Lecture Note 5
Internet Routing Protocols

Basic Routing

Routing Information Protocol (RIP)

Open Shortest Path First (OSPF)

Border Gateway Protocol (BGP)

Lecture 6
Internet Routing Protocols

Basic Routing

Autonomous Systems

- Global Internet viewed as collection of autonomous systems.
- Autonomous system (AS) is a set of routers or networks administered by a single organization
- Same routing protocol need not be run within the AS
- But, to the outside world, an AS should present a *consistent picture of what ASs are reachable* through it
- Stub AS: has only a single connection to the outside world.
- Multihomed AS: has multiple connections to the outside world, but refuses to carry transit traffic
- Transit AS: has multiple connections to the outside world, and can carry transit and local traffic.

AS Number

- For exterior routing, an AS needs a globally unique AS 16-bit integer number
- Currently, there are about 11,000 registered ASs in Internet (and growing)
- *Stub AS*, which is the most common type, does not need an AS number since the prefixes are placed at the provider's routing table
- *Transit AS* needs an AS number
- Request an AS number from one of the five RIRs (Regional Internet Registries)
  - ARIN: American Registry for Internet Numbers
  - RIPE NCC: Réseaux IP Européens Network Coordination Centre
  - APNIC: Asia Pacific Network Information Centre
  - LACNIC: Latin America and Caribbean Network Information Centre
  - AFRINIC: African Network Information Centre
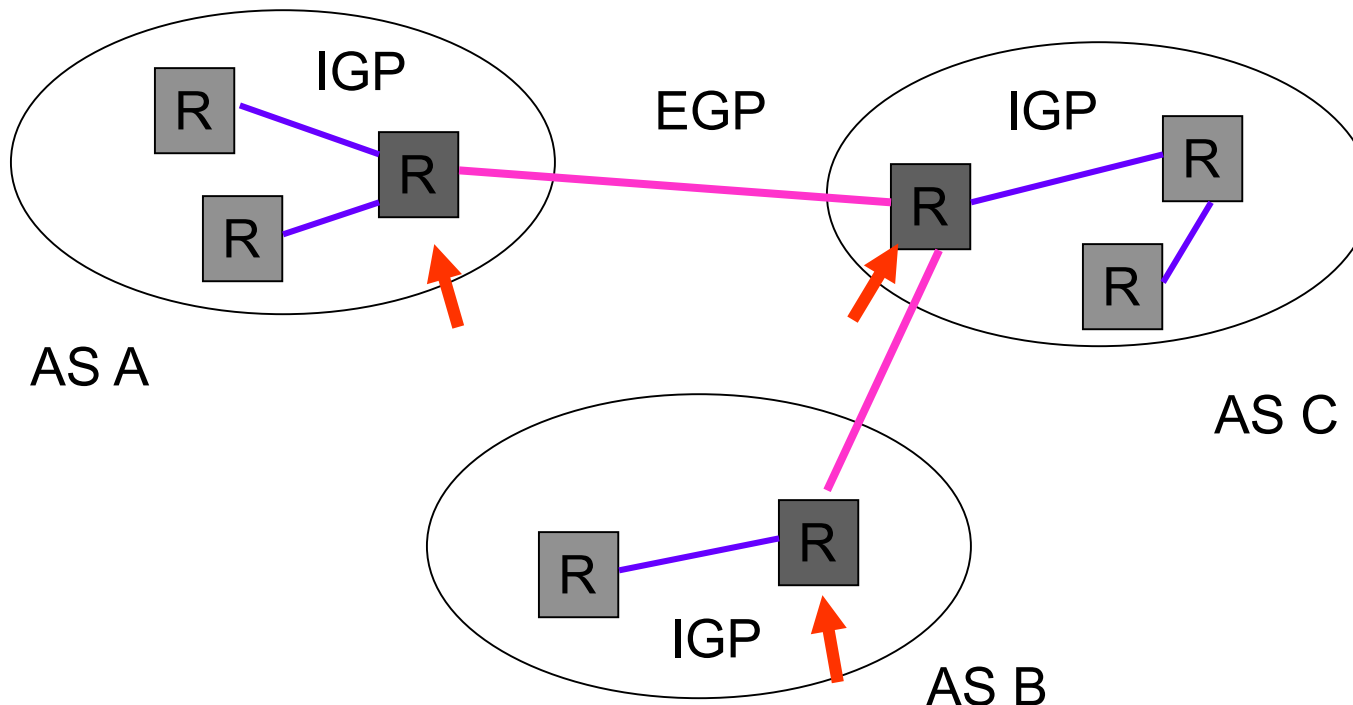
Inter and Intra Domain Routing

Interior Gateway Protocol (IGP): routing within AS
- RIP, OSPF, IGRP, EIGRP, IS-IS

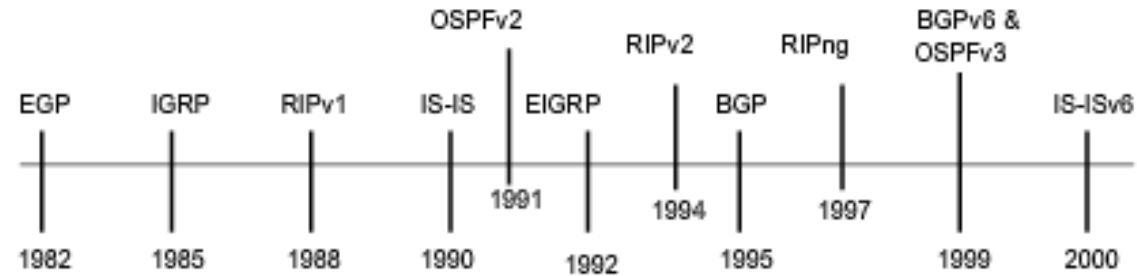*Exterior Gateway Protocol (EGP):* routing between AS' s
- BGPv4

Border Gateways perform IGP & EGP routing
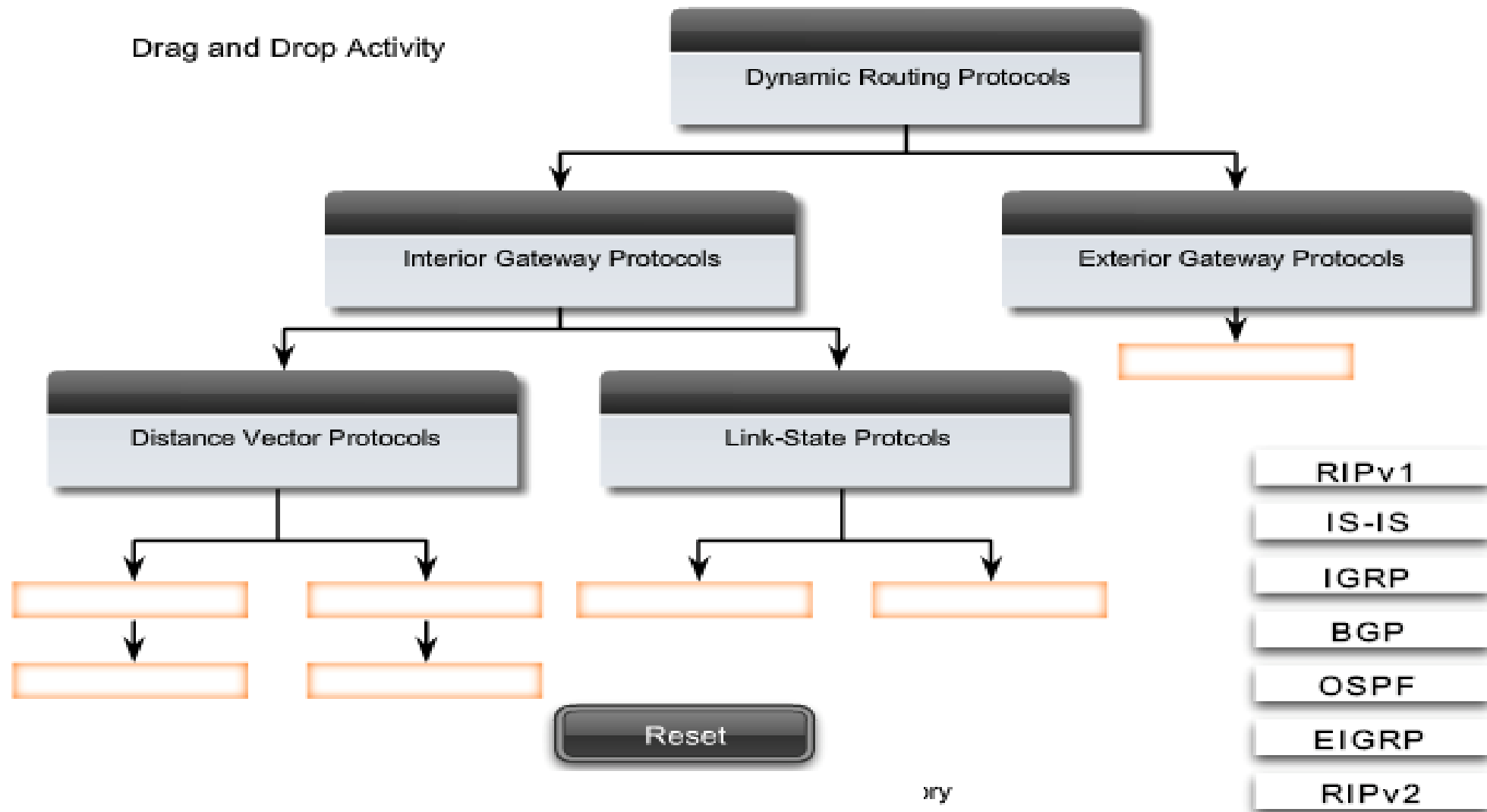
Inter and Intra Domain Routing

Routing Protocols Evolution and Classification

Interior Gateway Protocols

Exterior Gateway Protocols

	Distance Vector Routing Protocols		Link State Routing Protocols		Path Vector
Classful	RIP	IGRP			EGP
Classless	RIPv2	EIGRP	OSPFv2	IS-IS	BGPv4
IPv6	RIPng	EIGRP for IPv6	OSPFv3	IS-IS for IPv6	BGPv4 for IPv6

Highlighted routing protocols are the focus of this course.

Image Source: www.netacad.com

Inter and Intra Domain Routing

Drag and Drop Activity

Dynamic Routing Protocols

Interior Gateway Protocols

Exterior Gateway Protocols

Distance Vector Protocols

Link-State Protcols

Reset

RIPv1

IS-IS

IGRP

BGP

OSPF

EIGRP

RIPv2

Review of Routing

- **Distance Vector Shortest Path Routing Algorithm**
  - Bellman-Ford SPRA
  - Distributed SPRA

- **Link State Shortest Path Routing Algorithm**
  - Dijkstra SPRA
  - Centralized SPRA

- For more details, let's go to Lecture 2-2 Routing

Outline

- Basic Routing
- **Routing Information Protocol (RIP)**
- Open Shortest Path First (OSPF)
- Border Gateway Protocol (BGP)

Routing Information Protocol (RIP)

- RFC 1058
- Uses the distance-vector algorithm
- Runs on top of UDP, port number 520
- Metric: number of hops
- Max limited to 15
  - suitable for small networks (local area environments)
  - value of 16 is reserved to represent infinity
  - small number limits the *count-to-infinity* problem

RIP Operation

- Router sends update message to neighbors every 30 sec
- A router expects to receive an update message from each of its neighbors within 180 seconds in the worst case
- If router does not receive update message from neighbor X within this limit, it assumes the link to X has failed and sets the corresponding minimum cost to 16 (infinity)
- Uses *split horizon with poisoned reverse*

RIP Protocol

- Routers run RIP in active mode (advertise distance vector tables)
- Hosts can run RIP in passive mode (update distance vector tables, but do not advertise)
- RIP datagrams broadcast over LANs & specifically addressed on pt-pt or multi-access non-broadcast nets
- Two RIP packet types:
  - *request* to ask neighbor for distance vector table
  - *response* to advertise distance vector table
    - periodically; in response to request; triggered

Outline

- Basic Routing
- Routing Information Protocol (RIP)
- **Open Shortest Path First (OSPF)**
- Border Gateway Protocol (BGP)

Open Shortest Path First

- RFC 2328 (v2)
- Fixes some of the deficiencies in RIP
- Enables each router to learn complete network topology
- Each router monitors the *link state* to each neighbor and floods the link-state information to other routers
- Each router builds an identical *link-state database*
- Allows router to build shortest path tree with router as root
- OSPF typically converges faster than RIP when there is a failure in the network
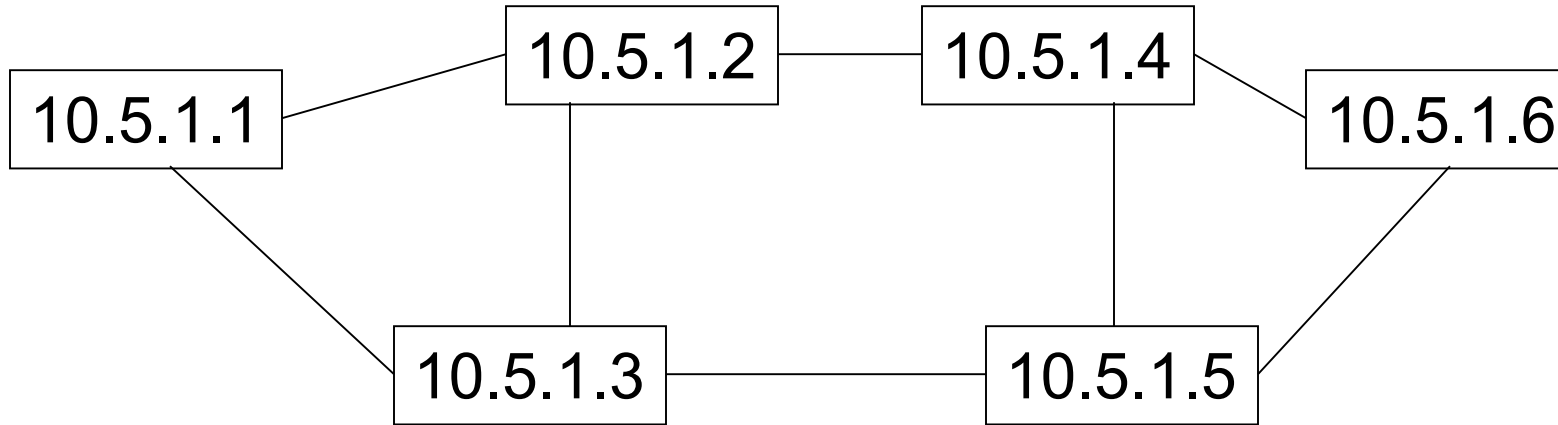
OSPF Features

- *Multiple routes* to a given destination, one per type of service
- Support for *variable-length subnetting* by including the subnet mask in the routing message
- Distribution of traffic over *multiple paths* of equal cost
- Uses *notion of area* to partition sites into subsets
- Support *host-specific routes* as well as net-specific routes
- *Designated router* to minimize table maintenance overhead

Flooding

- Used in OSPF to distribute link state (LS) information
- Forward incoming packet to all ports except where packet came in
- Packet eventually reaches destination as long as there is a path between the source and destination
- Generates exponential number of packet transmissions
- Approaches to limit # of transmissions:
  - Use a TTL at each packet; won't flood if TTL is reached
  - Each router adds its identifier to header of packet before it floods the packet; won't flood if its identifier is detected
  - Each packet from a given source is identified with a unique sequence number; won't flood if sequence number is same

Example OSPF Topology

```
                 ┌──────────┐      ┌──────────┐
                 │ 10.5.1.2 │──────│ 10.5.1.4 │        ┌──────────┐
   ┌──────────┐  └──────────┘      └──────────┘        │ 10.5.1.6 │
   │ 10.5.1.1 │╱        │                │              └──────────┘
   └──────────┘╲        │                │             ╱
                ╲       │                │            ╱
                 ┌──────────┐      ┌──────────┐     ╱
                 │ 10.5.1.3 │──────│ 10.5.1.5 │────╱
                 └──────────┘      └──────────┘
```

- **At steady state:**
- **All routers have same LS database**
- **Know how many routers in network**
- **Interfaces & links between routers**
- **Cost of each link**
- **Occasional Hello messages (10 sec) & LS updates sent (30 min)**

OSPF Network

- **To improve scalability, AS may be partitioned into areas**
  - Area is identified by 32-bit Area ID
  - Router in area only knows complete topology inside area & limits the flooding of link-state information to area
  - Area border routers summarize info from other areas
- **Each area must be connected to backbone area (0.0.0.0)**
  - Distributes routing info between areas
- **Internal router has all links to nets within the same area**
- **Area border router has links to more than one area**
- **Backbone router has links connected to the backbone**

OSPF Areas
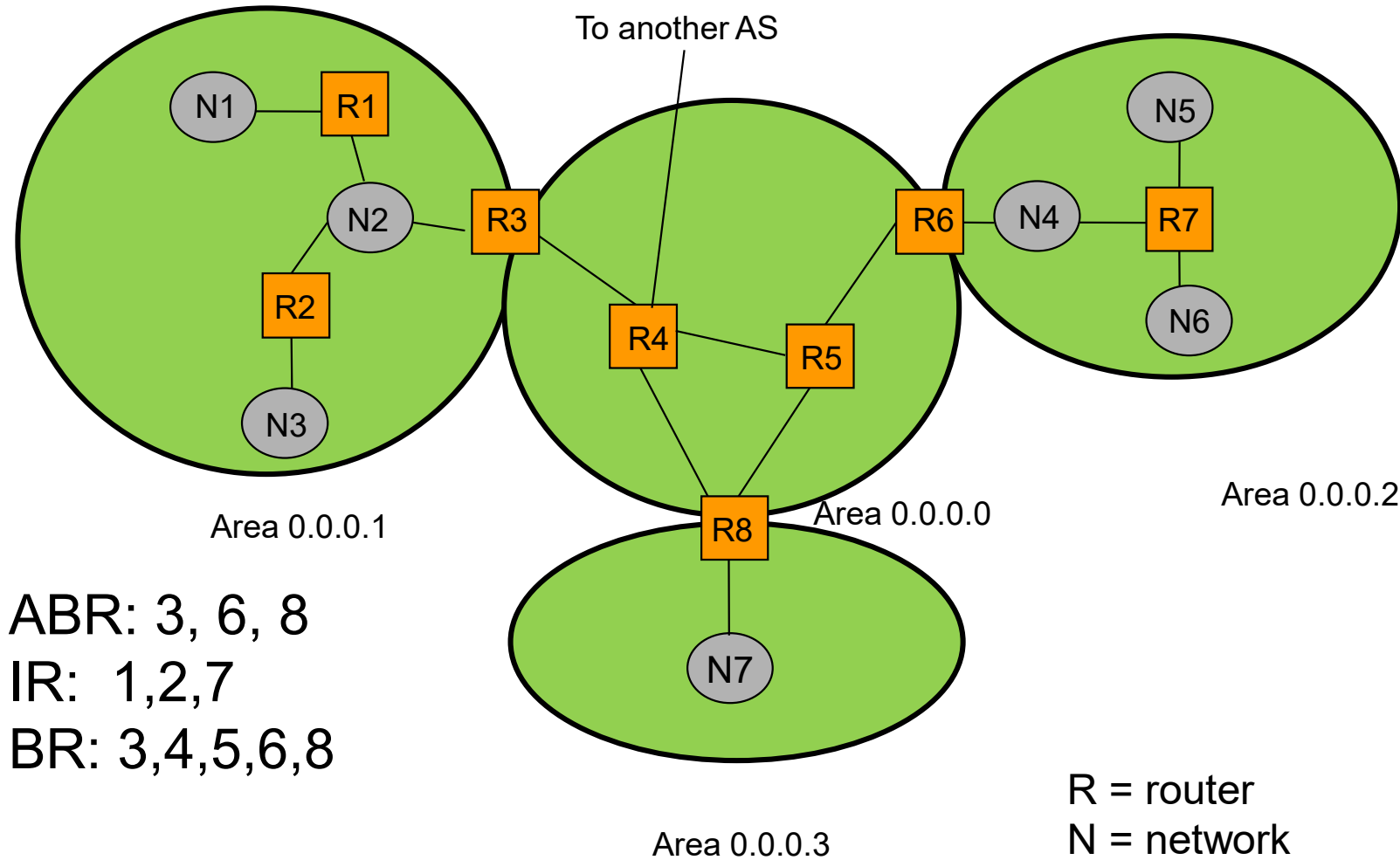
Neighbor, Adjacent & Designated Routers

- *Neighbor routers*: two routers that have interfaces to a common network
  - Neighbors are discovered dynamically by *Hello protocol*
- Each neighbor of a router described by a state
- *Adjacent router*: neighbor routers become adjacent when they synchronize topology databases by exchange of link state information
  - Neighbors on point-to-point links become adjacent
  - Routers on multiaccess nets become adjacent only to designated routers (DR) & backup designated routers (BDR)
    - Reduces size of topological database & routing traffic

Designated Routers

- Reduces number of adjacencies
- Elected by each multiaccess network after neighbor discovery by hello protocol
- Election based on priority & id fields
- Generates link advertisements that list routers attached to a multi-access network
- Forms adjacencies with routers on multi-access network
- Backup prepared to take over if designated router fails
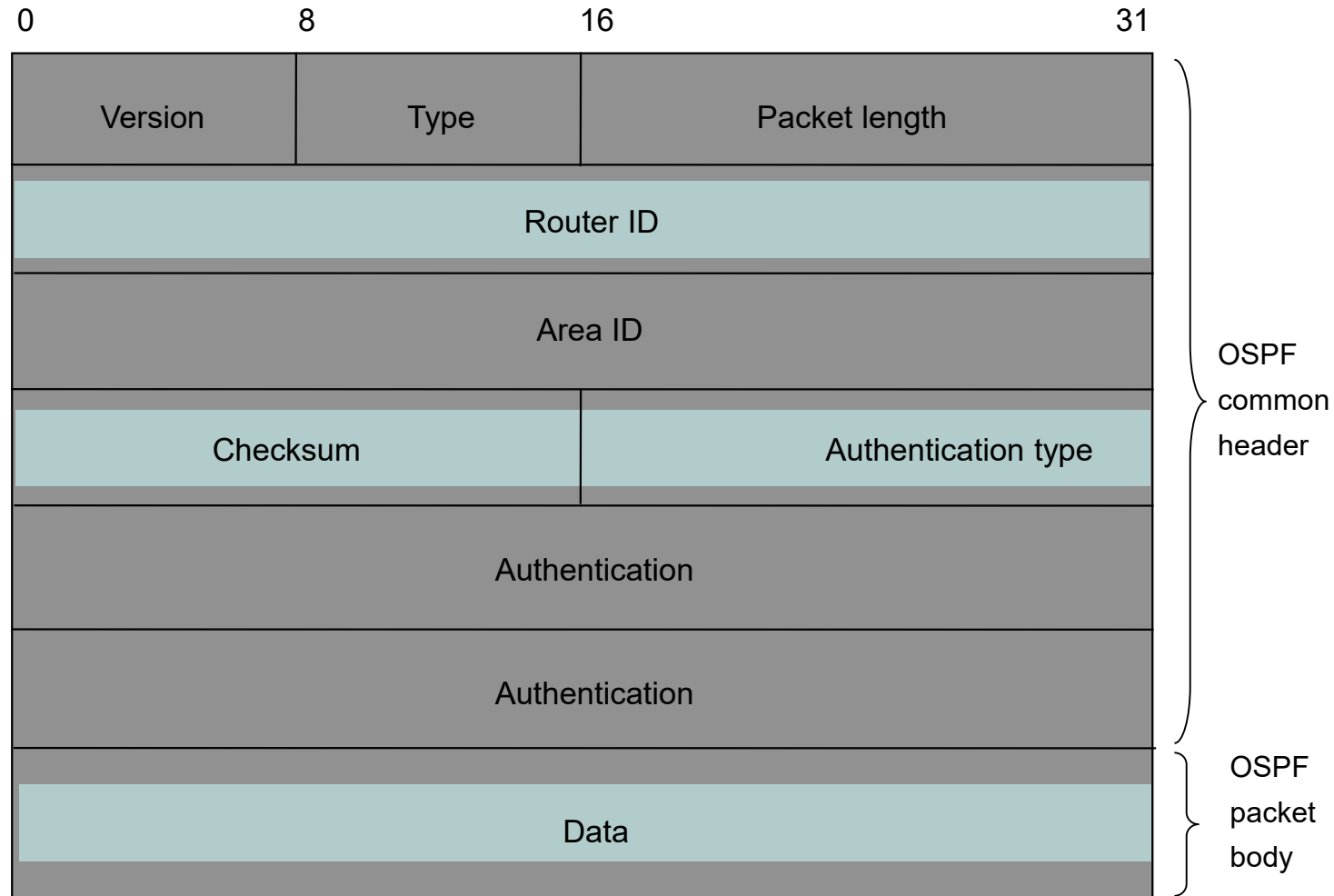
Link State Advertisements

- **Link state info exchanged by adjacent routers to allow**
  - area topology databases to be maintained
  - inter-area & inter-AS routes to be advertised

OSPF Protocol

- **OSPF packets transmitted directly on IP datagrams; Protocol ID 89**
- **OSPF packets sent to multicast address 224.0.0.5 (allOSPFRouters on pt-2-pt and broadcast nets)**
- **OSPF packets sent on specific IP addresses on non-broadcast nets**
- **Five OSPF packet types:**
  - Hello
  - Database description
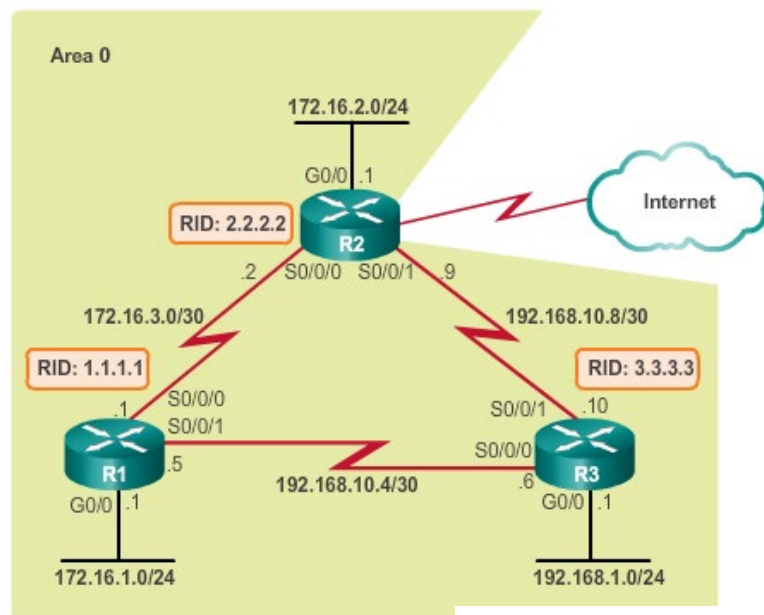 - Link state request; Link state update; Link state ack

OSPF Header

Version	Type	Packet length
Router ID		
Area ID		
Checksum		Authentication type
Authentication		
Authentication		
Data		

Bit positions: 0, 8, 16, 31

OSPF common header — Version, Type, Packet length, Router ID, Area ID, Checksum, Authentication type, Authentication, Authentication

OSPF packet body — Data

- **Type: Hello, Database description, Link state request, Link state update, Link state acknowledgements**

OSPF Stages

1. Discover neighbors by sending Hello packets (every 10 sec) and designated router elected in multiaccess networks

2. Adjacencies are established & link state databases are synchronized

3. Link state information is propagated & routing tables are calculated

Configuring Single-Area OSPF

Area 0

172.16.2.0/24

G0/0 .1

Internet

RID: 2.2.2.2

R2

.2 S0/0/0 S0/0/1 .9

172.16.3.0/30

192.168.10.8/30

RID: 1.1.1.1

RID: 3.3.3.3

S0/0/0

S0/0/1 .10

.1

S0/0/1

S0/0/0

R1

.5

R3

G0/0 .1

192.168.10.4/30

.6 G0/0 .1

172.16.1.0/24

192.168.1.0/24

1000000000/1544000 = 647
Cf. default: 100000000/1544000 = 64

```
R1(config)# interface GigabitEthernet0/0
R1(config-if)# bandwidth 1000000
R1(config-if)# exit
R1(config)# router ospf 10
R1(config-router)# router-id 1.1.1.1
R1(config-router)# auto-cost reference-bandwidth 1000
% OSPF: Reference bandwidth is changed.
        Please ensure reference bandwidth is consistent
across all routers.
R1(config-router)# network 172.16.1.0 0.0.0.255 area 0
R1(config-router)# network 172.16.3.0 0.0.0.3 area 0
R1(config-router)# network 192.168.10.4 0.0.0.3 area 0
R1(config-router)#
R1(config-router)# passive-interface g0/0
R1(config-router)#
```

```
R2(config)# interface GigabitEthernet0/0
R2(config-if)# bandwidth 1000000
R2(config-if)# exit
R2(config)# router ospf 10
R2(config-router)# router-id 2.2.2.2
R2(config-router)# auto-cost reference-bandwidth 1000
% OSPF: Reference bandwidth is changed.
        Please ensure reference bandwidth is consistent
across all routers.
R2(config-router)# network 172.16.2.1 0.0.0.0 area 0
R2(config-router)# network 172.16.3.2 0.0.0.0 area 0
R2(config-router)# network 192.168.10.9 0.0.0.0 area 0
R2(config-router)#
R2(config-router)# passive-interface g0/0
R2(config-router)#
```

```
R1# show ip ospf neighbor

Neighbor ID    Pri    State        Dead Time    Address        Interface
3.3.3.3         0     FULL/  -     00:00:32     192.168.10.6   Serial0/0/1
2.2.2.2         0     FULL/  -     00:00:38     172.16.3.2     Serial0/0/0
R1#
```

```
R1# show ip protocols
*** IP Routing is NSF aware ***

Routing Protocol is "ospf 10"
  Outgoing update filter list for all interfaces is not set
  Incoming update filter list for all interfaces is not set
  Router ID 1.1.1.1
  Number of areas in this router is 1. 1 normal 0 stub 0 nssa
  Maximum path: 4
  Routing for Networks:
    172.16.1.0 0.0.0.255 area 0
    172.16.3.0 0.0.0.3 area 0
    192.168.10.4 0.0.0.3 area 0
  Passive Interface(s):
    GigabitEthernet0/0
  Routing Information Sources:
    Gateway         Distance      Last Update
    3.3.3.3              110      00:12:14
    2.2.2.2              110      00:12:46
  Distance: (default is 110)

R1#v
```

Verifying Single-Area OSPF (cont.)

```
R1# show ip ospf
 Routing Process "ospf 10" with ID 1.1.1.1
 Start time: 00:06:18.952, Time elapsed: 00:39:56.400

<Output omitted>

 Number of areas in this router is 1. 1 normal 0 stub 0 nssa
 Number of areas transit capable is 0
 External flood list length 0
 IETF NSF helper support enabled
 Cisco NSF helper support enabled
 Reference bandwidth unit is 1000 mbps
 Area BACKBONE(0)
         Number of interfaces in this area is 3
Area has no authentication
SPF algorithm last executed 00:15:21.436 ago
SPF algorithm executed 6 times
Area ranges are
Number of LSA 3. Checksum Sum 0x023523
Number of opaque link LSA 0. Checksum Sum 0x000000
Number of DCbitless LSA 0
Number of indication LSA 0
Number of DoNotAge LSA 0
Flood list length 0

R1#
```

Verifying Single-Area OSPF (cont.)

```
R1# show ip ospf interface
GigabitEthernet0/0 is up, line protocol is up
  Internet Address 172.16.1.1/24, Area 0, Attached via Network
Statement
  Process ID 10, Router ID 1.1.1.1, Network Type BROADCAST, Cost: 1
  Topology-MTID    Cost     Disabled    Shutdown       Topology Name
          0          1         no          no             Base
  Transmit Delay is 1 sec, State DR, Priority 1
  Designated Router (ID) 1.1.1.1, Interface address 172.16.1.1
  No backup designated router on this network
  Timer intervals configured, Hello 10, Dead 40, Wait 40,
Retransmit 5
     oob-resync timeout 40
     No Hellos (Passive interface)
  Supports Link-local Signaling (LLS)
  Cisco NSF helper support enabled
  IETF NSF helper support enabled
  Index 1/1, flood queue length 0
  Next 0x0(0)/0x0(0)
  Last flood scan length is 0, maximum is 0
  Last flood scan time is 0 msec, maximum is 0 msec
  Neighbor Count is 0, Adjacent neighbor count is 0
  Suppress hello for 0 neighbor(s)
Serial0/0/1 is up, line protocol is up
  Internet Address 192.168.10.5/30, Area 0, Attached via Network
Statement
  Process ID 10, Router ID 1.1.1.1, Network Type POINT_TO_POINT,
Cost: 647

<Output omitted>
```

```
R1# show ip ospf interface brief
Interface    PID    Area    IP Address/Mask     Cost   State  Nbrs F/C
Gi0/0        10     0       172.16.1.1/24       1      DR     0/0
Se0/0/1      10     0       192.168.10.5/30     647    P2P    1/1
Se0/0/0      10     0       172.16.3.1/30       647    P2P    1/1
R1#
```

Outline

- Basic Routing
- Routing Information Protocol (RIP)
- Open Shortest Path First (OSPF)
- **Border Gateway Protocol (BGP)**

Exterior Gateway Protocols

- Within each AS, there is a consistent set of routes connecting the constituent networks
- The Internet is woven into a coherent whole by *Exterior Gateway Protocols (EGPs)* that operate between AS's
- EGP enables two AS's to exchange routing information about:
  - The networks that are contained within each AS
  - The AS's that can be reached through each AS
- EGP path selection guided by policy rather than path optimality
 - Trust, peering arrangements, etc

EGP Example

Only EGP routers are shown

AS2

N1 reachable through AS3

R1

AS1

R4

N1

AS3
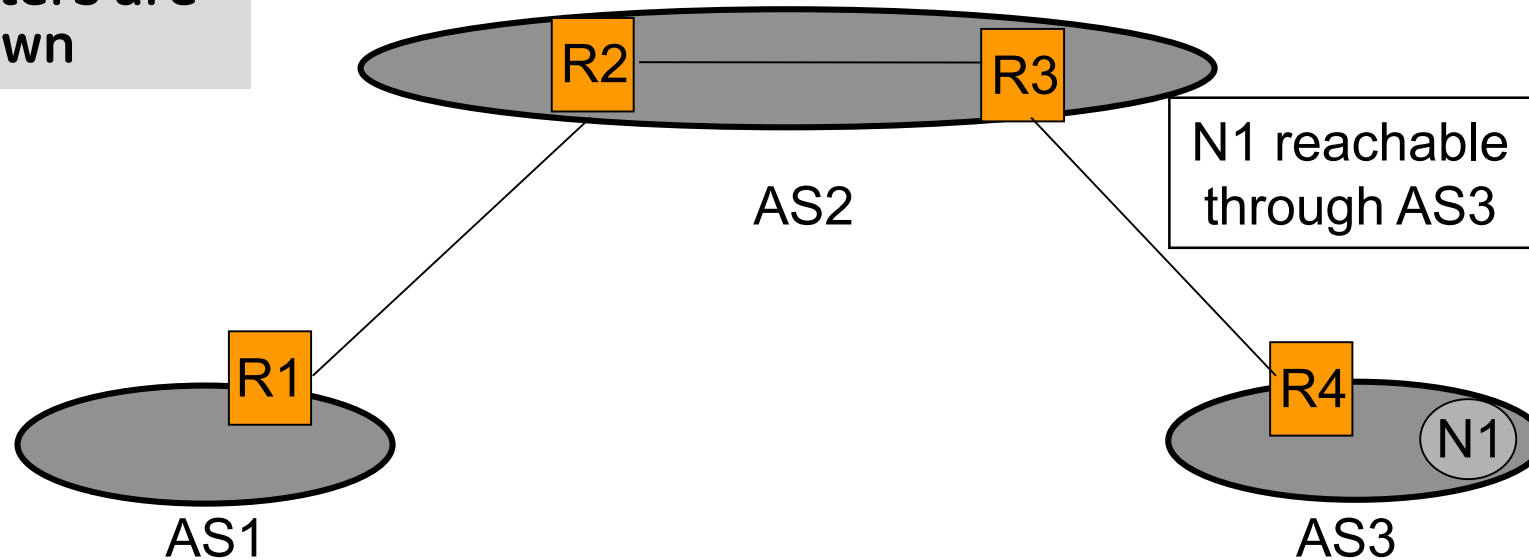
- R4 advertises that network N1 can be reached through AS3
- R3 examines announcement & applies policy to decide whether it will forward packets to N1 through R4
- If yes, routing table updated in R3 to indicate R4 as next hop to N1
- IGP propagates N1 reachability information through AS2

EGP Example
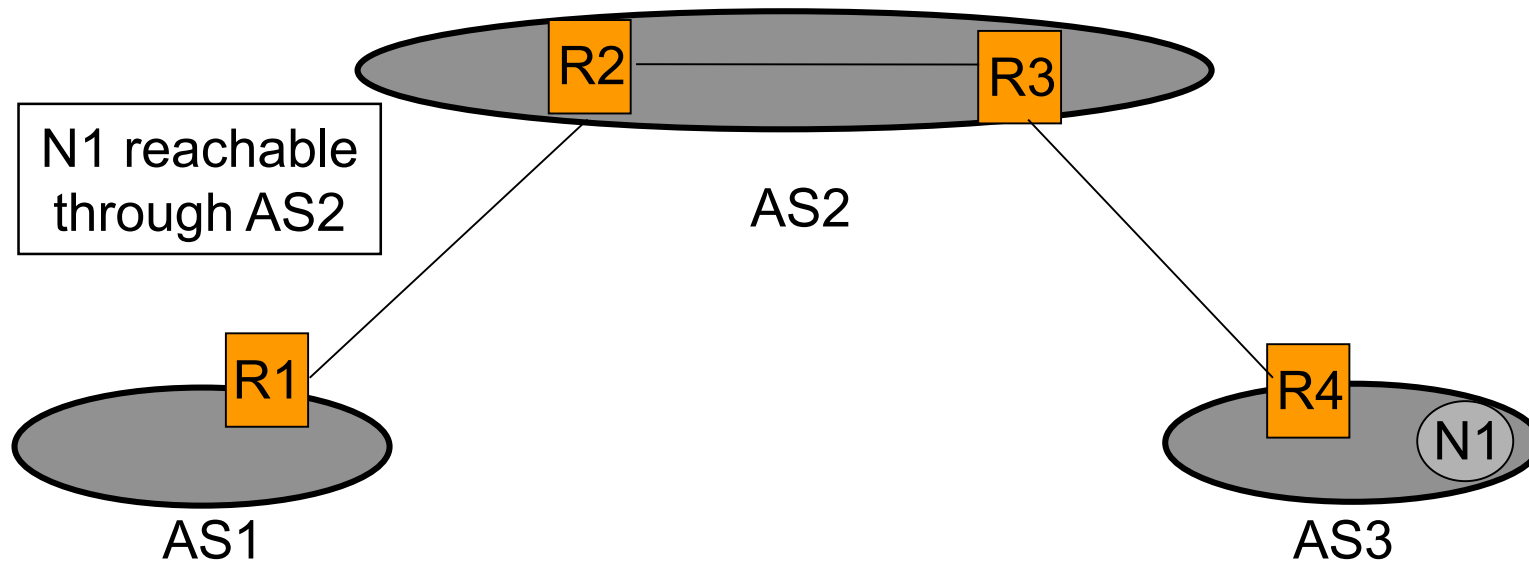
- EGP routers within an AS, e.g. R3 and R2, are kept consistent
- Suppose AS2 willing to handle *transit* packets from AS1 to N1
- R2 advertises to AS1 the reachability of N1 through AS2
- R1 applies its policy to decide whether to send to N1 via AS2

Peering and Inter-AS connectivity
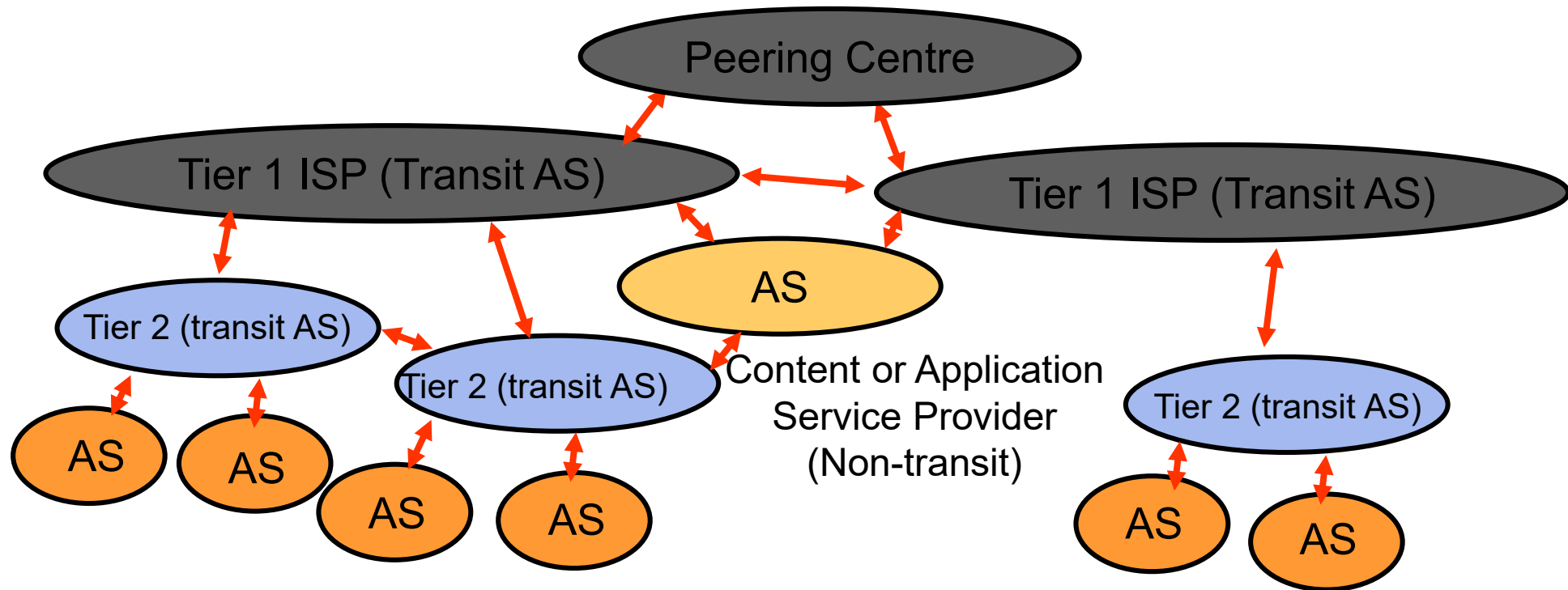
- **Non-transit AS's (stub & multihomed) do not carry transit traffic**
- **Tier 1 ISPs peer with each other, privately & peering centers**
- **Tier 2 ISPs peer with each other & obtain transit services from Tier 1s; Tier 1's carry transit traffic between their Tier 2 customers**
- **Client AS's obtain service from Tier 2 ISPs**

EGP Requirements

- **Scalability to global Internet**
  - Provide connectivity at global scale
  - Link-state does not scale
  - Should promote address aggregation
  - Fully distributed
- **EGP path selection guided by policy rather than path optimality**
  - Trust, peering arrangements, etc
  - EGP should allow flexibility in choice of paths

Border Gateway Protocol v4

AS2

AS1

AS6

AS3

AS5

AS4

AS7
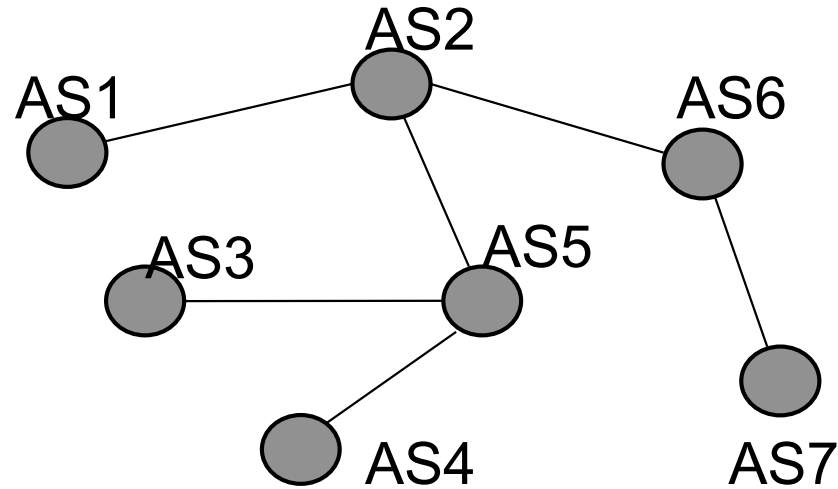
- BGP (RFC 1771) an EGP routing protocol to exchange network reachability information among BGP routers (also called BGP speakers)

- Network reachability info contains sequence of ASs that packets traverse to reach a destination network

- Info exchanged between BGP speakers allows a router to construct a graph of AS connectivity
  - Routing loops can be pruned
 - Routing policy at AS level can be applied

BGP Features

- BGP is *path vector protocol*: advertises sequence of AS numbers to the destination network
- Path vector info used to prevent routing loops
- BGP enforces policy through selection of different paths to a destination and by control of redistribution of routing information
- Uses CIDR to support aggregation & reduction of routing information

BGP Speaker & AS Relationship

- *BGP speaker*: a router running BGP
- *Peers or neighbors*: two speakers exchanging information on a connection
- BGP peers use TCP (port 179) to exchange messages
- Initially, BGP peers exchange entire BGP routing table
  - Incremental updates sent subsequently
  - Reduces bandwidth usage and processing overhead
  - Keepalive messages sent periodically (30 seconds)
- *Internal BGP* (iBPG) between BGP routers in same AS
- *External BGP* (eBGP) connections across AS borders

iBGP & eBGP
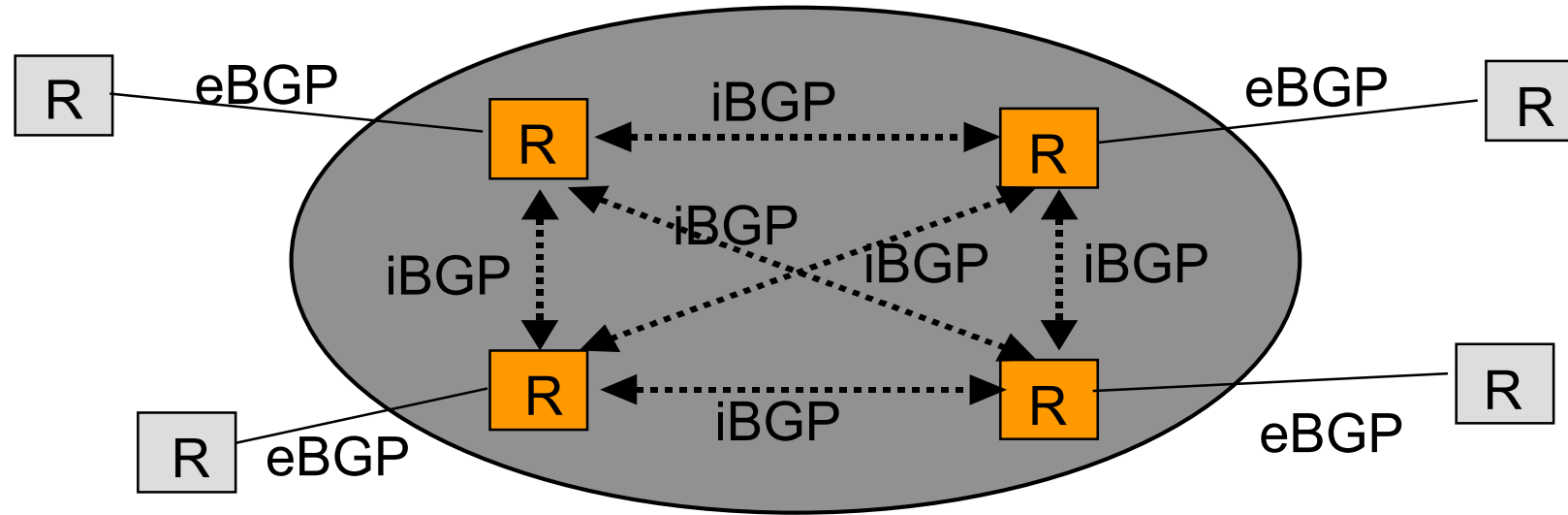
- eBGP to exchange reachability information in different AS's
  - **eBGP peers directly connected**
- iBGP to ensure net reachability info is consistent among the BGP speakers in the same AS
  - usually not directly connected
  - iBGP speakers exchange info learned from other iBGP speakers, and thus fully meshed

Path Selection

- **Each BGP speaker**
  - Evaluates paths to a destination from an AS border router
  - Selects the best that complies with policies
  - Advertises that route to all BGP neighbors
- **BGP assigns a preference order to each path & selects path with highest value; BGP does not keep a cost metric to any path**
- **When multiple paths to a destination exist, BGP maintains all of the paths, but only advertises the one with highest preference value**

BGP Policy

- **Examples of policy:**
  - Never use AS X
  - Never use AS X to get to a destination in AS Y
  - Never use AS X and AS Y in the same path
- *Import policies* to accept, deny, or set preferences on route advertisements from neighbors
- *Export policies* to determine which routes should be advertised to which neighbors
  - A route is advertised only if AS is willing to carry traffic on that route

(Ex) Typical BGP Policies

- Typical policies involve political, security, or economic considerations.
  - No transit traffic through certain ASes.
  - Never put country X on a route starting at the Pentagon.
  - Do not use the United States to get from British Columbia to Ontario.
  - Only transit country A if there is not alternative to the destination.
  - Traffic starting or ending at IBM should not transit Microsoft.

BGP Protocol

- Opening & confirming of a BGP connection with a neighbor router
- Maintaining the BGP connection
- Sending reachability information
- Notification of error conditions