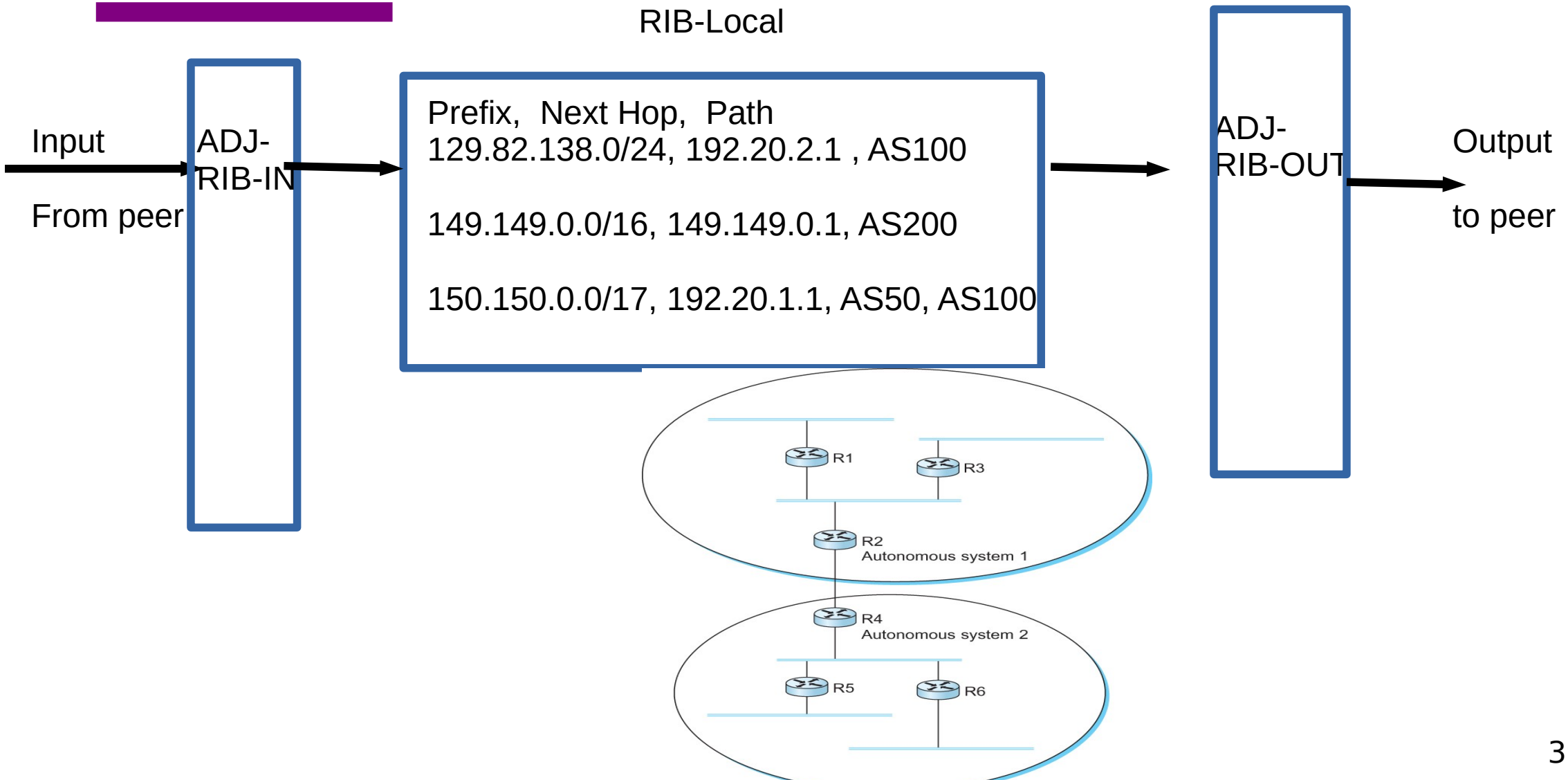


Default IP route should be sufficient

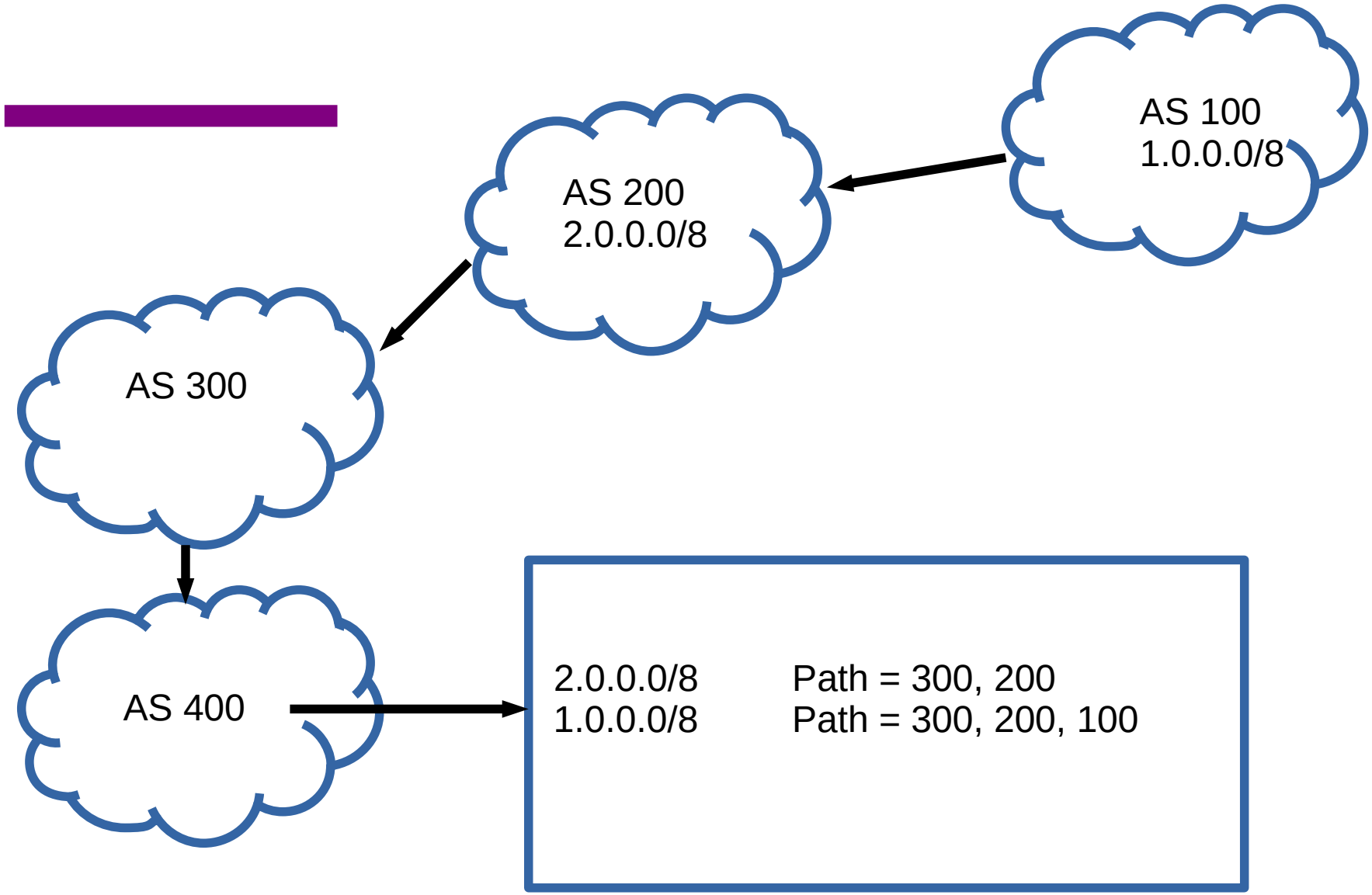
BGP Messages

- 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

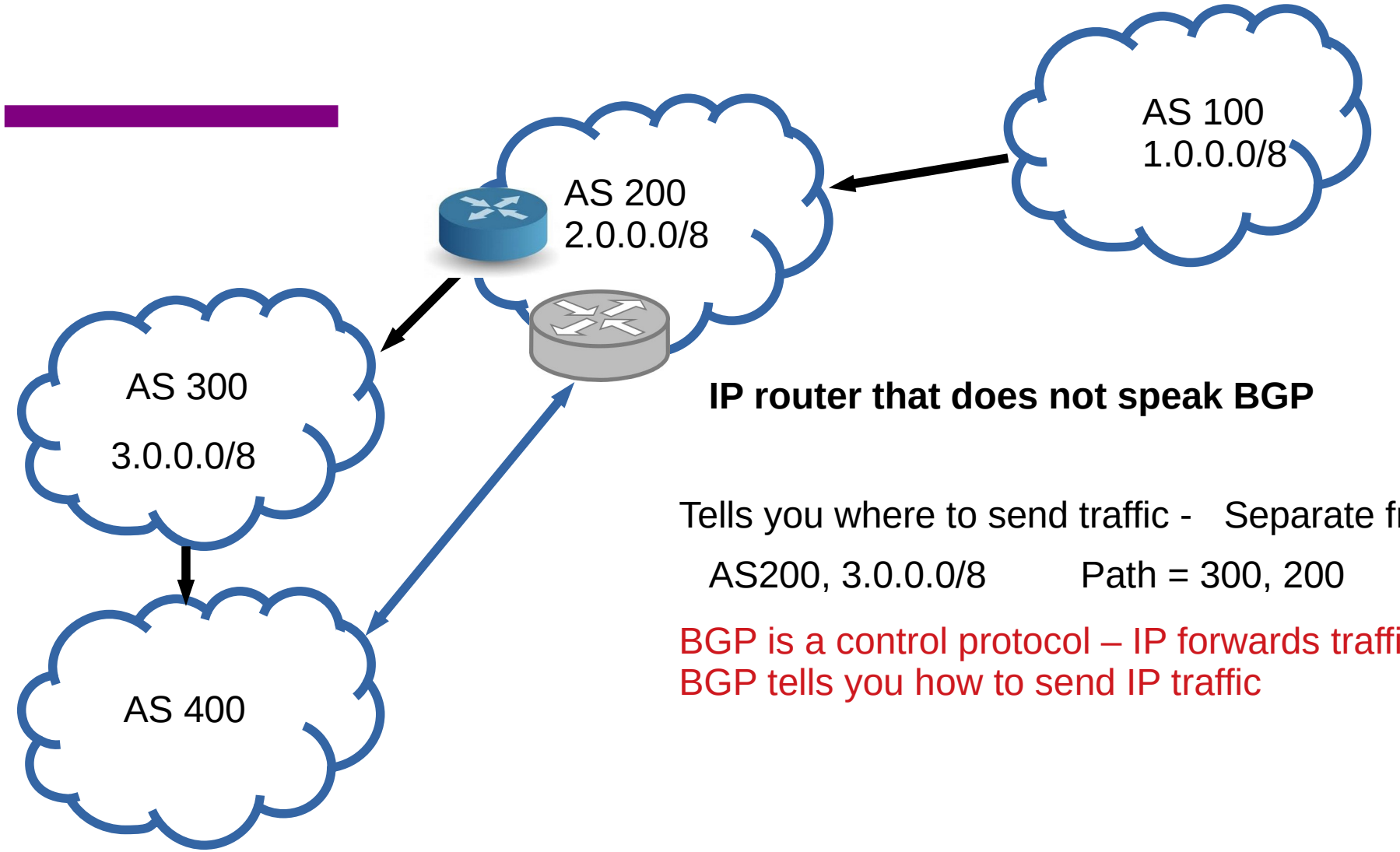
Routing Information Bases (RIB)



BGP Attributes - AS_PATH



BGP Attributes – Next hop?



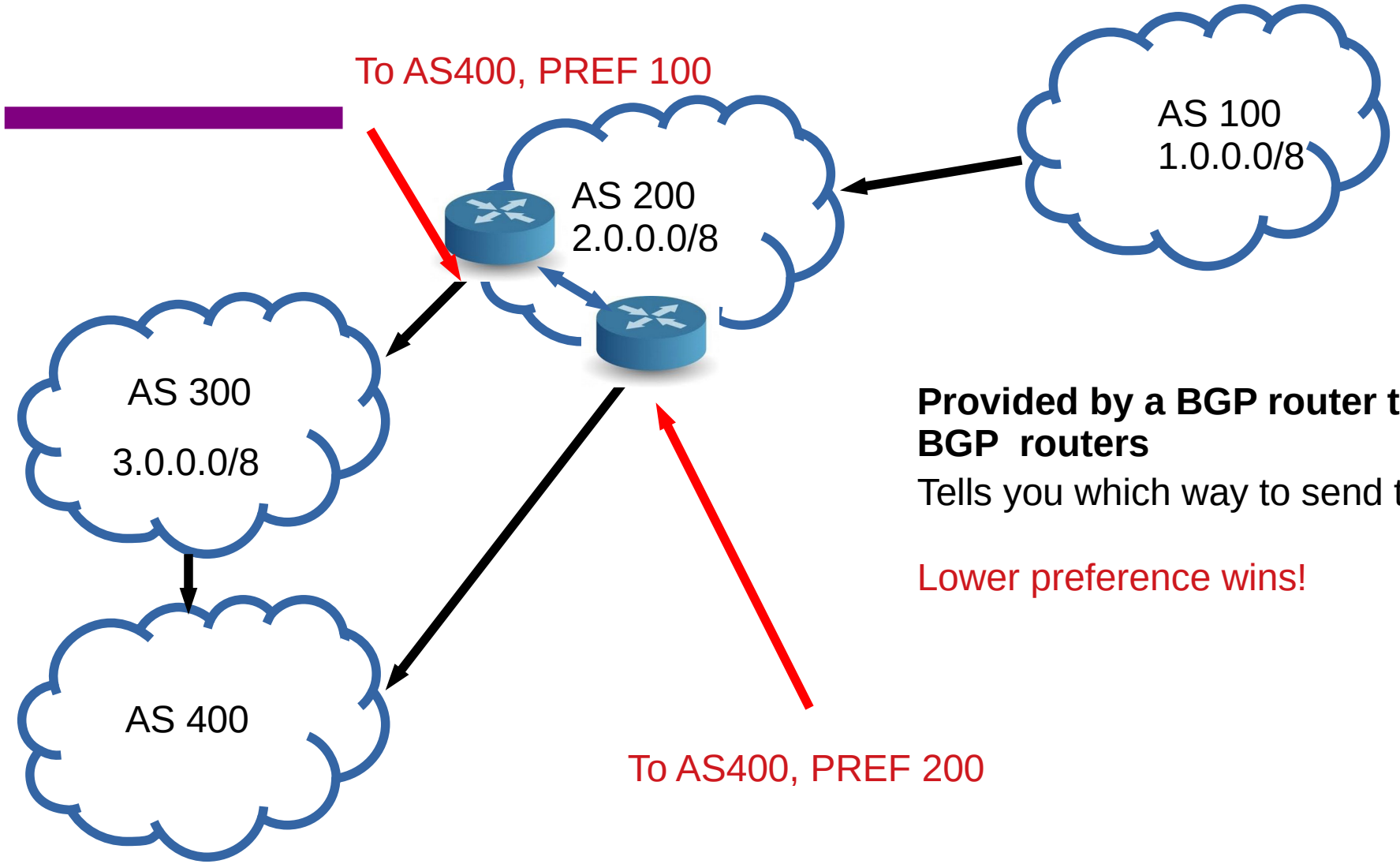
IP router that does not speak BGP

Tells you where to send traffic - Separate from BGP

AS200, 3.0.0.0/8 Path = 300, 200

BGP is a control protocol – IP forwards traffic
BGP tells you how to send IP traffic

BGP Attributes – LOCAL-PREF

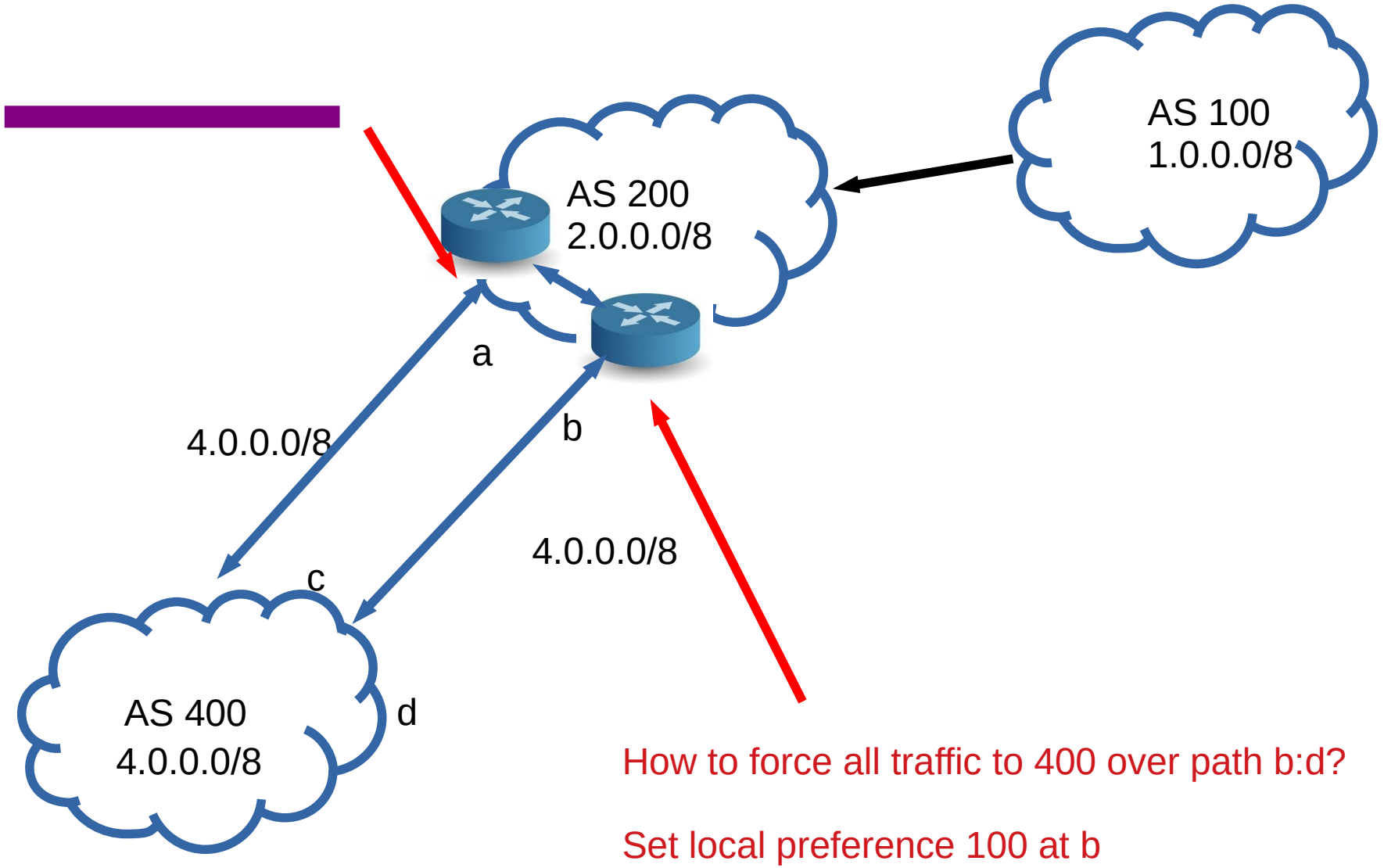


Provided by a BGP router to all other *internal* BGP routers

Tells you which way to send traffic

Lower preference wins!

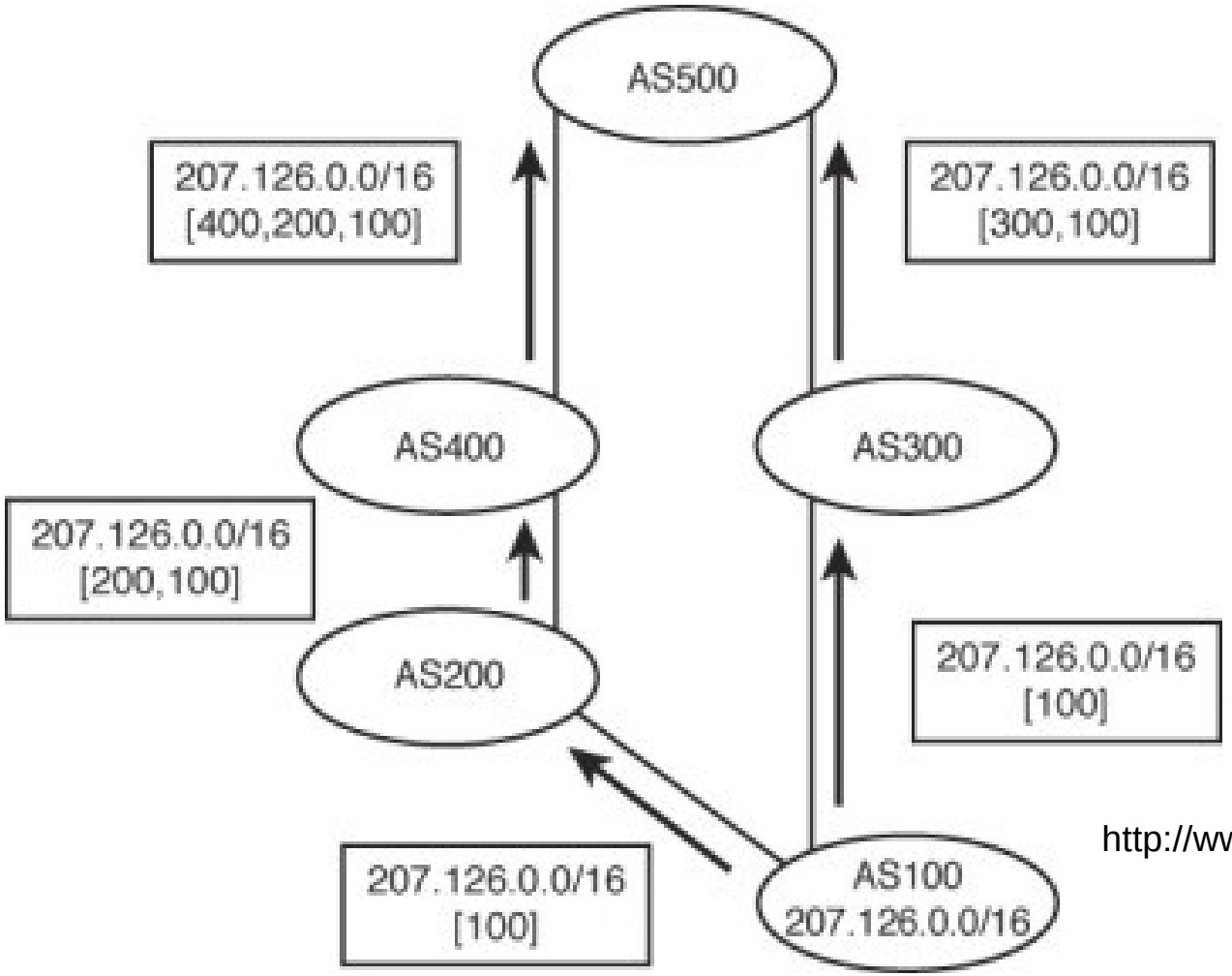
BGP Attributes – LOCAL-PREF Example



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

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

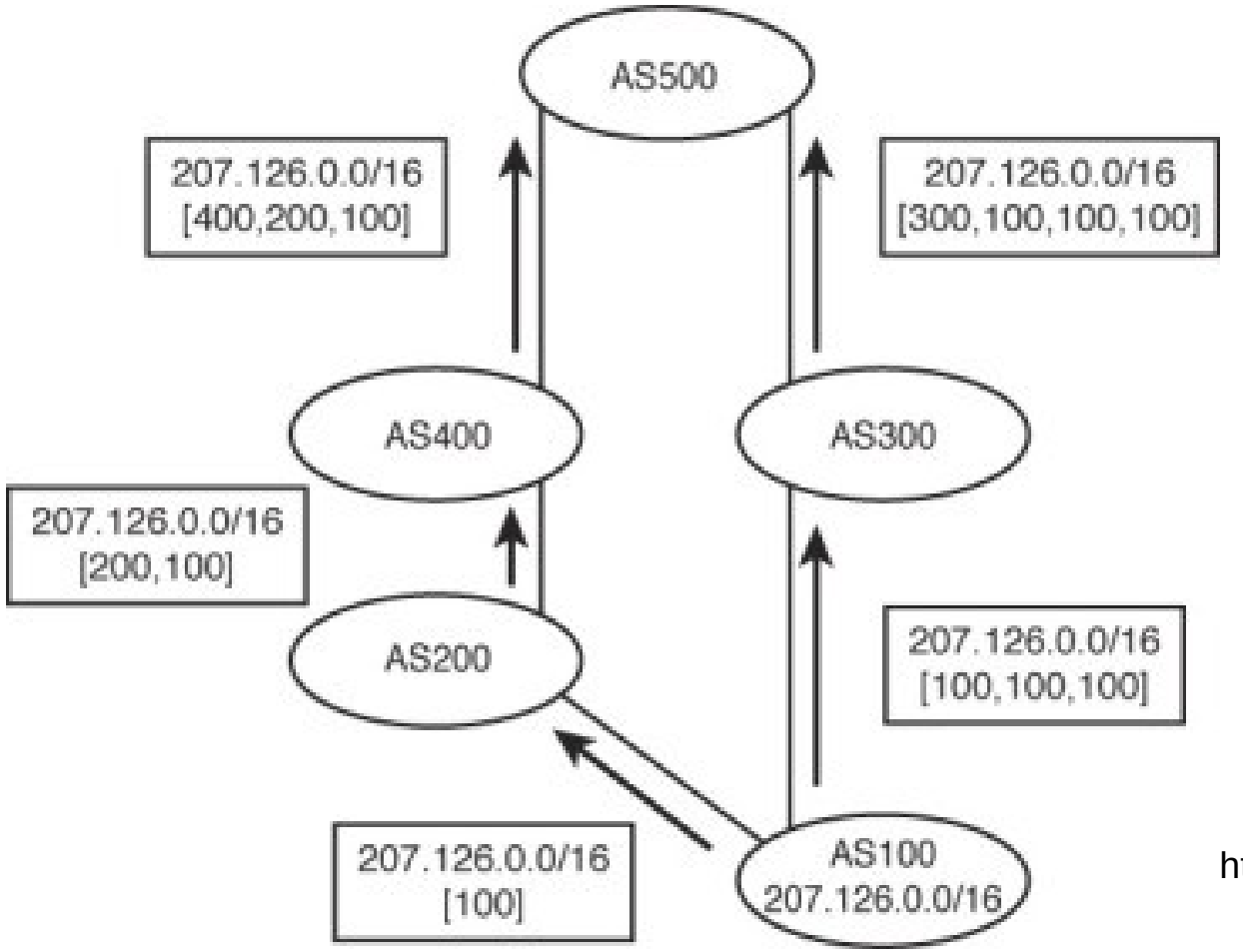
BGP Attribute – AS PATH



Each hop adds ASN to the path
-Only externally

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

BGP Attribute – AS PATH

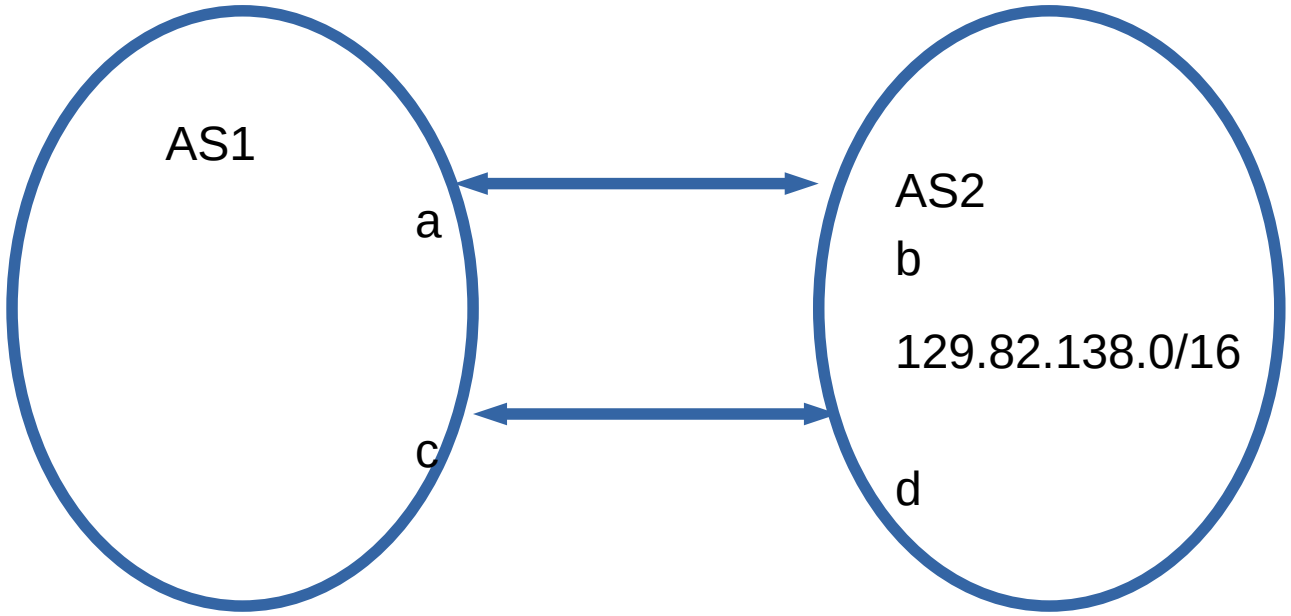


AS100 trying to influence path selection at AS500

- 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?

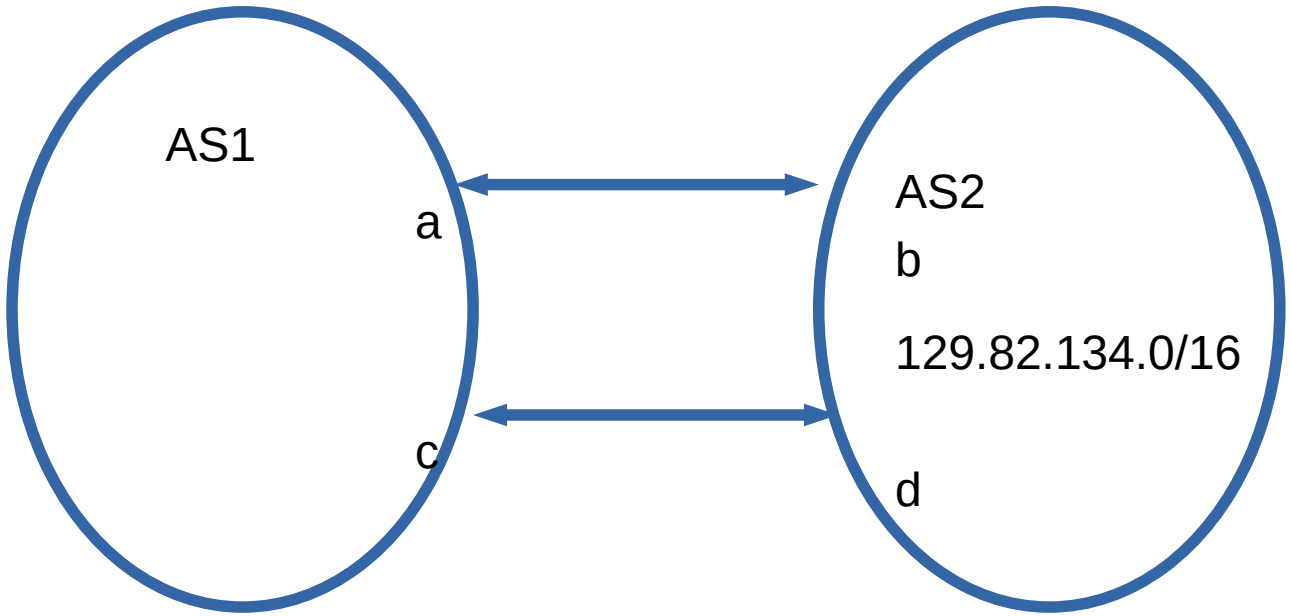
At A:

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

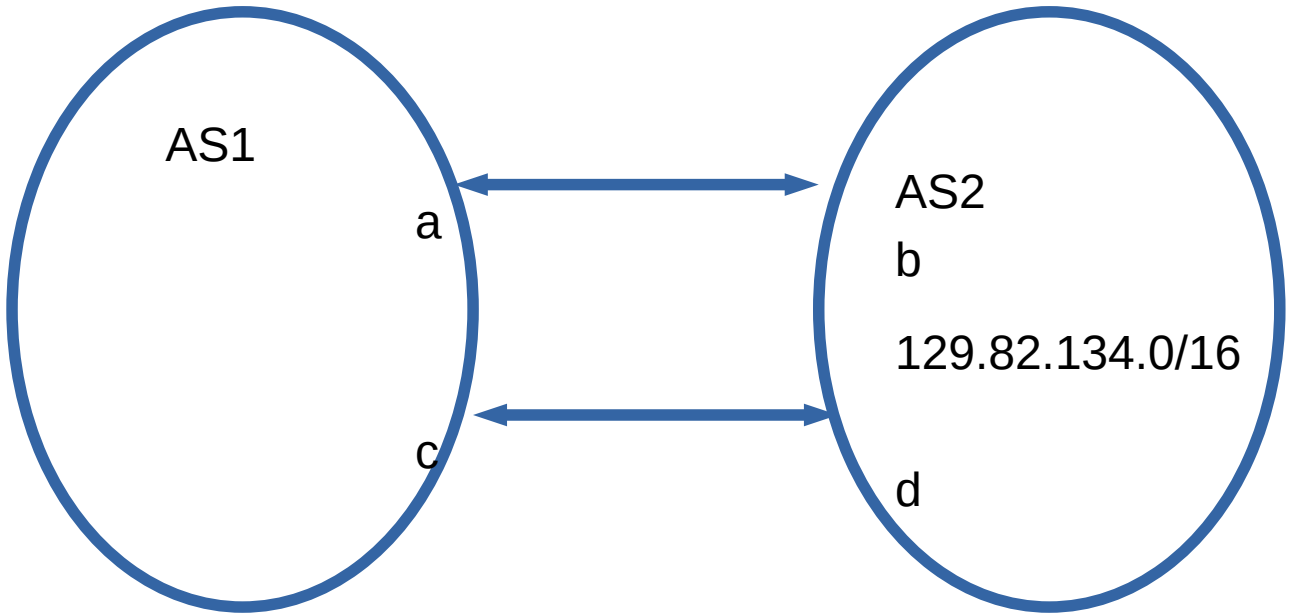
BGP Attribute – Local Preference



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

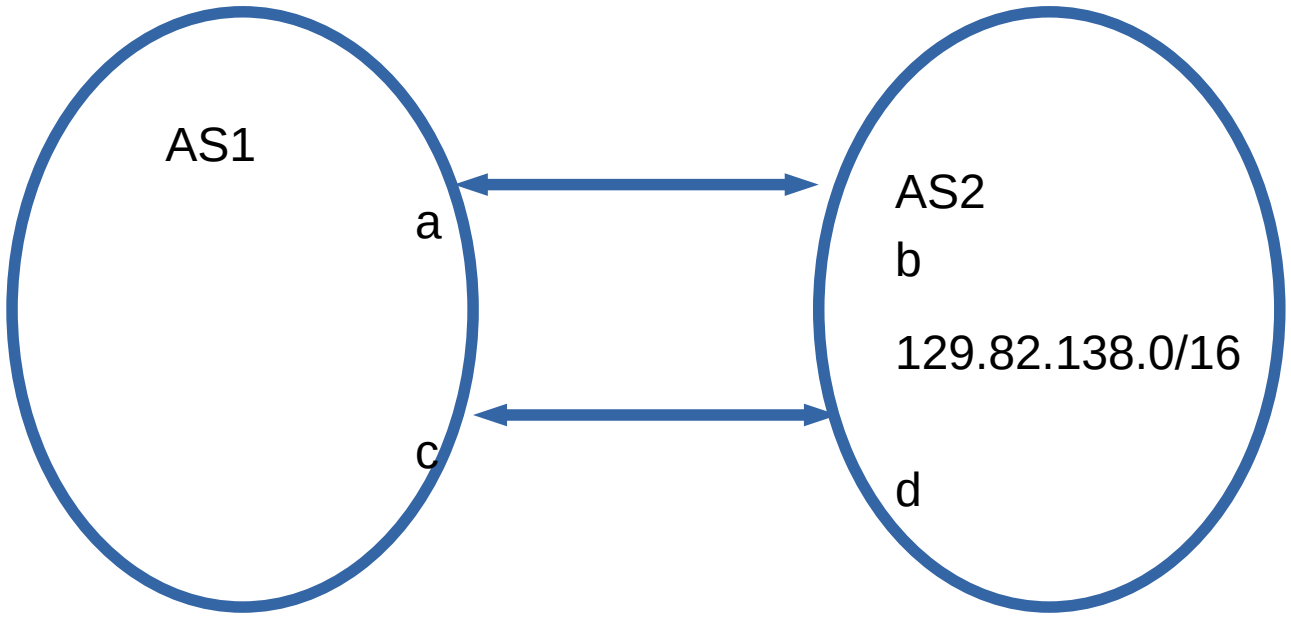


AS1 and AS2 has two paths between them

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

Lower cost wins

BGP Attribute – MED



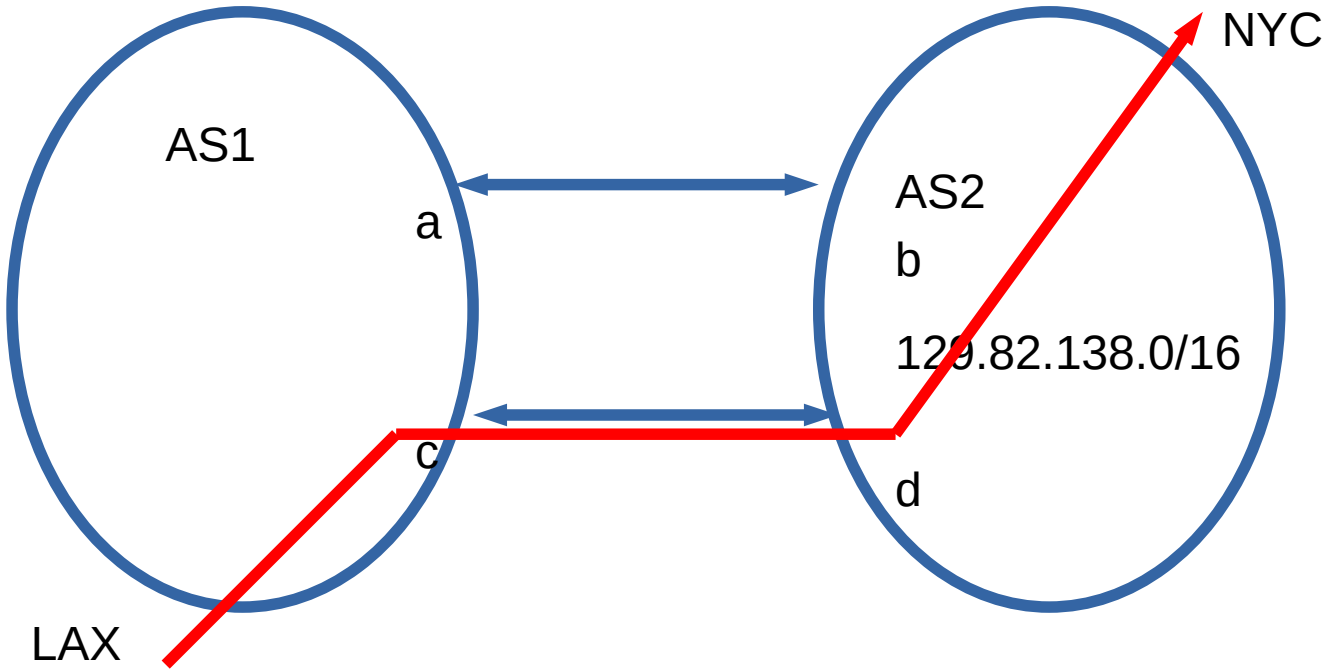
How would AS1 make AS2 send
129.82.138.0/17 over a-b
and
129.82.138.128/17 over c-d?

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

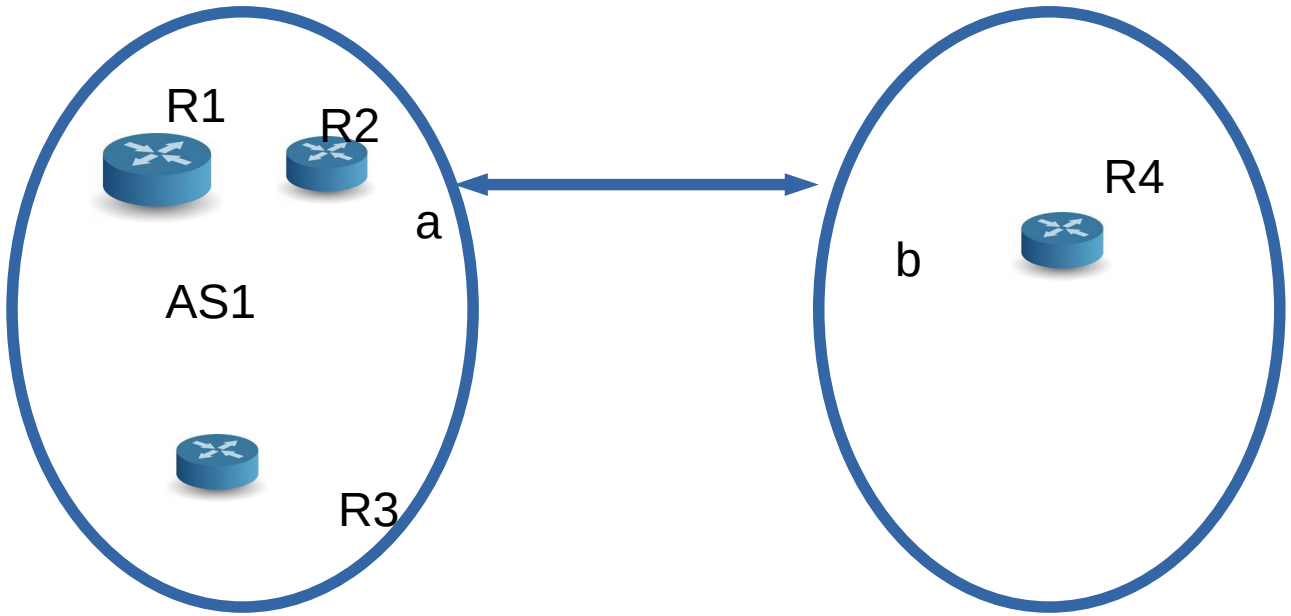
MED → External – you tell you neighbors which route you prefer
Neighbor is an autonomous system, so it can ignore you

BGP Attribute - Community

Put anything you want – between Ases, not known publicly

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

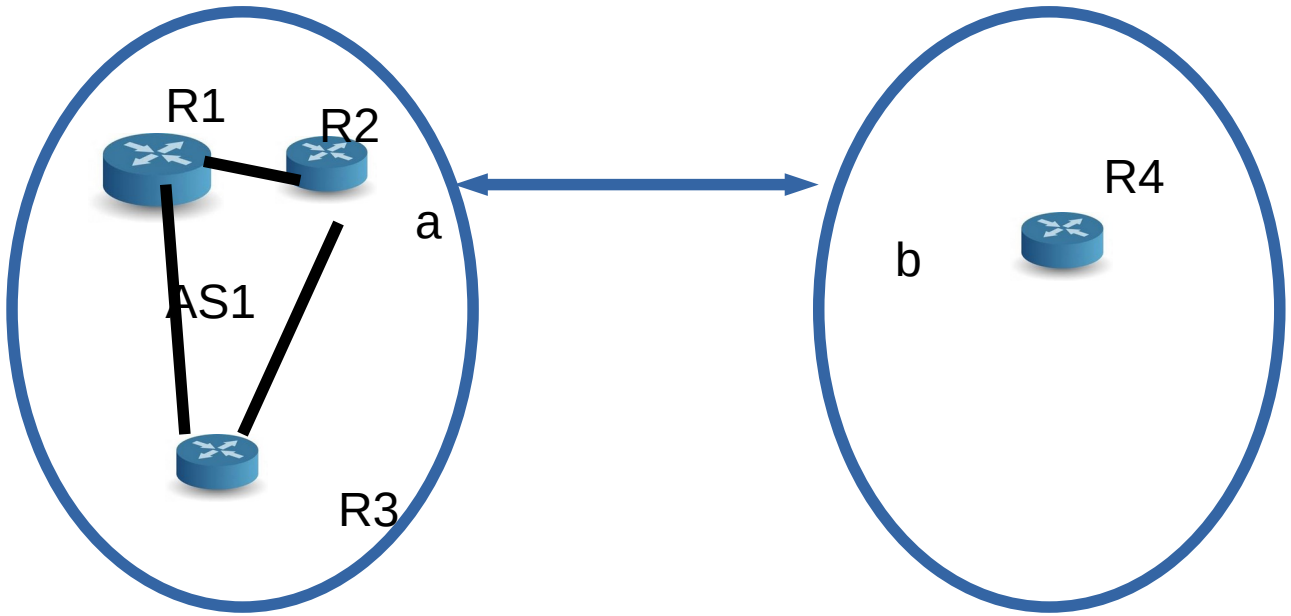
Internal vs External BGP



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

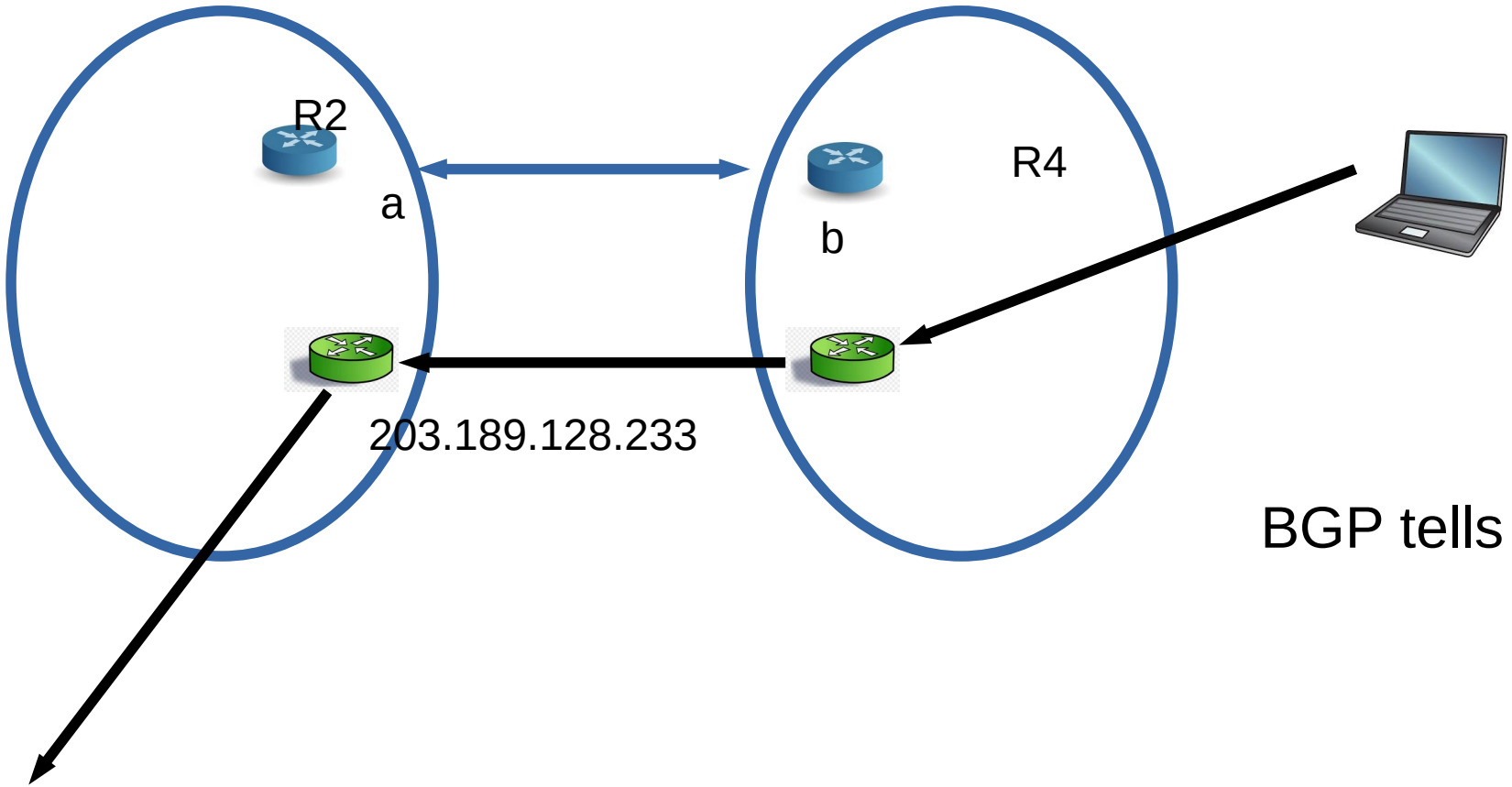
R2 can tell R3 and R1 about R4
R2 can not tell R1 about prefixes from R2 -loop!

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

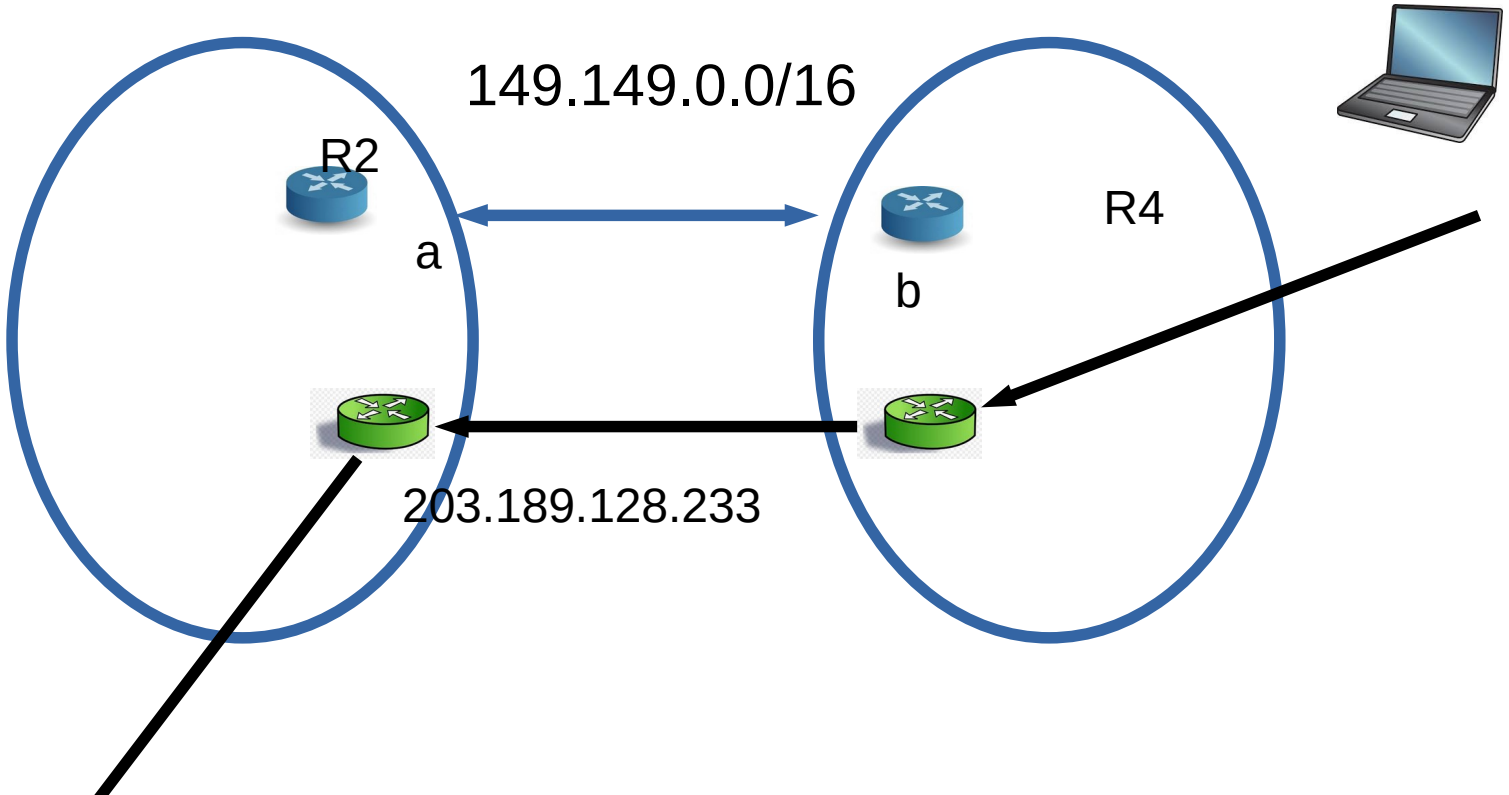


BGP tells you which IP router to use

BGP Decision process

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

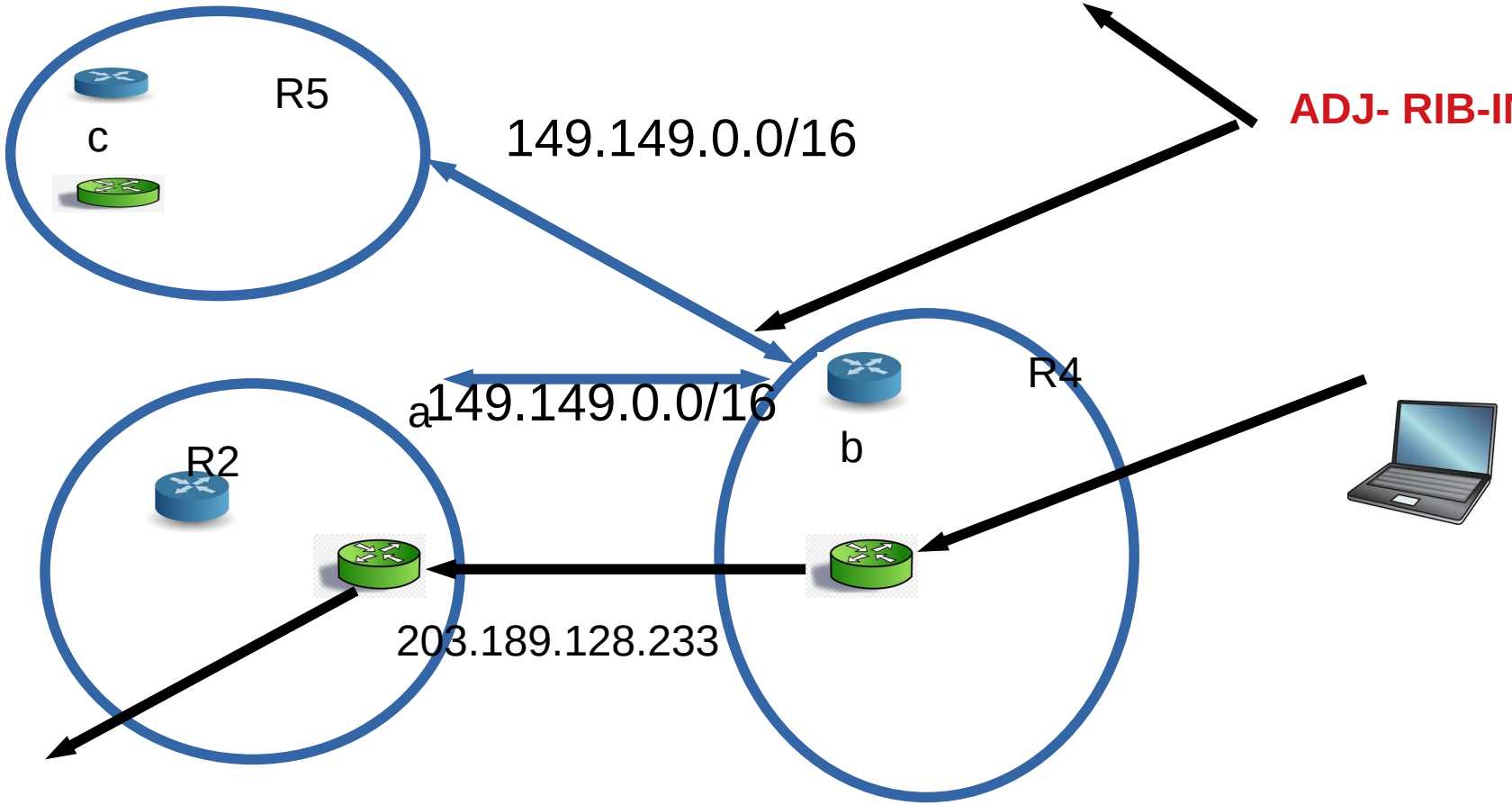
203.189.128.233 | 23673 | 149.149.0.0/16 | 23673 1299 | 10 | 5| 100
203.189.128.233 | 23673 | 149.149.0.0/16 | 23673 1299 | 100 | 50| 10



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

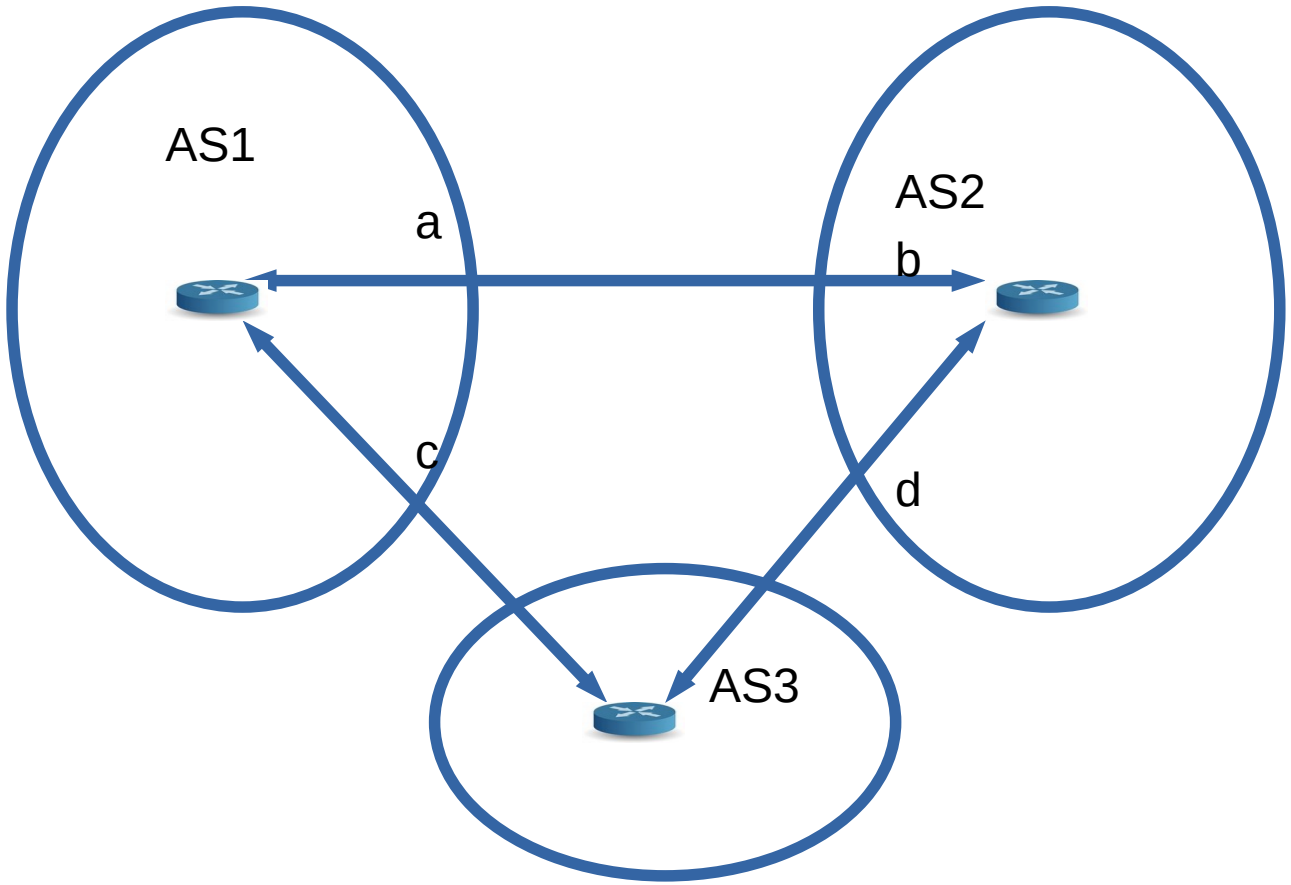
At ADJ-RIB-IN calculate degree of preference until one route for each destination remains!!

- 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 Assignments

- Scaling to billions:
  - https://book.systemsapproach.org/scaling/problem.html#problem-scaling-to-billions
    - ~2 minutes read

- Global Internet
  - https://book.systemsapproach.org/scaling/global.html#global-internet
  - Skip the routing areas section
  - Read until “Common AS Relationships and Policies”
  - ~40 minutes
- BGP - - https://book.systemsapproach.org/scaling/global.html#interdomain-routing-bgp
- - 30-40 minutes read