CS120: Computer Networks

Lecture 11. Routing 2

Zhice Yang

Routing Protocols

- Routing Information Protocol (RIP)
  - Algorithm: Distance Vector
- Open Shortest Path First (OSPF)
  - Algorithm: Link State
- Border Gateway Protocol (BGP)

Intradomain Routing Protocol

Interdomain Routing Protocol

Link State Routing

- Network topology is known to all routers
  - Accomplished via broadcasting link state packets (LSP) to all routers
- Routing Algorithm: computes shortest paths from one node ('source') to all other nodes
  - Based on Dijkstra's algorithm

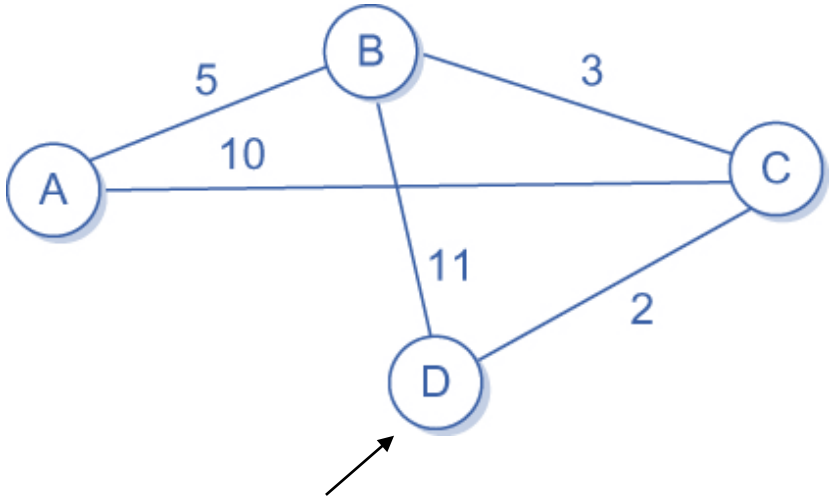
Dijkstra's Algorithm

Initialization:
M = {s}
for all nodes v
  if v adjacent to s
then $D_s(v) = c(u,v)$
else $D_s(v) = \inf$

Loop
  find w not in M such that $D_s(w)$ is a minimum
    add w to M
    update $D_s(v)$ for all v adjacent to w and not in M:
      $D_s(v) = \min(D_s(v), D_s(w) + c(w,v))$
until all nodes in M

- M: set of node processed
- S: node of the local router
- v: node of other routers
- $D_s(v)$ distance from s to v
- $c(u,v)$ link weight between node u and v

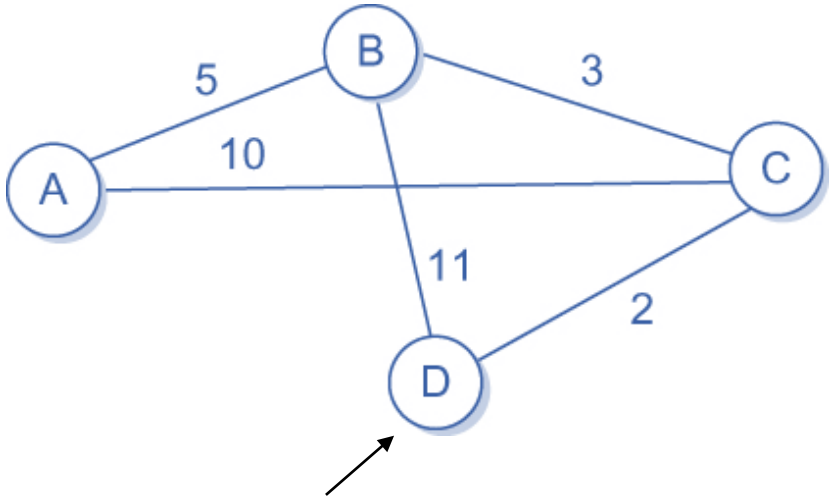
Dijkstra's Algorithm



Target Router

M	$D_D(A)$	$D_D(B)$	$D_D(C)$
{D}	Inf, from D	11, from D	2, from D

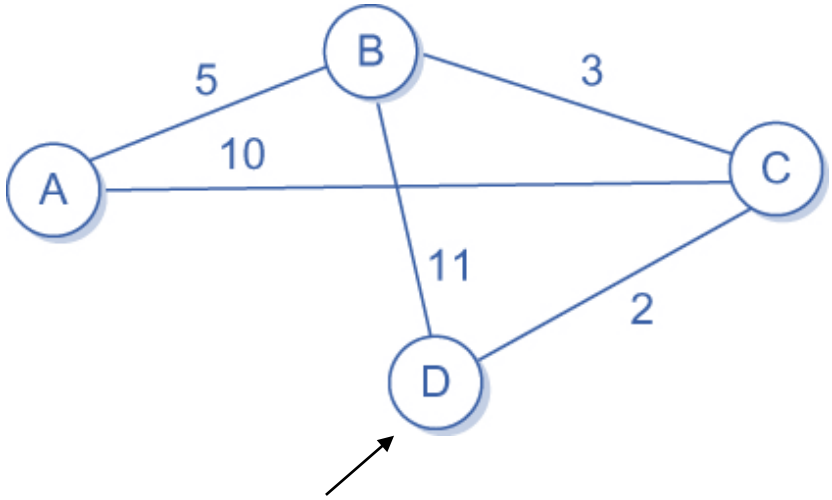
Dijkstra's Algorithm



Target Router

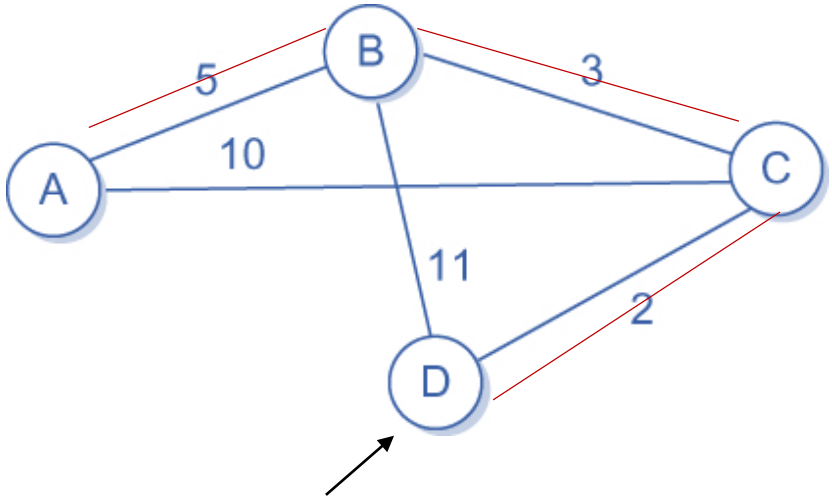
M	$D_D(A)$	$D_D(B)$	$D_D(C)$
{D}	Inf, from D	11, from D	2, from D
{D, C}	12, from C	5, from C	2, from D

Dijkstra's Algorithm



M	$D_D(A)$	$D_D(B)$	$D_D(C)$
{D}	Inf, from D	11, from D	2, from D
{D, C}	12, from C	5, from C	2, from D
{D, C, B}	10, from B	5, from C	2, from D

Dijkstra's Algorithm



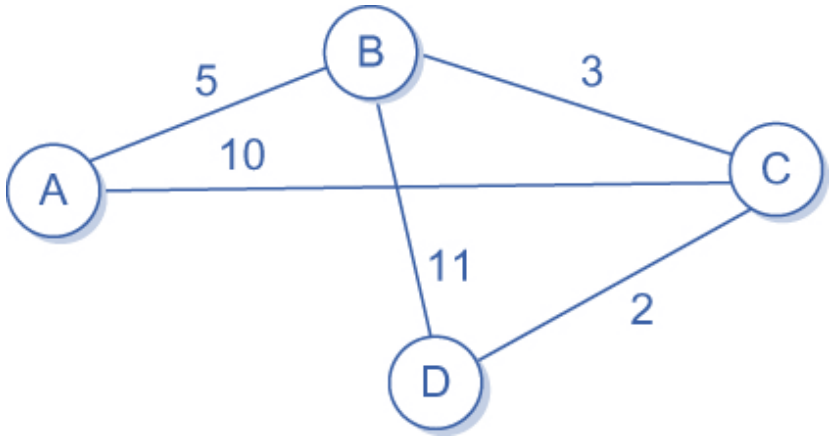
Target Router

M	$D_D(A)$	$D_D(B)$	$D_D(C)$
{D}	Inf, from D	11, from D	2, from D
{D, C}	12, from C	5, from C	2, from D
{D, C, B}	10, from B	5, from C	2, from D

Dijkstra's Algorithm (Another notation)

- <Destination, Cost, Nexthop>

Step	Confirmed	Tentative
1	(D,0,–)	
2	(D,0,–)	(B,11,B) (C,2,C)
3	(D,0,–) (C,2,C)	(B,11,B)
4	(D,0,–) (C,2,C)	(B,5,C) (A,12,C)
5	(D,0,–) (C,2,C) (B,5,C)	(A,12,C)
6	(D,0,–) (C,2,C) (B,5,C)	(A,10,C)
7	(D,0,–) (C,2,C) (B,5,C) (A,10,C)	



Open Shortest Path First (OSPF)

- "Open": nonproprietary standard created under Engineering Task Force (IETF).
  - Security: all OSPF messages authenticated (to prevent malicious intrusion)
  - Hierarchical routing: OSPF in large domains
  - Load balancing: multiple same-cost paths allowed (only one path in RIP)

OSPF: Link State Announcement

- Link State Announcement (LSA) contains key Information to describe the network topology
  - ID of the router
  - **Neighbors of the router, and the cost to the neighbor (link state of its neighbors)**
    - i.e., neighbor ID and link weight
- LSA is broadcasted to other routers
- Routers use the link states of other routers to construct the topology of entitle net
  - Based on that to calculate the shortest path

OSPF: Link State Announcement

- Flooding method
  - Basic Flooding
    - Transmit LSA to adjacent routers
    - Routers ACK the received LSA
    - Routers store the received (new) LSA
    - Forward LSA to adjacent routers
  - Key Designs: avoid duplication and out-of-date LSA
    - Seq: each LSA contains an increasing sequence number
      - Differentiate duplicated and old LSA.
        - Restart? Send out-of-order LSA to neighbors, neighbors ack the last stored seq
    - Aging: each LSA is associated with an increasing age
      - Flush LSA when reaching the max age (typically 1 hour)
      - Routers periodically broadcast LSA to refresh LSA (typically 0.5 hour)

Routing Protocols

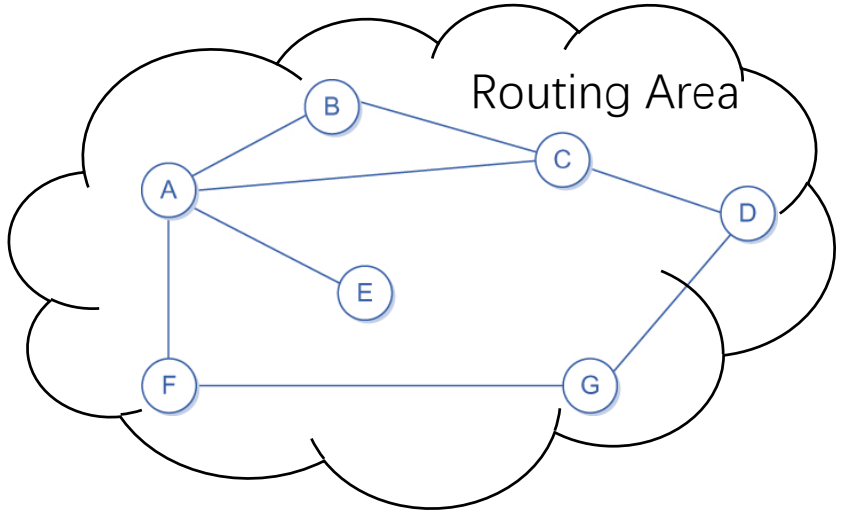
- Routing Information Protocol (RIP)
  - Algorithm: Distance Vector
- Open Shortest Path First (OSPF)
  - Algorithm: Link State
- Border Gateway Protocol (BGP)

Intradomain Routing Protocol

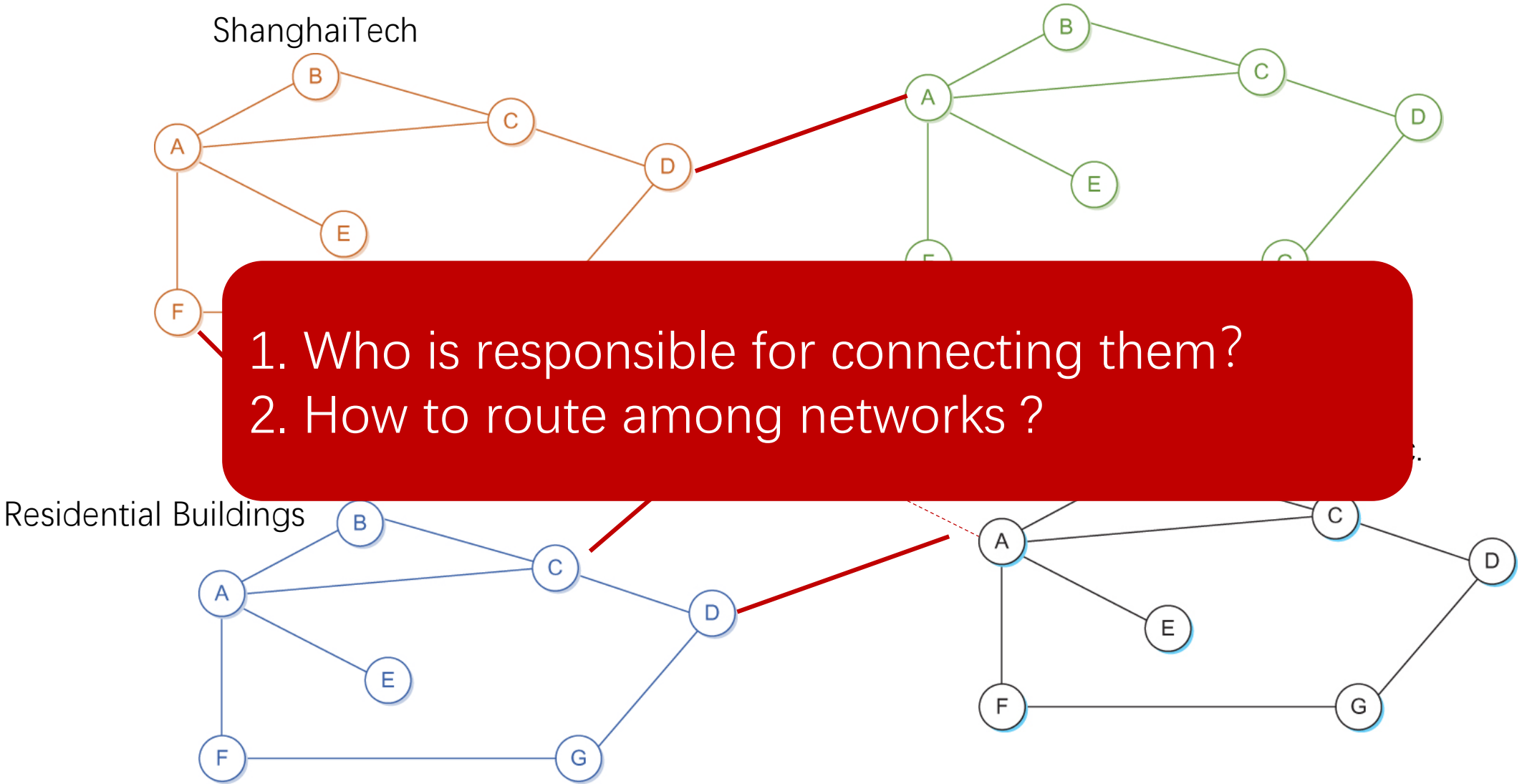
Interdomain Routing Protocol

The Discussion on Routing So Far …

- Routers in the network are managed by the same administrator
  - e.g., Residential building, Campus, Network of a big company, etc.
- Routers are running same routing protocol
  - e.g., OSPF or RIP.
  - These routers have certain coverage, called the routing area



The Real Internet: Network of Network

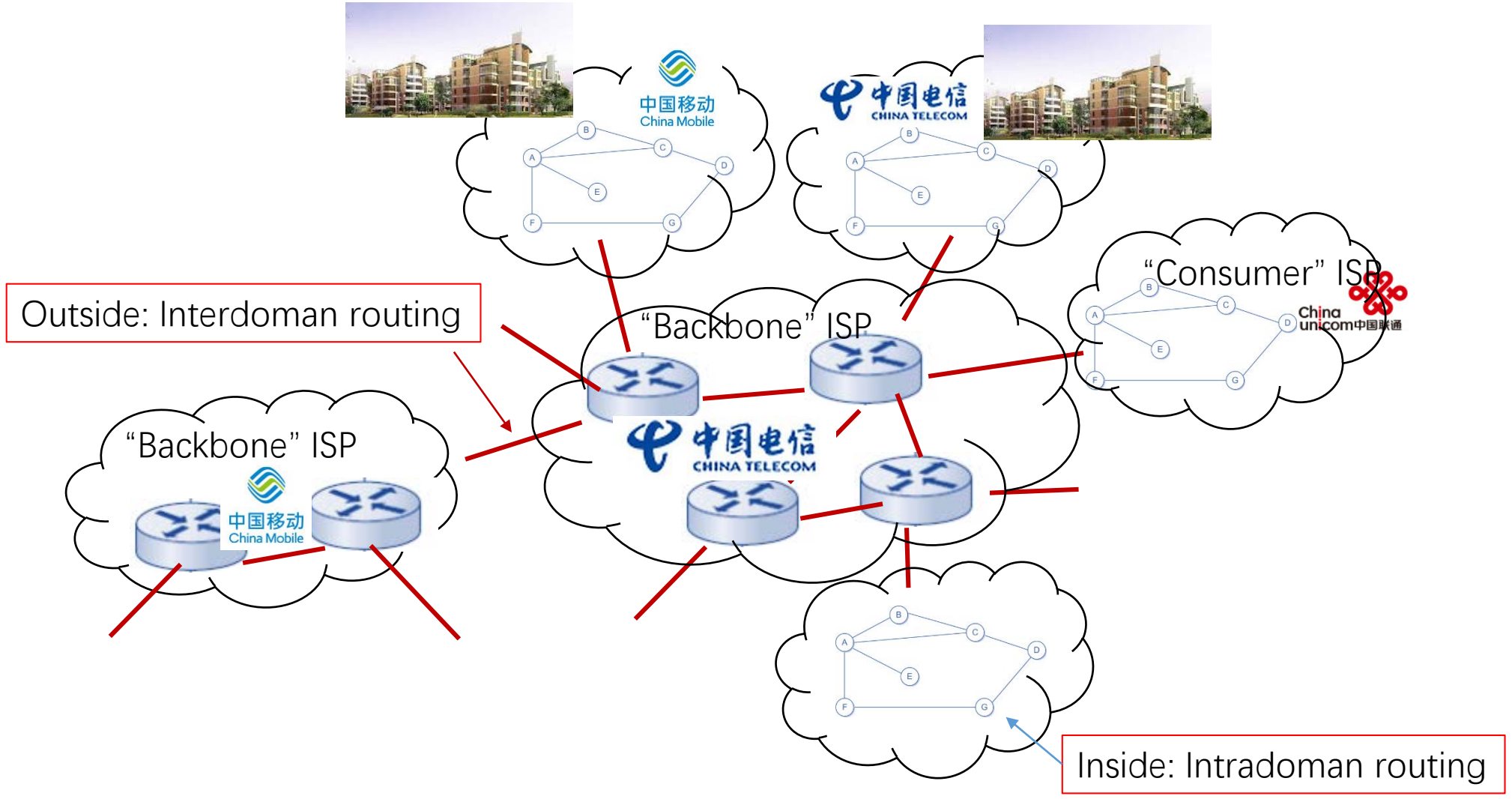


Internet Service Providers

- End systems connect to Internet via “Consumer” ISPs (Internet Service Providers)
  - Residential, company, and university ISPs
- “Consumer” ISPs are connected “Backbone” ISPs
  - Three Major Commercial ISPs in China

  - Other ISPs
    - e.g. cernet

Network of Networks

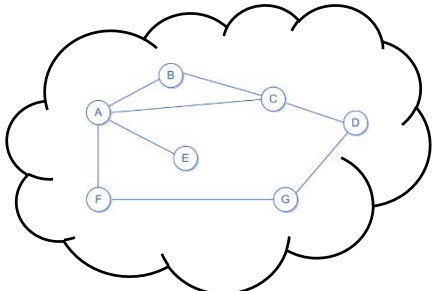


Interdomain Routing Problems

- Scalability: more than 600 million destinations
  - Storage overhead
    - Routing table
  - Calculation overhead
    - Shortest path
  - Communication overhead
    - Exchange routing information
- Routing Management: Complex Routing Polices

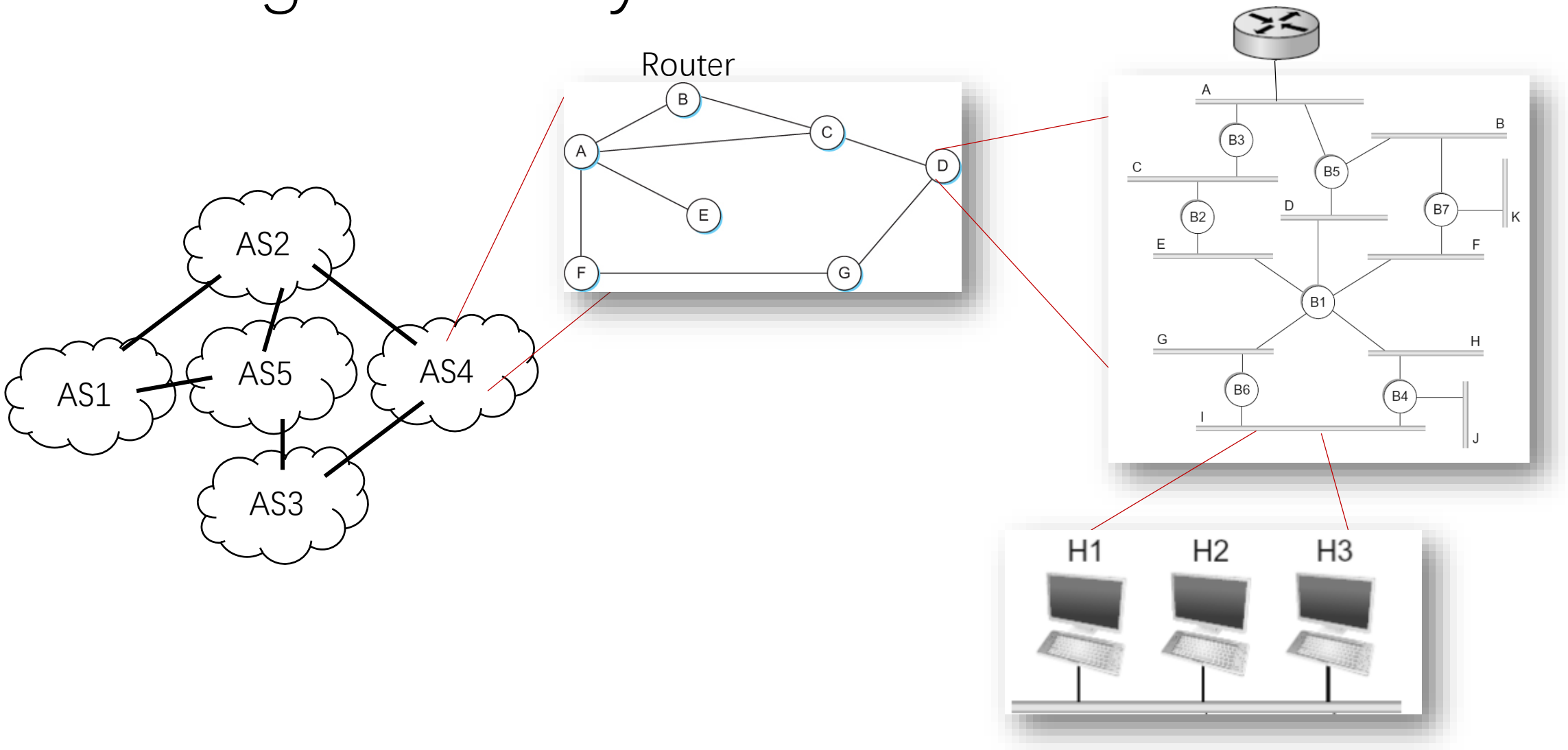
Interdomain Routing: New Hierarchy

- Aggregate Routers into Logical Areas: Autonomous System
- Autonomous System (AS)
  - Corresponds to an administrative domain
  - e.g. University, company, backbone network
- Routers in same AS run the same intradomain routing protocol
  - RIP, OSPF, etc.
- Routers in different AS run intrerdomain routing protocol
  - BGP, EGP
- Interdomain routing element: AS



Autonomous System (AS)

Routing Hierarchy



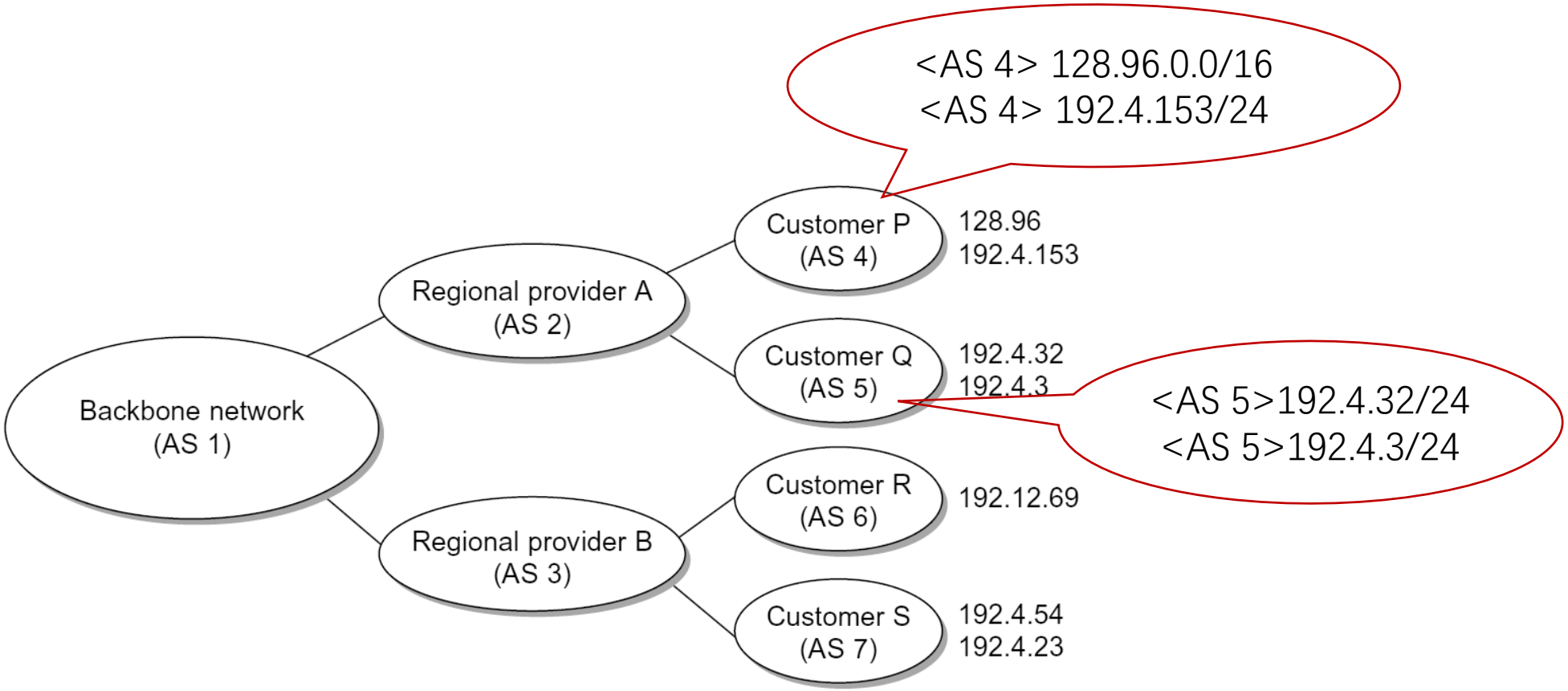
Border Gateway Protocol (BGP)

- Widely-used Interdomain Routing Protocol
- Not for small community or companies
- Routing Element: AS
- Routing Algorithm
  - Target on Reachability
    - Not the “shortest” route
  - Avoid Loops

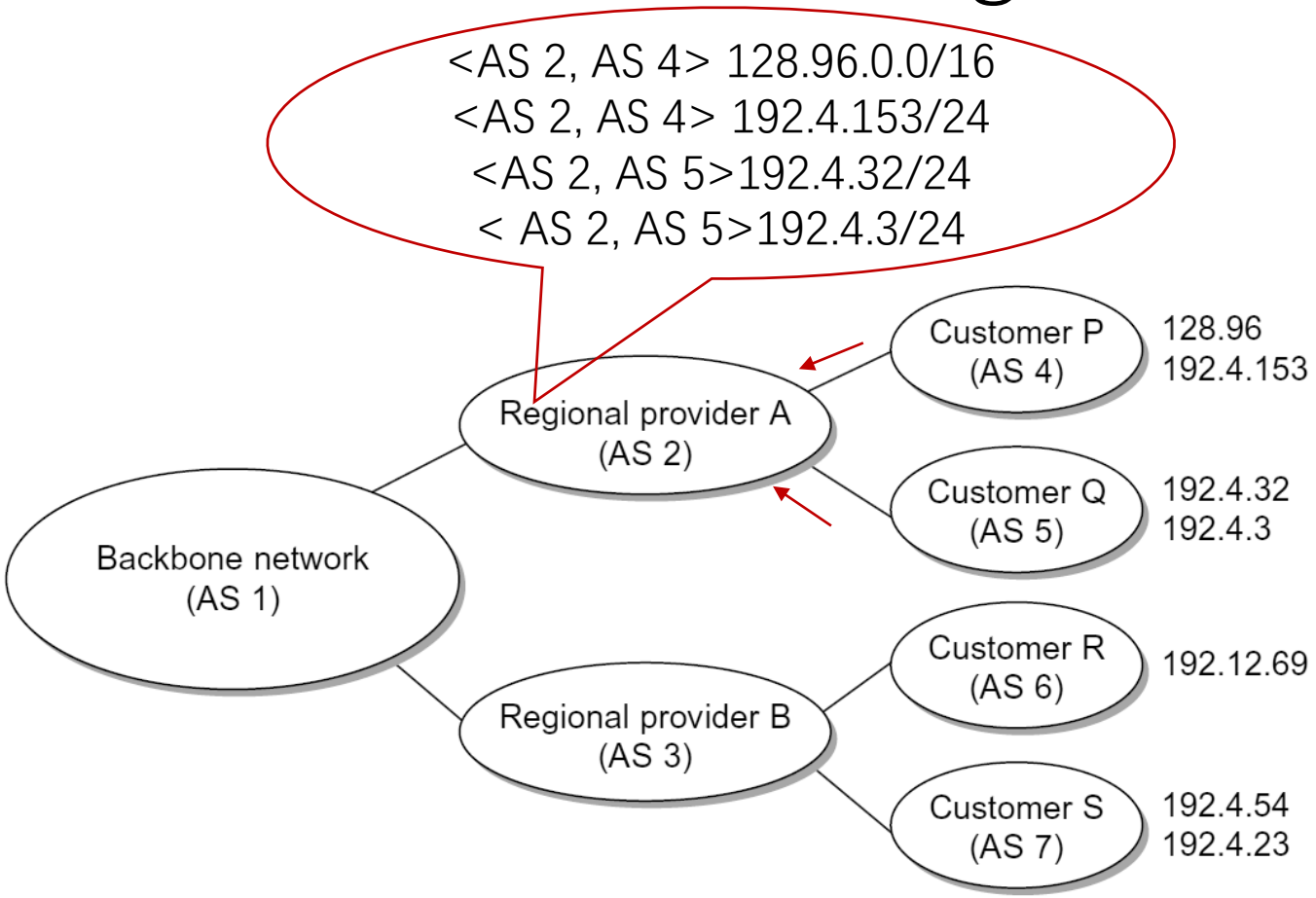
Border Gateway Protocol (BGP)

- Broadcast route entries to neighbors
  - Similar to RIP
  - BGP route entry
    - AS path + network prefix+ next hop
    - e.g., <AS a, AS b, AS c, ⋯> 128.96.0.0/16 12.5.6.1
  - AS number is used to detect loops

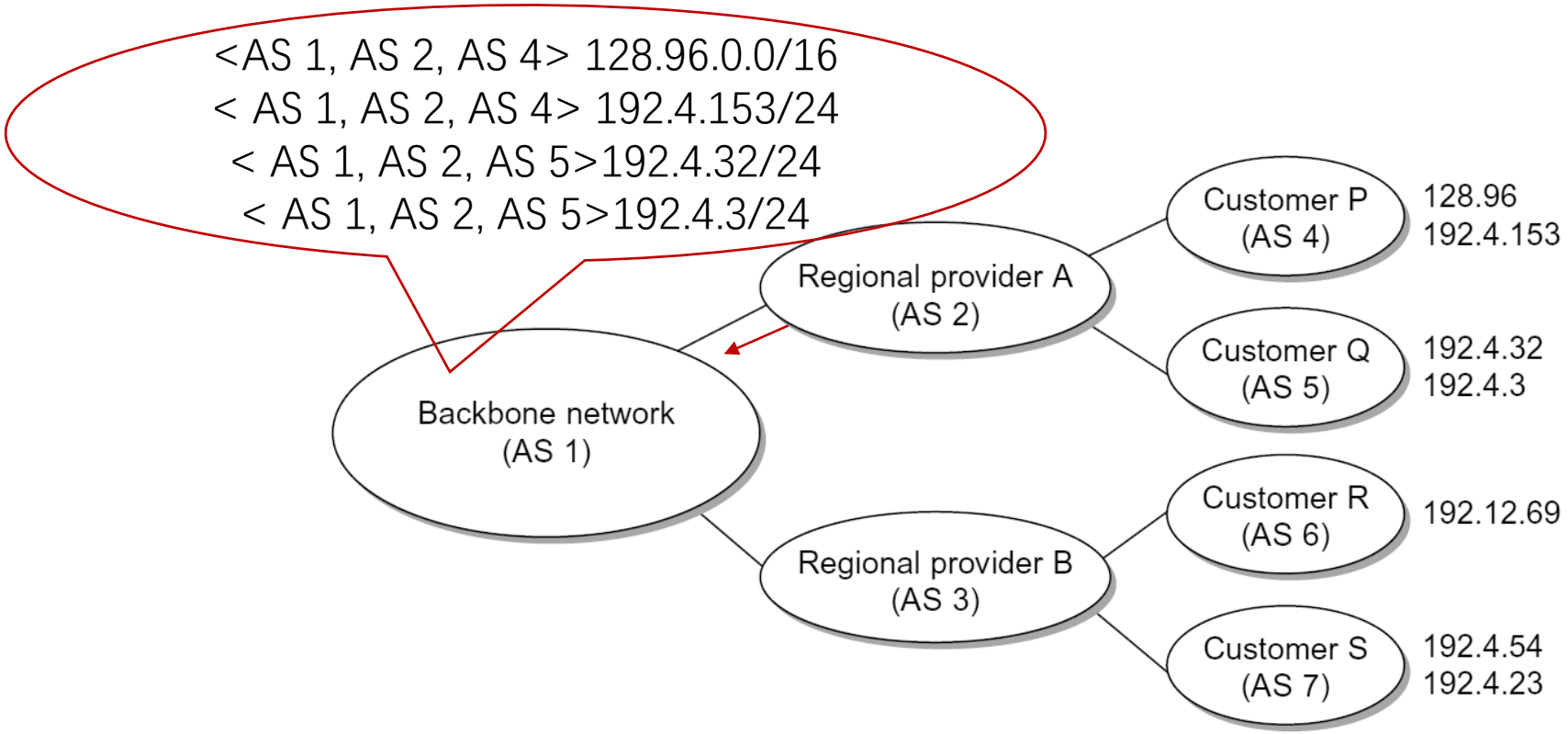
BGP: Path Vector Routing



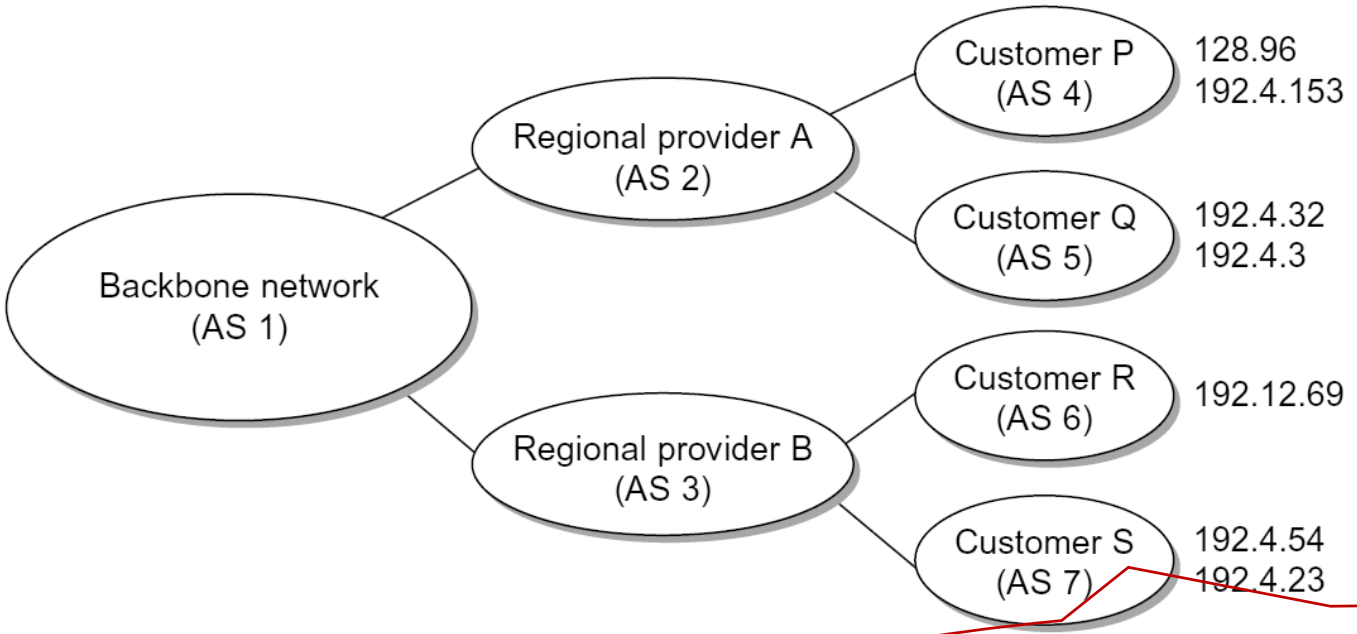
BGP: Path Vector Routing



BGP: Path Vector Routing

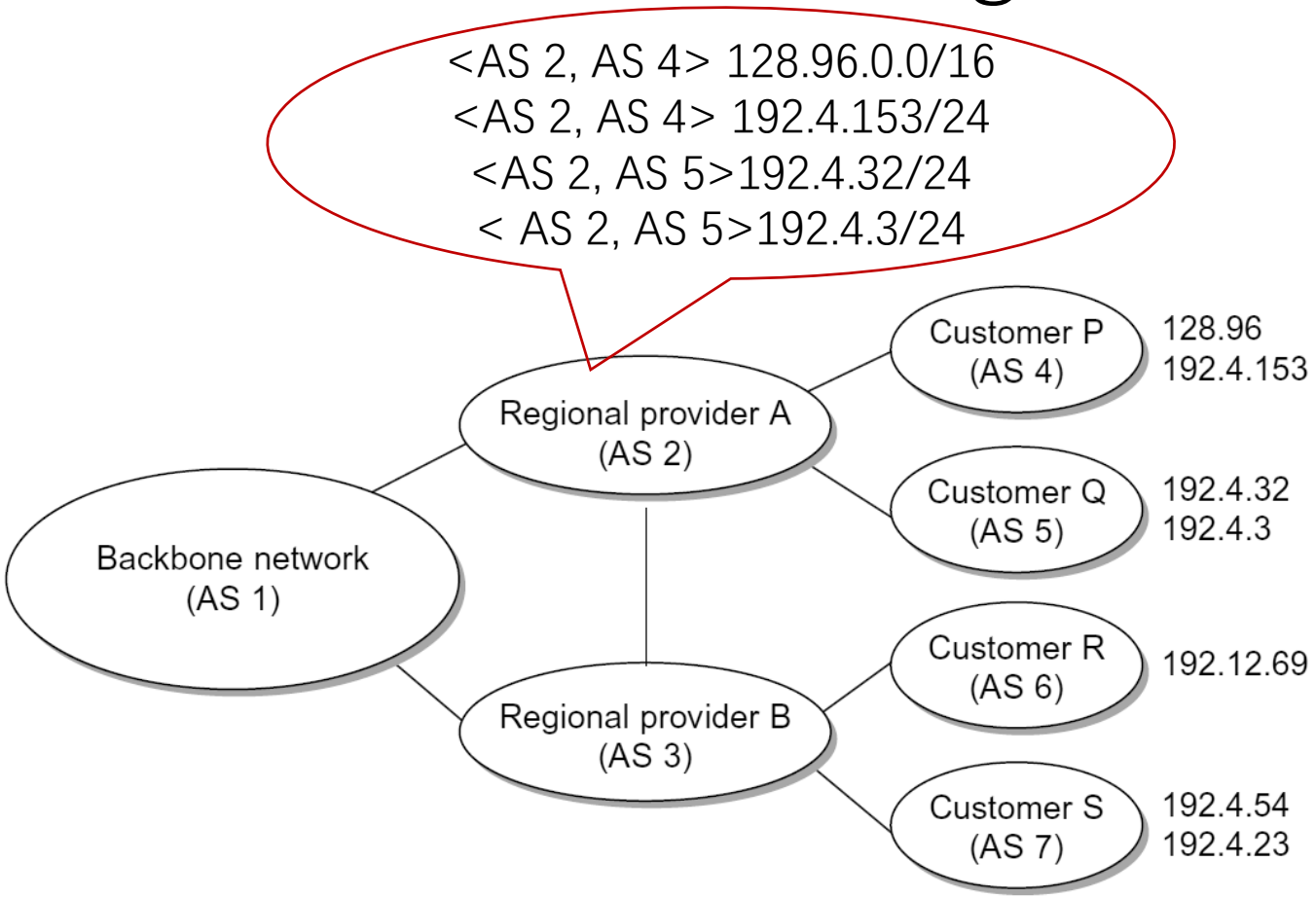


BGP: Path Vector Routing

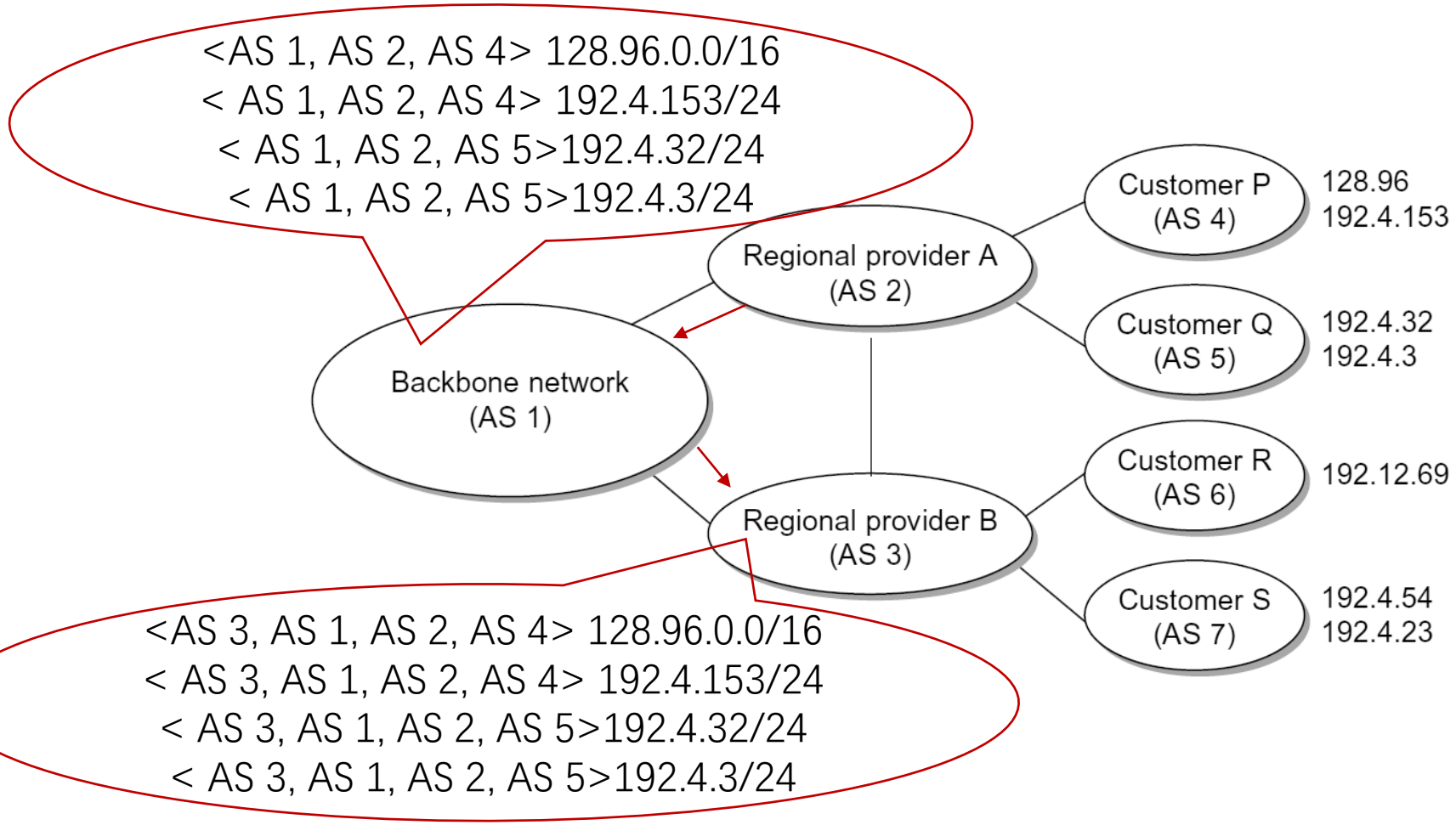


< AS 7, AS 3, AS 1, AS 2, AS 4> 128.96.0.0/16
< AS 7, AS 3, AS 1, AS 2, AS 4> 192.4.153/24
< AS 7, AS 3, AS 1, AS 2, AS 5>192.4.32/24
< AS 7, AS 3, AS 1, AS 2, AS 5>192.4.3/24

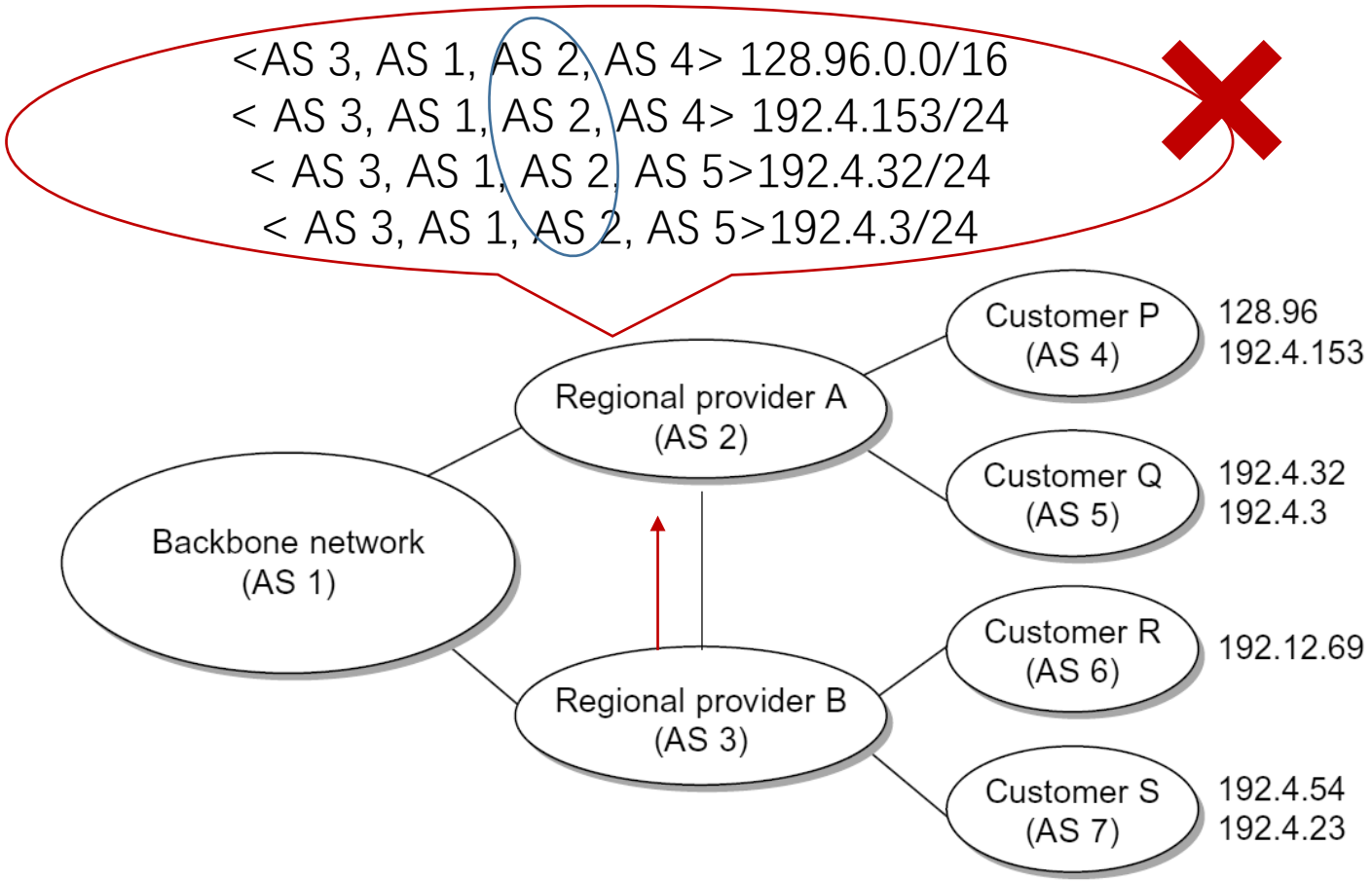
BGP: Path Vector Routing



BGP: Path Vector Routing



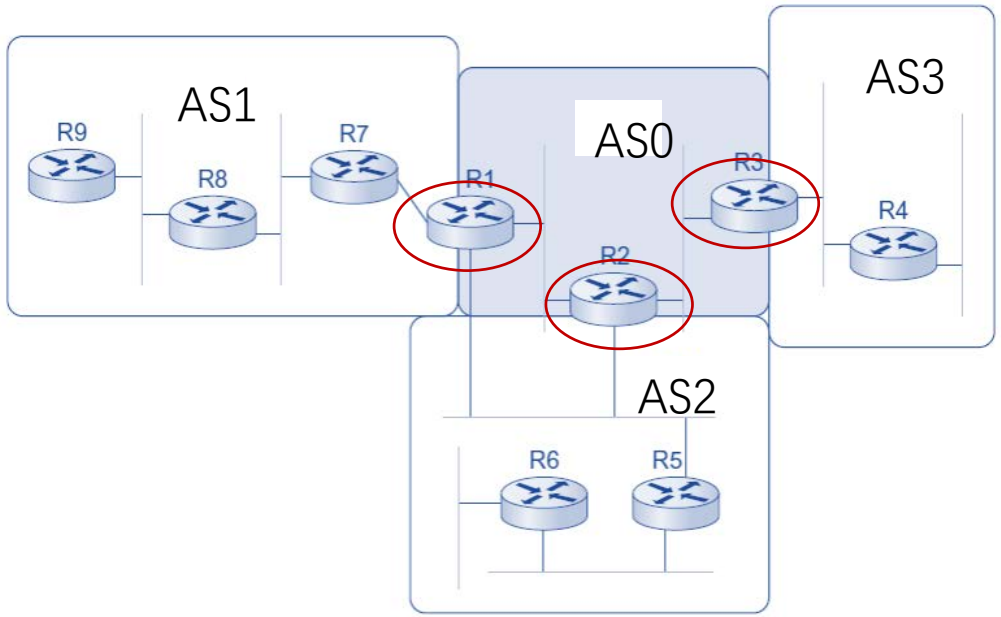
BGP: Path Vector Routing



It's not a useful path

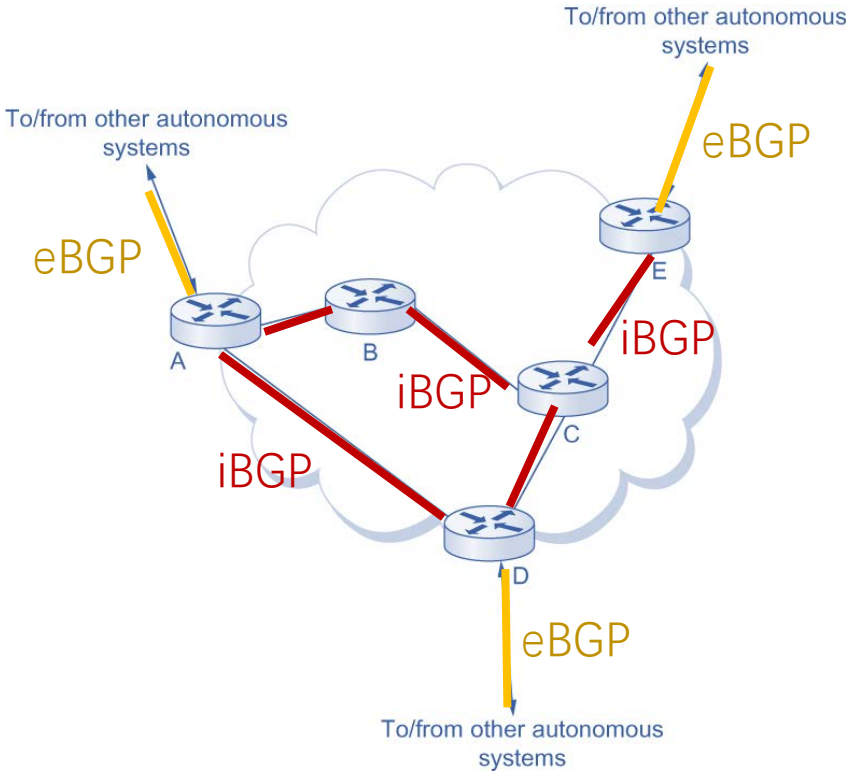
BGP: Border Router

- Border Routers: connecting more than one ASs

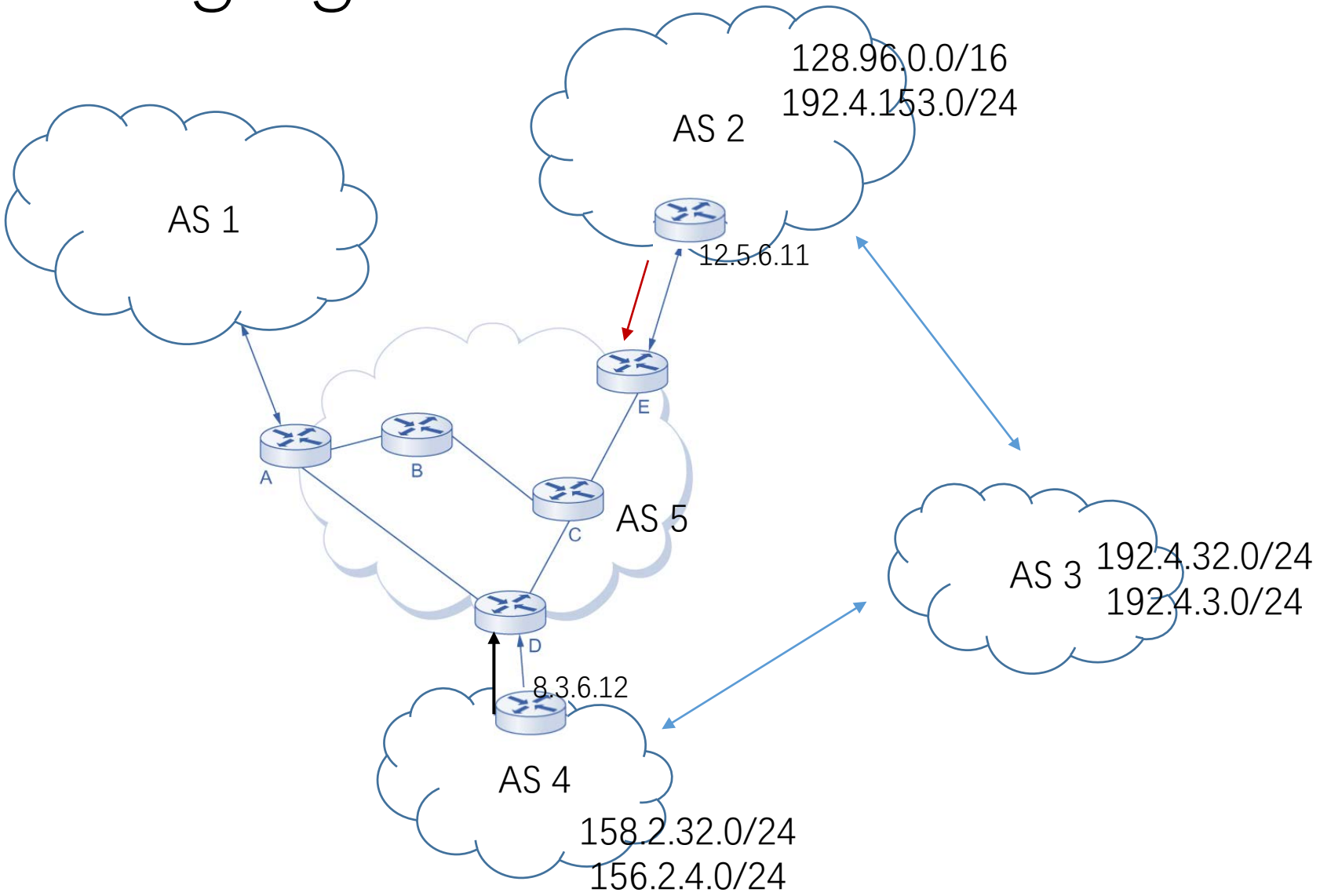


BGP: Border Router

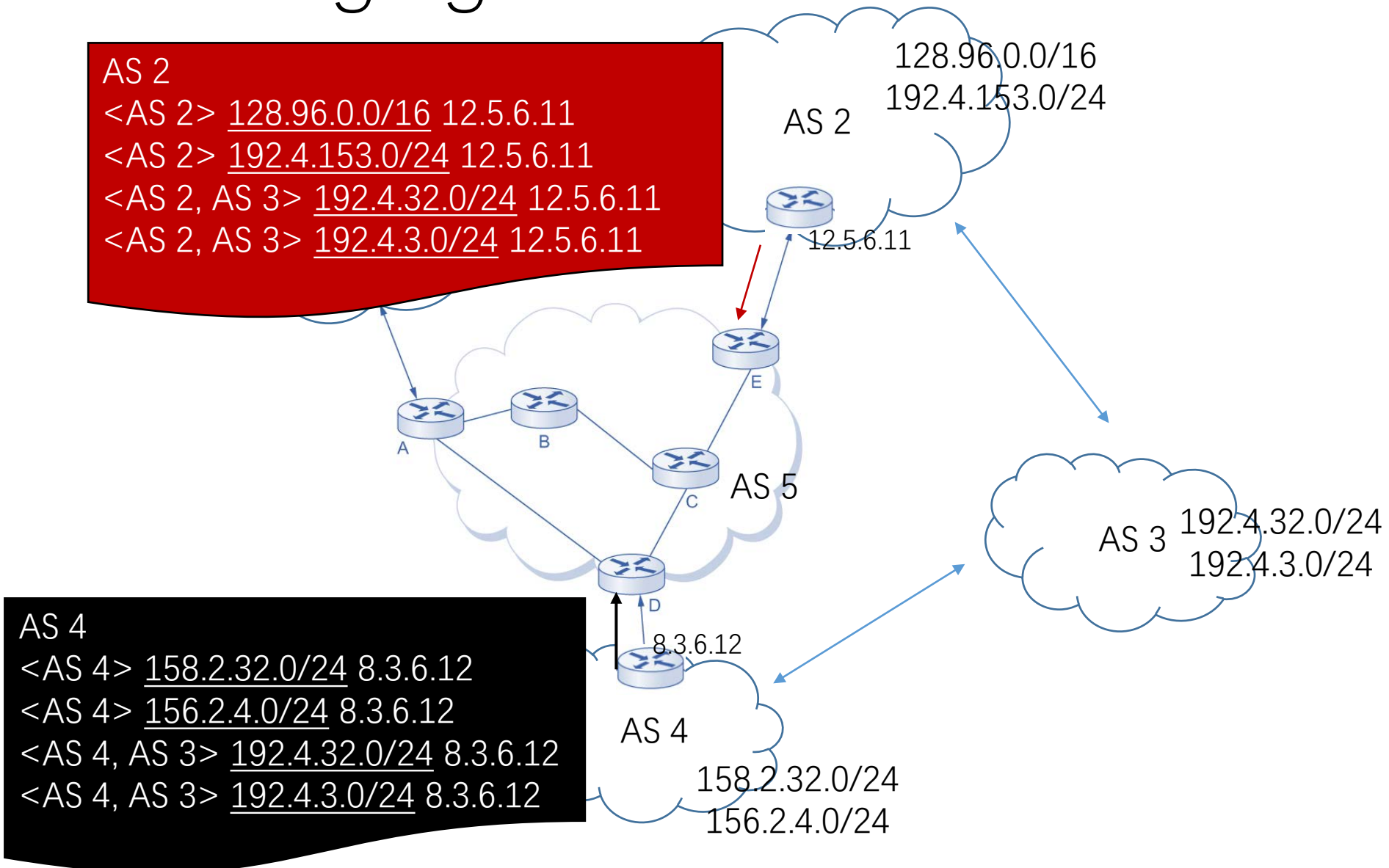
- Border Routers: connecting more than one ASs
  - Selected and configured by AS administrators
  - Routing entries are exchanged with other Boarder Routers through exterior BGP (eBGP)
  - Routing entries are exchanged with routers within the same AS through interior BGP (iBGP)



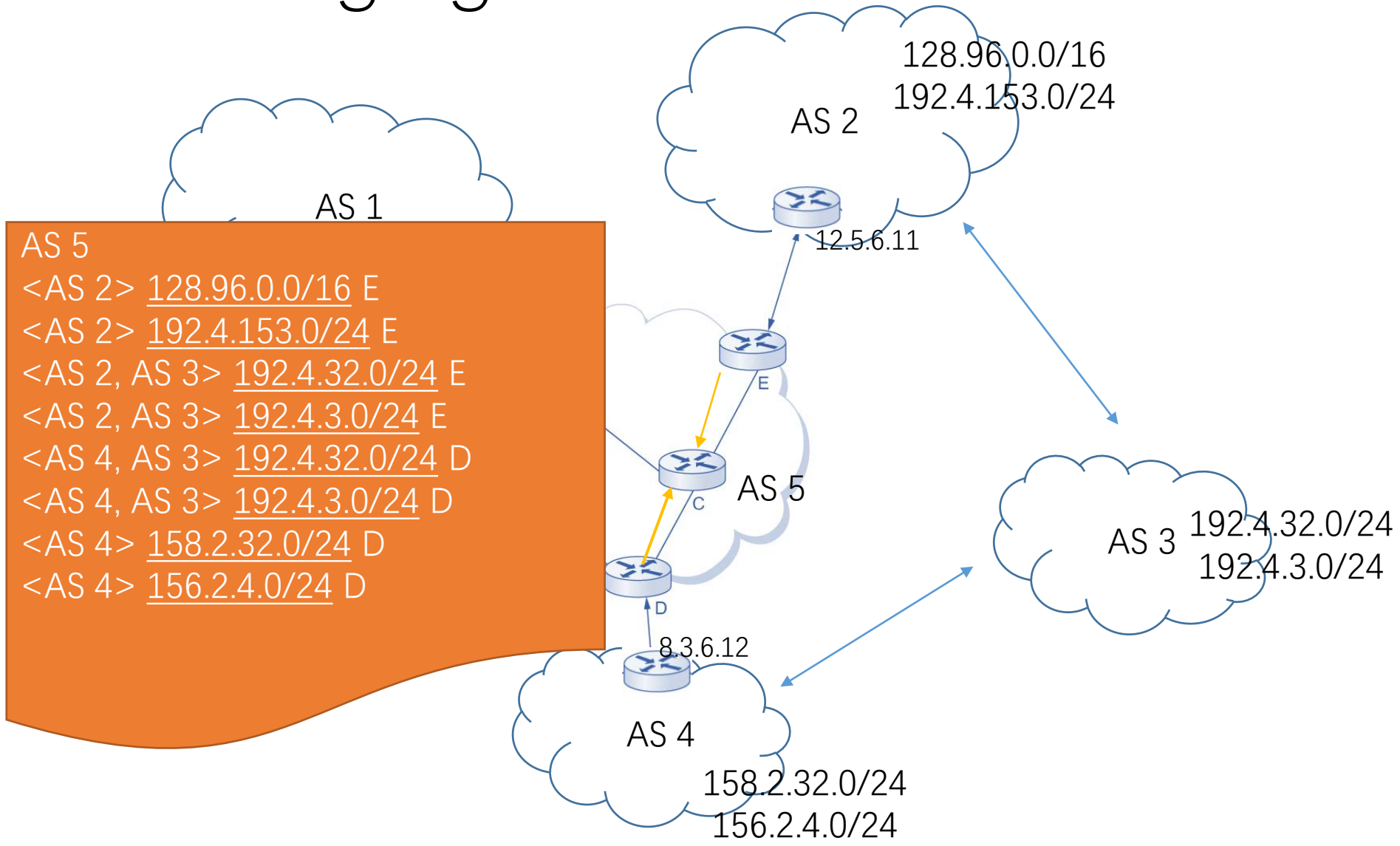
BGP: Exchanging BGP entries



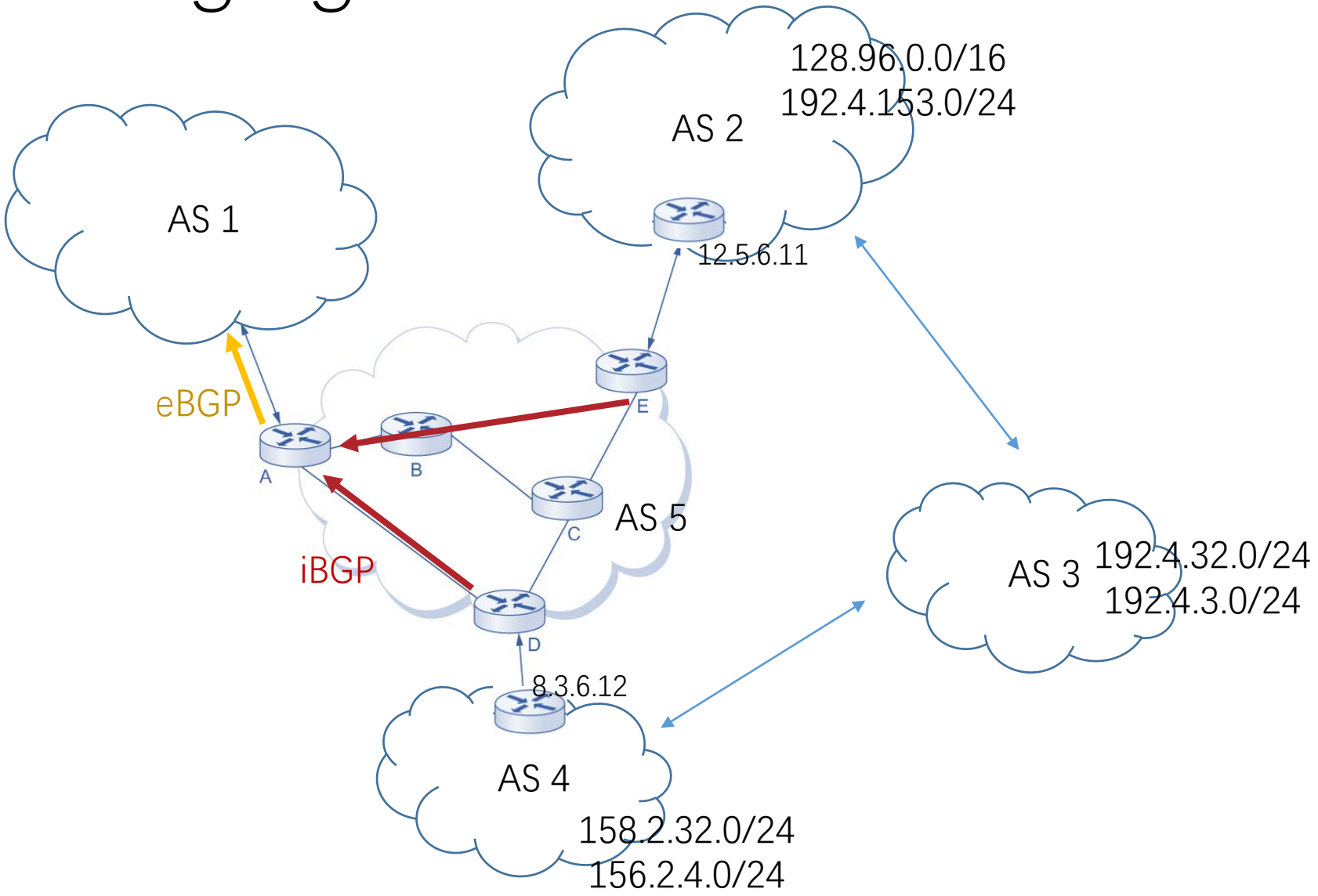
BGP: Exchanging BGP entries



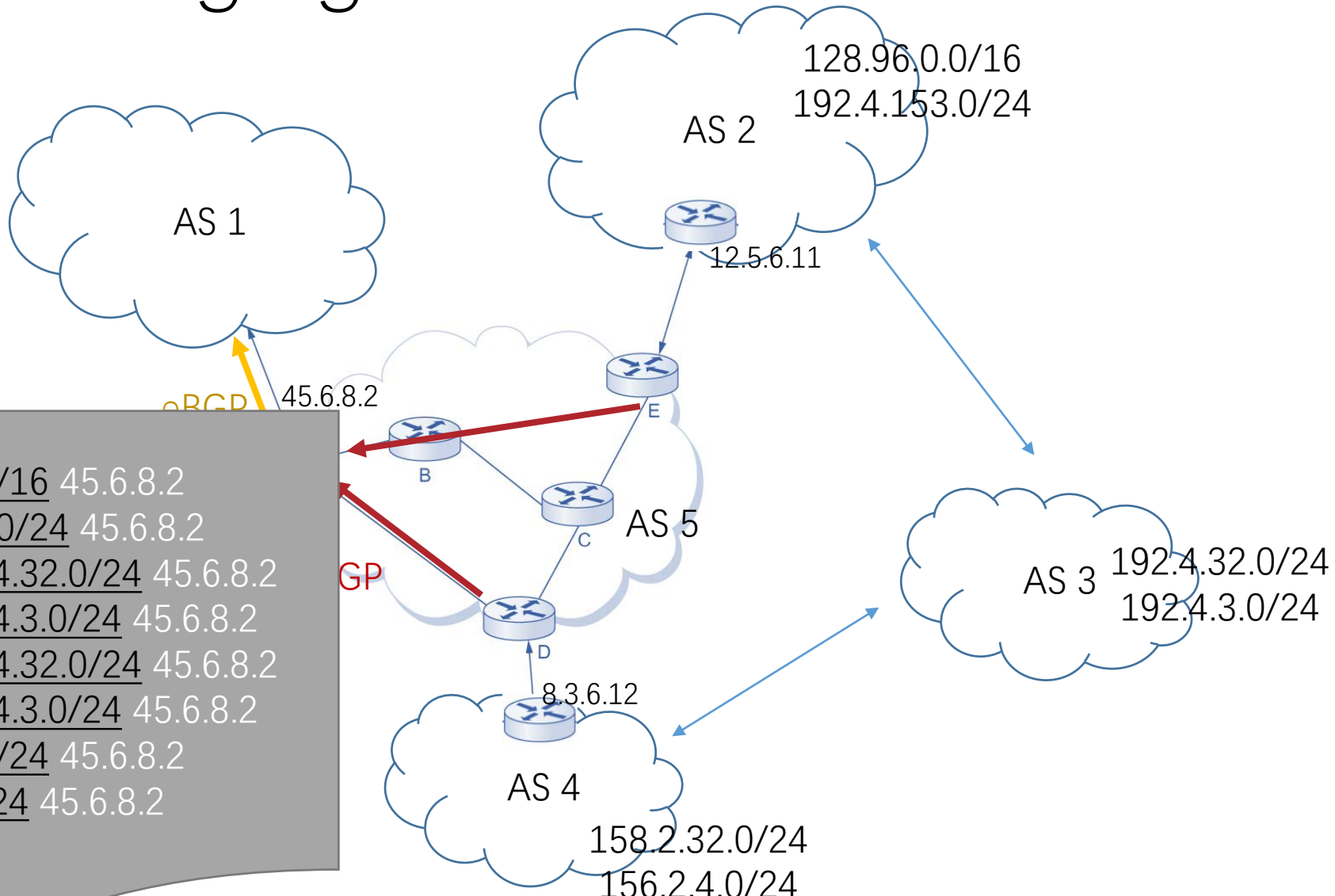
BGP: Exchanging BGP entries



BGP: Exchanging BGP entries

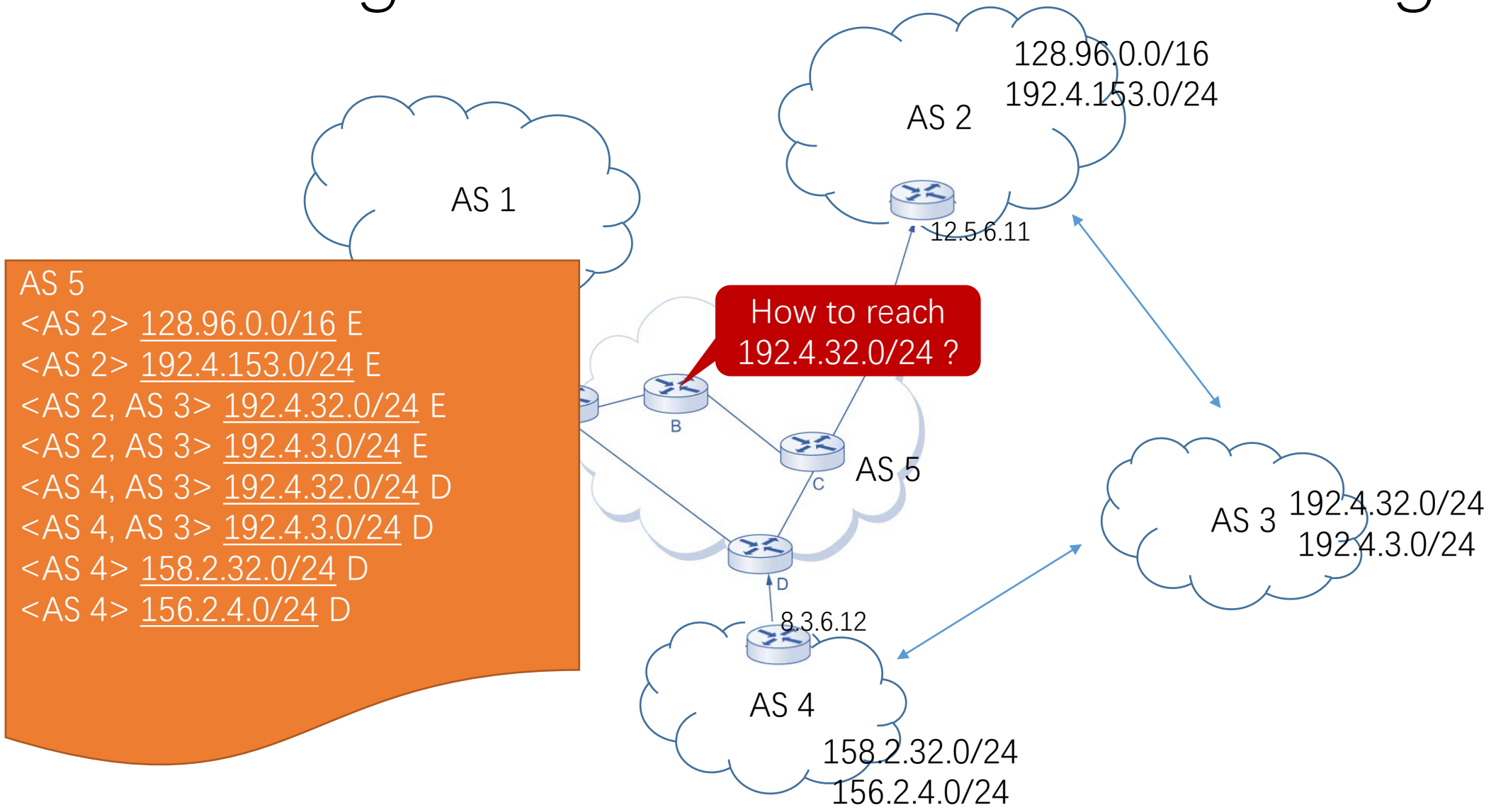


BGP: Exchanging BGP entries



AS1
<AS 5, AS 2> 128.96.0.0/16 45.6.8.2
<AS 5, AS 2> 192.4.153.0/24 45.6.8.2
<AS 5, AS 2, AS 3> 192.4.32.0/24 45.6.8.2
<AS 5, AS 2, AS 3> 192.4.3.0/24 45.6.8.2
<AS 5, AS 4, AS 3> 192.4.32.0/24 45.6.8.2
<AS 5, AS 4, AS 3> 192.4.3.0/24 45.6.8.2
<AS 5, AS 4> 158.2.32.0/24 45.6.8.2
<AS 5, AS 4> 156.2.4.0/24 45.6.8.2

BGP: Integrate with Intradomain Routing



BGP: Integrate with Intradomain Routing

- Combine BGP Table and Inteadomain Routing Table

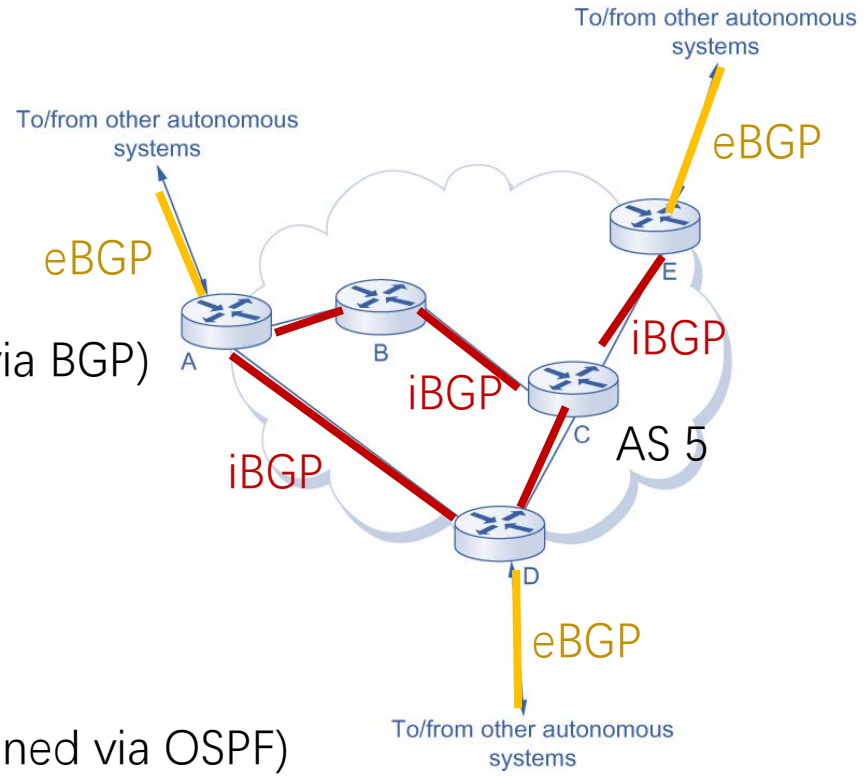
AS 5
<AS 2> 128.96.0.0/16 E
<AS 2> 192.4.153.0/24 E
<AS 2, AS 3> 192.4.32.0/24 E
<AS 2, AS 3> 192.4.3.0/24 E
<AS 4, AS 3> 192.4.32.0/24 D
<AS 4, AS 3> 192.4.3.0/24 D
<AS 4> 158.2.32.0/24 D
<AS 4> 156.2.4.0/24 D

which one ?

Router B's BGP Table (obtained via BGP)

Dest	Next
A	A
C	C
D	C
E	C

Router B's Routing Table (obtained via OSPF)



BGP: Integrate with Intradomain Routing

- Combine BGP Table and Inteadomain Routing Table

AS 5
<AS 2> 128.96.0.0/16 E
<AS 2> 192.4.153.0/24 E
<AS 2, AS 3> 192.4.32.0/24 E
<AS 2, AS 3> 192.4.3.0/24 E
<AS 4, AS 3> 192.4.32.0/24 D
<AS 4, AS 3> 192.4.3.0/24 D
<AS 4> 158.2.32.0/24 D
<AS 4> 156.2.4.0/24 D

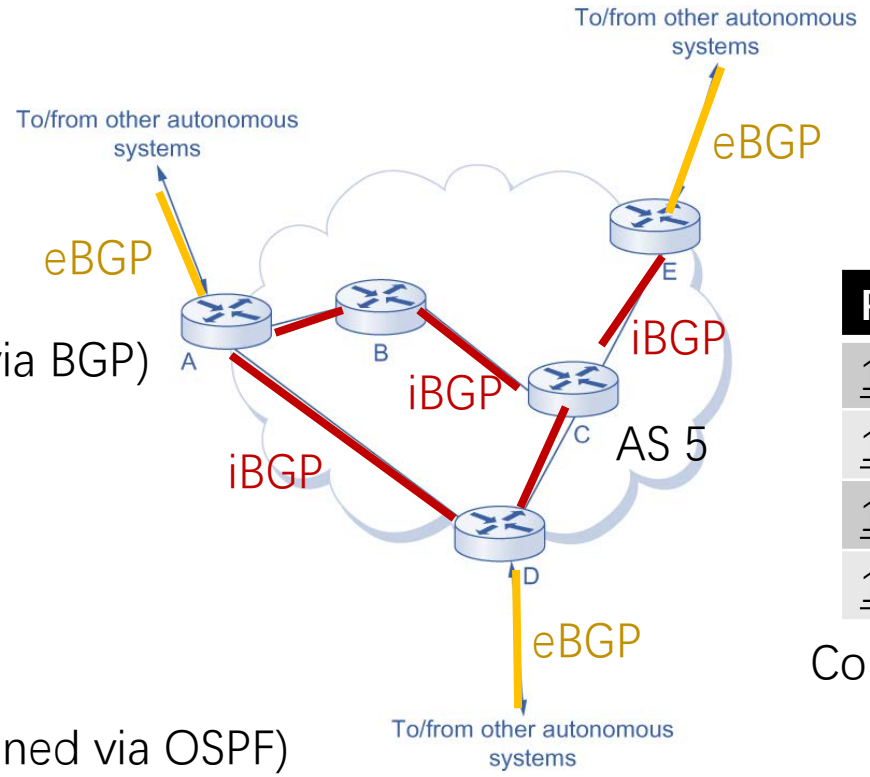
which one ?

Router B's BGP Table (obtained via BGP)

Dest	Next
A	A
C	C
D	C
E	C

Router B's Routing Table (obtained via OSPF)

+



Prefix	Next
128.96.0.0/16	C
192.4.153.0/24	C
158.2.32.0/24	C
156.2.4.0/24	C

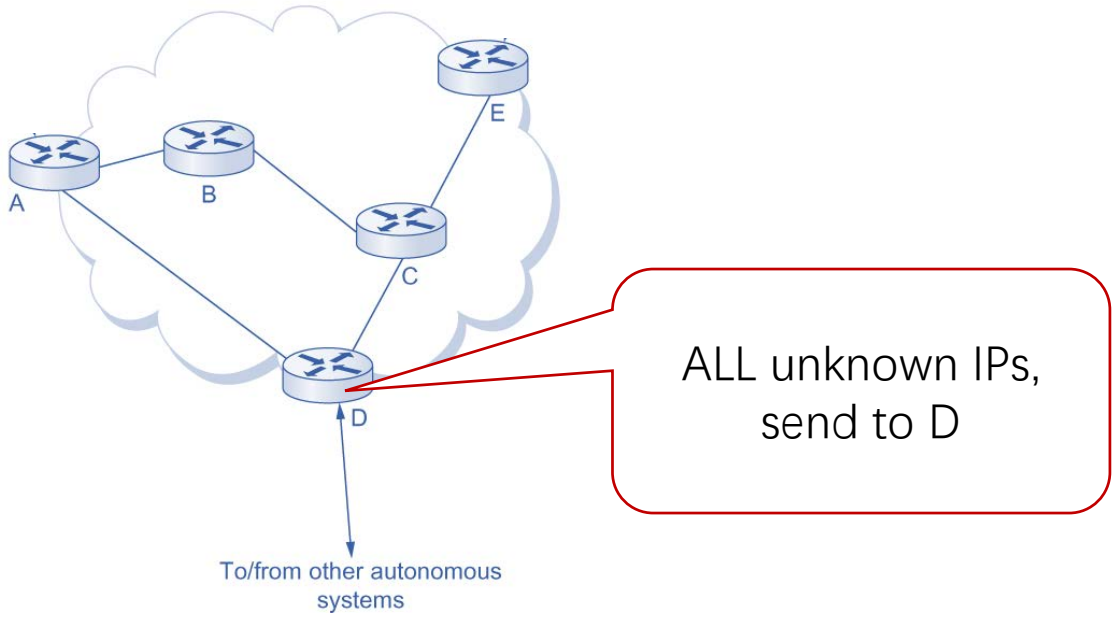
Combined Routing Table for B

BGP: Integrate with Intradomain Routing

- Roughly determine the best border router for certain prefix
  - Selection Priority
    - Local Preference
    - AS hops
    - Distance to the border router
    - BGP ID

BGP: Integrate with Intradomain Routing

- Other Methods: Default Gateway Router (Static Method)
  - Inject a default router entry into all routers in the AS through intradomain routing protocol



Demo

- BGP Entries
  - http://bgp.potaroo.net/as2.0/bgp-active.html
- BGP Looking Glass: check BGP entry from a certain AS to an IP address
  - eg: https://www.bgp4.as/looking-glasses
- AS Number Look Up: more information about AS number
  - eg: https://spyse.com/target/as

Example

- AS route from telia Sofia to my computer
  - https://lg.telia.net/?type=bgp&router=sfia-b2&address=59.78.171.135

Dest Network

```
59.78.171.0/24 (2 entries, 1 announced)

*BGP    Preference: 170/-201
        Source: 2.255.253.187
        Protocol next hop: 2.255.254.180
        State: <Active Int Ext>
        Local AS:  1299 Peer AS:  1299
        Age: 4d 11:16:10        Metric: 100     Metric2: 2682
        AS path: 4637 4637 4637 4637 4538 4538 24364 I  (Originator)
```

Src Router

Border Router

Src AS

AS Route

AS1299

```
Country: EU
Registration Date: 1993-09-01
Registrar: ripencc
Owner: TELIANET Telia Carrier, EU
```

AS4637

```
Country: HK
Registration Date: 1995-10-30
Registrar: apnic
Owner: ASN-TELSTRA-GLOBAL Telstra Global, HK
```

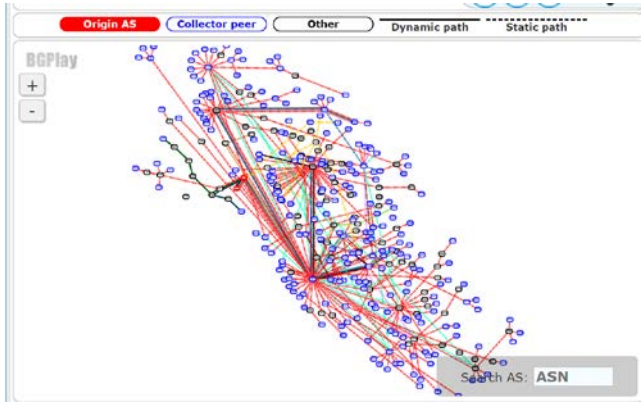
AS4538

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

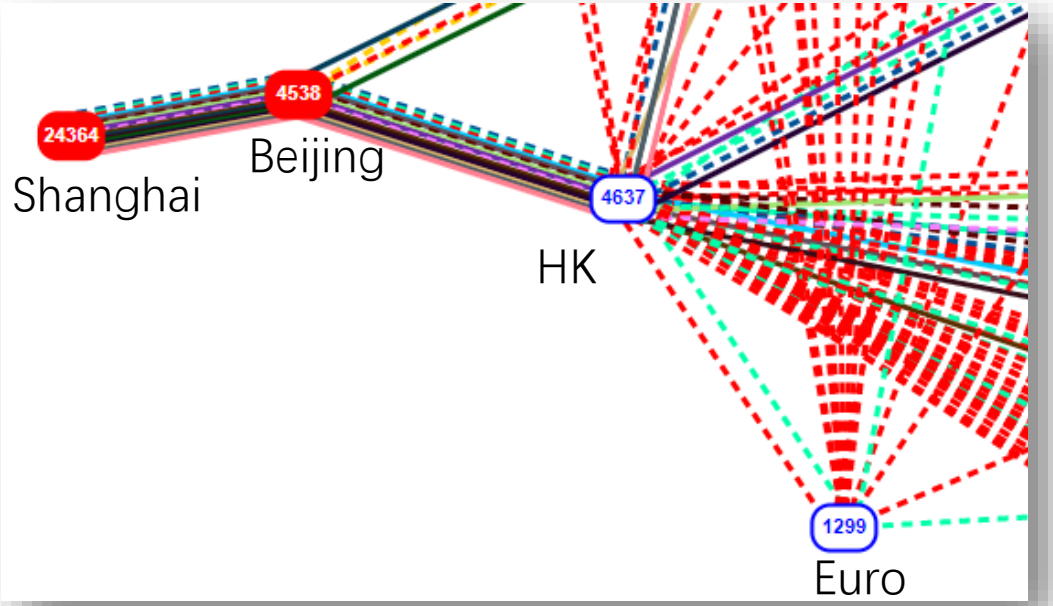
AS24364

```
Country: CN
Registration Date: 2005-03-03
Registrar: apnic
Owner: CNGI-SH-IX-AS-AP CERNET2 IX at Shanghai Jiaotong University, CN
```

Example



- AS route from multiple ASs
  - https://stat.ripe.net/special/bgplay#bgplay_fetch.resource=59.78.171.135




```
Network: AS1299 - Telia Carrier
Router: Sofia (sfia-b2)
Command: traceroute 59.78.171.135 as-number-lookup

traceroute to 59.78.171.135 (59.78.171.135), 30 hops max, 52 byte packets
 1  win-bb3-link.telia.net (62.115.121.36)  49.289 ms  47.015 ms  47.068 ms
 2  ffm-bb1-link.telia.net (62.115.137.202)  46.935 ms  47.882 ms  48.332 ms
 3  prs-bb3-link.telia.net (62.115.123.13)  47.737 ms  46.954 ms  47.681 ms
 4  ldn-bb3-link.telia.net (62.115.123.68)  47.194 ms ldn-bb3-link.telia.net (62.115.13
 5  ldn-b7-link.telia.net (62.115.138.151)  46.976 ms  48.680 ms  46.756 ms
 6  telstra-ic-324829-ldn-b7.c.telia.net (62.115.154.237)  47.813 ms  48.194 ms  47.571
 7  i-91.ulco-core02.telstraglobal.net (202.40.148.33) [AS  4637]  47.833 ms  51.549 ms
 8  i-0-1-1-1.gfr4-core01.telstraglobal.net (202.84.141.121) [AS  4637]  283.475 ms i-1
298.512 ms
 9  202.84.153.26 (202.84.153.26) [AS  4637]  282.290 ms  281.651 ms 202.84.157.37 (202
10  CER-0003.10026.telstraglobal.net (61.8.59.38) [AS  4637]  274.015 ms  294.241 ms  2
11  * * 101.4.114.181 (101.4.114.181) [AS  4538]  317.416 ms
12  101.4.118.121 (101.4.118.121) [AS  4538]  294.191 ms 101.4.114.238 (101.4.114.238)
13  101.4.114.58 (101.4.114.58) [AS  4538]  323.583 ms  317.161 ms  323.827 ms
14  101.4.116.85 (101.4.116.85) [AS  4538]  318.175 ms  294.344 ms  318.948 ms
15  101.4.115.106 (101.4.115.106) [AS  4538]  322.209 ms  321.473 ms  343.786 ms
16  202.112.27.2 (202.112.27.2) [AS  4538]  340.422 ms  343.047 ms  340.877 ms
17  * * *
18  * * *
19  * * *
20  * * *

(Timeout)
```

Interdomain Routing Problems

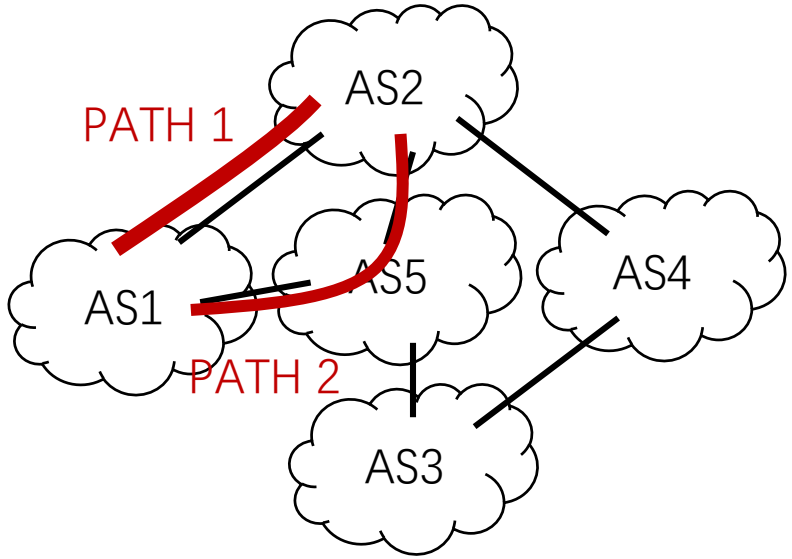
- Scalability: More than 600 million destinations
  - Storage
    - Routing Table
  - Calculation
    - Shortest Path
  - Communication
    - Exchanges Routing Information

➢Routing Management: Complex Routing Polices

Local Preference

- AS 1 can reach AS2 through AS5
  - AS5 helps AS1 to forward traffic
    - ISP negotiation
    - Priority
      - e.g., AS path prepending
  - AS5 blocks AS1 traffic to AS2
    - Does not broadcast AS2 entry to AS1

```
AS path: 4637 4637 4637 4637 4538 4538 24364
```



Reference

- Textbook 3.3
- Textbook 4.1
- http://www.ciscopress.com/articles/article.asp?p=24090