NANYANG TECHNOLOGICAL UNIVERSITY

CE3005: Computer Networks

Module 2-3:
Network Layer - IP Routing Protocols

Semester 1 2016-2017

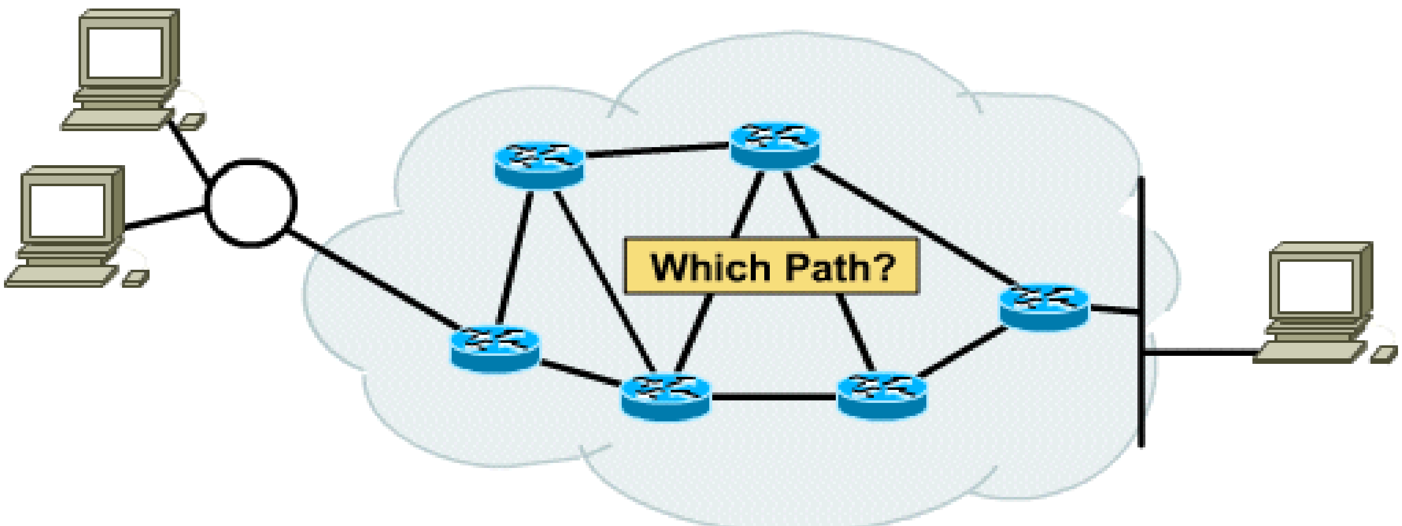
School of Computer Engineering

Contents

- **Internet Routing**
  - **Concept of Autonomous System (AS)**
  - **Intra-AS and Inter-AS Routing**
- **Intra-AS Routing**
  - **Distance Vector Routing**
    - **e.g. Routing Information Protocol (RIP)**
  - **Link State Routing**
    - **e.g. Open Shortest Path First (OSPF)**
- **Inter-AS Routing**
  - **Path Vector Routing**
    - **e.g. Border Gateway Protocol (BGP)**

Router

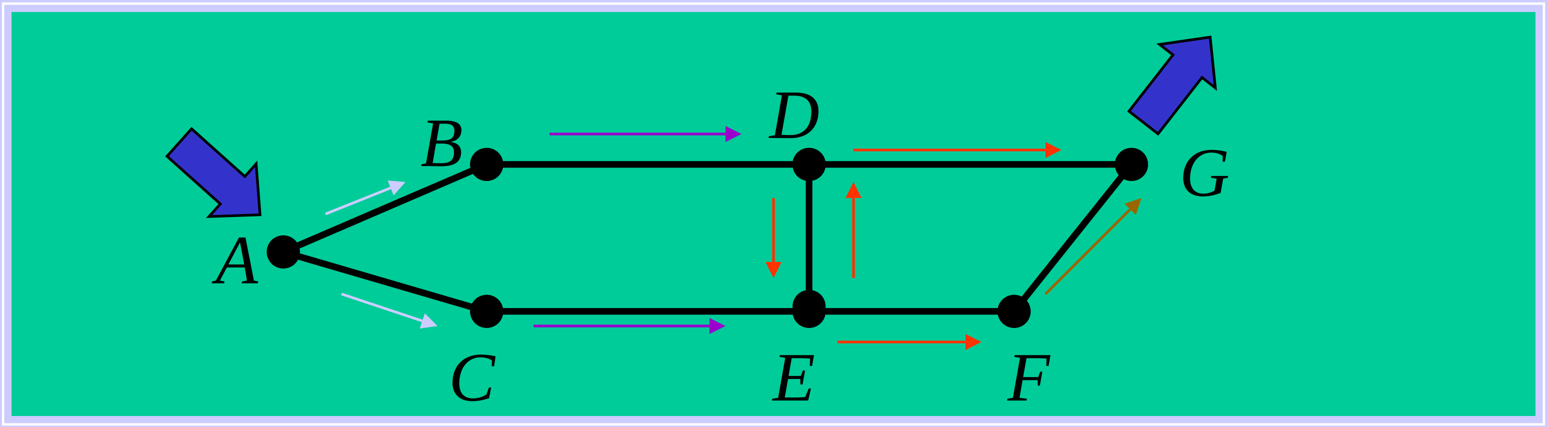
A **router** is a device used to interconnect networks, and to forward packets by **examining** the **destination address** in the **IP header** of each packet.



How does a router decide which path to forward?
>> Routing Protocol, used to initialize/update routing table so that a route/path can be determined.

Routing: Flooding

- A straight forward way of routing is by **flooding**.
- When a node receives a packet, it will forward the packet to all other links except the incoming link. The packet will be labeled with a unique identifier
- Should the same packet return, the packet will be discarded



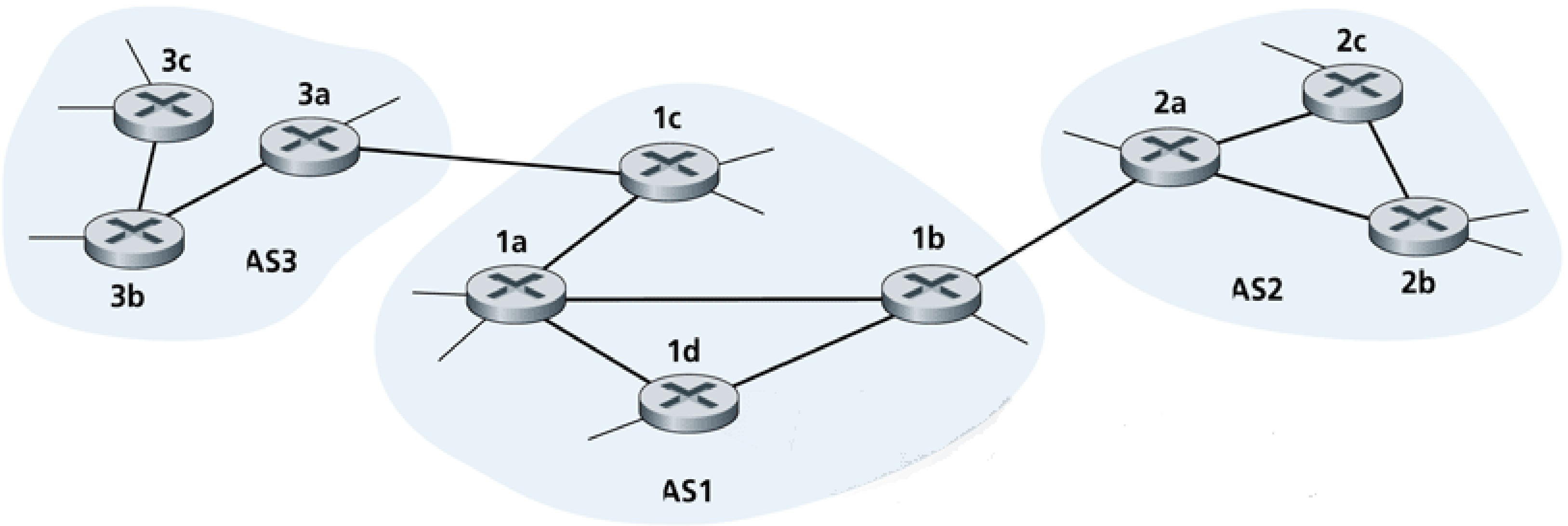
Packets transmitted using flooding is 9

Routing: Flooding

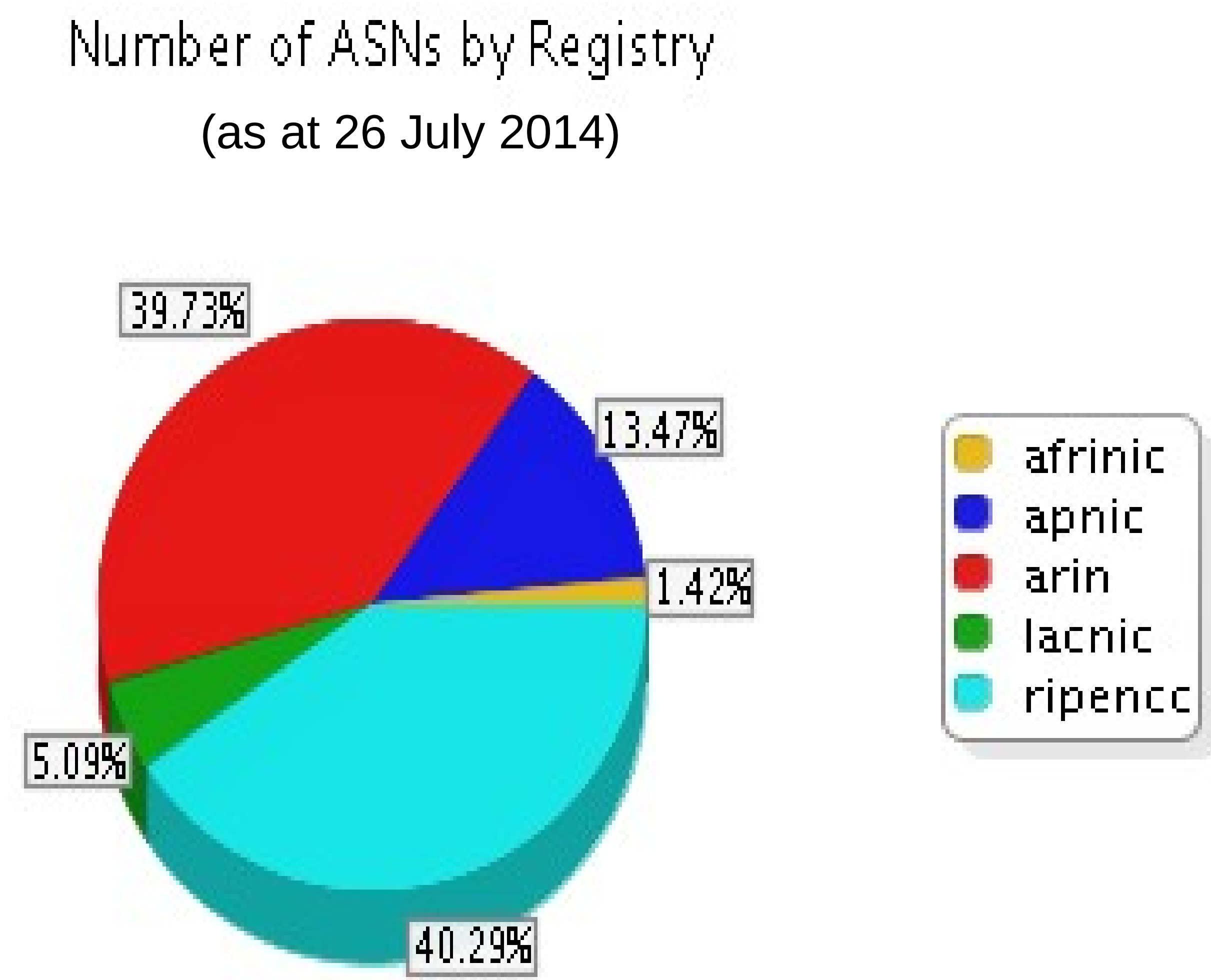
- **Advantages:**
  - **A packet will always get through if one or more path exists (very robust)**
- **Disadvantages:**
  - **Very wasteful of bandwidth, may cause serious congestion, hence not used in the Internet**
- **Applications:**
  - **Military applications (routers may be destroyed anytime)**
  - **Ad hoc wireless networks (nodes may be turned off or moved away anytime)**

Routing in the Internet

For routing purpose, Internet is divided into Autonomous Systems (AS). An AS is a group of routers under the authority of a single administration; e.g. an ISP.



Each AS is uniquely identified by a 2-byte (0 - 65,535) or 4-byte (65,536 - 4,294,967,295) ASN (AS number), which is assigned by IANA; e.g. ASN of NTU: AS9419.



(as at 26 July 2014)

http://www.apnic.net/publications/research-and-insights/stats/asn-geographic

Intra-AS and Inter-AS Routing

In practice, routing in Internet is done in a hierarchical manner, which includes intra-AS and inter-AS routings.

Intra-AS Routing:

- Routing within an AS
- Protocols for Intra-AS routing are also called Interior Gateway Protocols (IGPs)
- Different AS can choose to run their preferred protocols
- e.g. intra-AS routers: 1a, 1b, 1c, 1d in AS1

Inter-AS Routing:

- Routing between AS
- Protocols for Inter-AS routing are also called Exterior Gateway Protocols (EGPs)
- All AS must run the same standard protocol
- e.g. inter-AS routers: 1b, 2a, 1c, 3a

Why different Intra- and Inter-AS Routing?

- **Policy:**
  - **Inter-AS: different admin wants control over how its traffic are forwarded, who routes through its network**
  - **Intra-AS: single admin, so no policy decision needed**
- **Performance:**
  - **Intra-AS: can focus on performance**
  - **Inter-AS: policy may dominate over performance**
- **Scale:**
  - **Internet is too large to be treated as a single routing domain**

REGISTERED USERS: 17,857

UltraTools[SM]
Powered by Neustar® UltraDNS®

Email Address | Password | Login
Remember me | Forgot your password?

Home | Domain Health Report | NEW! WHOIS+ | NEW! Monitoring | UltraTools | Statistics | Community | UltraTools Mobile | Create Free Account

Email | Share

IP Tools

- Decimal IP Calculator
- **ASN Information**
- CIDR/Netmask
- What's your IP
- IP Geo-location Lookup
- IPWHOIS Lookup

ASN Information

The ASN Information tool provides complete autonomous system (AS) information.

Enter an AS number, IP address, or a Company name.

AS4538 | Go »

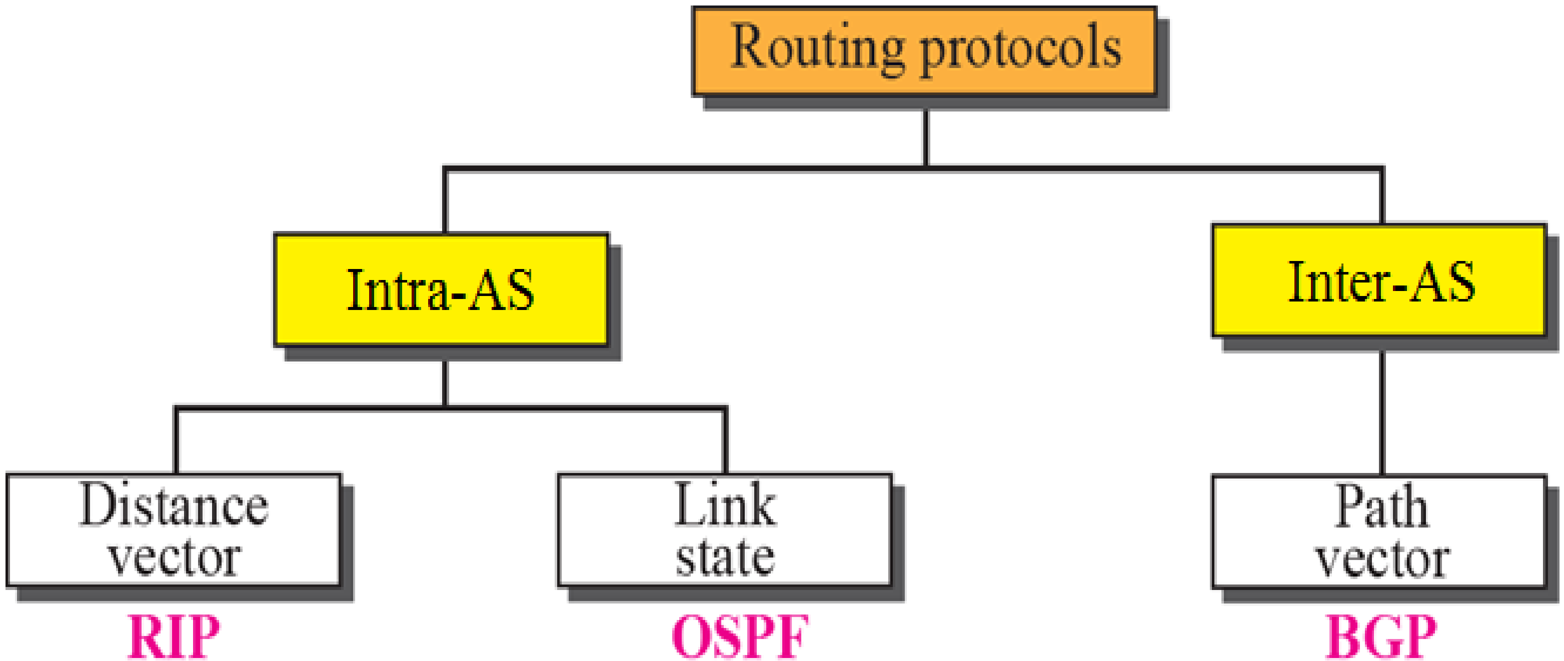
Related Tools: CIDR/Netmask What's your IP Decimal IP Calculator

AS4538

```
          Country: CN
Registration Date: 2002-08-01
        Registrar: apnic
            Owner: ERX-CERNET-BKB China Education and Research Network Center
```

www.ultratools.com/

Intra-AS and Inter-AS Routing Protocols



Intra-AS Routing:
Distance Vector Routing

Distance Vector routing also known as "*Bellman-Ford*" or "*old ARPANET*" routing.

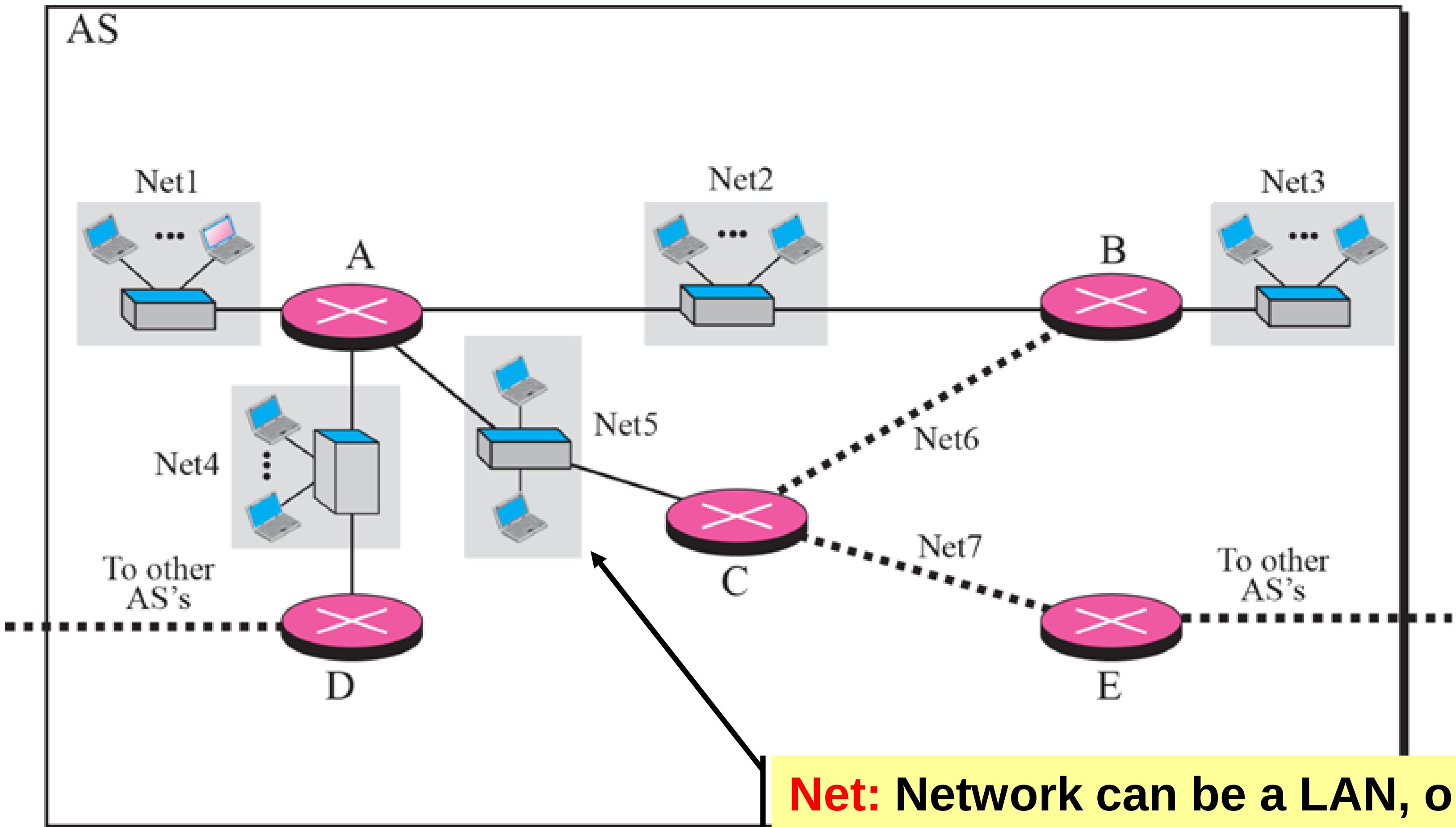
Essentially, consists of 3 main stages:

➊ **Discover neighbors by multicasting request**

➋ **Exchange distance vectors (routing information) with immediate neighbors only**
- **Response to request**
- **Periodic updates (typically 30s interval)**
- **Triggered updates due to changes**

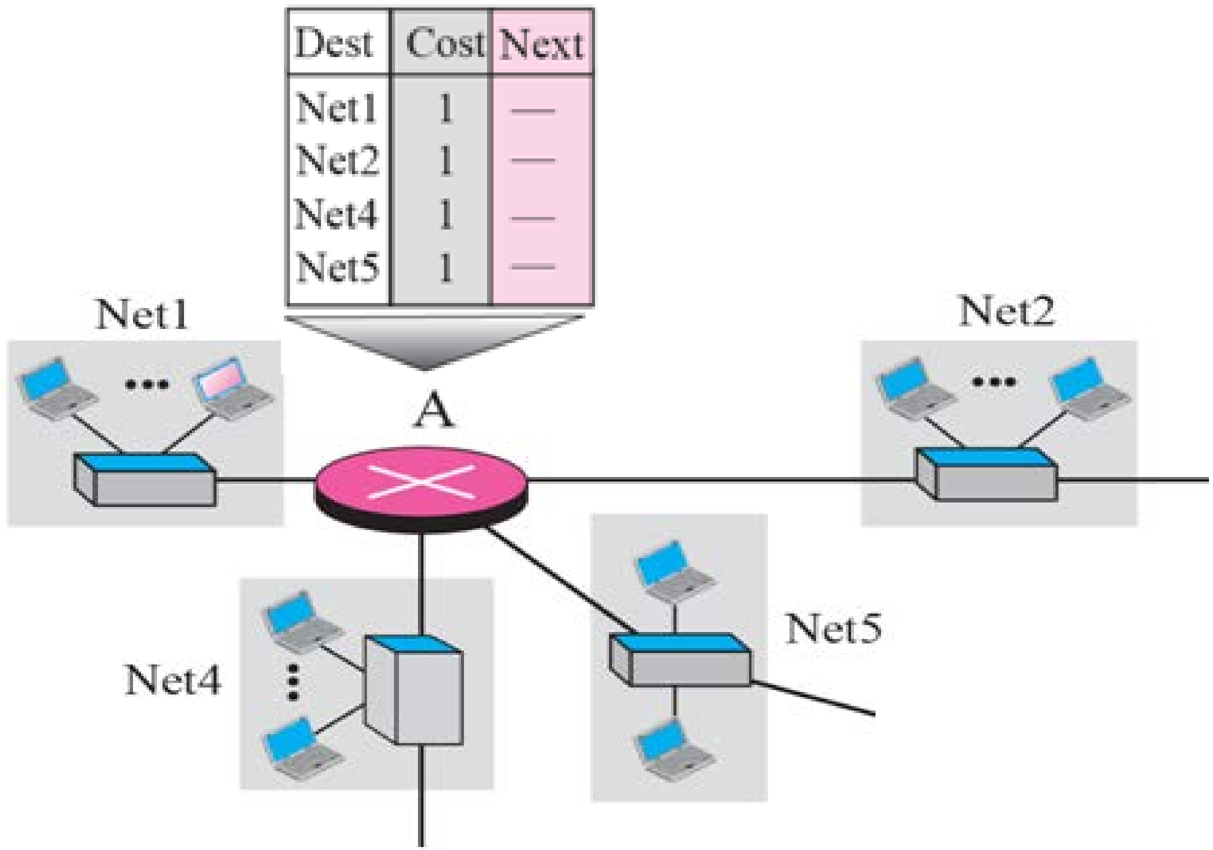
➌ **Compute shortest-path routes (using Bellman-Ford algorithm)**

An example of Distance Vector Routing



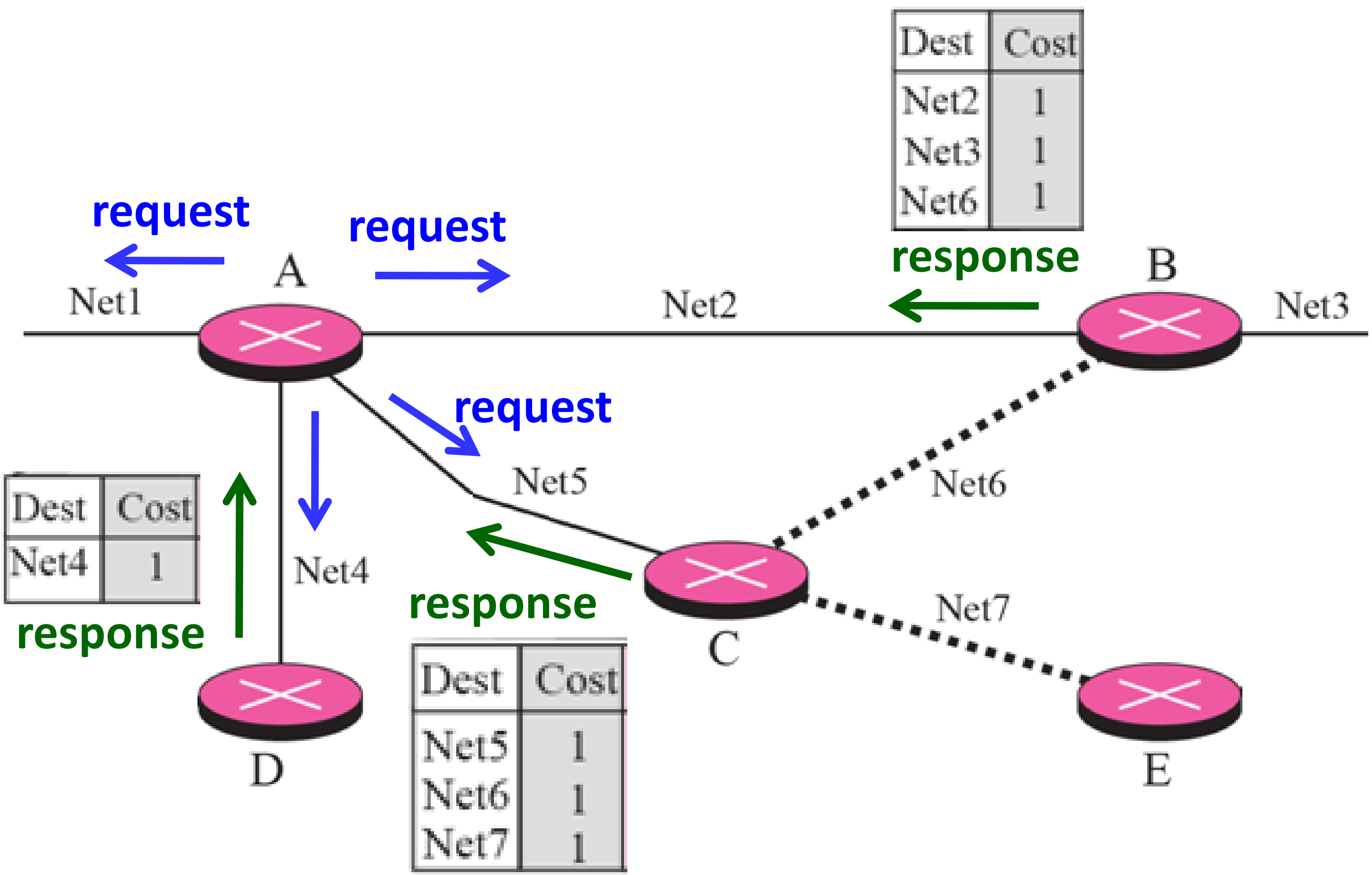
Net: **Network can be a LAN, or a point-to-point link**

Distance Vector Routing: Initially, a router only has its own configured routing table.

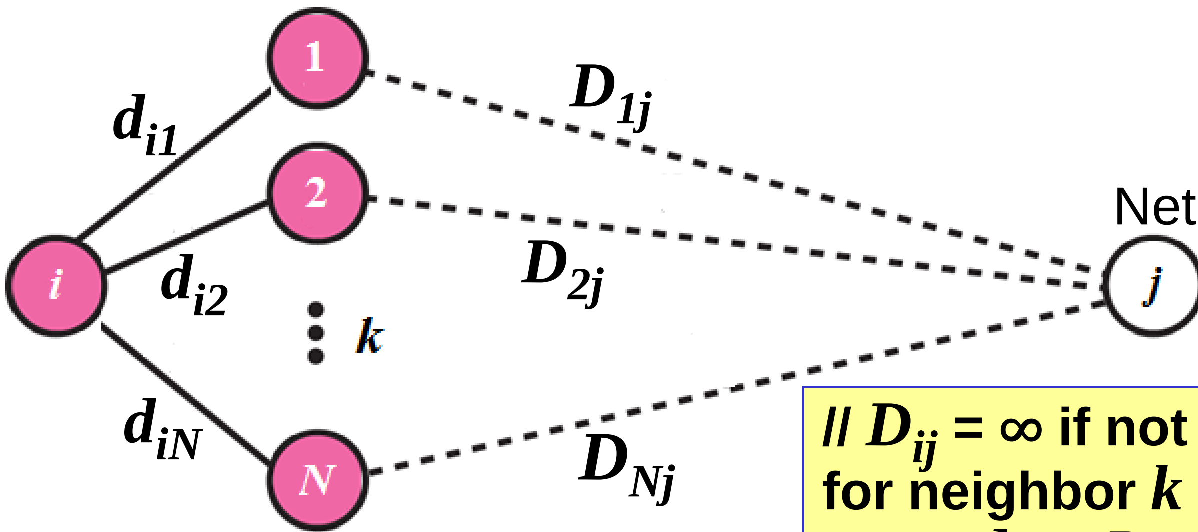


For simplicity but without loss of generality, assume each link has a cost of 1.

Distance Vector Routing: Discover Adjacent Neighbors and exchange distance vectors.



Distance Vector Routing: Computing Shortest-Path using Bellman-Ford Algorithm



// D_{ij} = ∞ if not reachable initially
for neighbor k = 1 to N {
if ($d_{ik} + D_{kj} < D_{ij}$) {
$D_{ij} = d_{ik} + D_{kj}$;
}}

d_{ik} = cost of going directly from node i to adjacent neighbor node k ;

D_{ij} = least total cost of going from node i to destination j

Distance Vector Routing: Computing Shortest-Path using Bellman-Ford Algorithm

from B:

Dest	Cost
Net2	1
Net3	1
Net6	1

from C:

Dest	Cost
Net5	1
Net6	1
Net7	1

from D:

Dest	Cost
Net4	1

Dest	Cost	Next
Net1	1	—
Net2	1	—
Net4	1	—
Net5	1	—

original routing table of A

Dest	Cost	Next
Net1	1	—
Net2	1	—
Net3	2	B
Net4	1	—
Net5	1	—
Net6	2	C
Net7	2	C

updated routing table of A

Assume Cost of 1 between node

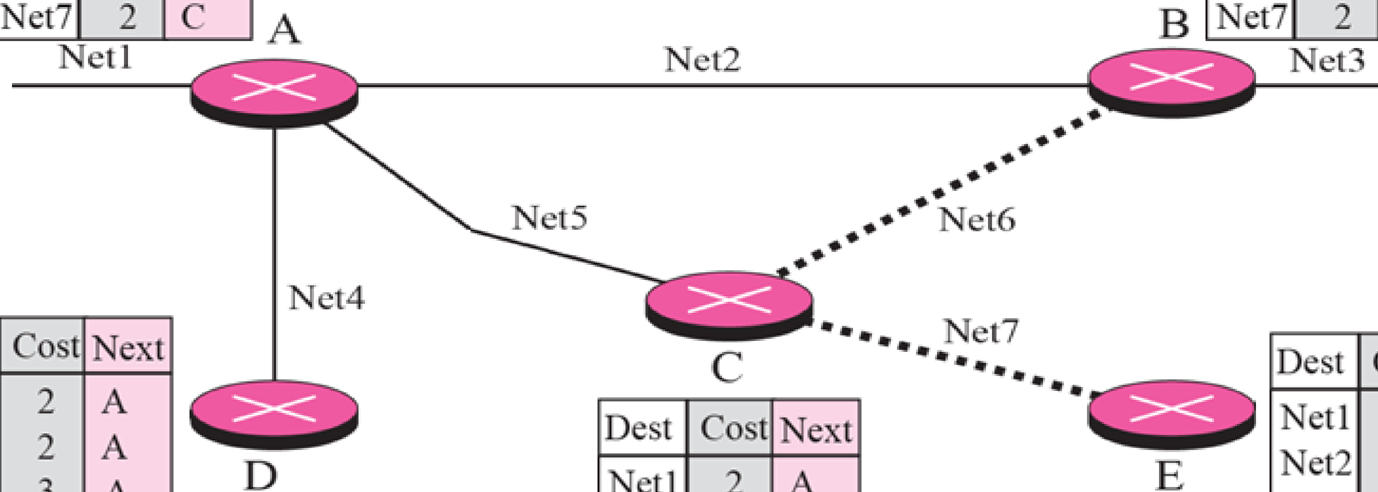
Distance Vector Routing: Resulting routing tables at each router after convergence.

A

Dest	Cost	Next
Net1	1	—
Net2	1	—
Net3	2	B
Net4	1	—
Net5	1	—
Net6	2	C
Net7	2	C

B

Dest	Cost	Next
Net1	2	A
Net2	1	—
Net3	1	—
Net4	2	A
Net5	2	A
Net6	1	—
Net7	2	C



D

Dest	Cost	Next
Net1	2	A
Net2	2	A
Net3	3	A
Net4	1	—
Net5	1	A
Net6	3	A
Net7	3	A

C

Dest	Cost	Next
Net1	2	A
Net2	2	A
Net3	2	B
Net4	2	A
Net5	1	—
Net6	1	—
Net7	1	—

E

Dest	Cost	Next
Net1	3	C
Net2	3	C
Net3	3	C
Net4	3	C
Net5	2	C
Net6	2	C
Net7	1	—

Distance Vector Routing: Count-to-Infinity problem

Before failure:

Dest	Cost	Next
Net1	1	-

Dest	Cost	Next
Net1	2	A

After failure:

Periodic update from B

Dest	Cost	Next
Net1	∞	-

Dest	Cost
Net1	2

Dest	Cost	Next
Net1	2	A

Distance Vector Routing: Count-to-Infinity problem

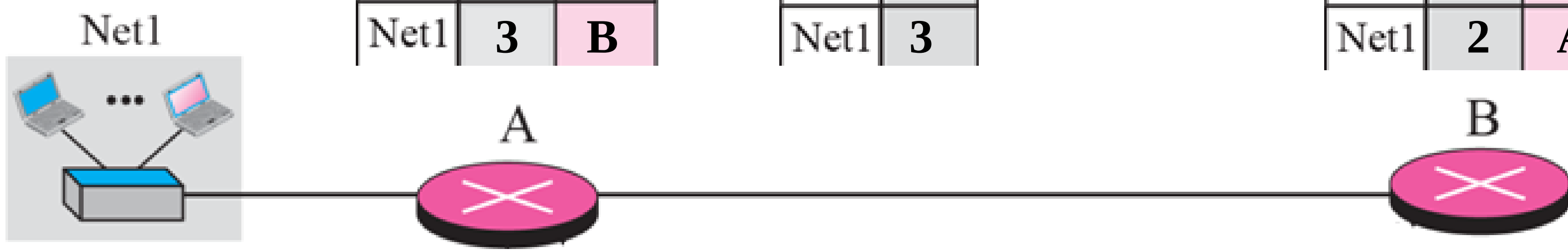
After A received update from B:

Triggered update from A →

Dest	Cost	Next
Net1	3	B

Dest	Cost
Net1	3

Dest	Cost	Next
Net1	2	A



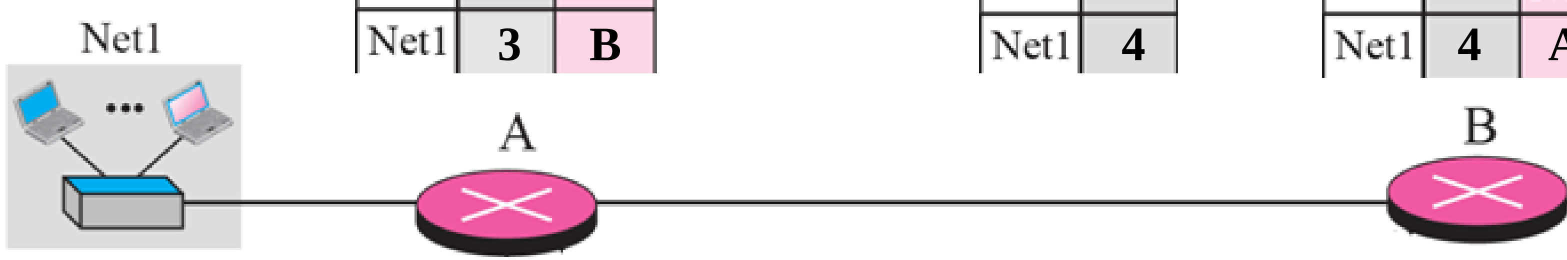
After B received update from A:

Triggered update from B ←

Dest	Cost	Next
Net1	3	B

Dest	Cost
Net1	4

Dest	Cost	Next
Net1	4	A



Distance Vector Routing: Count-to-Infinity problem

⋮

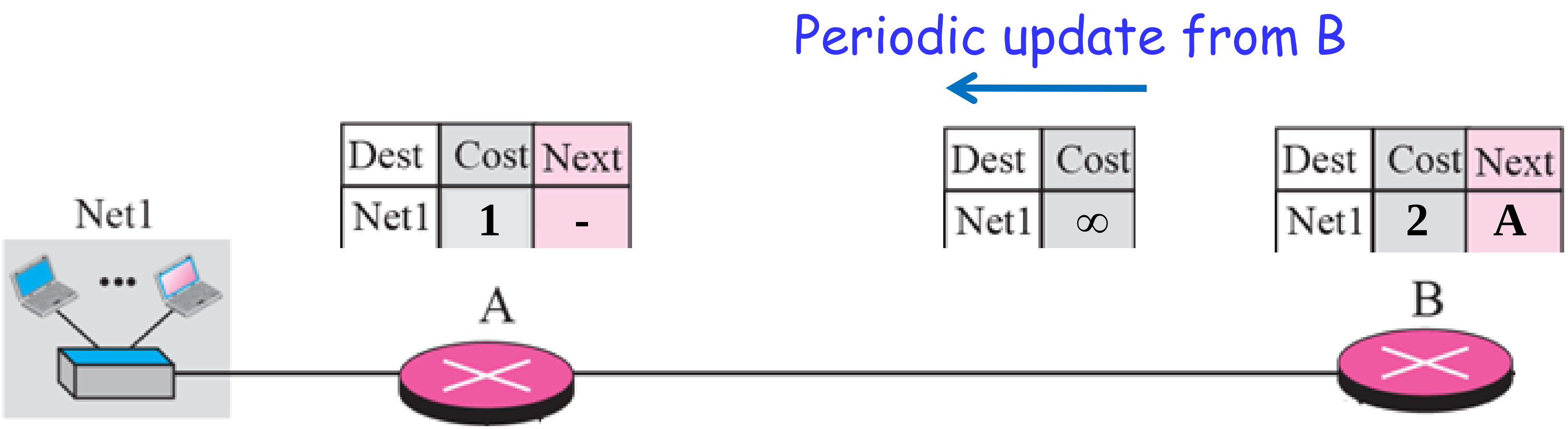
Eventually:

Dest	Cost	Next
Net1	∞	**B**

Dest	Cost	Next
Net1	∞	**A**

Distance Vector Routing: Solving Count-to-Infinity Problem - Split Horizon with Poisoned Reverse (RFC 2453)

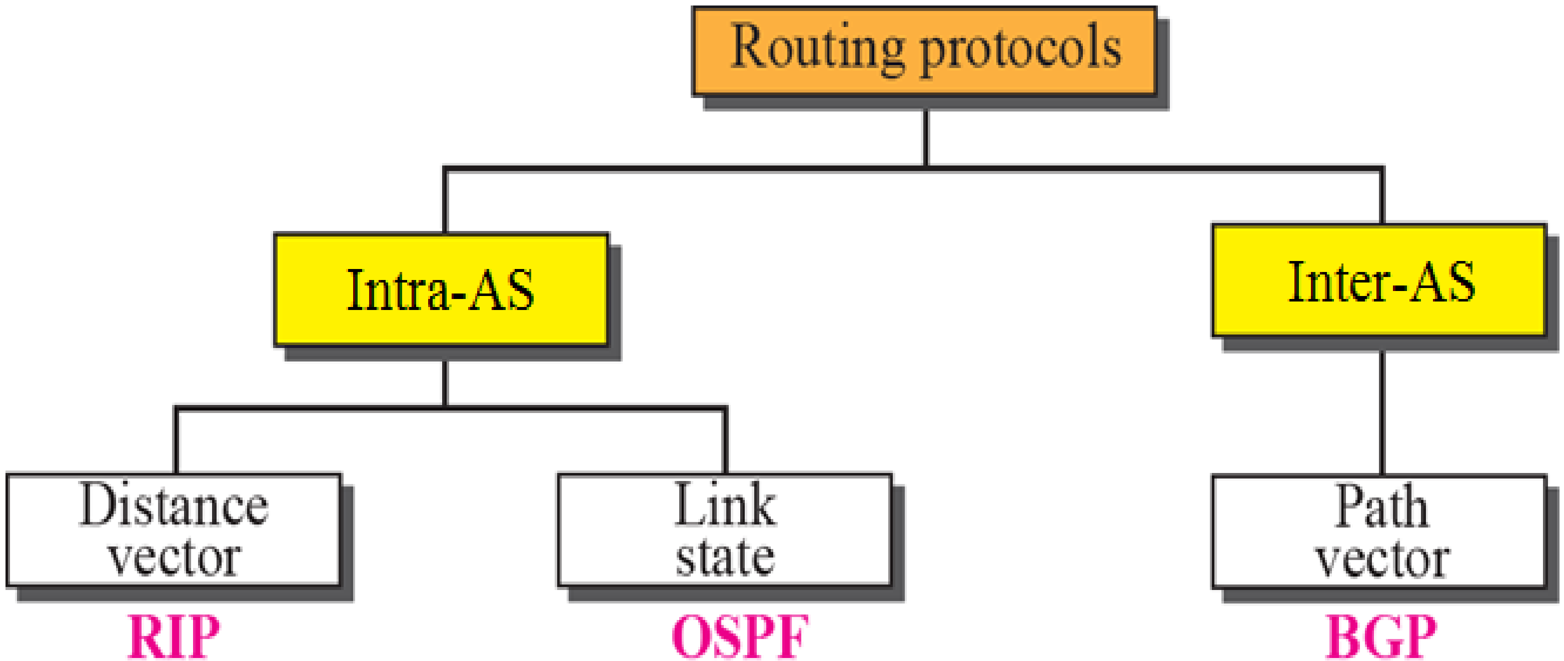
If B gets to Net1 via A, then its update to A should indicate that Net1 is unreachable.



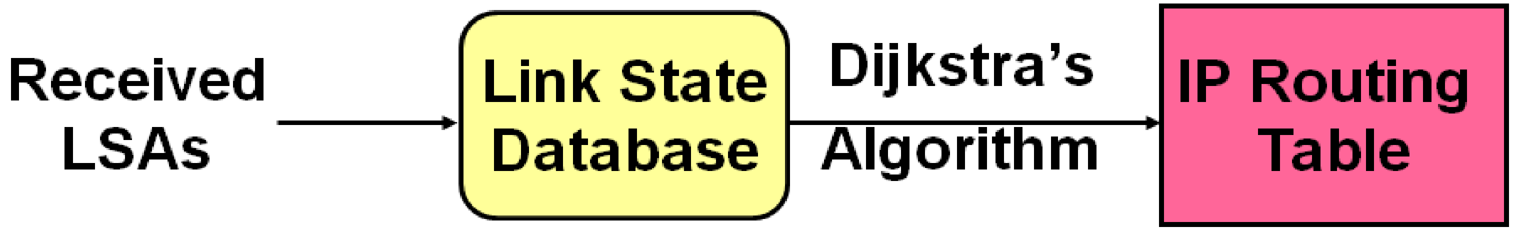
Intra-AS Distance Vector Routing: Routing Information Protocol

- **RIP uses Distance Vector algorithm, cost is simply based on the number of hops**
- **Allows maximum 15 hops, 16 indicates** ∞
- **Routing information exchanged every 30 sec via Response Message**
- **If no advertisement heard after 180 seconds → neighbor/link declared dead**
- **RIP related RFC Documentations:**
  - **RFC 1058, 1387, 1388, 1723 (RIP version 2)**

Intra-AS and Inter-AS Routing Protocols



Intra-AS Routing:
Link State Routing



Consists of Five stages:

❶ **Discover Neighbors by multicasting Hello**

❷ **Construct Link State Advertisement Packet (LSA/LSP)**

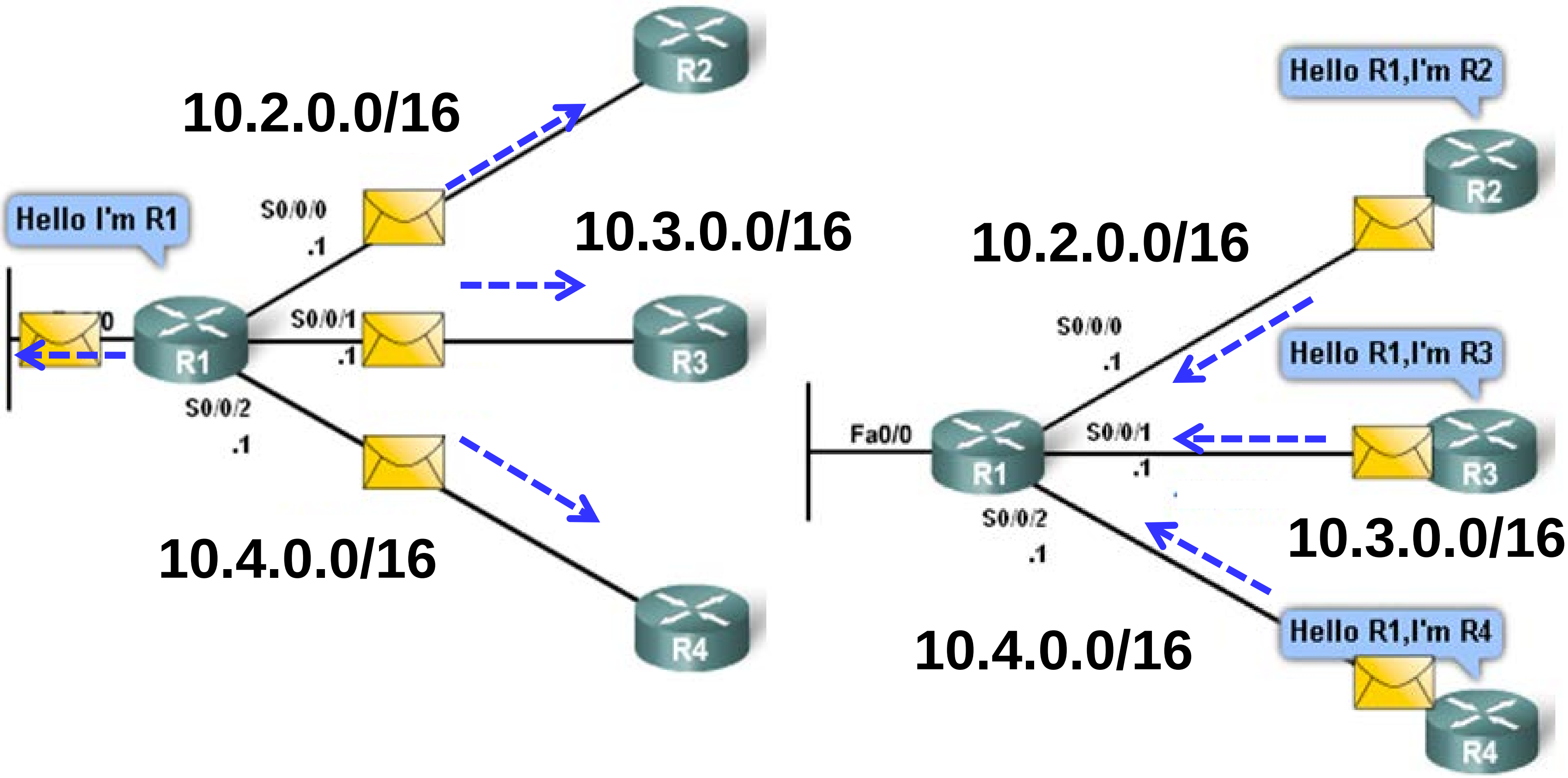
❸ **Flood LSA/LSP to ALL Routers**

- **During initial start-up**
- **When there is a change in topology**

❹ **Construct Link State Database**

❺ **Compute Shortest-Path Routes (using Dijkstra's algorithm)**

Link State Routing: Discover Neighbors

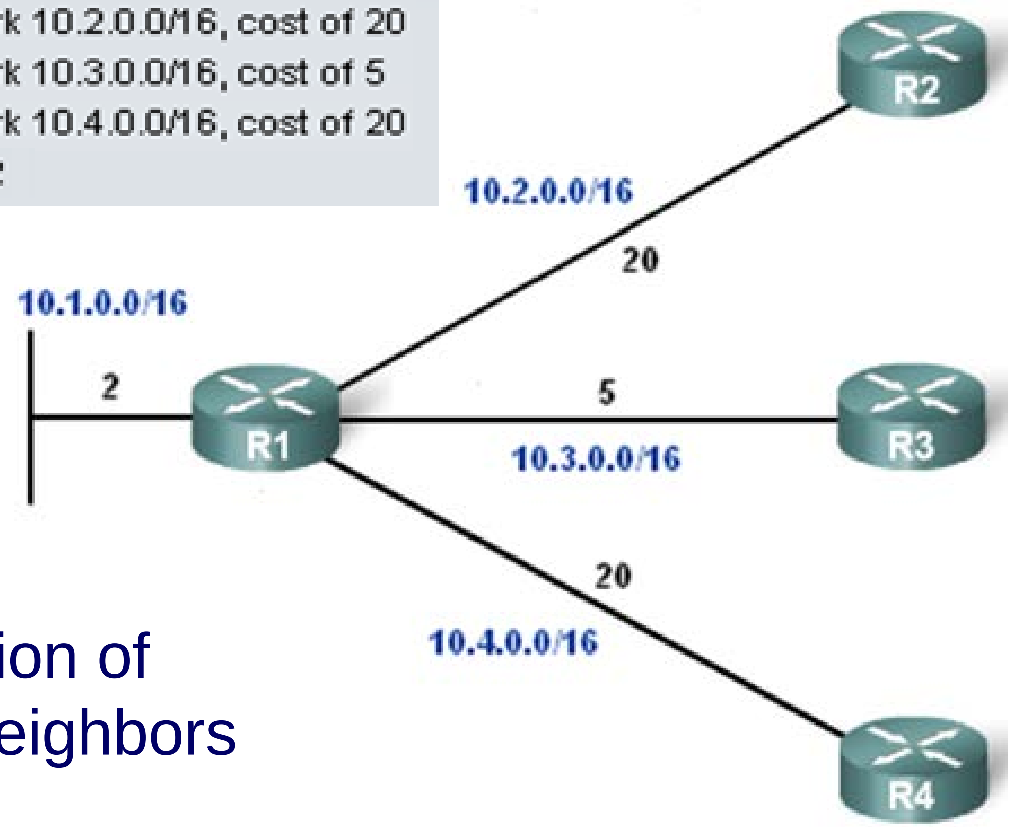


- *A* transmits HELLO packet on each of its links
- *A* 's neighbors identify themselves to *A*

Link State Routing: Construct Link State Advertisement Packet (LSA/LSP)

R1 Link-State

- Connected to neighbor R2 on network 10.2.0.0/16, cost of 20
- Connected to neighbor R3 on network 10.3.0.0/16, cost of 5
- Connected to neighbor R4 on network 10.4.0.0/16, cost of 20
- Has a network 10.1.0.0/16, cost of 2

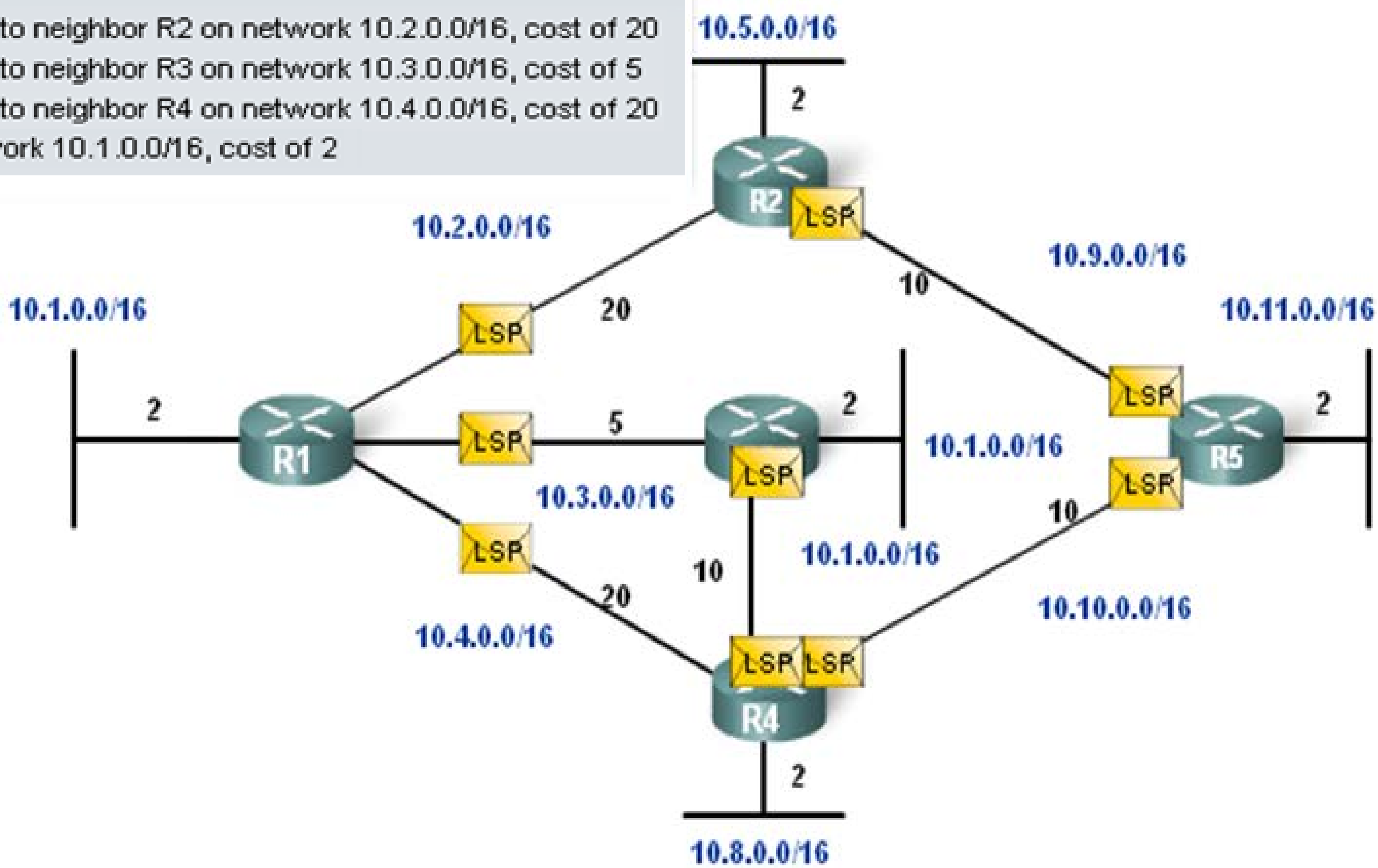


- *LSA* contains information of itself and immediate neighbors only.

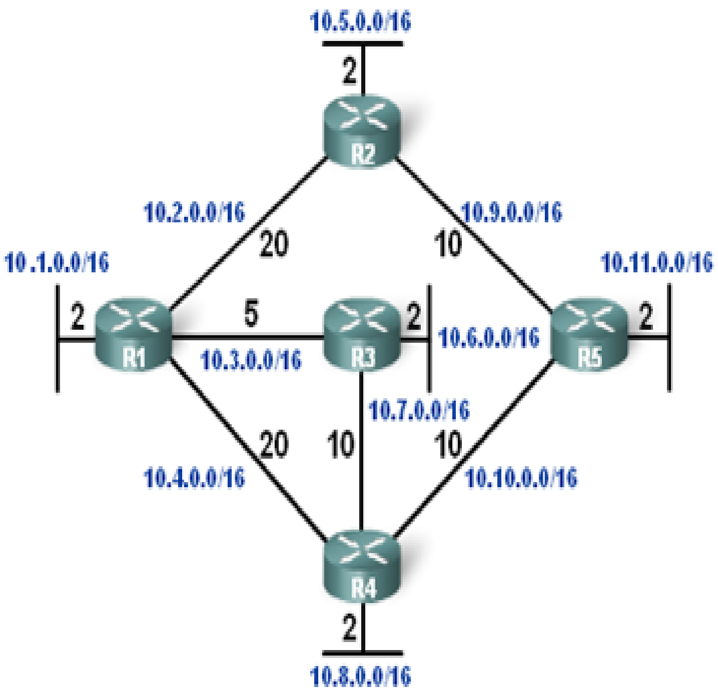
Link State Routing: Flood LSP to ALL routers

R1 Link-State

- Connected to neighbor R2 on network 10.2.0.0/16, cost of 20
- Connected to neighbor R3 on network 10.3.0.0/16, cost of 5
- Connected to neighbor R4 on network 10.4.0.0/16, cost of 20
- Has a network 10.1.0.0/16, cost of 2



Link State Routing: Build Link State Database



Each router builds its own link state database to have a complete topology of the whole network.

R1 Link-State Database

R1 Link-states:

- Connected to neighbor R2 on network 10.2.0.0/16, cost of 20
- Connected to neighbor R3 on network 10.3.0.0/16, cost of 5
- Connected to neighbor R4 on network 10.4.0.0/16, cost of 20
- Has a network 10.1.0.0/16, cost of 2

LSPs from R2:

- Connected to neighbor R1 on network 10.2.0.0/16, cost of 20
- Connected to neighbor R5 on network 10.9.0.0/16, cost of 10
- Has a network 10.5.0.0/16, cost of 2

LSPs from R3:

- Connected to neighbor R1 on network 10.3.0.0/16, cost of 5
- Connected to neighbor R4 on network 10.7.0.0/16, cost of 10
- Has a network 10.6.0.0/16, cost of 2

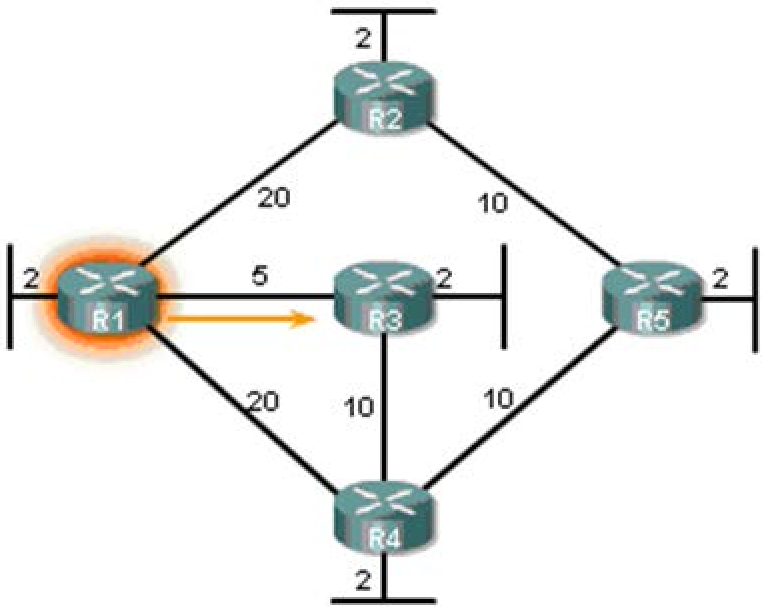
LSPs from R4:

- Connected to neighbor R1 on network 10.4.0.0/16, cost of 20
- Connected to neighbor R3 on network 10.7.0.0/16, cost of 10
- Connected to neighbor R5 on network 10.10.0.0/16, cost of 10
- Has a network 10.8.0.0/16, cost of 2

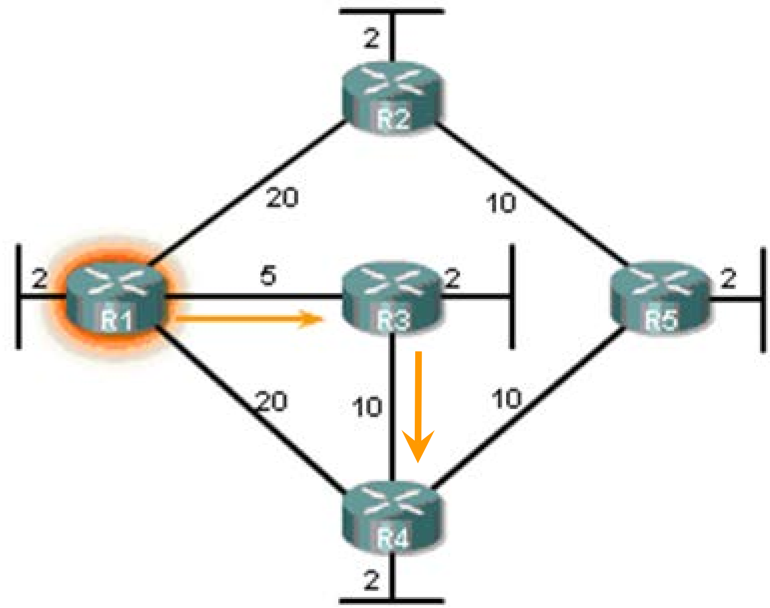
LSPs from R5:

- Connected to neighbor R2 on network 10.9.0.0/16, cost of 10
- Connected to neighbor R4 on network 10.10.0.0/16, cost of 10
- Has a network 10.11.0.0/16, cost of 2

Link State Routing: Compute shortest-path routes using Dijkstra's Algorithm

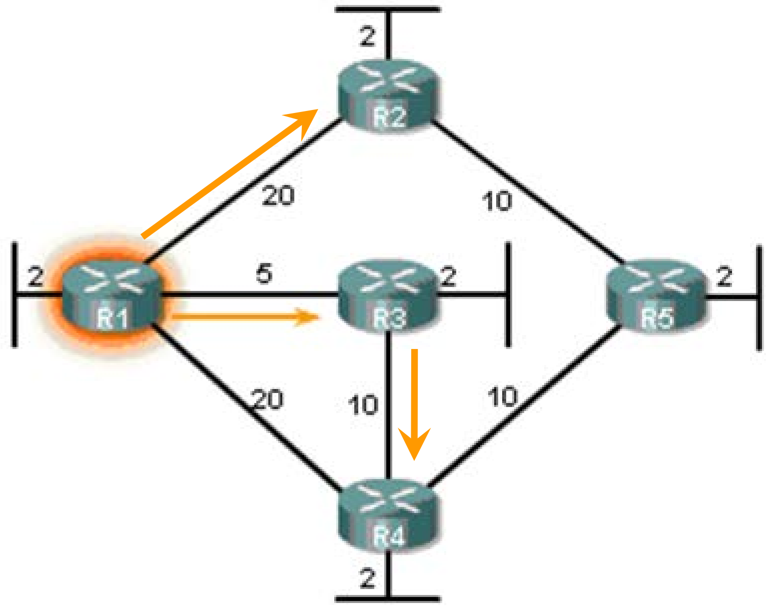


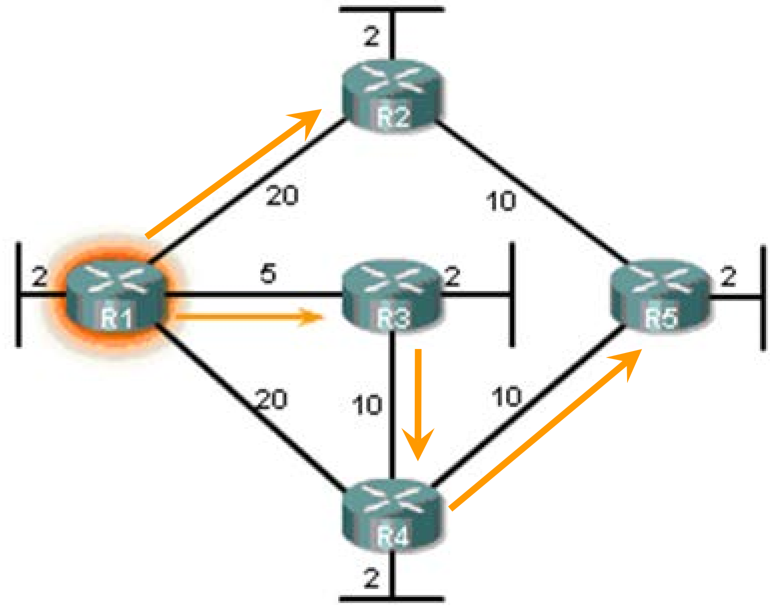
Iteration	D_{12}	D_{13}	D_{14}	D_{15}
{1}	20	5 ✓	20	∝



Iteration	D_{12}	D_{13}	D_{14}	D_{15}
{1}	20	5 ✓	20	∝
{1,3}	20	5	15 ✓	∝

Link State Routing: Compute shortest-path routes using Dijkstra's Algorithm

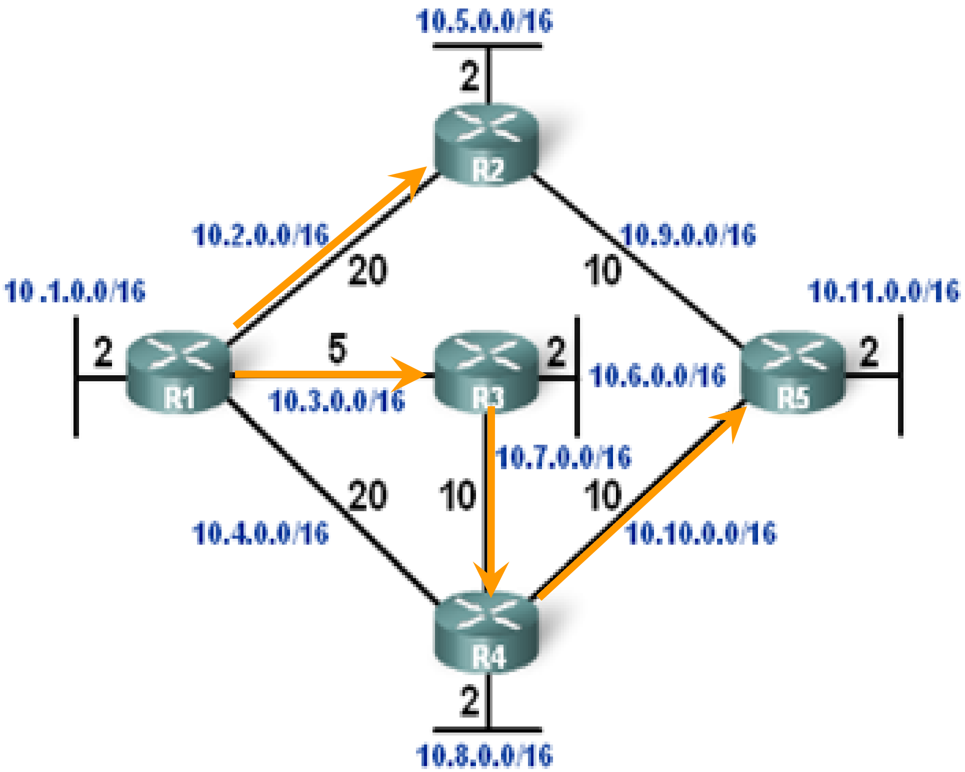




Iteration	D_{12}	D_{13}	D_{14}	D_{15}
{1}	20	5 ✓	20	∝
{1,3}	20	5	15✓	∝
{1,3,4}	20✓	5	15	25

Iteration	D_{12}	D_{13}	D_{14}	D_{15}
{1}	20	5 ✓	20	∝
{1,3}	20	5	15✓	∝
{1,3,4}	20✓	5	15	25
{1,3,4,2}	20	5	15	25 ✓

Link State Routing: Update routing table



R1 Routing Table

Directly Connected Networks

- 10.1.0.0/16 Directly Connected Network
- 10.2.0.0/16 Directly Connected Network
- 10.3.0.0/16 Directly Connected Network
- 10.4.0.0/16 Directly Connected Network

Remote Networks

- 10.5.0.0/16 via R2 , cost = 22
- 10.6.0.0/16 via R3 , cost = 7
- 10.7.0.0/16 via R3 , cost = 15
- 10.8.0.0/16 via R3 , cost = 17
- 10.9.0.0/16 via R2 , cost = 30
- 10.10.0.0/16 via R3 , cost = 25
- 10.11.0.0/16 via R3 , cost = 27

Intra-AS Link State Routing: Open Shortest Path First

- "Open": publicly available
- Uses Link State algorithm
  - LSA packet dissemination
  - Topology map at each node
  - Route computation using Dijkstra's algorithm

- Advertisements disseminated to entire AS (via flooding)
- OSPF related RFC documentations:
  - RFC 1131, 1247, 1583 (OSPF version 2)

Inter-AS Path Vector Routing: Border Gateway Protocol (BGP)

Consists of 3 main stages:

❶ **Configure border router to know its neighbors**

❷ **Exchange path vectors (routing information) with neighbors**

③ **Select path based on policy**

- **BGP: the *de facto* standard, current ver 4 (RFC 4271)**

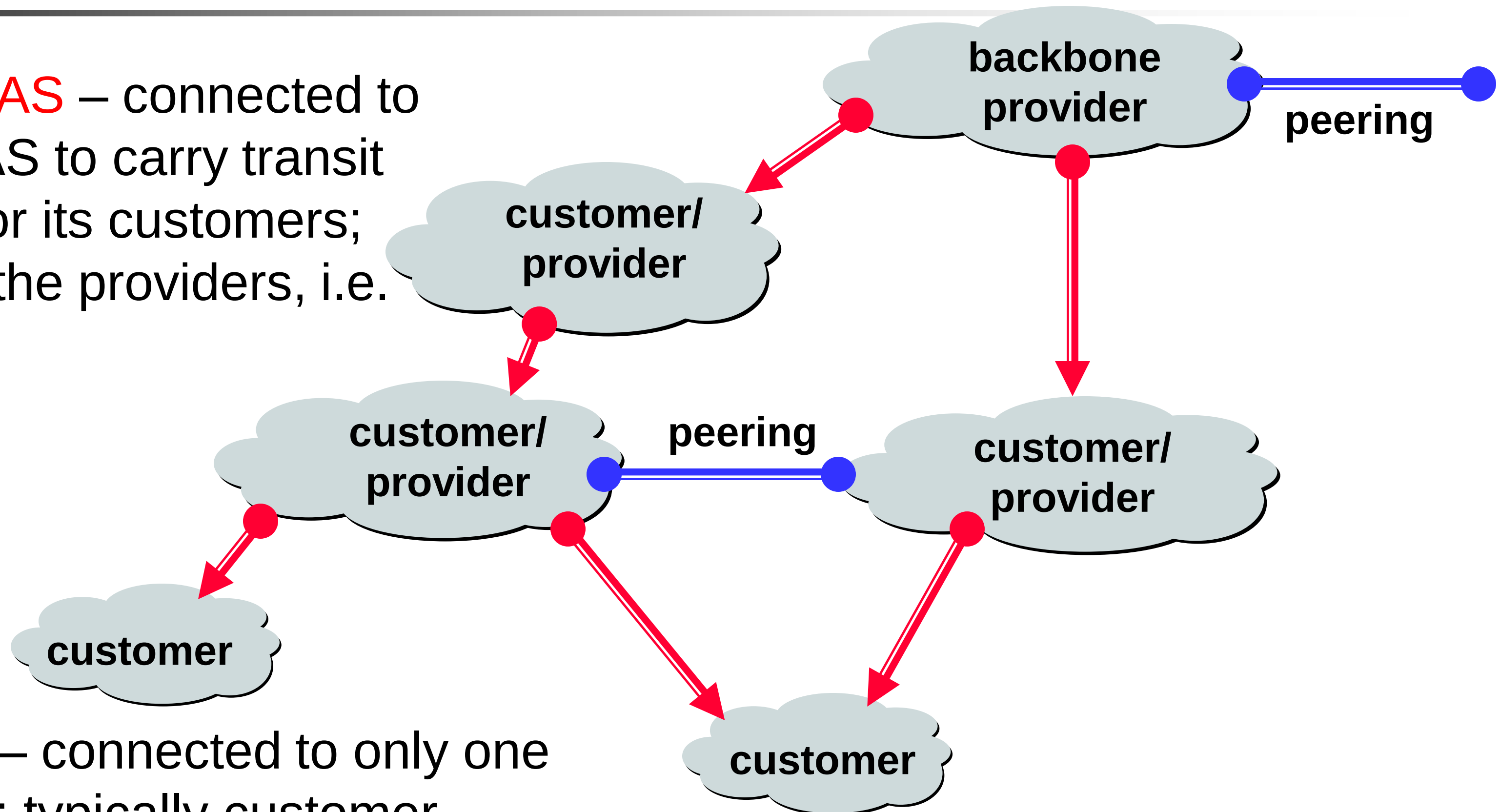
IP header	TCP header	BGP Packet

port: 179

- After configuration, BGP routers establish TCP connections with their neighbors to exchange routing information.

Broadly, AS can be classified into stub AS, multi-homed AS or transit AS.

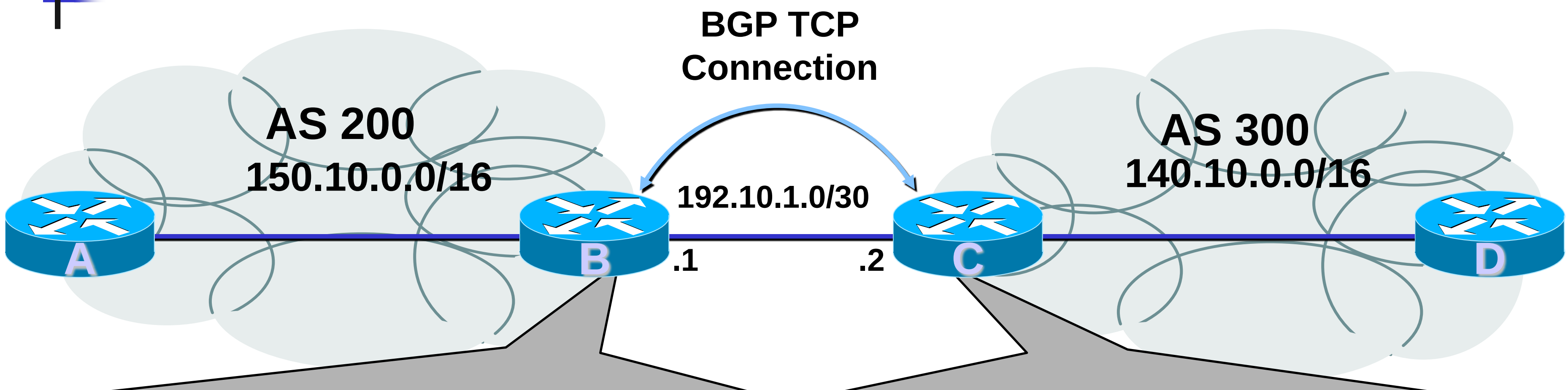
Transit AS – connected to other AS to carry transit traffic for its customers; mainly the providers, i.e. ISPs



Stub AS – connected to only one other AS; typically customer connected to its provider. In fact, it's not necessary for stub AS to run BGP since it has only 1 path to its ISP.

Multi-homed AS – connected to more than one AS, but does not carry transit traffic; typically for customer requiring reliability

Configuring BGP routers to know their neighbors (peers) to establish TCP connections.

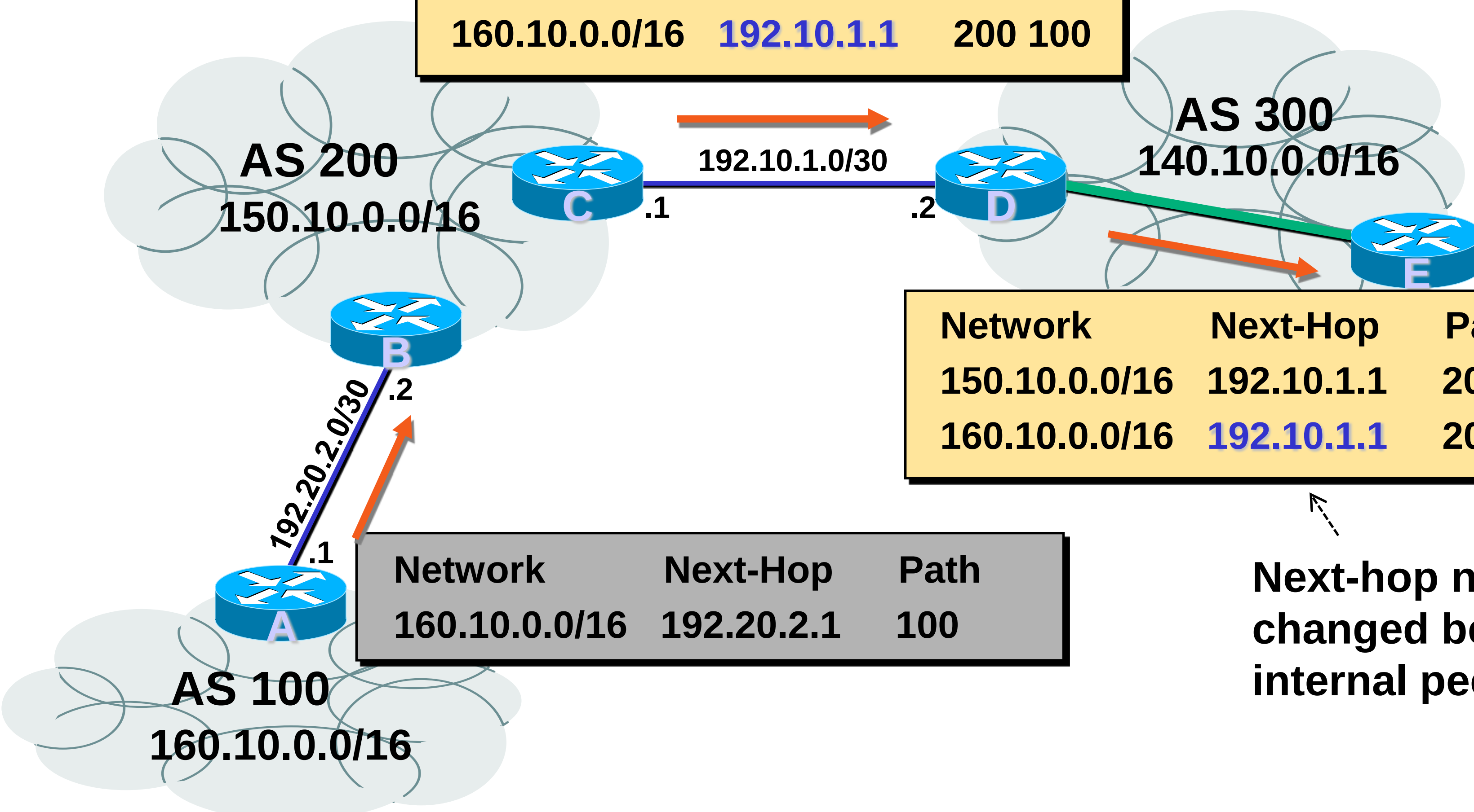



```
interface Serial 0
ip address 192.10.1.1 255.255.255.252

router bgp 200
 network 150.10.0.0 mask 255.255.0.0
 neighbor 192.10.1.2 remote-as 300
```

```
interface Serial 0
ip address 192.10.1.2 255.255.255.252

router bgp 300
 network 140.10.0.0 mask 255.255.0.0
 neighbor 192.10.1.1 remote-as 200
```

- External peers (between different AS) are normally adjacent to each other and share a subnet.
- Internal peers may be in any subnet within the same AS.

Peers exchange routing information containing complete AS Path to detect loop problem.



Network	Next-Hop	Path
150.10.0.0/16	192.10.1.1	200
160.10.0.0/16	192.10.1.1	200 100

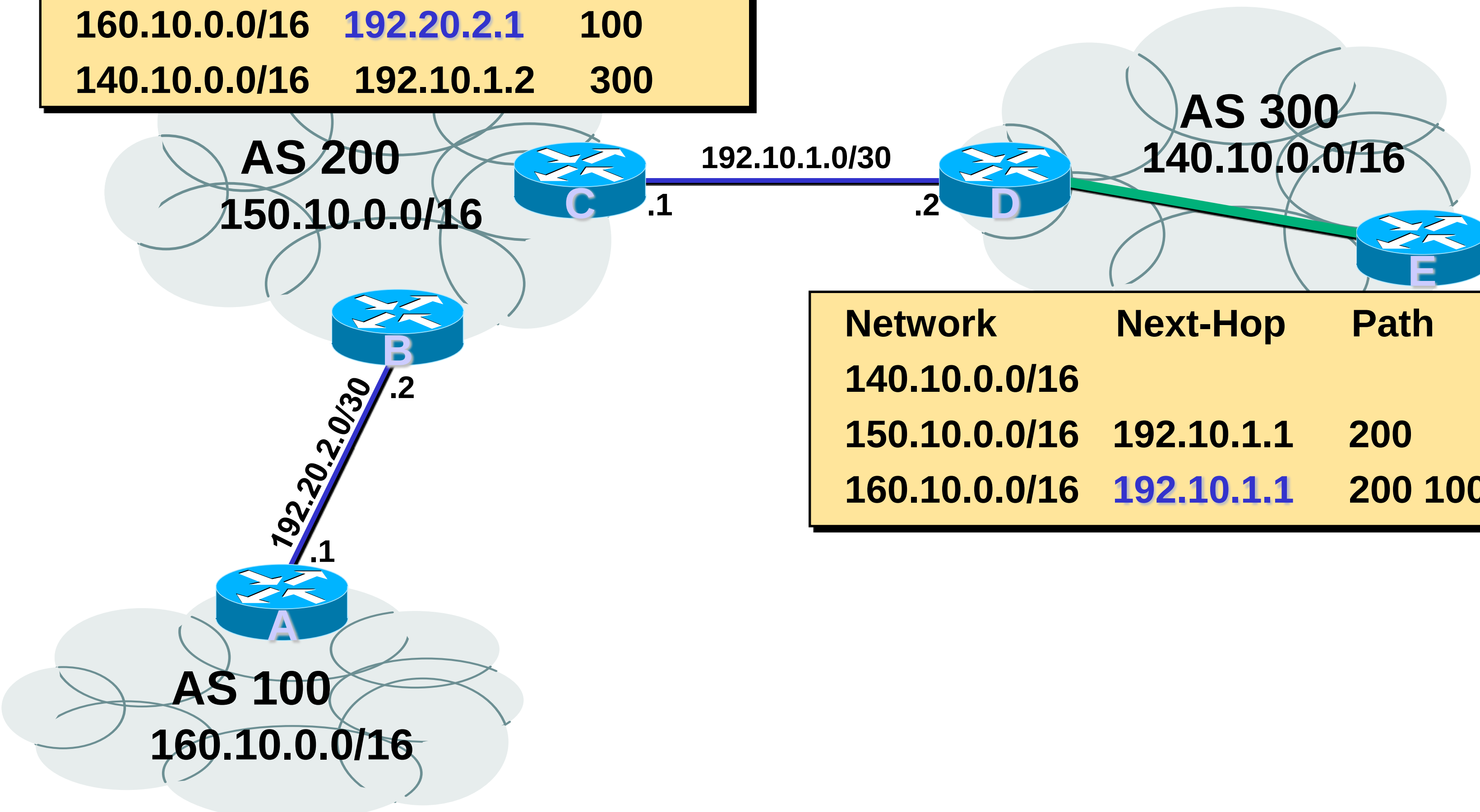
Network	Next-Hop	Path
150.10.0.0/16	192.10.1.1	200
160.10.0.0/16	192.10.1.1	200 100

Network	Next-Hop	Path
160.10.0.0/16	192.20.2.1	100

Next-hop not changed between internal peers.

AS Table content

Network	Next-Hop	Path
150.10.0.0/16		
160.10.0.0/16	192.20.2.1	100
140.10.0.0/16	192.10.1.2	300

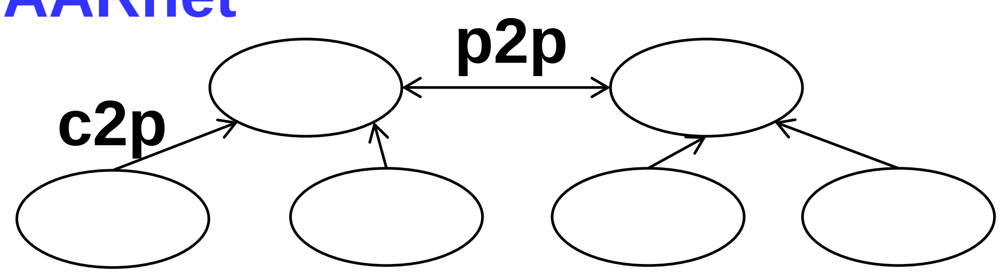


Network	Next-Hop	Path
140.10.0.0/16		
150.10.0.0/16	192.10.1.1	200
160.10.0.0/16	192.10.1.1	200 100

There is also the commercial relationship between neighbouring ASes.

ASes relationship:

- **Customer-Provider: Customer pays the provider to send and receive traffic**
  - **e.g. NTU and SingTel**
- **Peer-to-peer: Two Ases agree to transit each other traffic.**
  - **eg. SingAREN and AARnet**



Based on policy, BGP routers can decide to accept/decline offered paths; and to drop/ advertise paths to their neighbors.

Path selection (policy-based):

- **Import policy: may or may not select path offered**
  - **e.g. cost, business relationships, don't route through competitors, loop prevention reasons.**
- **Export policy: can filter routes you don't want to tell neighbors**
  - **eg. don't want to route traffic to *Z* → don't advertise any route to *Z***

BGP Path Attributes

- AS-PATH. A sequence of AS along the path.
- NEXT-HOP. IP address of the border router that should be the next hop to the destination.
- MULT-EXIT-DISC. Used by administrator to determine the exit interface to neighbouring AS. (Lower the higher priority)
- LOCAL-PREF. Assigned by administrator to select between multiple path to an AS. This is used to informed other BGP speakers in the same AS of the preference.
- ATOMIC-AGGREGATOR. It is set when the BGP speaker select a less specific route, when presented with overlapping route from its peers.
- AGGREGATOR. Contains the AS and IP address of the BGP speaker that performs route aggregation.

BGP route propagation 4

- Check all incoming filters defined by BGP session. Ignore routes that are not allowed.
- Insert route in the BGP table.
- Compare the new routes with other routes in the BGP table with the same destination prefix, and execute AS path selection process.
- If the new route is the best. Set as preferred route.
- Propagate the new best route to BGP neighbours.

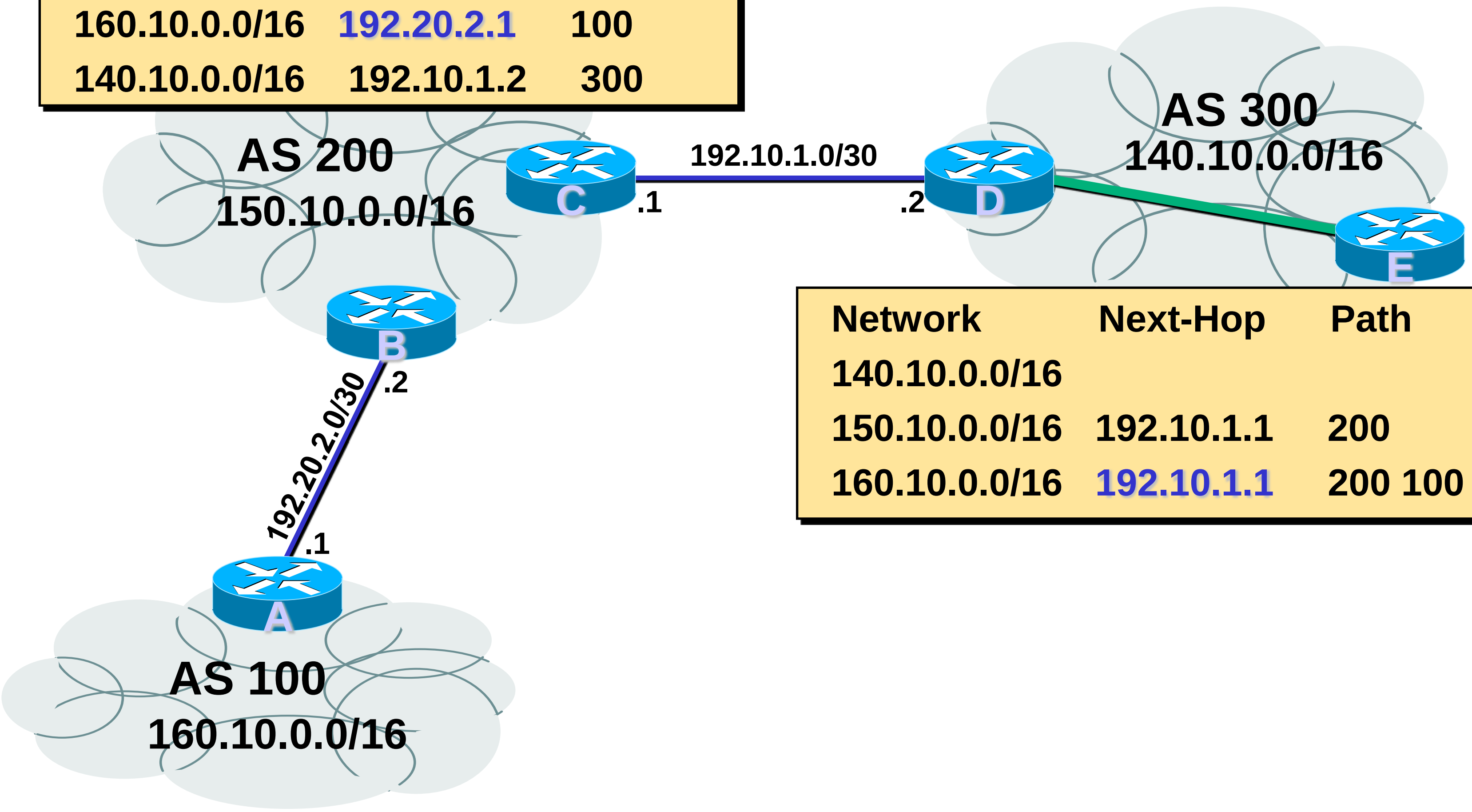
Path Selection

- **Uses Shortest Path(assuming no preference is assigned)**

- **Distance Metric: ASes Hop Count(NOT number of routers).**

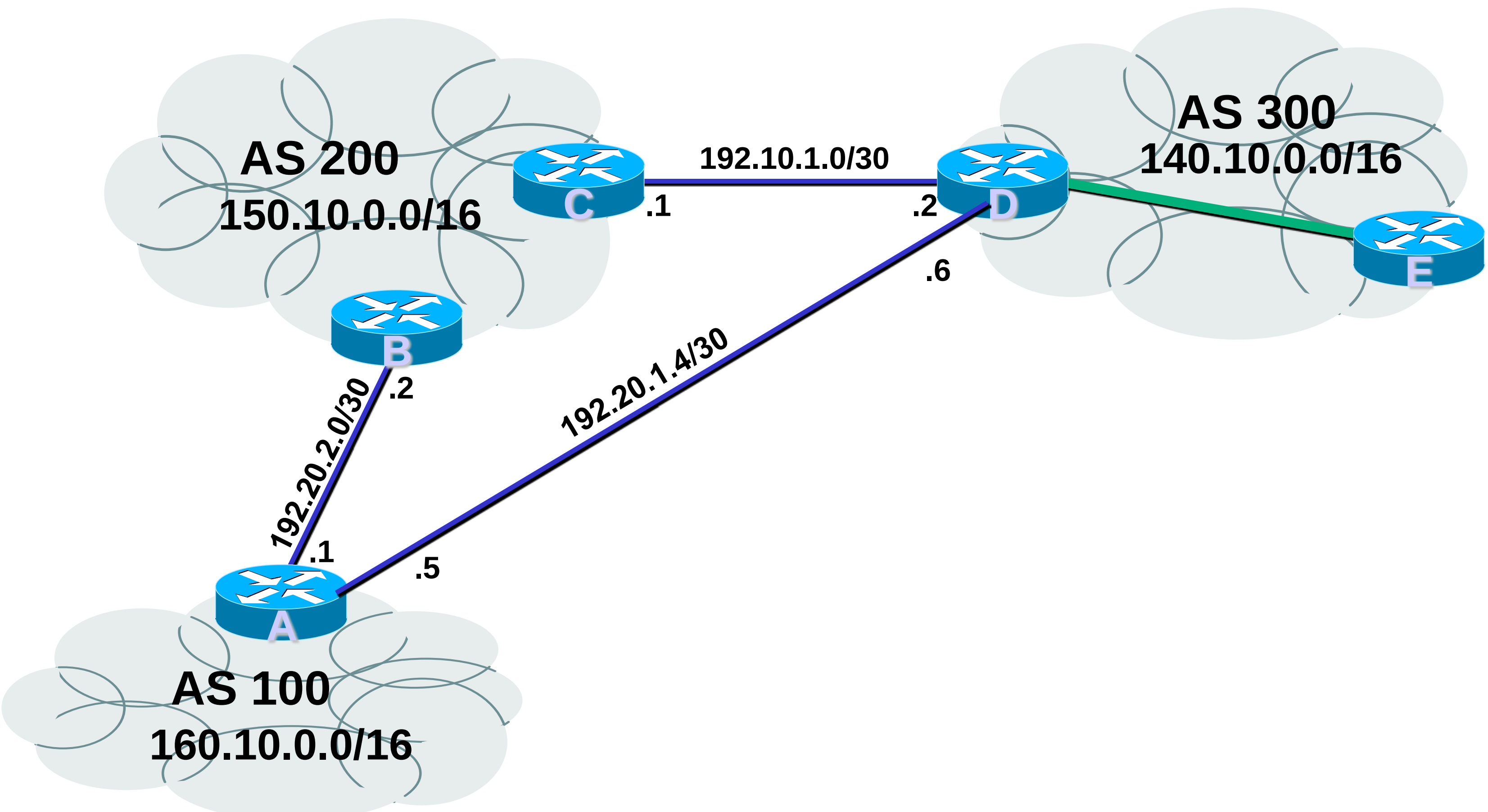
BGP Table content

Network	Next-Hop	Path
150.10.0.0/16		
160.10.0.0/16	192.20.2.1	100
140.10.0.0/16	192.10.1.2	300



Network	Next-Hop	Path
140.10.0.0/16		
150.10.0.0/16	192.10.1.1	200
160.10.0.0/16	192.10.1.1	200 100

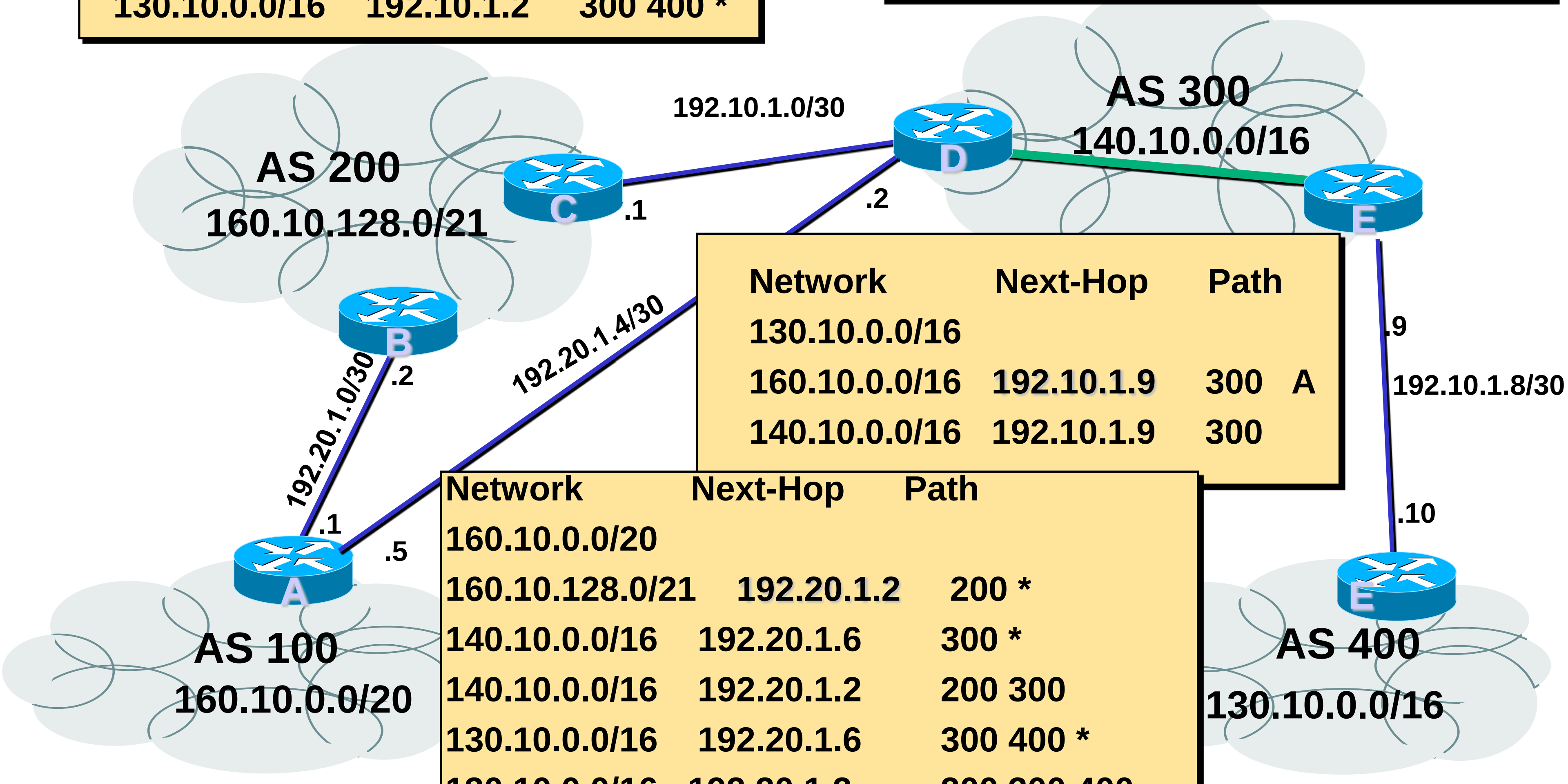
Path Selection(based on AS Hops)



Route Aggregation

Network	Next-Hop	Path
160.10.128.0/21		
160.10.0.0/20	192.20.1.1	100 *
160.10.0.0/20	192.10.1.2	300 100
140.10.0.0/16	192.10.1.2	300 *
130.10.0.0/16	192.10.1.2	300 400 *

Network	Next-Hop	Path
140.10.0.0/16		
130.10.0.0/16	192.10.1.10	400 *
160.10.0.0/20	192.20.1.5	100 *
160.10.0.0/20	192.10.1.1	200 100
160.10.128.0/21	192.10.1.1	200 *



Network	Next-Hop	Path
130.10.0.0/16		
160.10.0.0/16	192.10.1.9	300 A
140.10.0.0/16	192.10.1.9	300

Network	Next-Hop	Path
160.10.0.0/20		
160.10.128.0/21	192.20.1.2	200 *
140.10.0.0/16	192.20.1.6	300 *
140.10.0.0/16	192.20.1.2	200 300
130.10.0.0/16	192.20.1.6	300 400 *
130.10.0.0/16	192.20.1.2	200 300 400

http://as-rank.caida.org/?mode0=as-ranking | AS Rank: AS Ranking - CAI...

File Edit View Favorites Tools Help

caida Center for Applied Internet Data Analysis

Google™ Custom Search

DONATE | CON

HOME | RESEARCH | DATA | TOOLS | INTERACTIVE | PUBLICATIONS | WORKSHOPS | PROJECTS | FUNDING

http://as-rank.caida.org >

AS Rank: AS Ranking

This page presents the BETA version of AS Rank, CAIDA's ranking of Autonomous Systems (AS) (which approximately map to Internet Service Providers) and organizations (Orgs) (which are a collection of one or more ASes). This ranking is derived from topological data collected by CAIDA's Archipelago Measurement Infrastructure and Border Gateway Protocol (BGP) routing data collected by the Route Views Project and RIPE NCC.

ASes and Orgs are ranked by their customer cone size, which is the number of their direct and indirect customers. Note: We do *not* have data to rank ASes (ISPs) by traffic, revenue, users, or any other non-topological metric.

AS Ranking | Org Ranking | Information for a single AS | Information for a single Org | Background | Data Sources | Help | AS Ranking H

The top ASes ranked by customer cone size are displayed below.
For information about a specific AS, enter its AS name, its AS number, or the name of the Org of which the AS is a member.

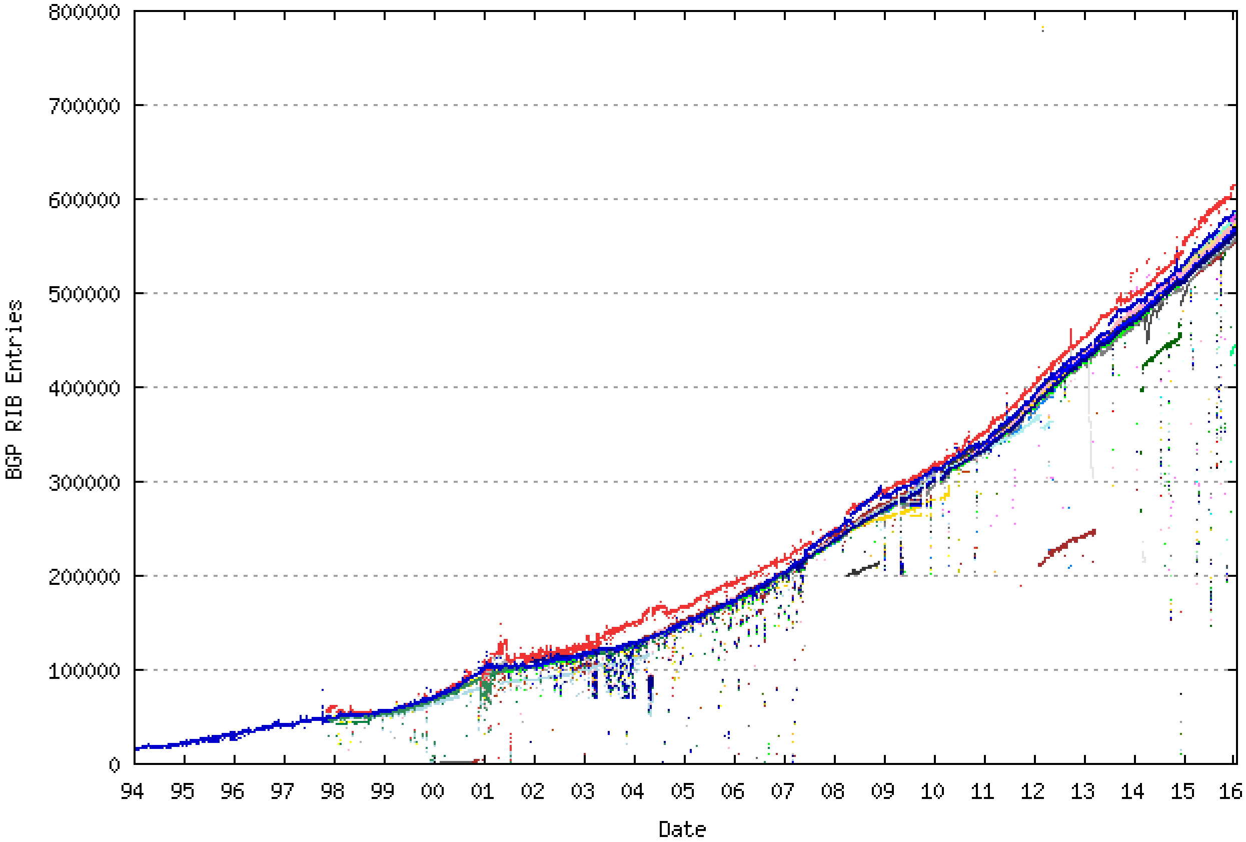
Dataset: 2014-03-01 | Change datas

Look up an AS by number or name | Search

Table shows 10 of 45658 ASes, sorted by number of ASes in customer cone | update view

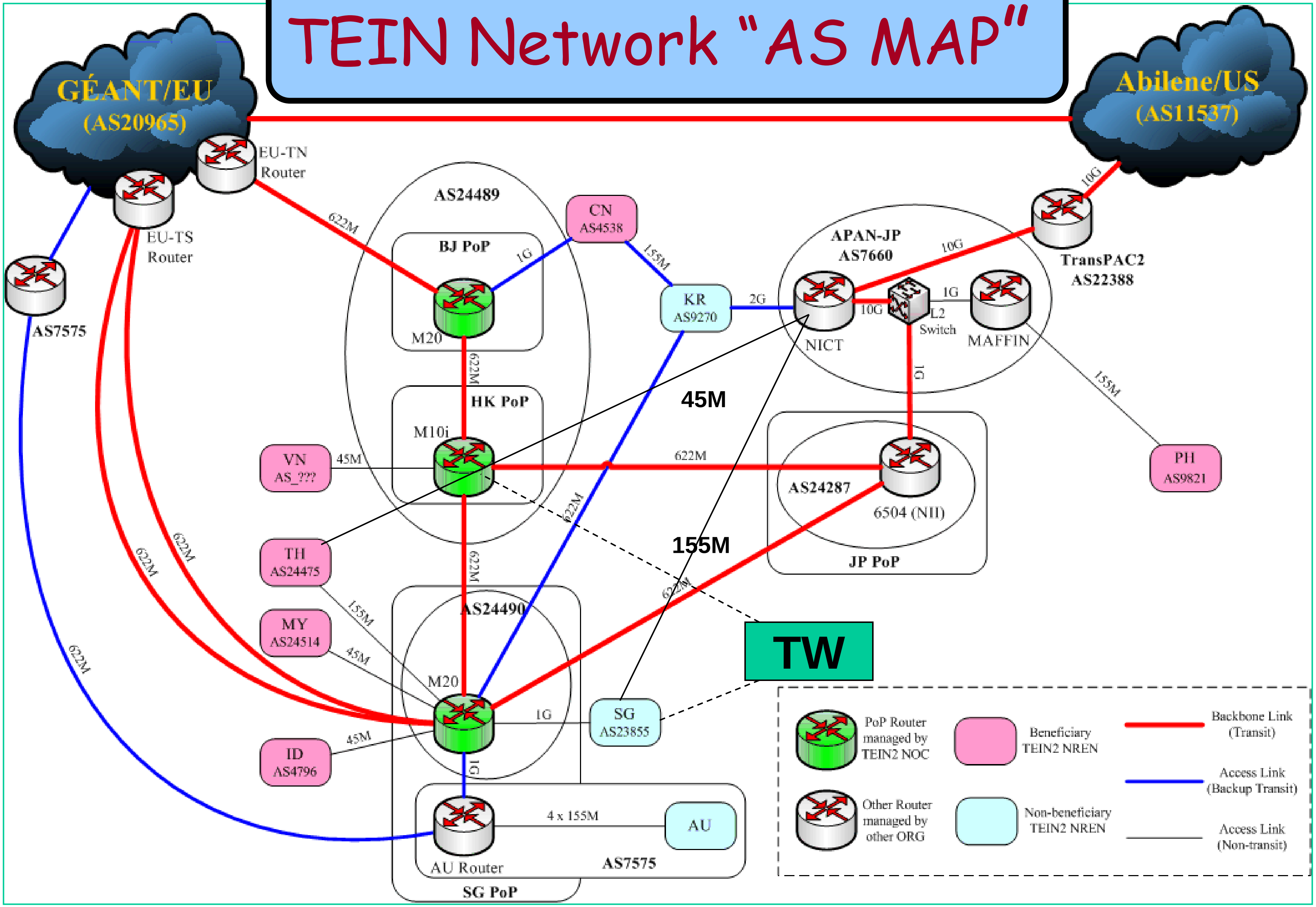
AS rank	AS number	AS name	Org name	AS Type (s)	customer cone						AS transit degree
					Number of			Percentages of all			
					ASes	IPv4 Prefixes	IPv4 Addresses	ASes	IPv4 Prefixes	IPv4 Addresses	
1	3356	LEVEL3	Level 3 Communications, Inc.	Tr Co	25,318	322,403	1,562,430,335	55%	64%	72%	3971
2	174	COGENT-174	Cogent Communications	Tr	17,484	208,063	744,220,957	38%	41%	34%	4212
3	3257	TINET-BACK...	Tinet SpA	Tr Co	15,623	222,392	846,663,937	34%	44%	39%	975
4	1299	TELIANET	TeliaSonera International Carrier	Tr Co	15,178	228,540	785,632,128	33%	45%	36%	812
5	2914	NTT-COMMUN...	NTT America, Inc.	Tr Co	14,876	224,278	929,277,565	32%	45%	43%	1047
6	3549	LVLT-3549	Level 3 Communications, Inc.	Tr Co	10,586	172,217	560,436,792	23%	34%	26%	3575
7	6453	AS6453	Tata Communications	Tr Co	10,229	167,716	610,754,120	22%	33%	28%	616

BGP routing table from 1994 - present



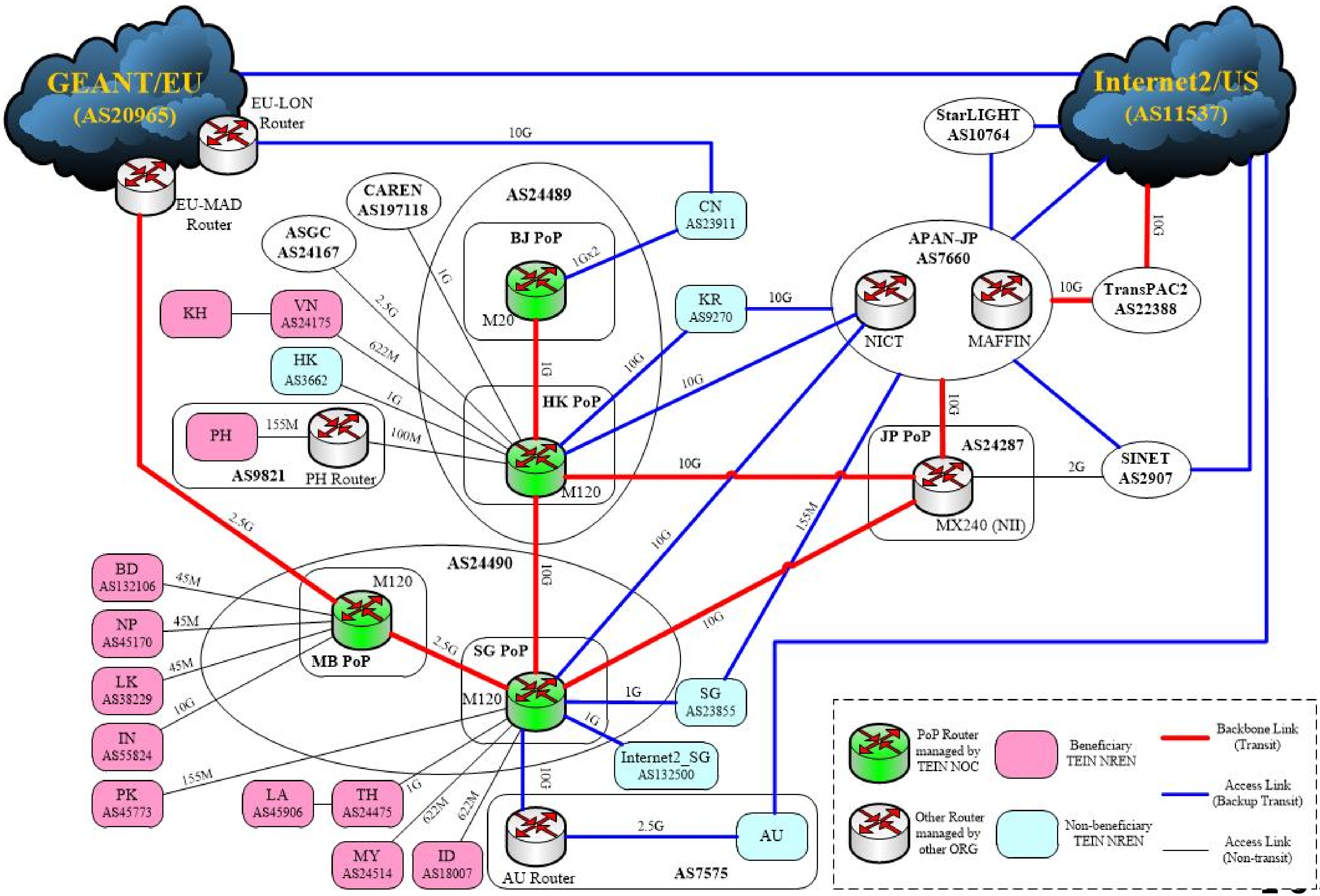
Source : http://bgp.potaroo.net/

TEIN Network “AS MAP”



Created 2006

TEIN Topology
(~ 31 Jan 2015)



Summary of Routing Protocols