Lecture 6
Communication Networks and Services

Internet Routing Protocols
DHCP, NAT, and Mobile IP

Lecture 6
Communication Networks and Services

Internet Routing Protocols

Outline

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

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

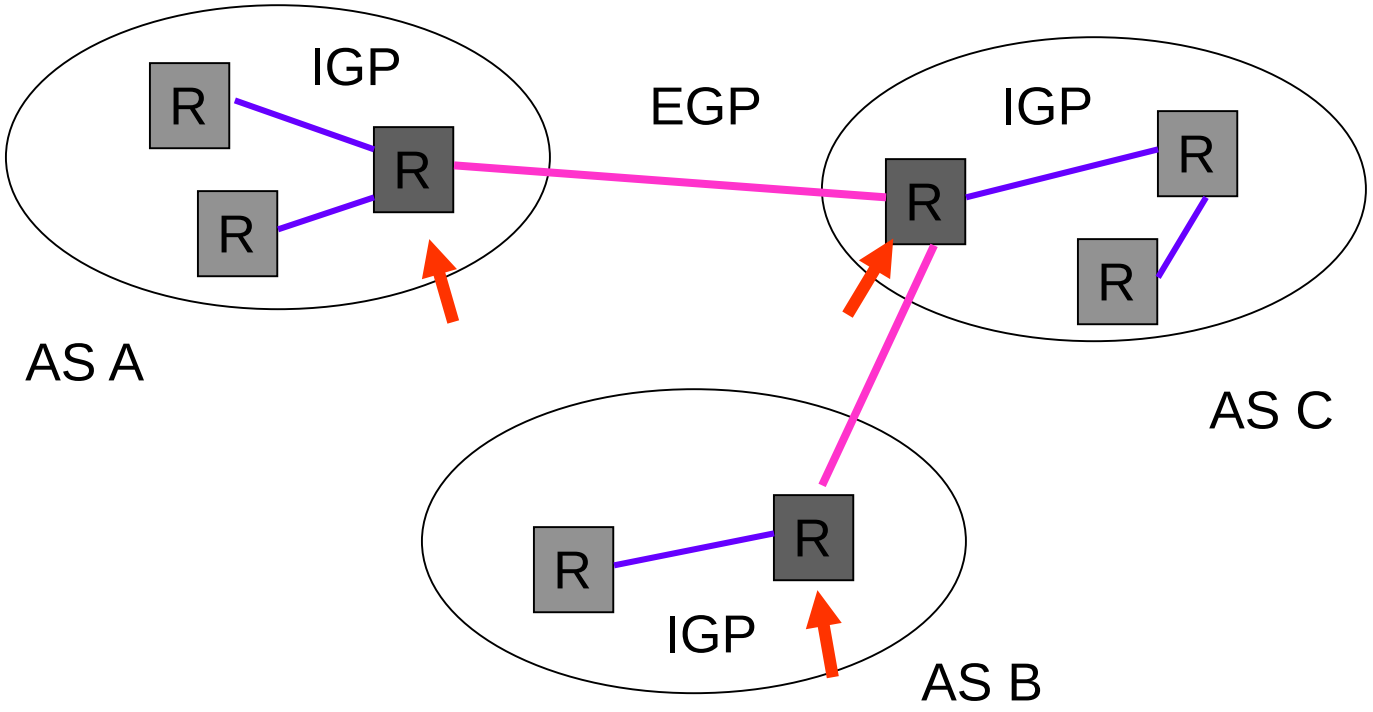
Interior Gateway Protocol (IGP): routing within AS

- 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

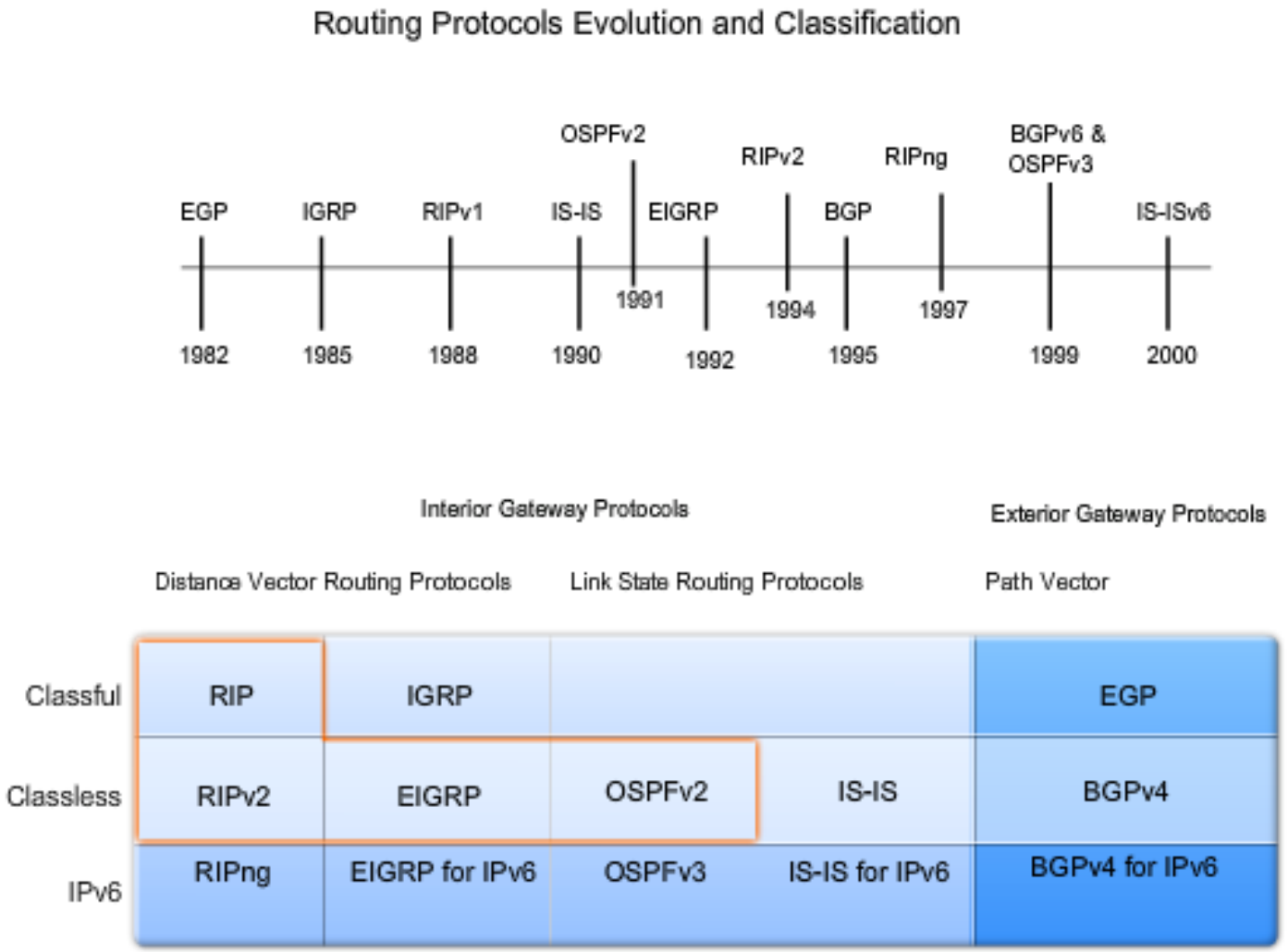
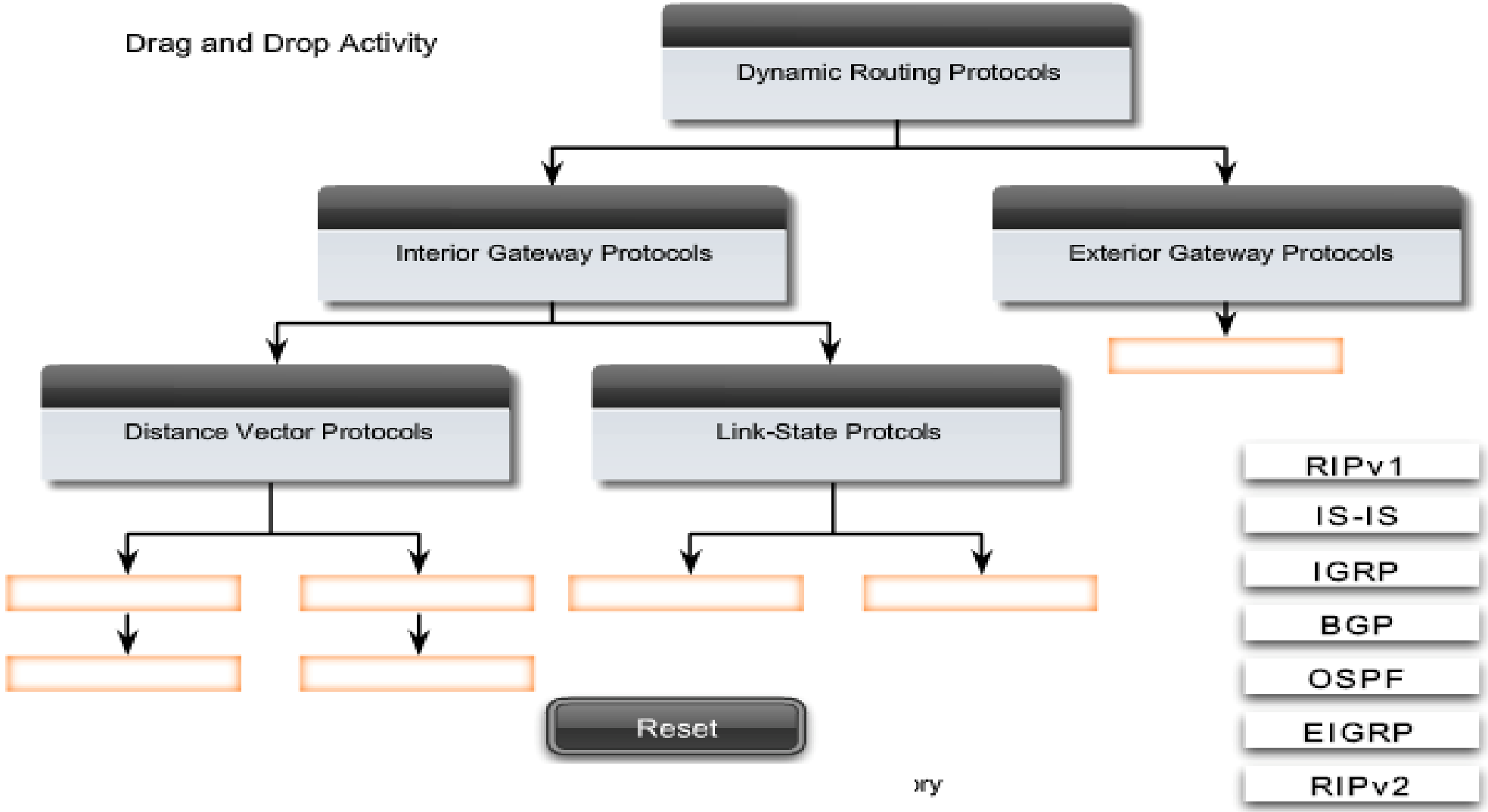


Image Source: www.netacad.com

Inter and Intra Domain 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

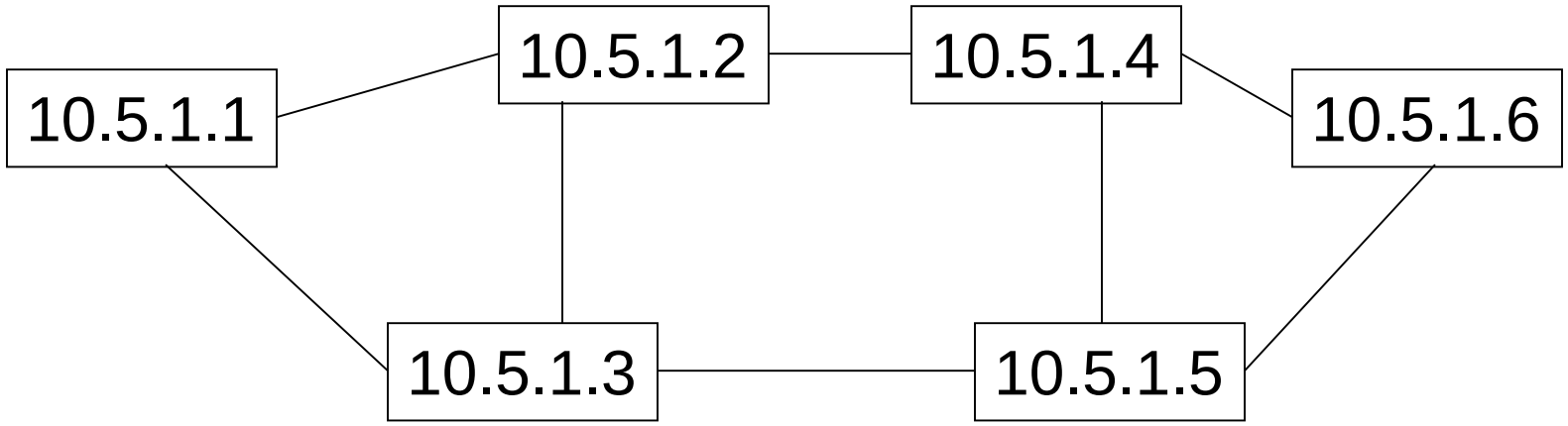
OSPF Features

- *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

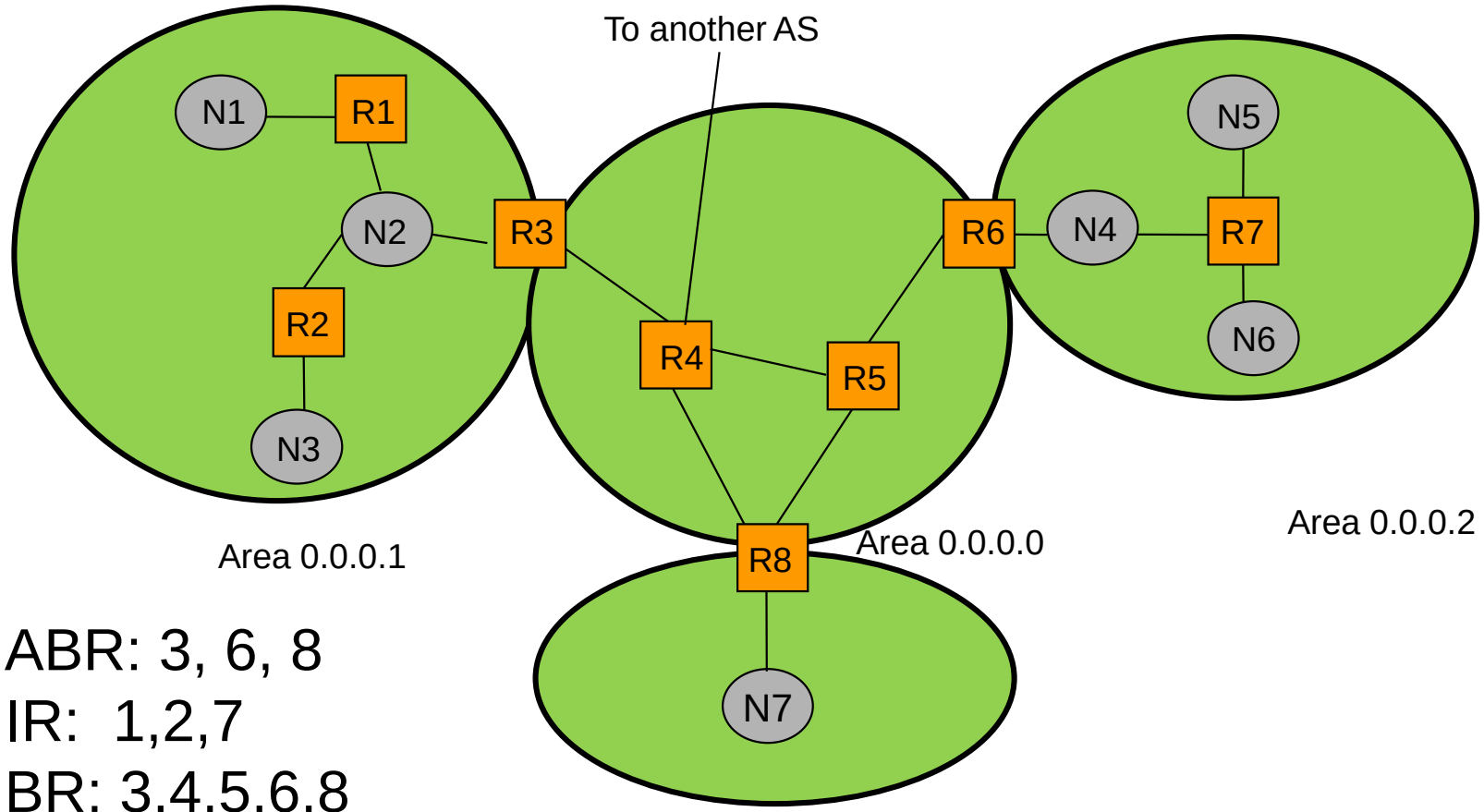


- At steady state:
- All routers have same LS database
- Know how many routers in network
- Interfaces & links between routers
- Cost of each link
- Occasional Hello messages (10 sec) & LS updates sent (30 min)

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



ABR: 3, 6, 8
IR: 1,2,7
BR: 3,4,5,6,8

R = router
N = network

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 router (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

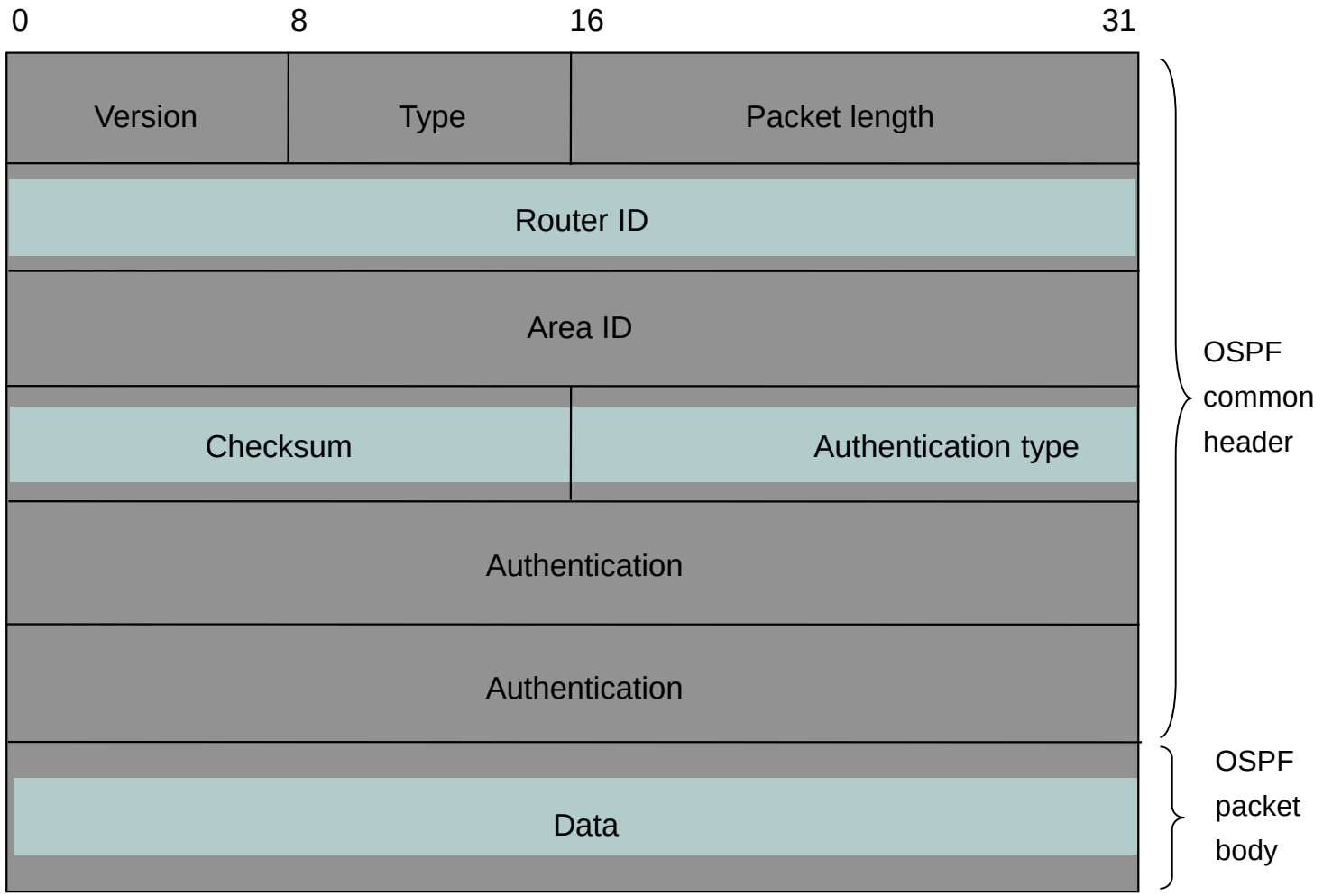
Link State Advertisements

- **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



- **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

Outline

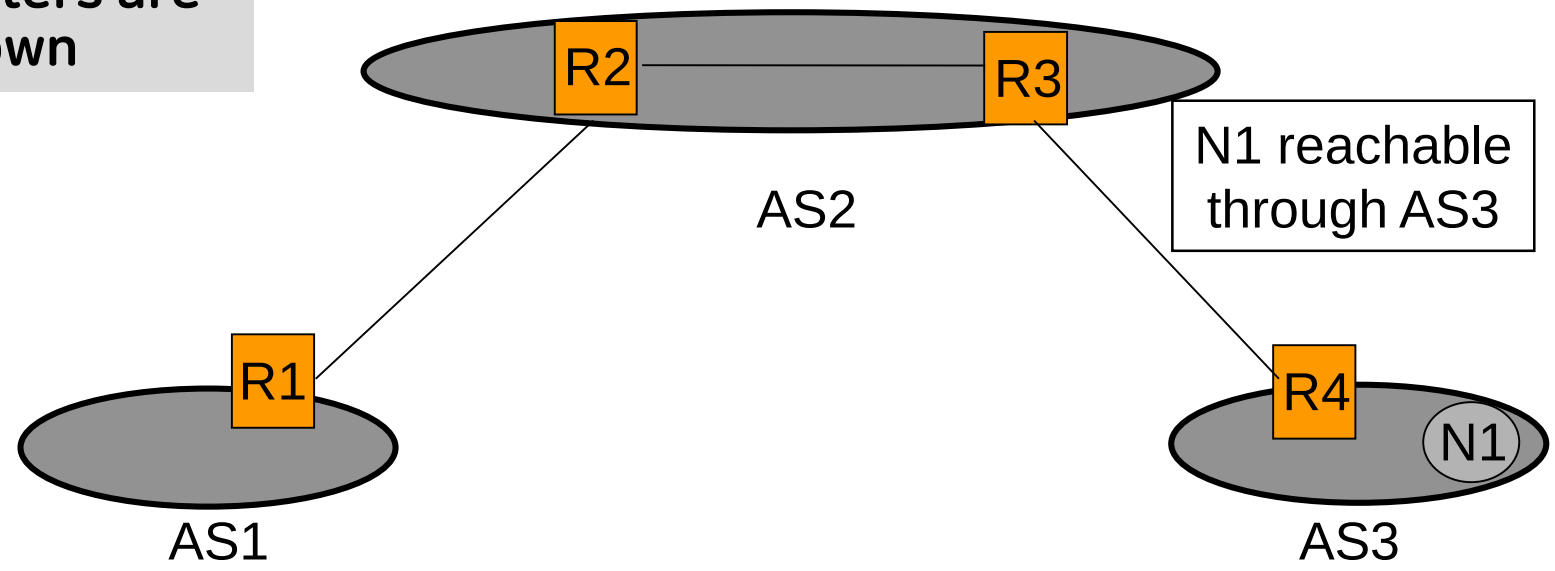
- Basic Routing
- Routing Information Protocol (RIP)
- Open Shortest Path First (OSPF)
- **Border Gateway Protocol (BGP)**

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

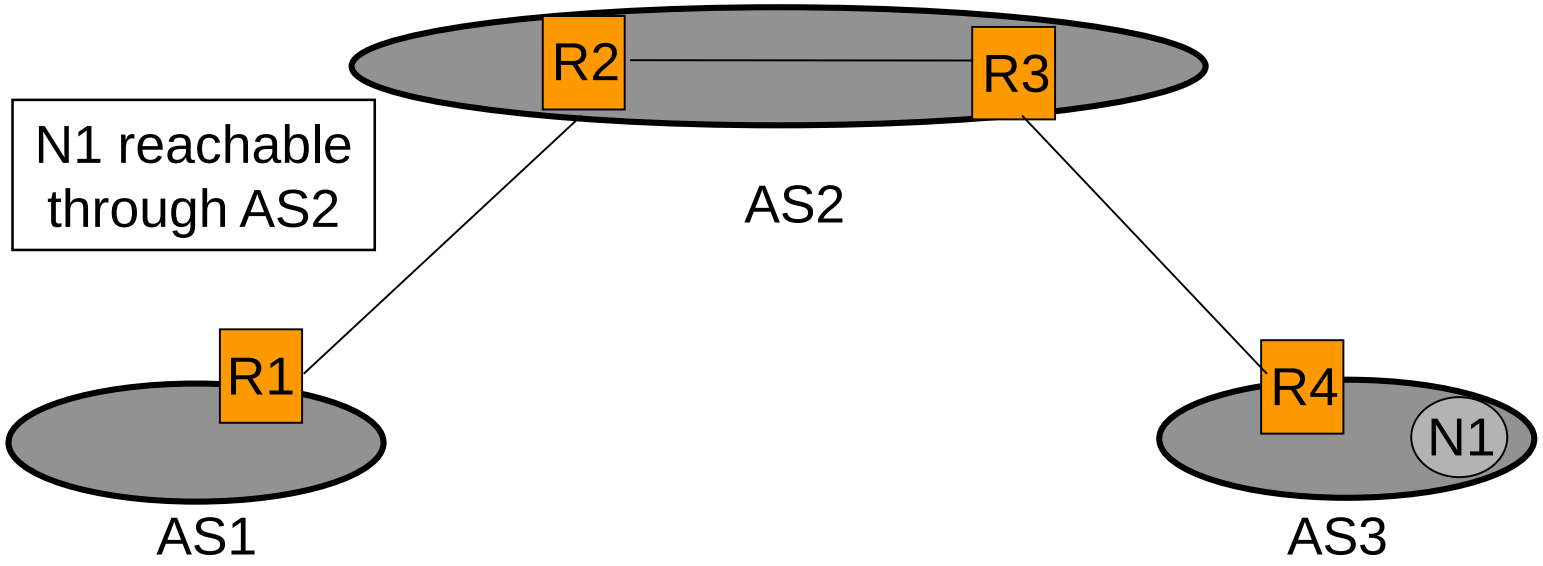
EGP Example

Only EGP routers are shown



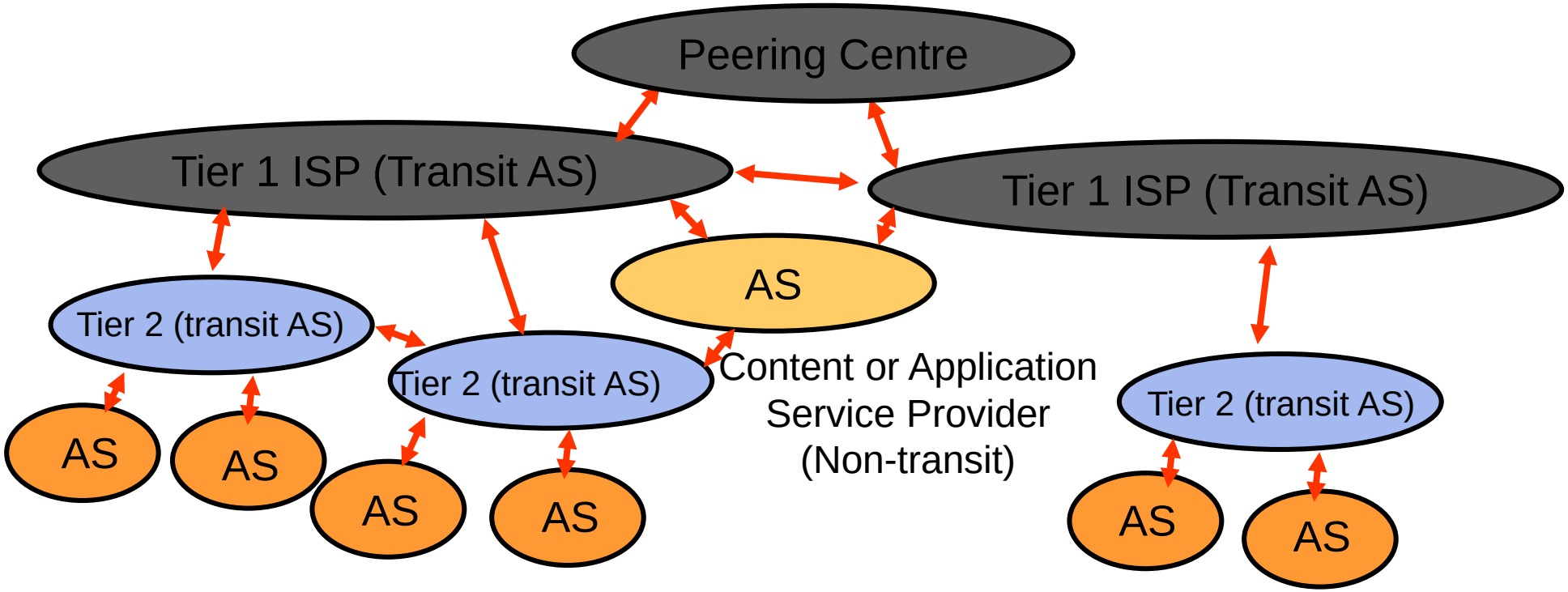
- R4 advertises that network N1 can be reached through AS3
- R3 examines announcement & applies policy to decide whether it will forward packets to N1 through R4
- If yes, routing table updated in R3 to indicate R4 as next hop to N1
- IGP propagates N1 reachability information through AS2

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

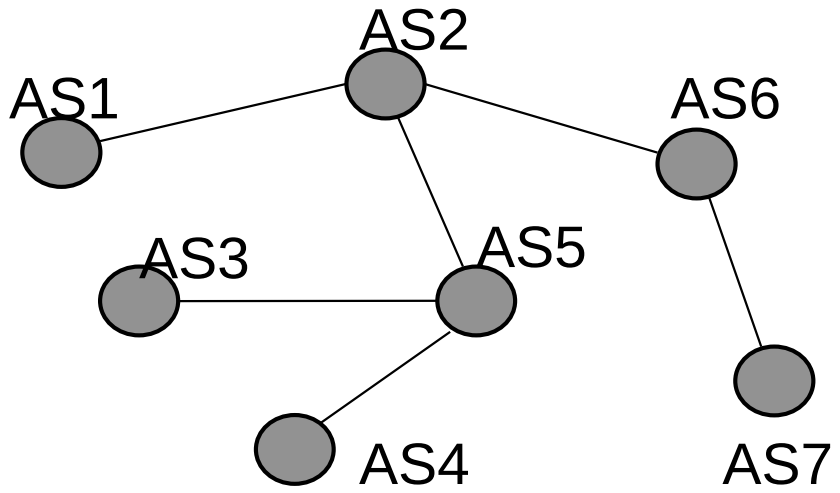


- Non-transit AS's (stub & multihomed) do not carry transit traffic
- Tier 1 ISPs peer with each other, privately & peering centers
- Tier 2 ISPs peer with each other & obtain transit services from Tier 1s; Tier 1's carry transit traffic between their Tier 2 customers
- Client AS's obtain service from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



- 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

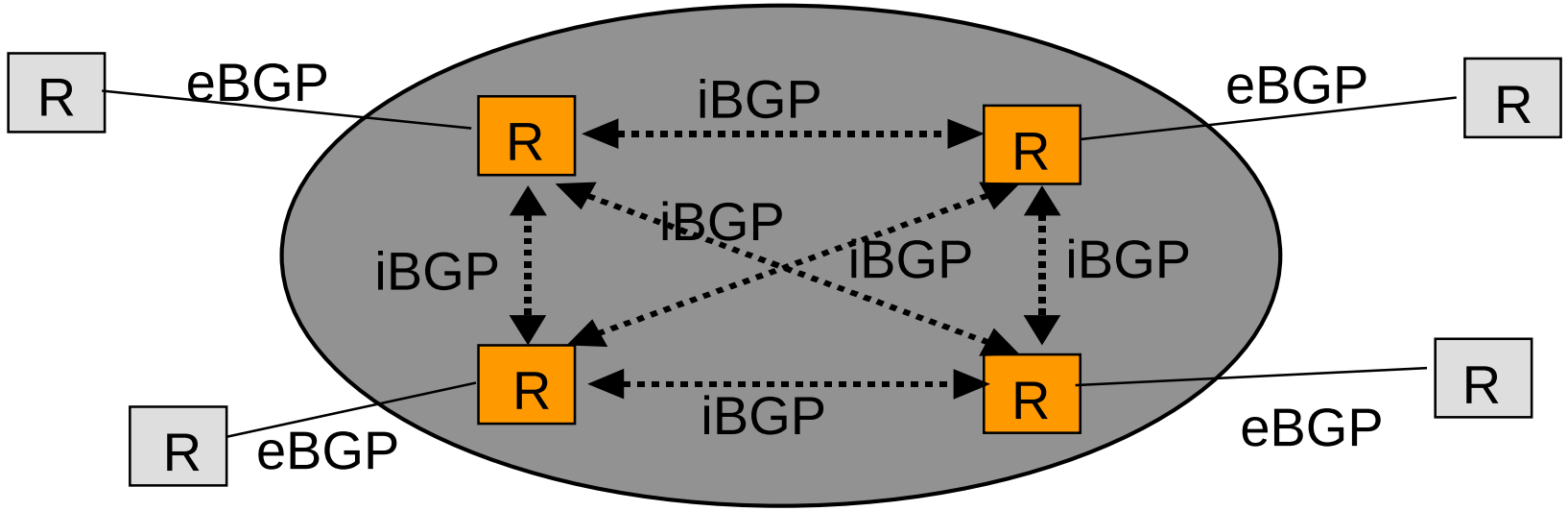
BGP Features

- 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 Iraq on a route starting at the Pentagon.
  - Do not use the United States to get from British Columbia to Ontario.
  - Only transit Albania 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

Lecture 6
Communication Networks and Services

Internet Routing Protocols
DHCP, NAT, and Mobile IP

Lecture 6
Communication Networks and Services

DHCP, NAT, and Mobile IP

DHCP

- Dynamic Host Configuration Protocol (RFC 2131)
- BOOTP (RFC 951, 1542) allows a diskless workstation to be remotely booted up in a network
  - UDP port 67 (server) & port 68 (client)
- DHCP builds on BOOTP to allow servers to deliver configuration information to a host
  - Used extensively to assign temporary IP addresses to hosts
  - Allows ISP to maximize usage of their limited IP addresses

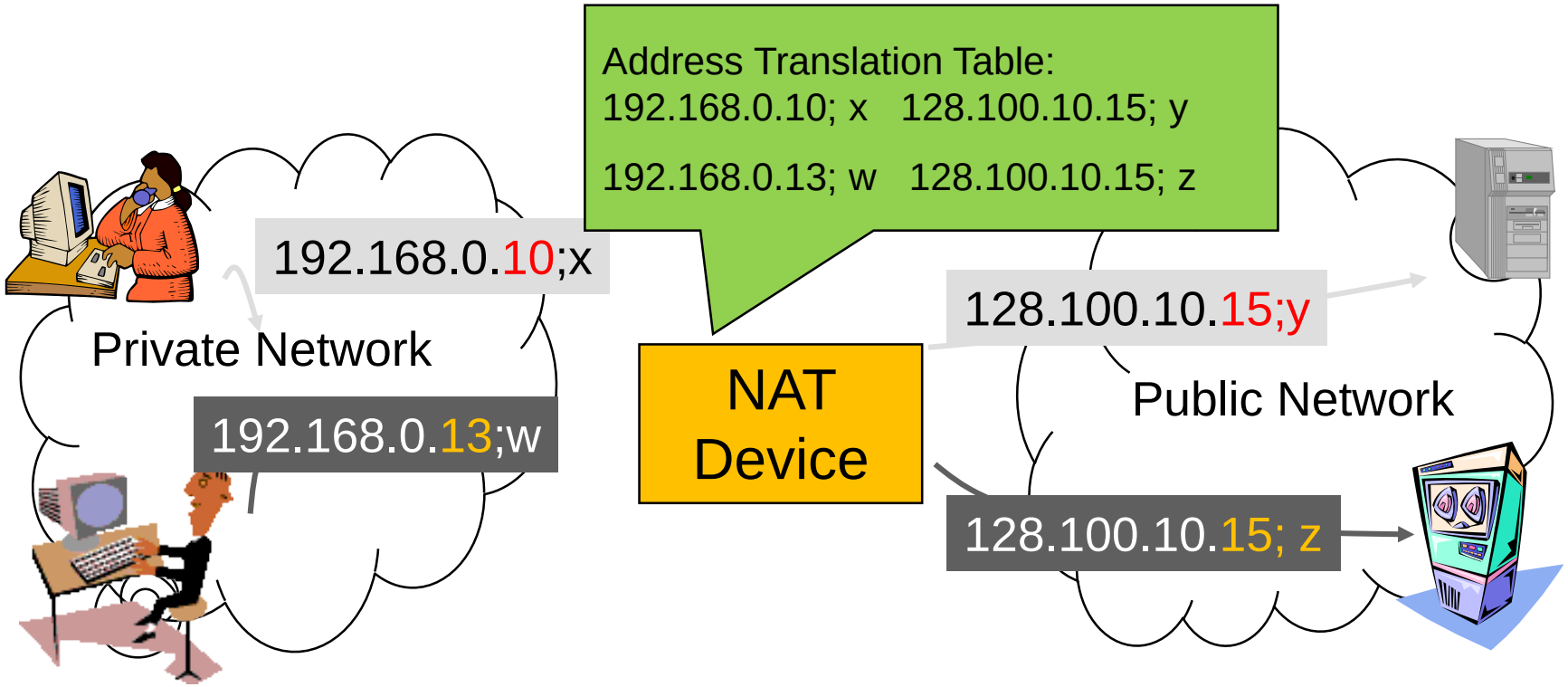
DHCP Operation

- Host broadcasts DHCP *Discover* message on its physical network
- Server replies with *Offer* message (IP address + configuration information)
- Host selects one offer and broadcasts *DHCP Request* message
- Server allocates IP address for lease time T
  - Sends DHCP ACK message with T, and threshold times T1 (=1/2 T) and T2 (=.875T)
- At T1, host attempts to renew lease by sending DHCP Request message to original server
- If no reply by T2, host broadcasts DHCP Request to *any* server
- If no reply by T, host must relinquish IP address and start from the beginning

Network Address Translation (NAT)

- Class A, B, and C addresses have been set aside for use within private internets
  - Packets with private ("unregistered") addresses are discarded by routers in the global Internet
- NAT (RFC 1631): method for mapping packets from hosts in private internets into packets that can traverse the Internet
  - A device (computer, router, firewall) acts as an agent between a private network and a public network
  - A number of hosts can share a limited number of registered IP addresses
    - Static/Dynamic NAT: map unregistered addresses to registered addresses
    - Overloading: maps multiple unregistered addresses into a single registered address (e.g. Home LAN)

NAT Operation (Overloading)



- **Hosts inside private networks generate packets with private IP address & TCP/UDP port #s**
- **NAT maps each private IP address & port # into shared global IP address & available port #**
- **Translation table allows packets to be routed unambiguously**

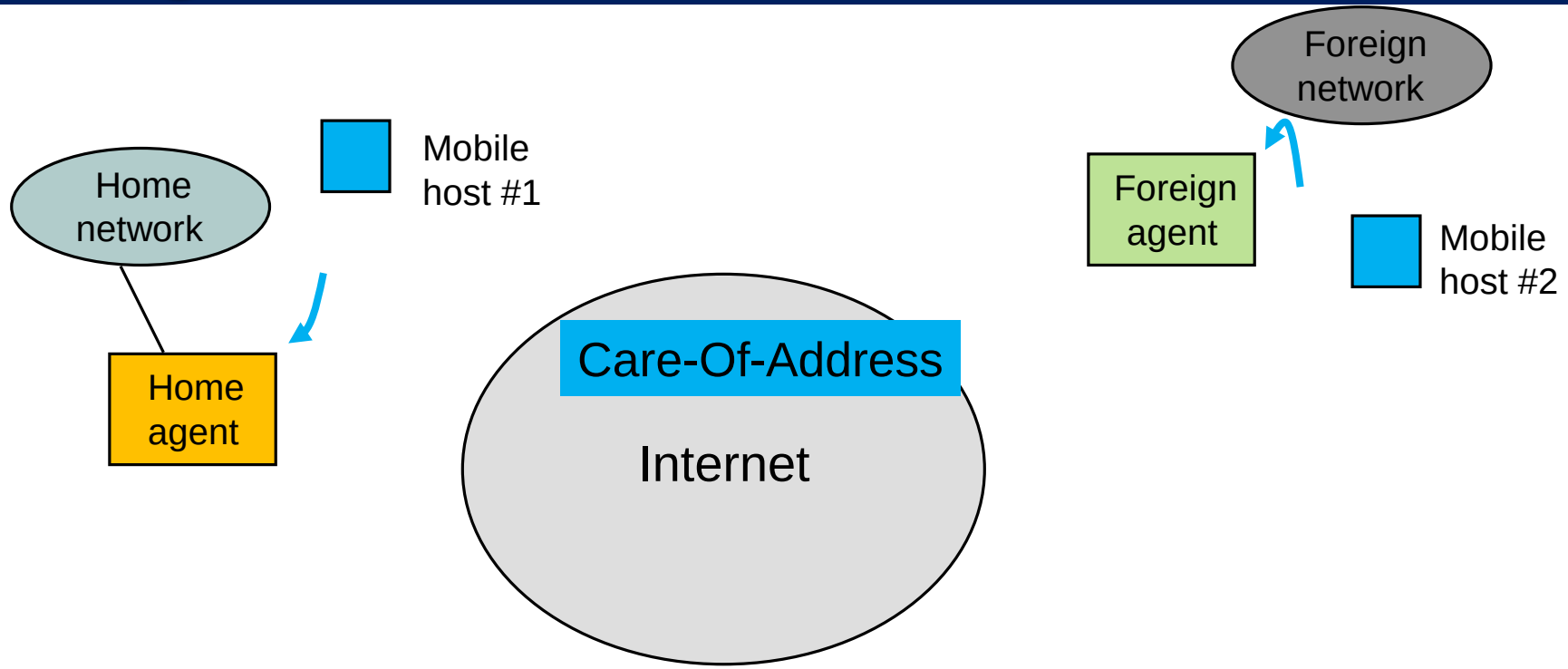
Review: Routable and Nonroutable Addresses

- Nonroutable Address [RFC 1918]
  - Internet Router ignore the following addresses.
    - 10.0.0.0 – 10.255.255.255
    - 172.16.0.0 – 172.31.255.255
    - 192.168.0.0 – 192.168.255.255
  - Millions of networks can exist with the same nonroutable address.
  - “Intranet” : Internal Internet
  - NAT (Network Address Translation) router
  - Side benefit : “Security”

Mobile IP

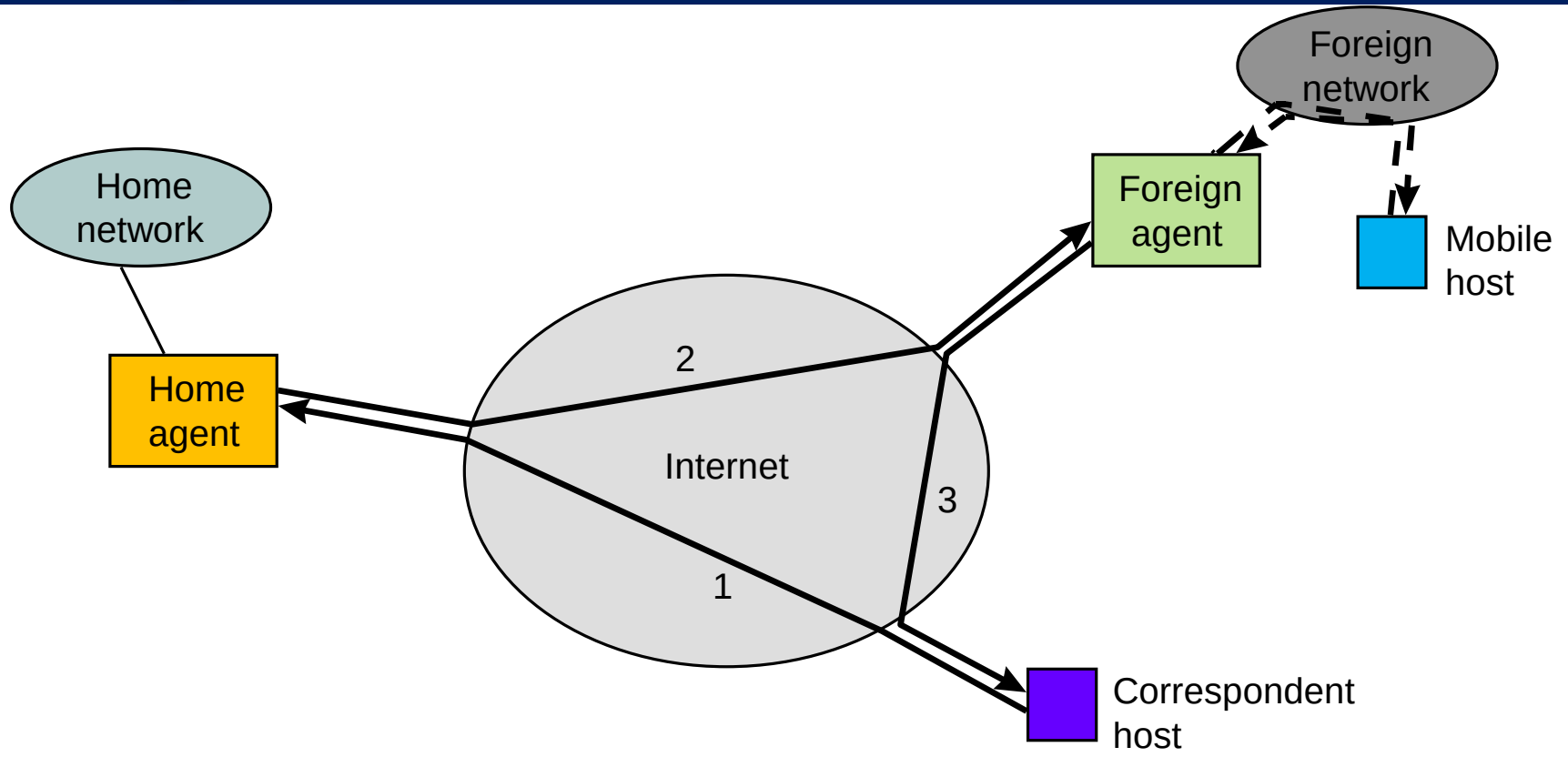
- Proliferation of mobile devices: PDAs, laptops, cellphones, …
- As user moves, point-of-attachment to network necessarily changes
- Problem: IP address specifies point-of-attachment to Internet
  - Changing IP address involves terminating all connections & sessions
- *Mobile IP (RFC 2002)*: device can change point-of-attachment while retaining IP address and maintaining communications

Routing in Mobile IP



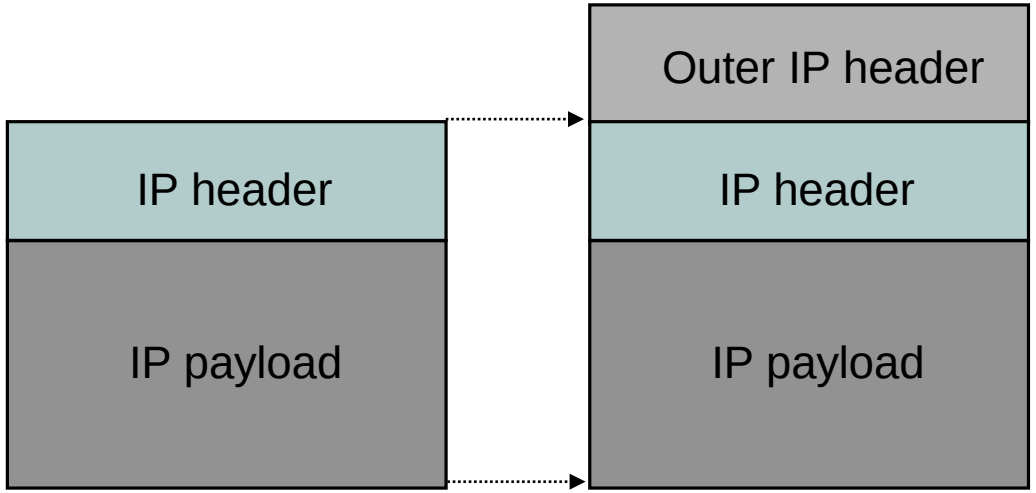
- Home Agent (HA) keeps track of location of each Mobile Host (MH) in its network; HA periodically announces its presence
- If an MH is in home network, e.g. MH#1, HA forwards packets directly to MH
- When an MH moves to a Foreign network, e.g. MH#2, MH obtains a care-of-address from foreign agent (FA) and registers this new address with its HA

Routing in Mobile IP



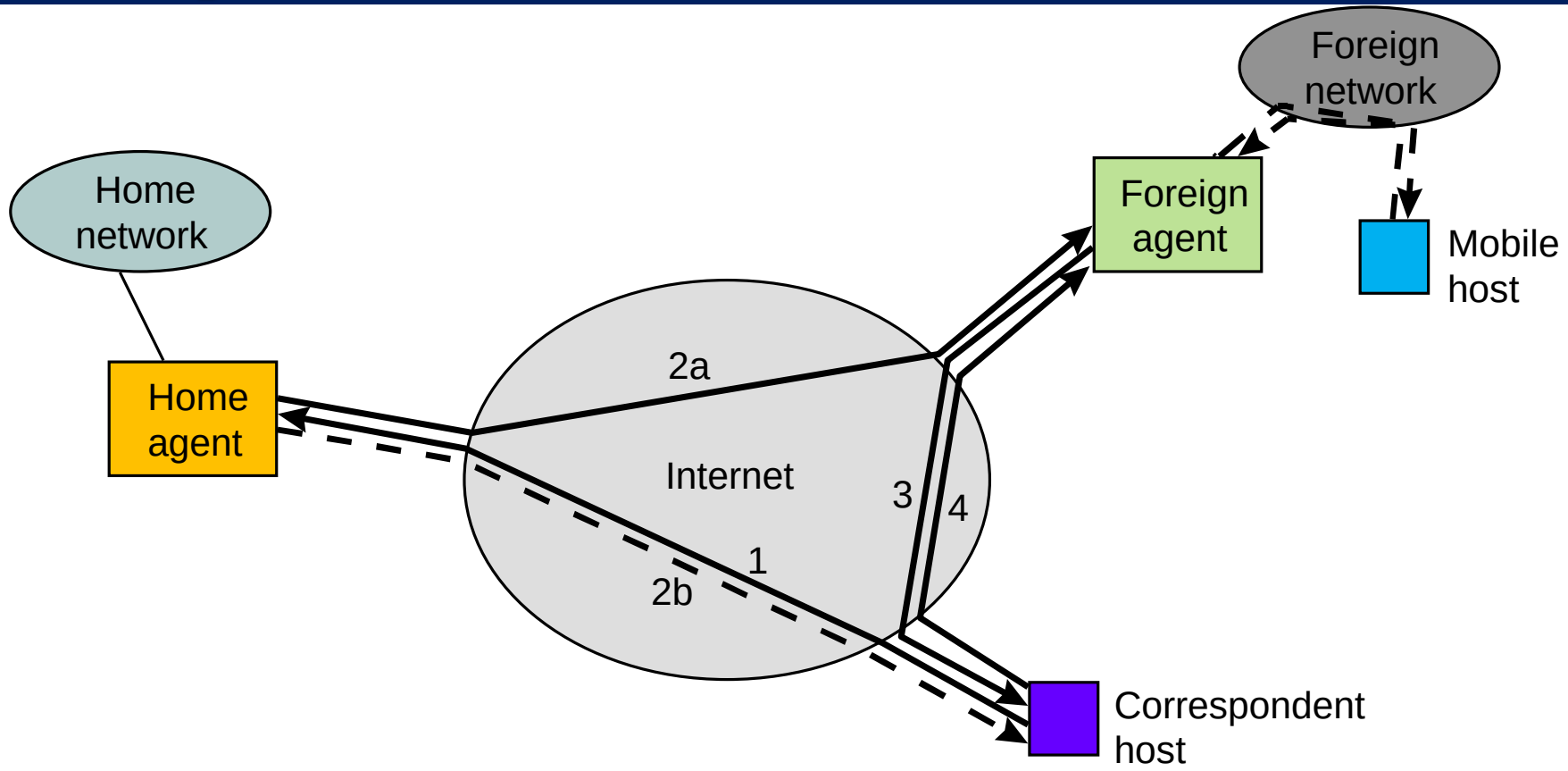
- Correspondent Host (CH) sends packets as usual (1)
- Packets are intercepted by HA which then forwards to Foreign Agent (FA) (2)
- FA forwards packets to the MH
- MH sends packet to CH as usual (3)
- How does HA send packets to MH in foreign network?

IP-to-IP Encapsulation



- HA uses IP-to-IP encapsulation
- IP packet has MH IP address
- Outer IP header has HA's address as source address and care-of-address as destination address
- FA recovers IP packet and delivers to MH

Route Optimization



- Going to HA inefficient if CH and MH are in same foreign network
- When HA receives pkt from CH (1), it tunnels using care-of-address (2a); HA also sends care-of-address to CH (2b)
- CH can then send packets directly to care-of-address (4)