BLM6196 COMPUTER NETWORKS AND COMMUNICATION PROTOCOLS

Prof. Dr. Hasan Hüseyin BALIK

(8-10th Week)

8. Internetwork Operation

8-10.Outline

- Multicasting
- Software Defined Networks
- OpenFlow
- Mobile IP

Multicasting

- The act of sending a packet from a source to the members of a multicast group
- Multicast addresses
  - Addresses that refer to a group of hosts on one or more networks

- Has a number of practical applications

Multimedia “broadcast”

Teleconferencing

Database

Distributed computing

Real time workgroups

LAN Multicast

- LAN multicast is easy
  - Send to IEEE 802 multicast MAC address
  - Those in multicast group will accept it
  - Only single copy of packet is needed
- A transmission from any one station is received by all other stations on LAN

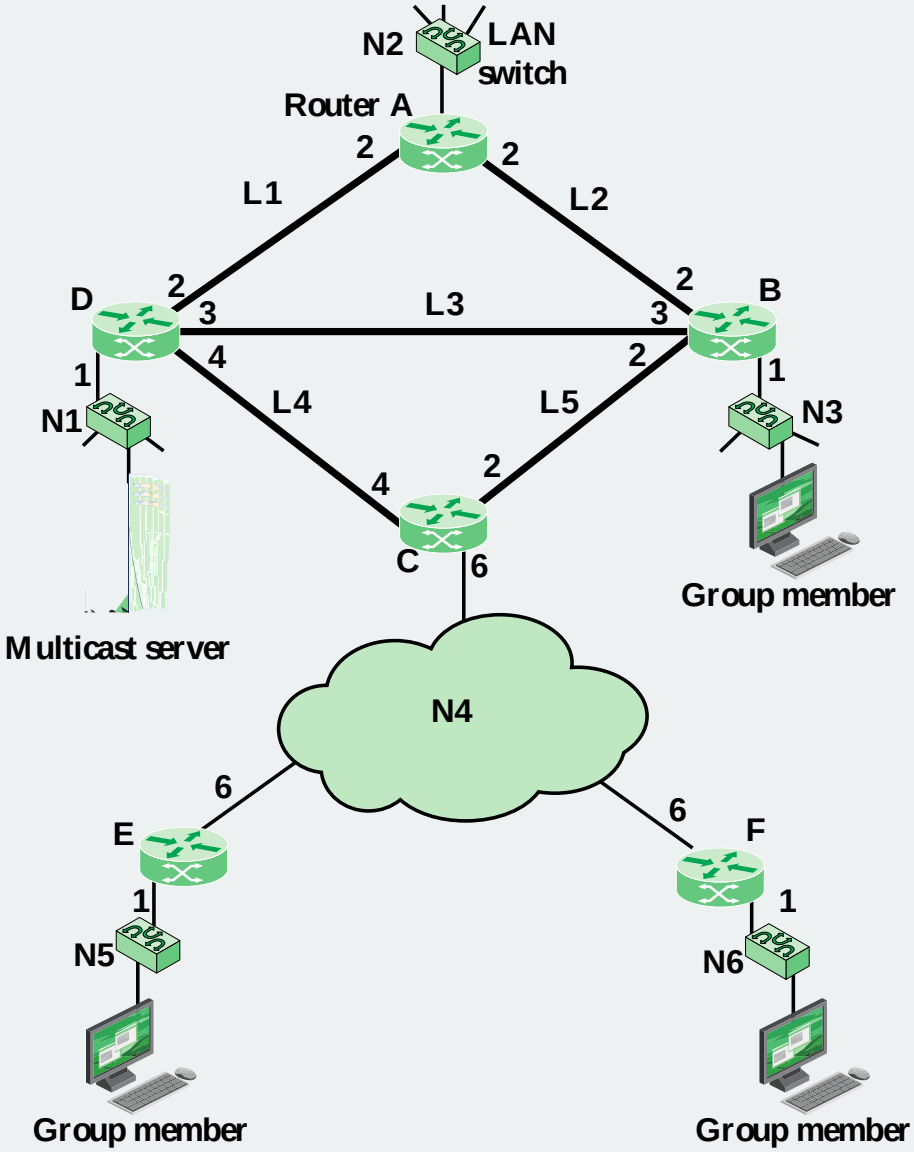


Figure 21.1 Example Configuration

Multicasting Strategies

Broadcast packet to each network

- If server does not know members of group
- Requires 13 packets

Could send multiple unicast packets

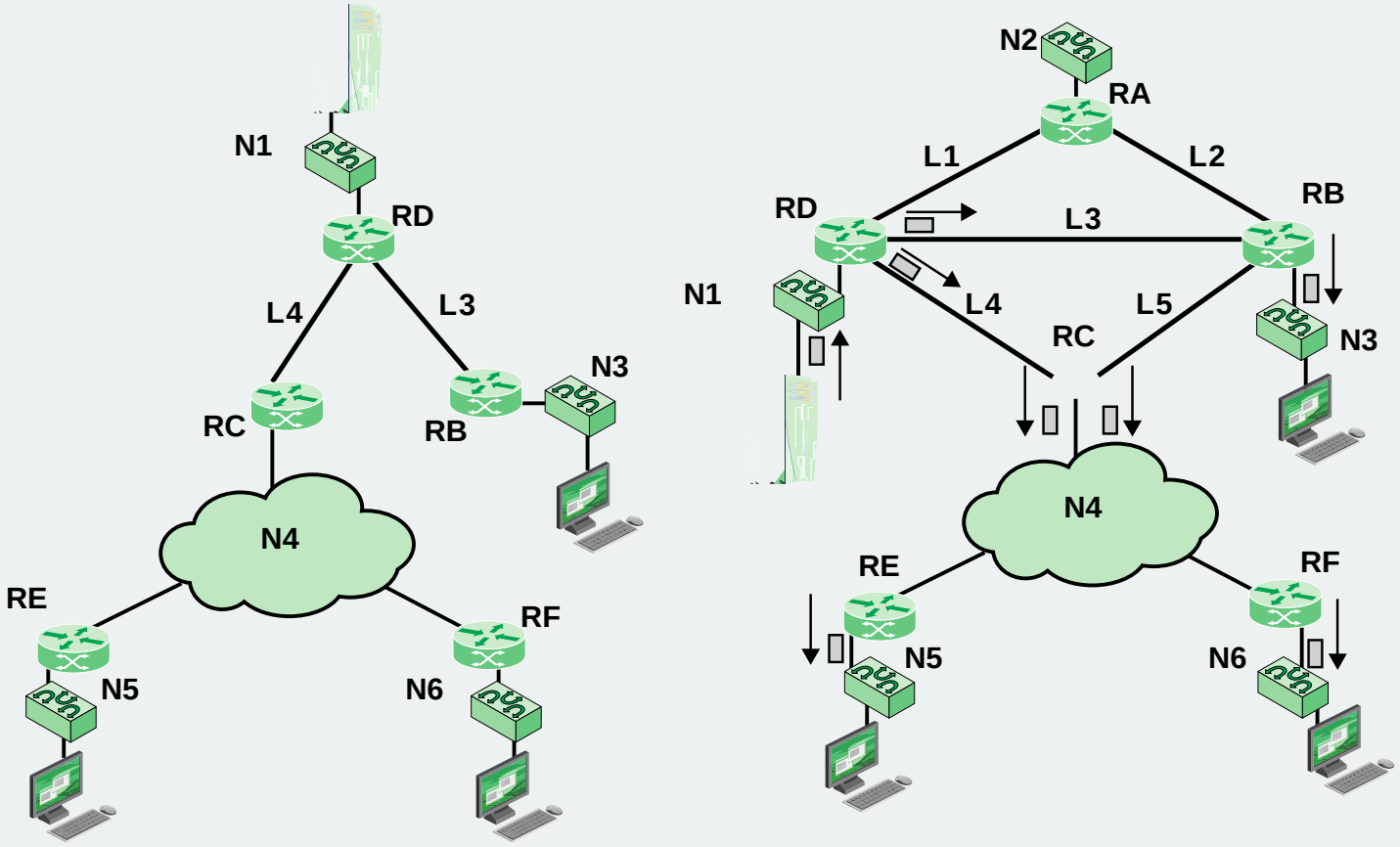
- To each network with members in multicast group
- Requires 11 packets

True multicast

- Spanning tree
- Replicated by routers at branch points
- Requires 8 packets

Traffic Generated by Various Multicasting Strategies

	Broadcast					Multiple Unicast				Multicast
	S → N2	S → N3	S → N5	S → N6	Total	S → N3	S → N5	S → N6	Total	
N1	1	1	1	1	4	1	1	1	3	1
N2										
N3		1			1	1			1	1
N4			1	1	2		1	1	2	2
N5			1		1		1		1	1
N6				1	1			1	1	1
L1	1				1					
L2										
L3		1			1	1			1	1
L4			1	1	2		1	1	2	1
L5										
Total	2	3	4	4	13	3	4	4	11	8



(a) Spanning tree from source to multicast group

(b) Packets generated for multicast transmission

Figure 21.2 Multicast Transmission Example

Requirements for Multicasting

- Router may have to forward more than one copy of packet
- Need convention to identify multicast addresses (IPv4, Class D, IPv6)
- Nodes translate between IP multicast addresses and list of networks containing group members
- Router must translate between IP multicast address and network multicast address

Cont…

Requirements for Multicasting

- Mechanism required for hosts to join and leave multicast group
- Routers must exchange information
  - Which networks include members of given group
  - Sufficient information to work out shortest path to each network
- Routing algorithm to calculate shortest path
- Routers must determine routing paths based on source and destination addresses

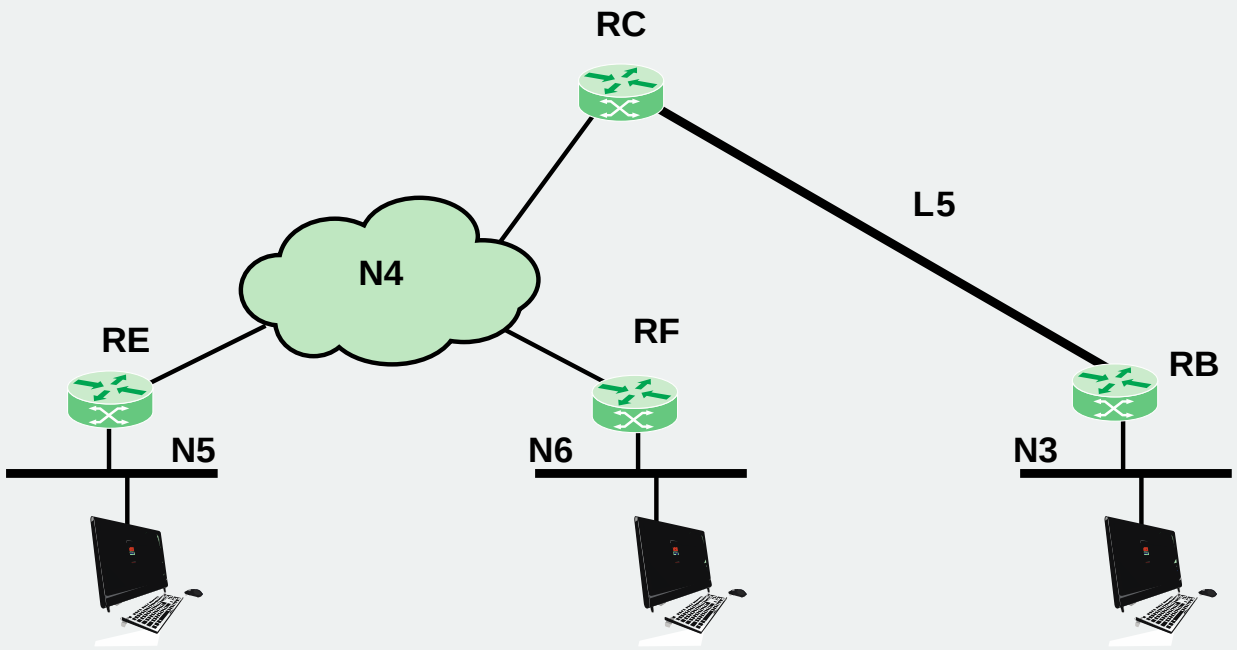


Figure 21.3 Spanning Tree from Router C to Multicast Group

Internet Group Management Protocol (IGMP)

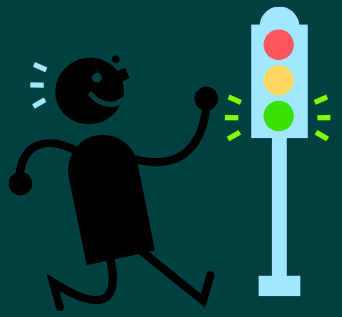
- Defined in RFC 3376
- Used to exchange multicast group information between hosts and routers on a LAN
- Hosts send messages to routers to subscribe and unsubscribe from multicast group
- Routers check which multicast groups are of interest to which hosts
- IGMP currently at version 3

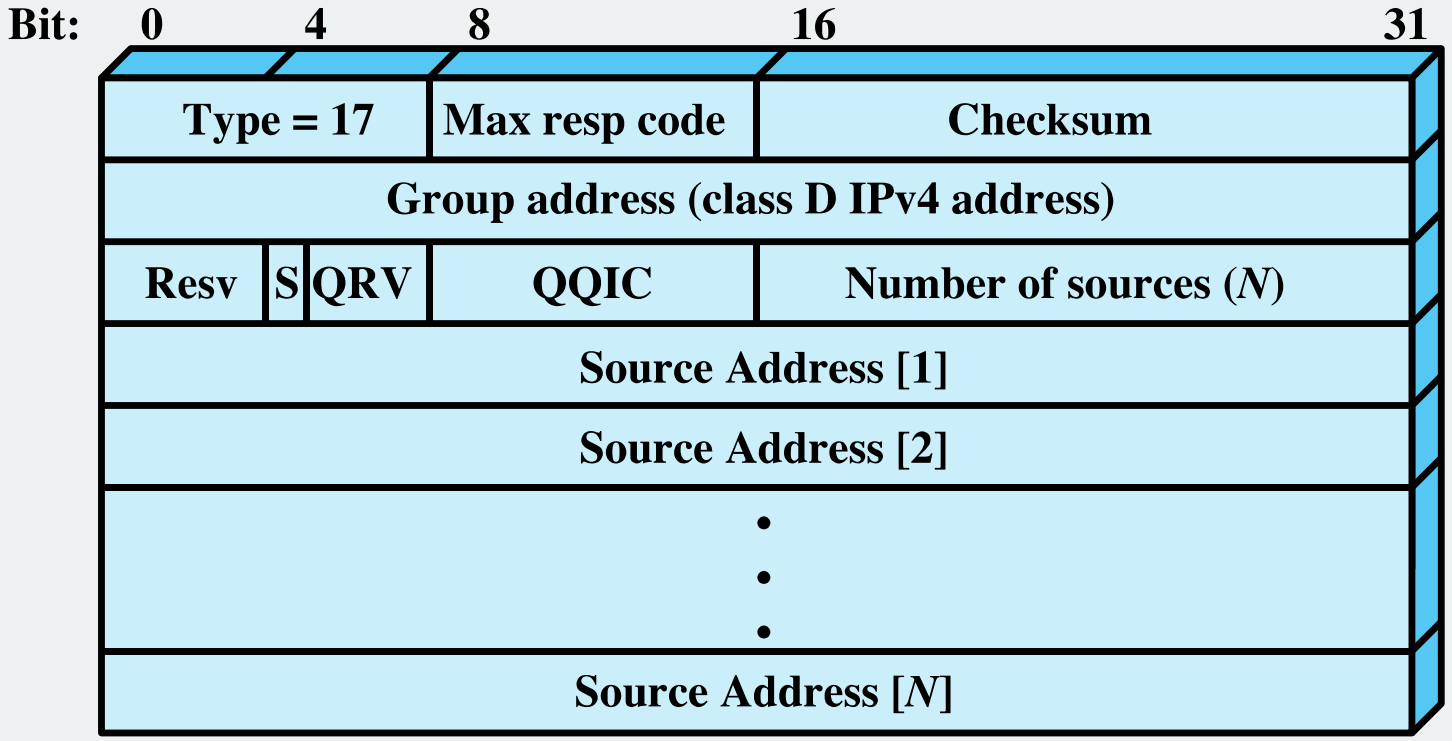
Operation of IGMP v1 and v2

- IGMPv1
  - Hosts could join group
  - Routers used timer to unsubscribe members
- IGMPv2 enabled hosts to unsubscribe
- Operational model:
  - Receivers have to subscribe to groups
  - Sources do not have to subscribe to groups
  - Any host can send traffic to any multicast group
- Problems:
  - Spamming of multicast groups
  - Establishment of distribution trees is problematic
  - Finding globally unique multicast addresses is difficult

IGMP v3

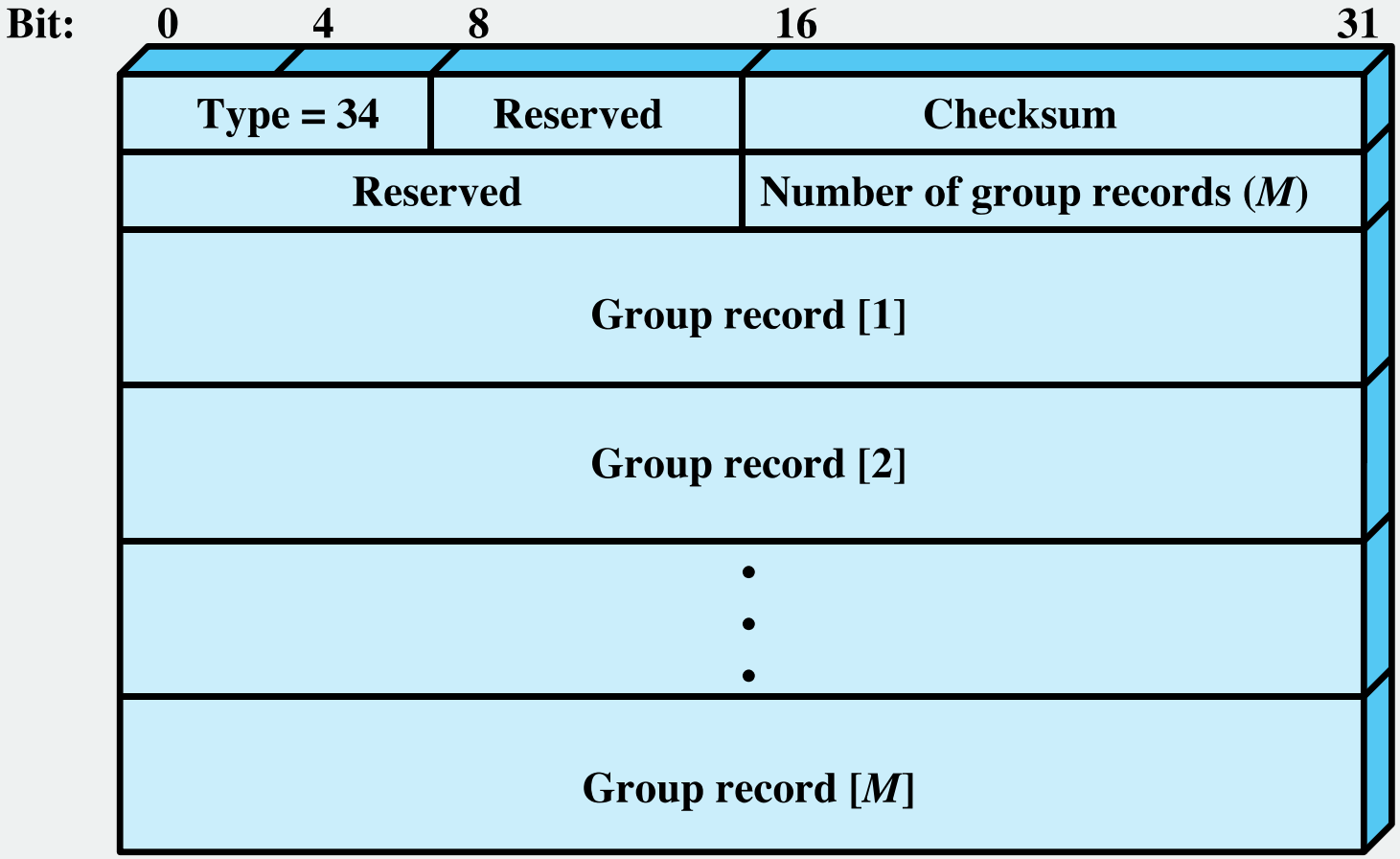
- Addresses weaknesses by:
  - Allowing hosts to specify list from which they want to receive traffic
  - Blocking traffic from other hosts at routers
  - Allowing hosts to block packets from sources that send unwanted traffic





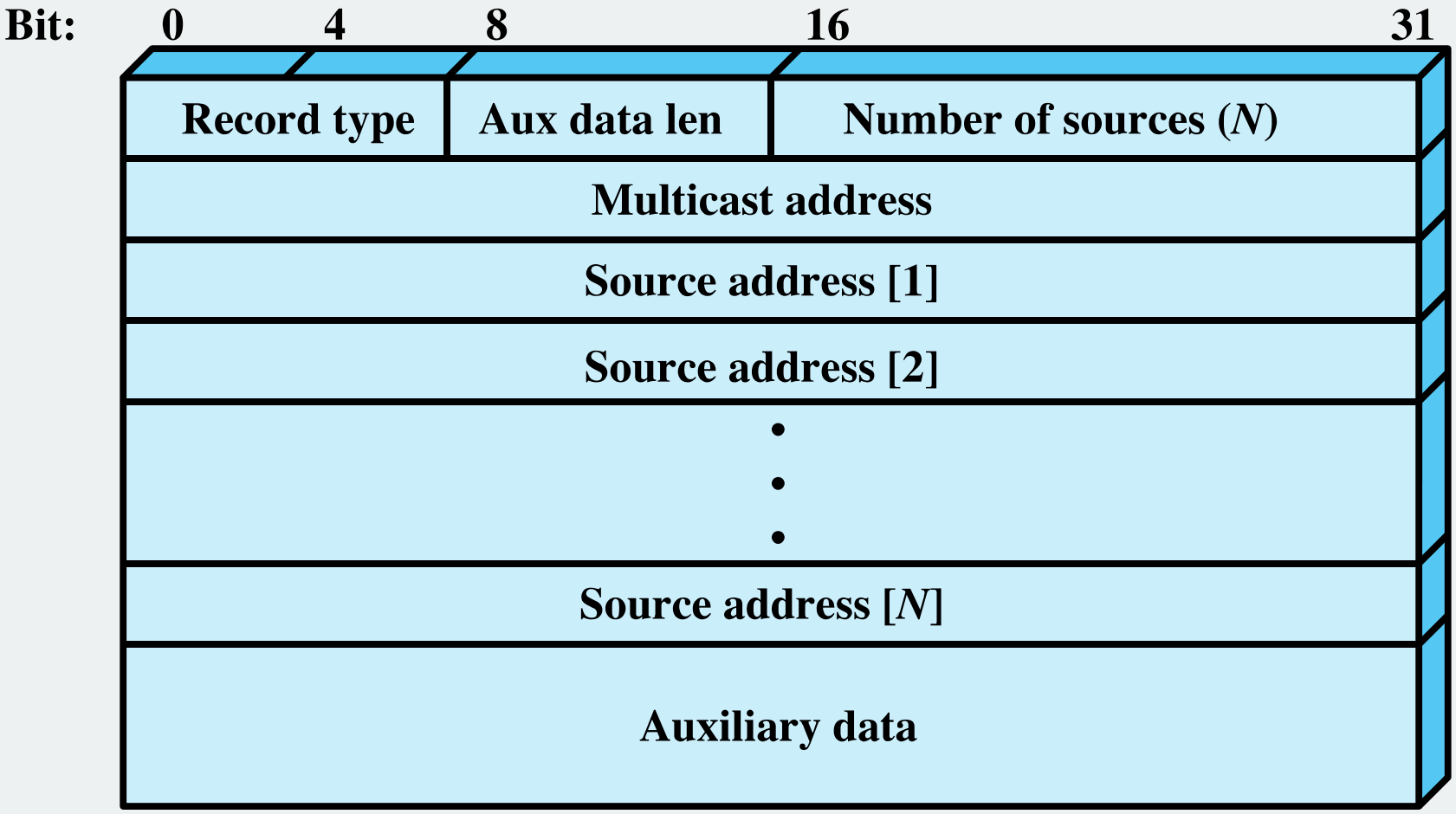
(a) Membership query message

IGMPv3 Message Formats



(b) Membership report message

IGMPv3 Message Formats



(c) Group record

IGMPv3 Message Formats

IGMP Operation - Joining

- IGMP host wants to make itself known as group member to other hosts and routers on LAN
- IGMPv3 can signal group membership with filtering capabilities with respect to sources
  - EXCLUDE mode – all members except those listed
  - INCLUDE mode – only from group members listed

To join a group a host sends an IGMP membership report message

- Address field is the multicast address of group
- Sent in an IP datagram with the same multicast destination address
- Current group members receive and learn new member
- Routers listen to all IP multicast addresses to hear all reports

IGMP Operation
Keeping Lists Valid

Routers periodically issue IGMP general query message

- In datagram with all-hosts multicast address
- Hosts must read such datagrams
- Hosts respond with report message

Router doesn't know every host in a group

- Needs to know at least one group member still active
- Each host in group sets timer with random delay
- Host hearing another report cancels own
- If timer expires, host sends report
- Only one member of each group reports to router

IGMP Operation - Leaving

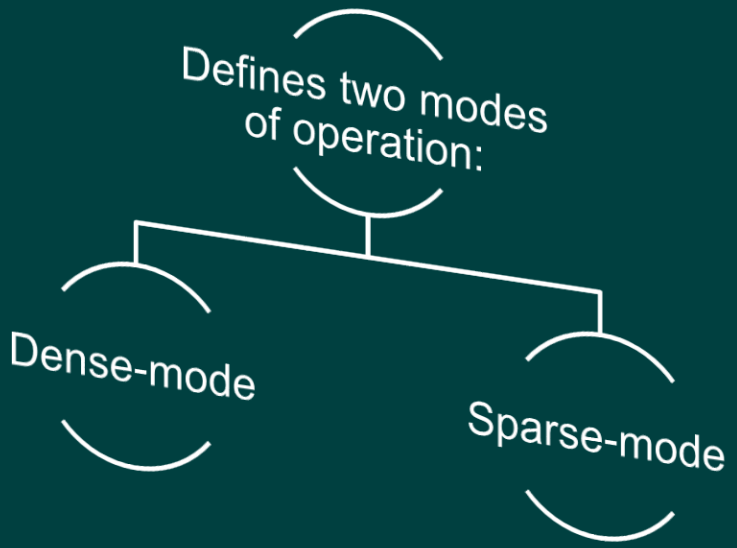
- Host leaves group by sending a leave group message to the all-routers static multicast address
  - Sends a membership report message with EXCLUDE option and null list of source addresses
- Router determines if any group members using group-specific query message remain

Group Membership with IPv6

- IGMP defined for IPv4
  - Uses 32-bit addresses
- IPv6 internets need functionality
- IGMP functions included in Internet Control Message Protocol v6 (ICMPv6)
  - ICMPv6 has functionality of ICMPv4 & IGMP
- ICMPv6 includes group-membership query and group-membership report message

Protocol Independent Multicast (PIM)

- A separate routing protocol, independent of any existing unicast routing protocol
- Designed to extract needed routing information from any unicast routing protocol
- Recognizes that a different approach may be needed to multicast routing depending on the concentration of multicast group members



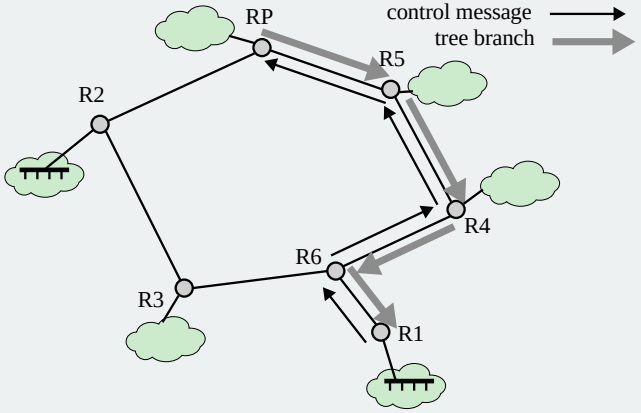
PIM

Dense-Mode PIM

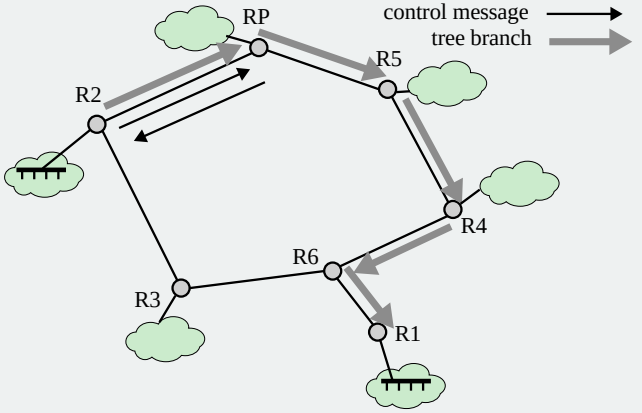
- Appropriate for intra-AS multicast routing
- May be viewed as a potential alternative to a multicast version of OSPF known as MOSPF

Sparse-Mode PIM

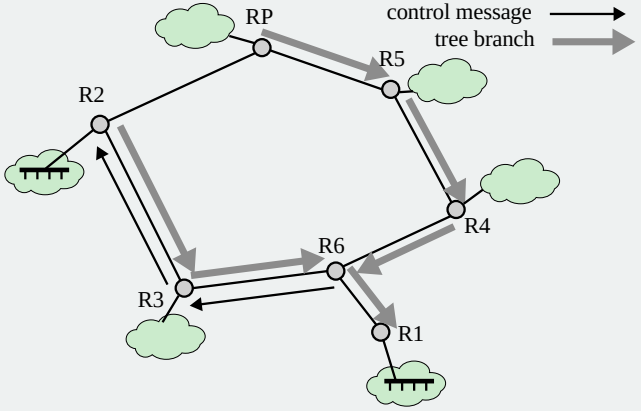
- Suited for inter-AS multicast routing
- RFC 2362
- Defined as a group in which:
  - The number of networks/domains with group members present is significantly smaller than the number of networks/domains in the internet
  - The internet spanned by the group is not sufficiently resource rich to ignore the overhead of current multicast routing schemes



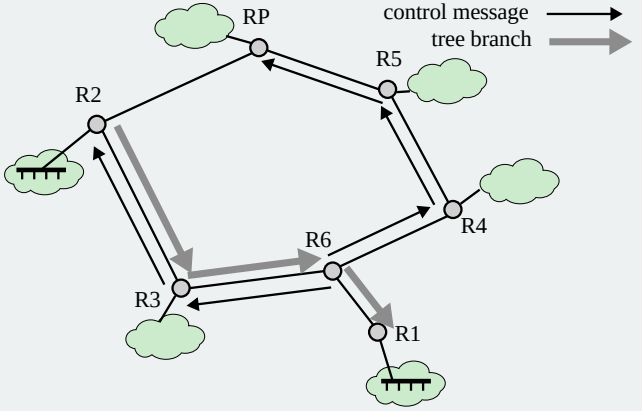
(a) R1 sends Join toward RP; RP adds path to distribution tree



(b) R2 sends Register to RP; RP returns Join; R2 builds path to RP



(c) R1 sends Join to R2; R2 prunes path to RP



(d) R6 sends Prune to RP; RP prunes path to R1

Figure 21.5 Example of PIM Operation

Evolving Network Requirements

- Increasingly widespread use of server virtualization
  - Makes it possible to partition a single machine into multiple independent servers, conserving hardware resources
  - Also makes it possible to quickly migrate a server from one machine to another for load balancing or for dynamic switchover in the case of machine failure
  - Has become a central element in dealing with “big data” applications and in implementing cloud computing infrastructures
- Increasing use by employees of mobile devices, such as smartphones, tablets, and notebooks to access enterprise resources
  - Network managers must be able to respond to rapidly changing resource, QoS, and security requirements

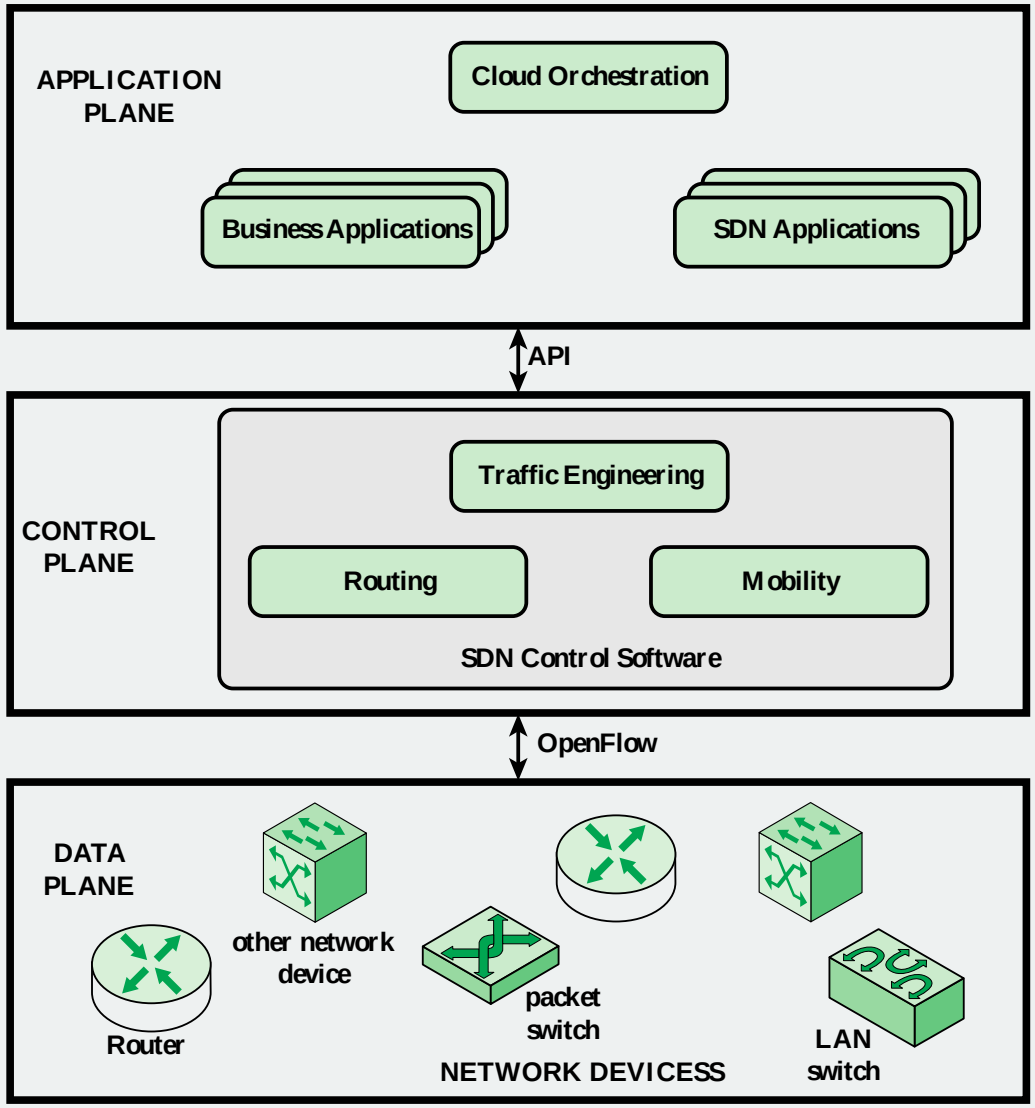


Figure 21.6 SDN Logical Structure

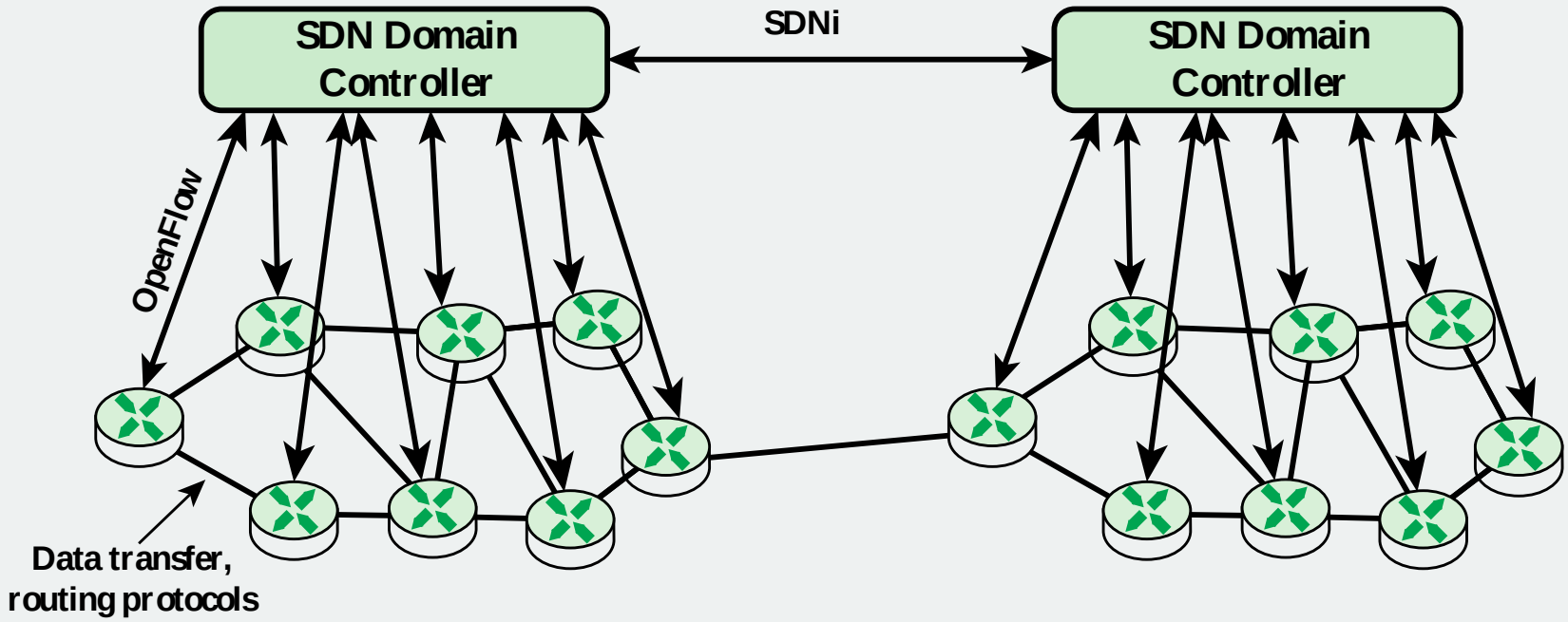


Figure 21.7 SDN Domain Structure

SDN Domains

- Reasons for using SDN domains:

Scalability

The number of devices an SDN controller can feasibly manage is limited

A reasonably large network may need to deploy multiple SDN controllers

Privacy

A carrier may choose to implement different privacy policies in different SDN domains

Incremental deployment

A carrier's network may consist of portion of legacy and nonlegacy infrastructure

Dividing the network into multiple, individually manageable SDN domains allows for flexible incremental deployment

SDNi

- Protocol being developed for interfacing SDN domain controllers
- Functionality includes:
  - Coordinate flow setup originated by application, continuing information such as path requirement, QoS, and service level agreements across multiple SDN domains
  - Exchange reachability information to facilitate inter-SDN routing

Message types tentatively include:

- Reachability update
- Flow setup/tear-down/update request
- Capability update

OpenFlow

- Defined in the *OpenFlow Switch Specification* published by the Open Networking Foundation (ONF)
  - ONF is a consortium of software providers, content delivery networks, and networking equipment vendors whose purpose is to promote software-defined networking
- Is both a protocol between SND controllers and network devices and a specification of the logical structure of the network switch functionality
- To turn the concept of SND into practical implementation two requirements must be met:

1. There must be a common logical architecture in all switches, routers, and other network devices to be managed by an SDN controller
2. A standard, secure protocol is needed between the SDN controller and the network device

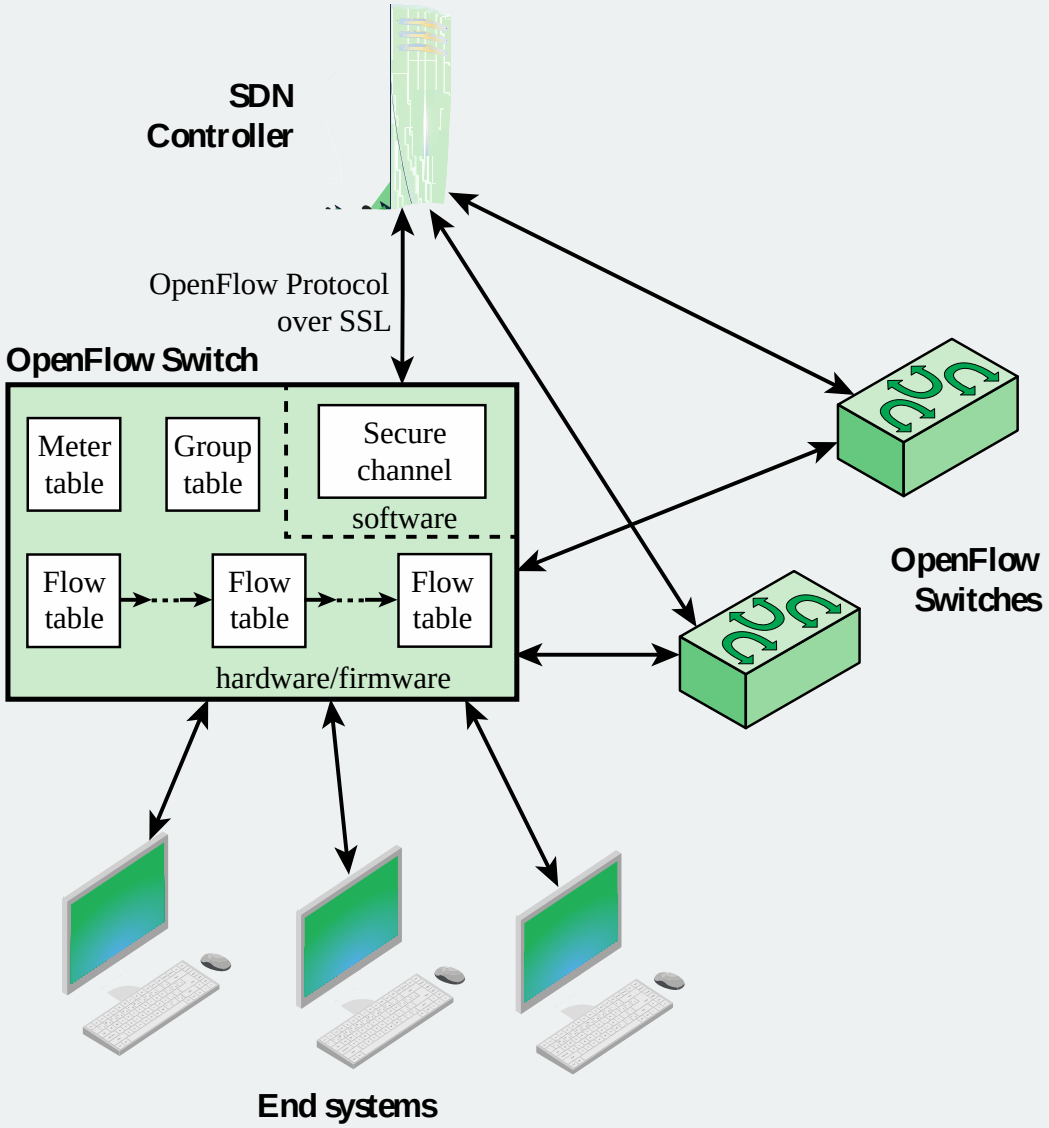
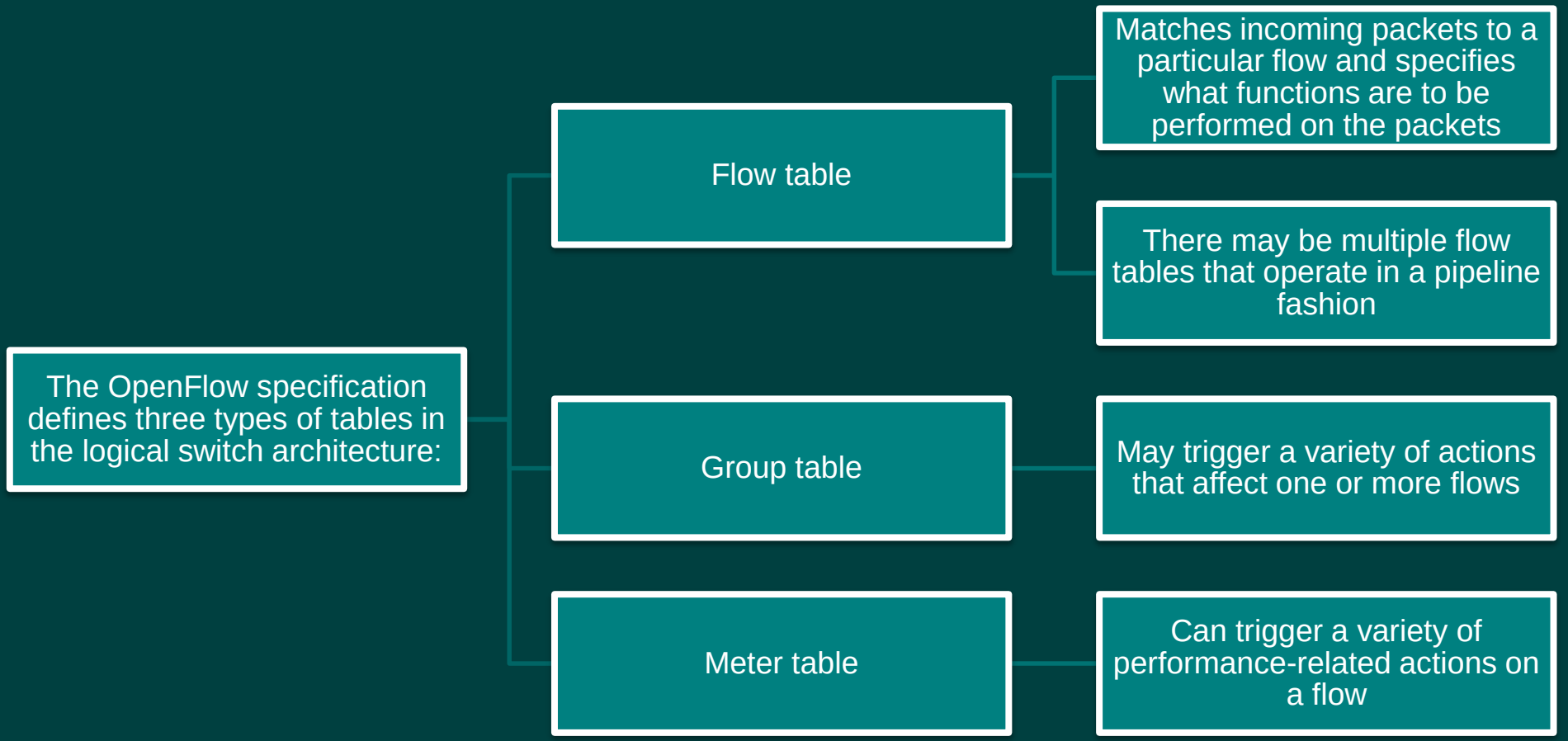


Figure 21.8 OpenFlow Switch

Tables



- A flow is a sequence of packets traversing a network that share a set of header field values

Flow Table Components

- Basic building block of the logical switch architecture
- May include a table-miss flow entry, which wildcards all match fields (regardless of value) and has the lowest priority (0)

Contains entries consisting of six components:	Match fields
	Priority
	Counters
	Instructions
	Time-outs
	Cookie

Match Fields Component

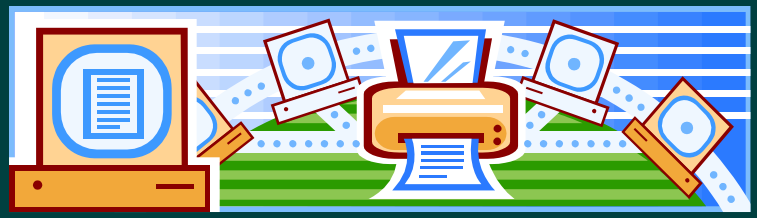
- Consists of the following required fields which must be supported by any OpenFlow-compliant switch:
  - Ingress port
  - Ethernet source and destination addresses
  - IPv4 or IPv6 protocol number
  - IPv4 or IPv6 source address, and destination address
  - TCP source and destination ports
  - UDP source and destination ports

The following fields may be optionally supported:

- Physical port
- Metadata
- Ethernet type
- VLAN ID and VLAN user priority
- IPv4 or IPv6 DS and ECN
- SCTP source and destination ports
- ICMP type and code fields
- ARP opcodes
- Source and target IPv4 addresses in ARP payload
- IPv6 flow label
- ICMPv6 type and code fields
- IPv6 neighbor discovery source and target addresses
- MPLS label value, traffic class, and BoS

Instructions Component

- Consist of a set of instructions that are executed if the packet matches the entry
- “Actions” describe packet forwarding, packet modification, and group table processing operations
- The OpenFlow specification includes the following actions
  - Output
  - Set-Queue
  - Group
  - Push-Tag/Pop-Tag
  - Set-Field
  - Change-TTL



Instructions Component

- “Action set” is a list of actions associated with a packet that are accumulated while the packet is processed by each table and that are executed when the packet exits the processing pipeline
- Instructions are of four types

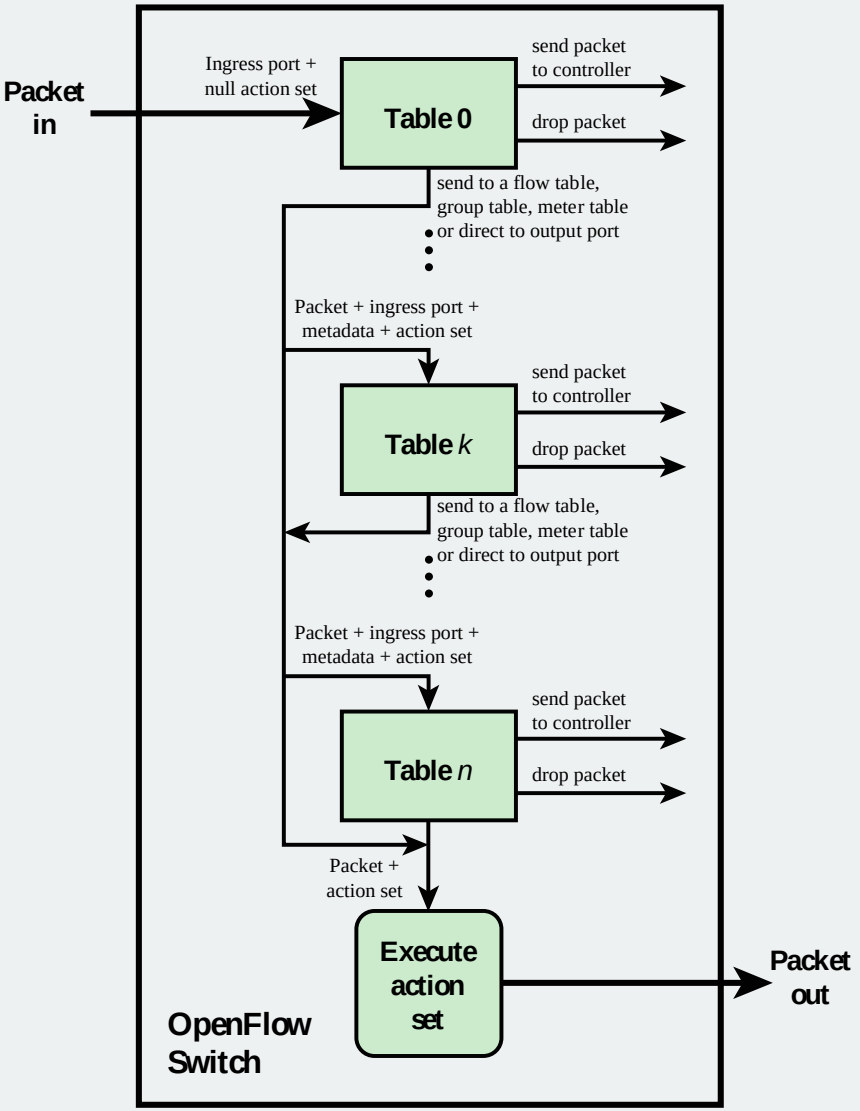


Figure 21.9 Packet Flow through OpenFlow-Compliant Switch

OpenFlow Messages

Message	Description
Controller-to-Switch	
Features	Request the capabilities of a switch. Switch responds with a features reply that specifies its capabilities.
Configuration	Set and query configuration parameters. Switch responds with parameter settings
Modify-State	Add, delete, and modify flow/group entries and set switch port properties.
Read-State	Collect information from switch, such as current configuration, statistics, and capabilities.
Packet-out	Direct packet to a specified port on the switch.
Barrier	Barrier request/reply messages are used by the controller to ensure message dependencies have been met or to receive notifications for completed operations.
Role-Request	Set or query role of the OpenFlow channel. Useful when switch connects to multiple controllers.
Asynchronous-Configuration	Set filter on asynchronous messages or query that filter. Useful when switch connects to multiple controllers.
Asynchronous	
Packet-in	Transfer packet to controller.
Flow-Removed	Inform the controller about the removal of a flow entry from a flow table.
Port-Status	Inform the controller of a change on a port.
Error	Notify controller of error or problem condition.
Symmetric	
Hello	Exchanged between the switch and controller upon connection startup.
Echo	Echo request/reply messages can be sent from either the switch or the controller, and must return an echo reply.
Experimenter	For additional functionality.

Mobile IP

- Enables computers to maintain Internet connectivity while moving from one Internet attachment point to another
- Particularly suited for wireless connections
- ***Mobile*** implies:
  - A user is connected to one or more applications across the Internet
  - The user's point of attachment changes dynamically
  - All connections are automatically maintained despite the change

Operation of Mobile IP

- Routers use the IP address in an IP datagram to perform routing
- Network portion is used to move a datagram to the network the target computer is attached to
- Final router uses the host portion to deliver to the destination
- With a mobile host the IP address may change while one or more TCP connections are active

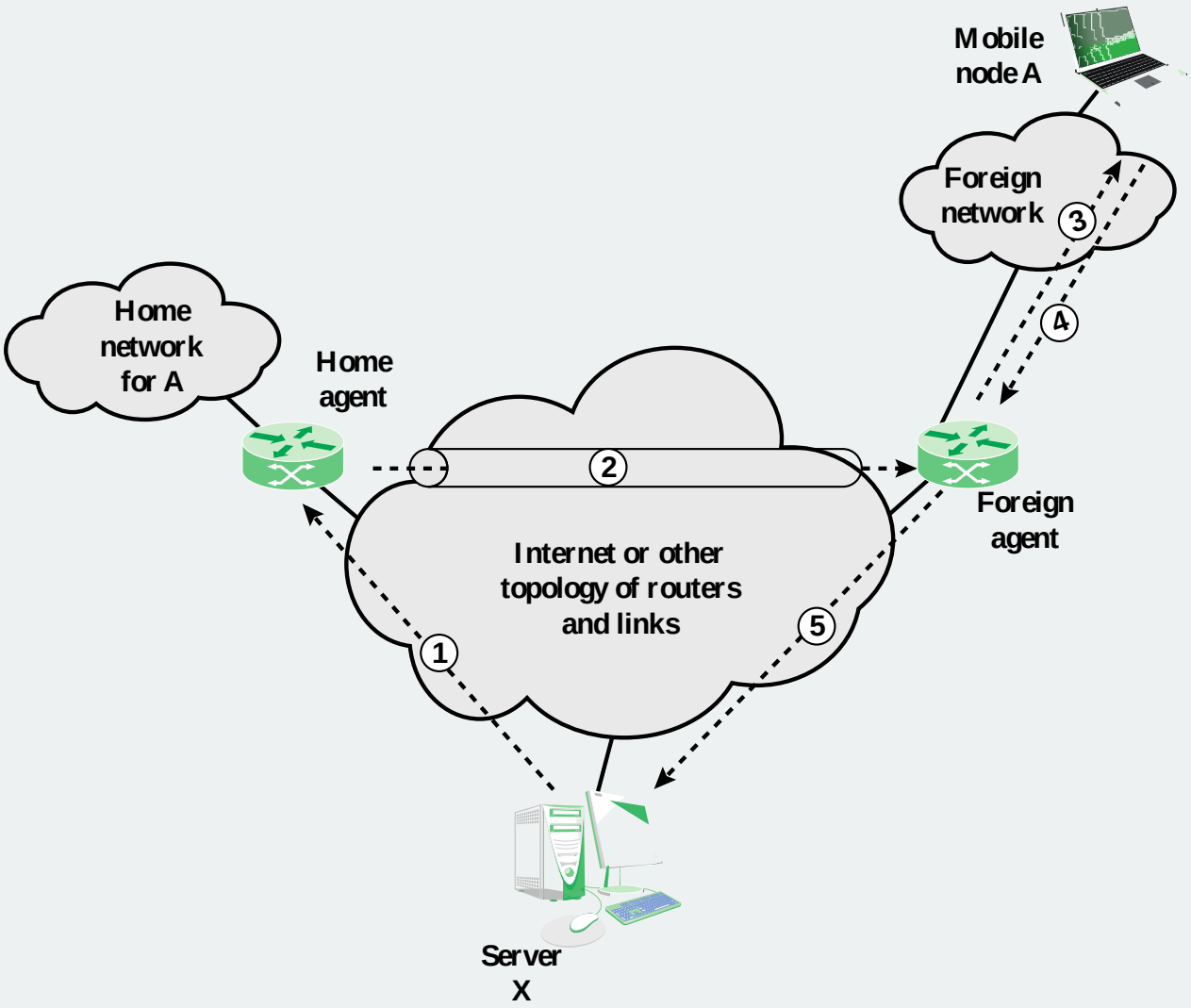
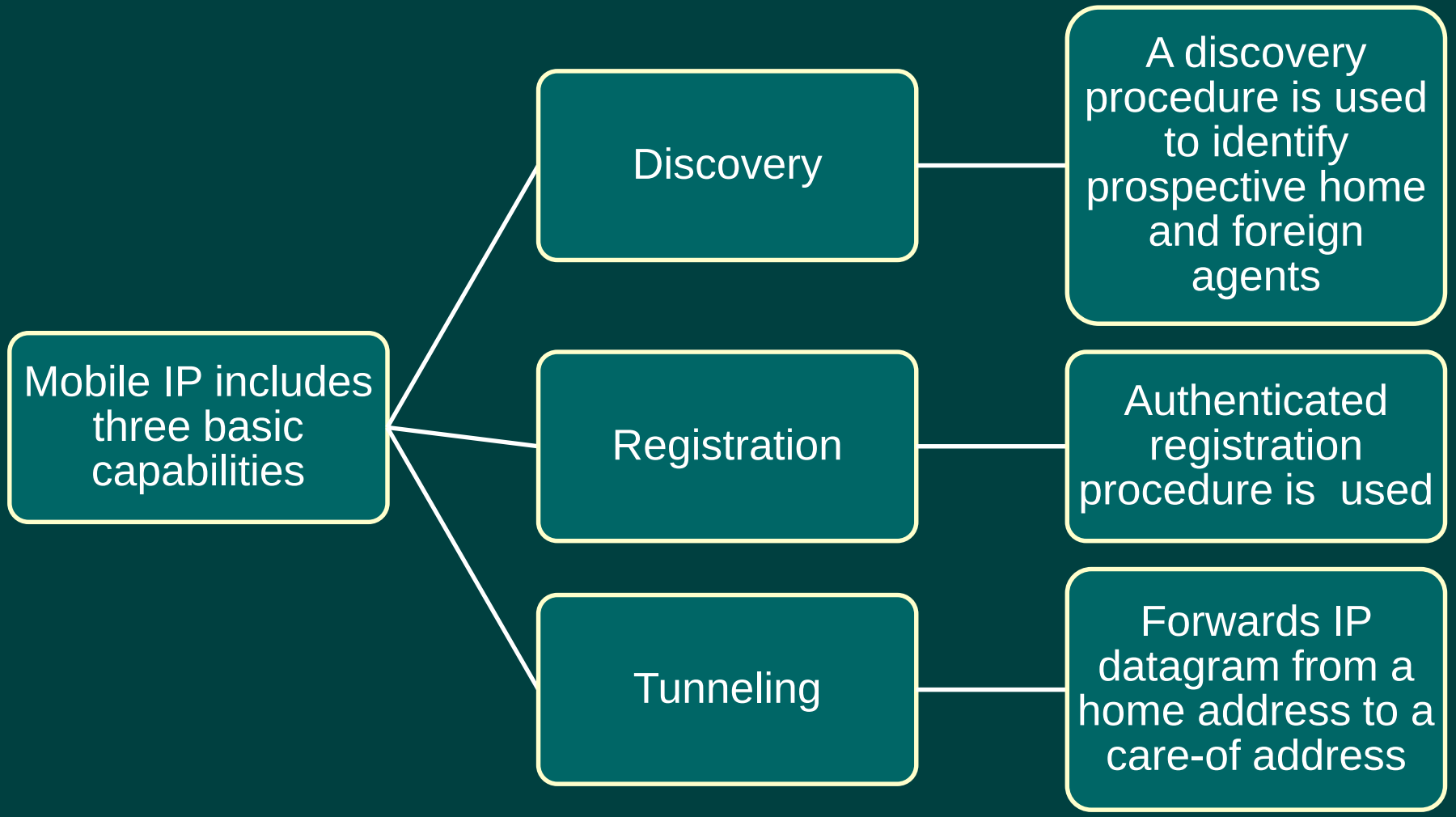


Figure 21.10 Mobile IP Scenario

Basic Capabilities of Mobile IP



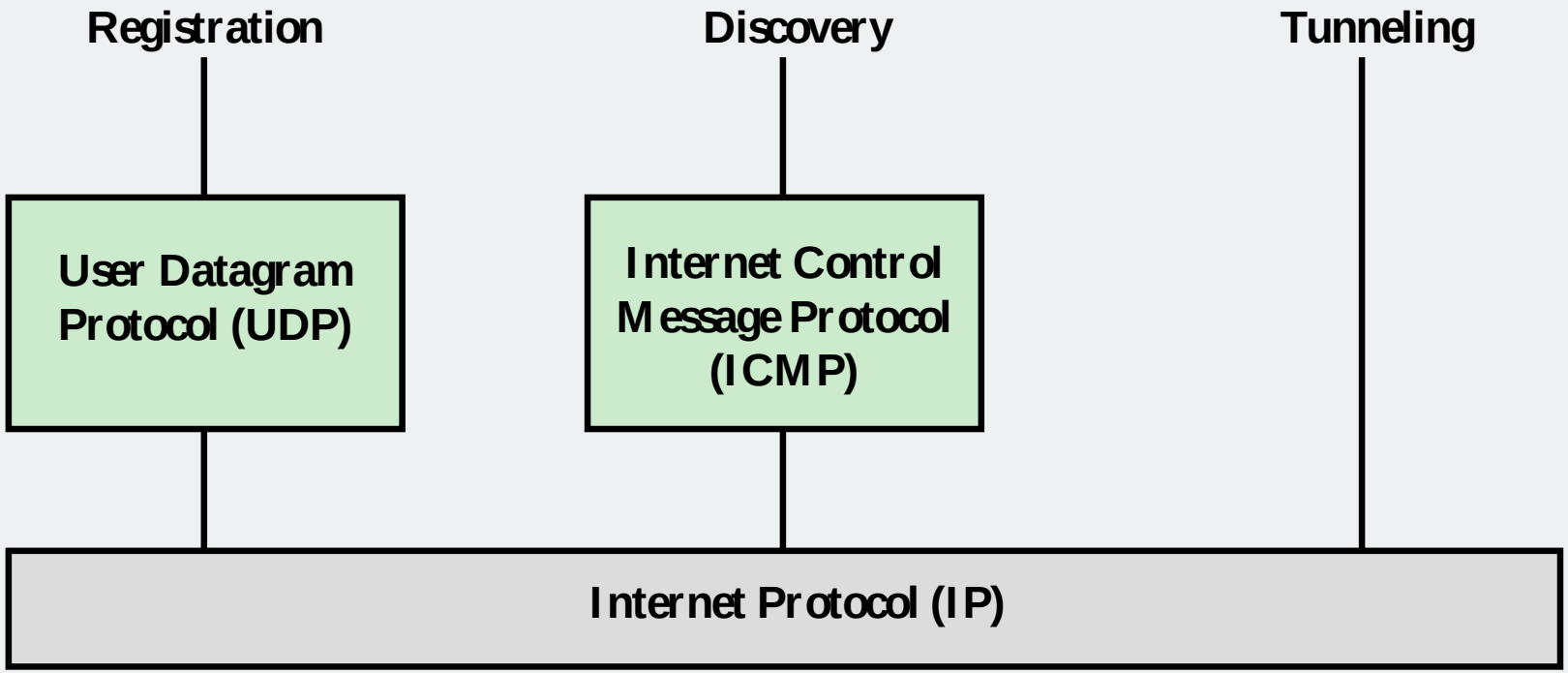


Figure 21.11 Protocol Support for Mobile IP

Mobile IP Terminology (RFC 3334)

Mobile node	A host or router that changes its point of attachment from one network or subnetwork to another. A mobile node may change its location without changing its IP address; it may continue to communicate with other Internet nodes at any location using its (constant) IP address, assuming link-layer connectivity to a point of attachment is available
Home address	An IP address that is assigned for an extended period of time to a mobile node. It remains unchanged regardless of where the node is attached to the Internet.
Home agent	A router on a mobile node's home network that tunnels datagrams for delivery to the mobile node when it is away from home, and maintains current location information for the mobile node.
Home network	A network, possibly virtual, having a network prefix matching that of a mobile node's home address. Note that standard IP routing mechanisms will deliver datagrams destined to a mobile node's Home Address to the mobile node's Home Network.
Foreign agent	A router on a mobile node's visited network that provides routing services to the mobile node while registered. The foreign agent detunnels and delivers datagrams to the mobile node that were tunneled by the mobile node's home agent. For datagrams sent by a mobile node, the foreign agent may serve as a default router for registered mobile nodes.
Foreign network	Any network other than the mobile node's Home Network.
Care-of address	The termination point of a tunnel toward a mobile node, for datagrams forwarded to the mobile node while it is away from home. The protocol can use two different types of care-of address: a "foreign agent care-of address" is an address of a foreign agent with which the mobile node is registered, and a "co-located care-of address" is an externally obtained local address which the mobile node has associated with one of its own network interfaces.
Correspondent node	A peer with which a mobile node is communicating. A correspondent node may be either mobile or stationary.
Link	A facility or medium over which nodes can communicate at the link layer. A link underlies the network layer.
Node	A host or a router.
Tunnel	The path followed by a datagram while it is encapsulated. The model is that, while it is encapsulated, a datagram is routed to a knowledgeable decapsulating agent, which decapsulates the datagram and then correctly delivers it to its ultimate destination.

Discovery

- Similar to the router advertisement process defined in ICMP
- Mobile node is responsible for an ongoing discovery process
- Home or foreign network
- Listens for agent advertisement message
- Compares IP address with home address
- If these do not match the mobile node is on a foreign network

Use of Lifetime Field

Upon receipt of an agent advertisement from a foreign agent the mobile node records the lifetime field as a timer

Otherwise, node uses agent solicitation to find an agent

If timer expires before receipt of another advertisement, node assumes it lost contact

If node has received an advertisement that is not expired, node registers with the new agent

Use of Network Prefix

Mobile node checks if newly received agent advertisement is on the same network as the node's current care-of address

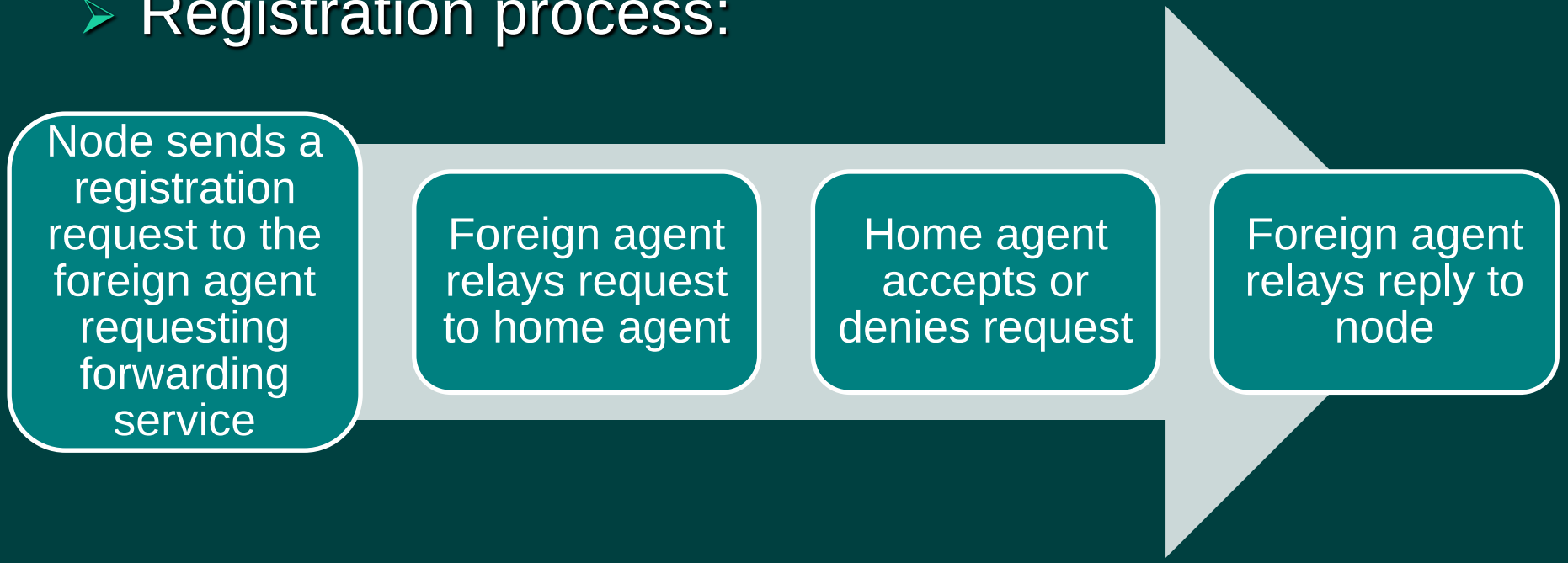
If it is not, the node assumes it moved and registers with advertisement the node has just received

Co-Located Address

- Node may move to a network that has no foreign agents or foreign agents are busy
  - May act as its own foreign agent by using a co-located care-of address
- Co-located care-of address is an IP address that is associated with the node's current interface to a network
  - Can dynamically acquire a temporary IP address
  - Co-located address may be owned by the node

Registration

- Once care-of address is acquired the mobile node needs to request the home agent forward its IP traffic
- Registration process:



- If node is using a co-located care-of address it registers directly with its home agent

Registration Messages

- Registration operation uses two types of messages carried in UDP segments
- **Registration request message includes**:
  - One-bit flags
  - Home address field
  - Home agent field
  - Care-of address field
  - Identification field
- **Registration reply message includes:**
  - Acceptance code
  - Reason for denial

Registration Security

- Mobile IP is designed to resist two types of attacks:
  - Node pretends to be a foreign agent and sends registration request to home agent to divert traffic
  - Malicious agent may replay old registration messages effectively cutting node from the network

Message Authentication

- Message authentication is used to protect against registration message attacks
- Authentication extension includes the following fields:
  - Security parameter index (SPI)
  - Authenticator

Three types of authentication extensions:

- Mobile-home
- Mobile-foreign
- Foreign-home

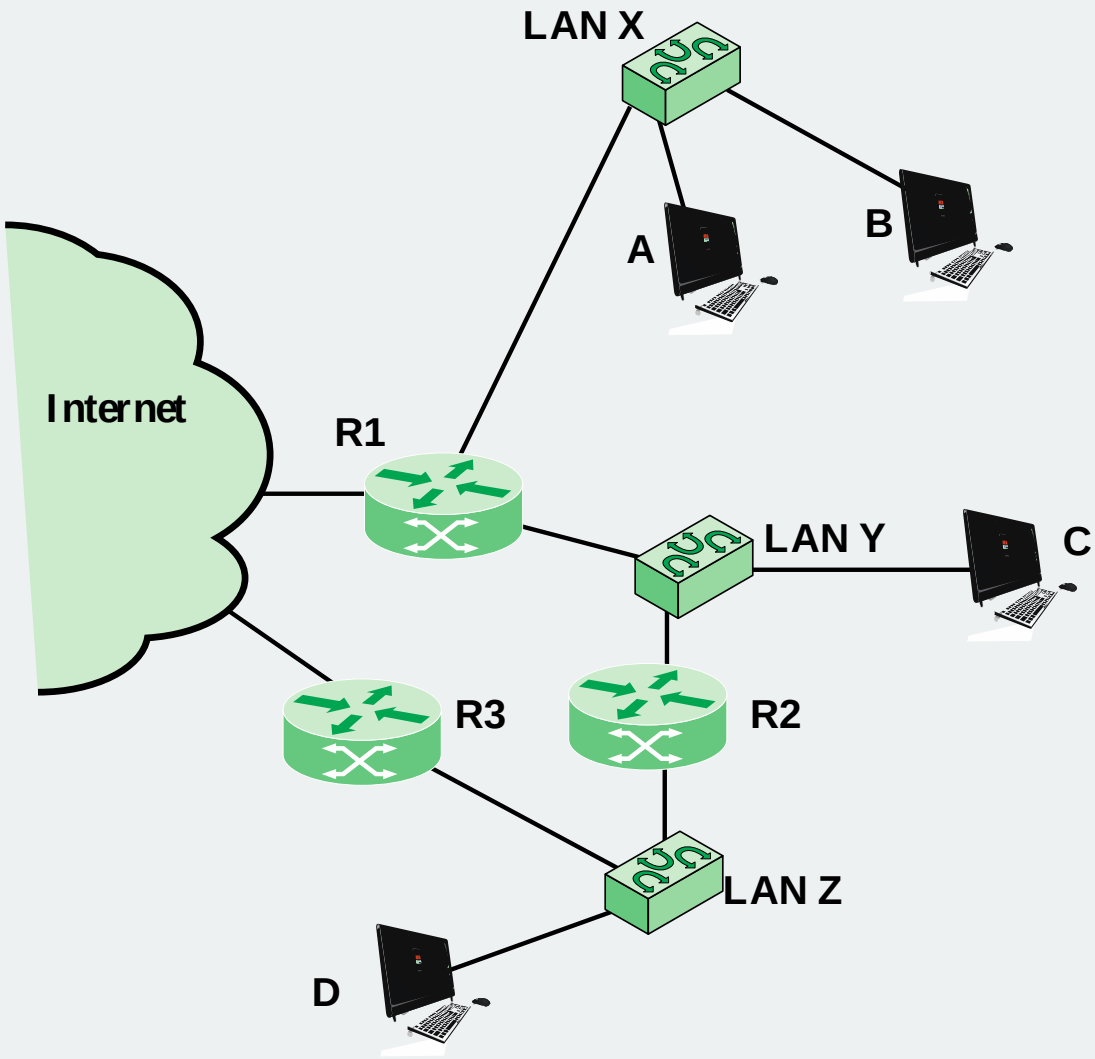
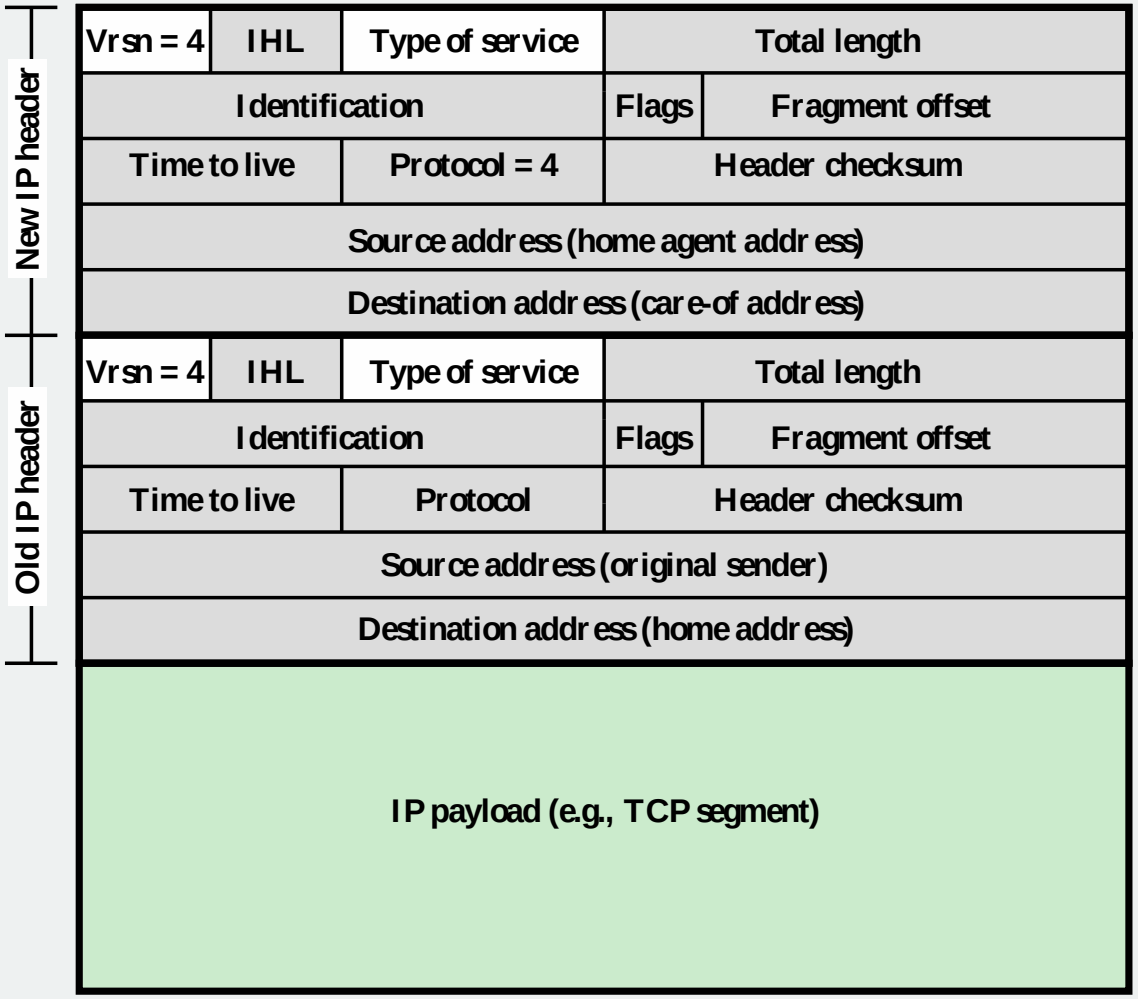


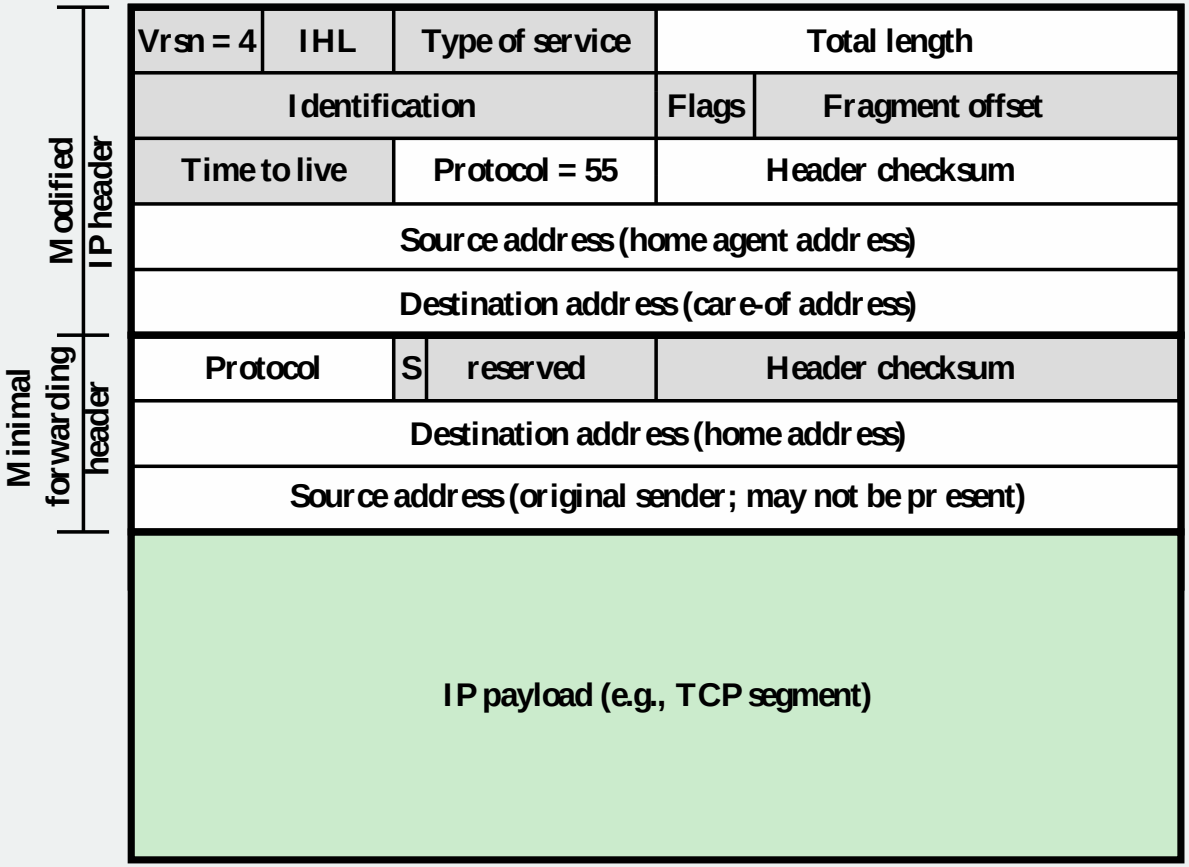
Figure 21.12 A Simple Internetworking Example



(a) IP-within-IP encapsulation

Unshaded fields are copied from the inner IP header to the outer IP header.

Mobile IP Encapsulation



(b) Minimal encapsulation

Unshaded fields in the inner IP header are copied from the original IP header. Unshaded fields in the outer IP header are modified from the original IP header.

Mobile IP Encapsulation

Dynamic Host Configuration Protocol (DHCP)

Internet protocol that enables dynamic allocation of IP addresses to hosts

Defined in RFC 2131

Was developed to deal with the shortage of IP addresses

Enables a local network to assign IP addresses from a pool of available IP addresses to hosts currently in use

- When a host is not in use, its IP address is returned to the pool managed by a DHCP server

Can also assign permanent IP addresses to some systems, such as servers, so that the address remains the same when the system is rebooted

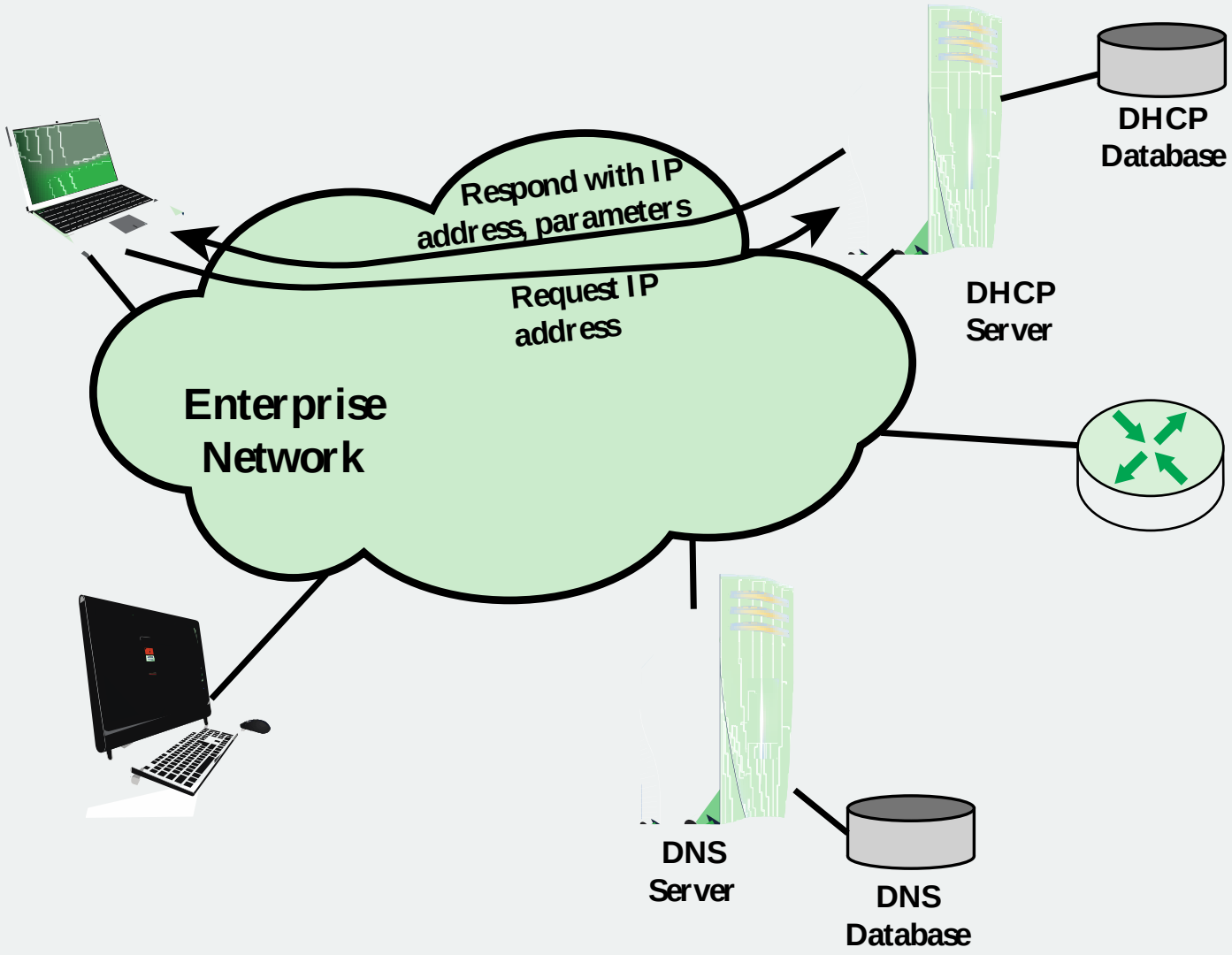


Figure 21.14 DHCP Role

The following DHCP messages are used for protocol operation:

DHCPDISCOVER

- Client broadcast to locate available servers

DHCPOFFER

- Server to client in response to DHCPDISCOVER with offer of configuration parameters

DHCPREQUEST

- Client message to servers either (a) requesting offered parameters from one server and implicitly declining offers from all others, (b) confirming correctness of previously allocated address after, for example, system reboot, or (c) extending the lease on a particular network address

DHCPACK

- Server to client with configuration parameters, including committed network address

DHCPNACK

- Server to client indicating client's notion of network address is incorrect (e.g., client has moved to new subnet) or client's lease has expired

DHCPDECLINE

- Client to server indicating network address is already in use. DHCP server should then notify sysadmin

DHCPRELEASE

- Client to server relinquishing network address and canceling remaining lease

DHCPINFORM

- Client to server, asking only for local configuration parameters client already has externally configured network address

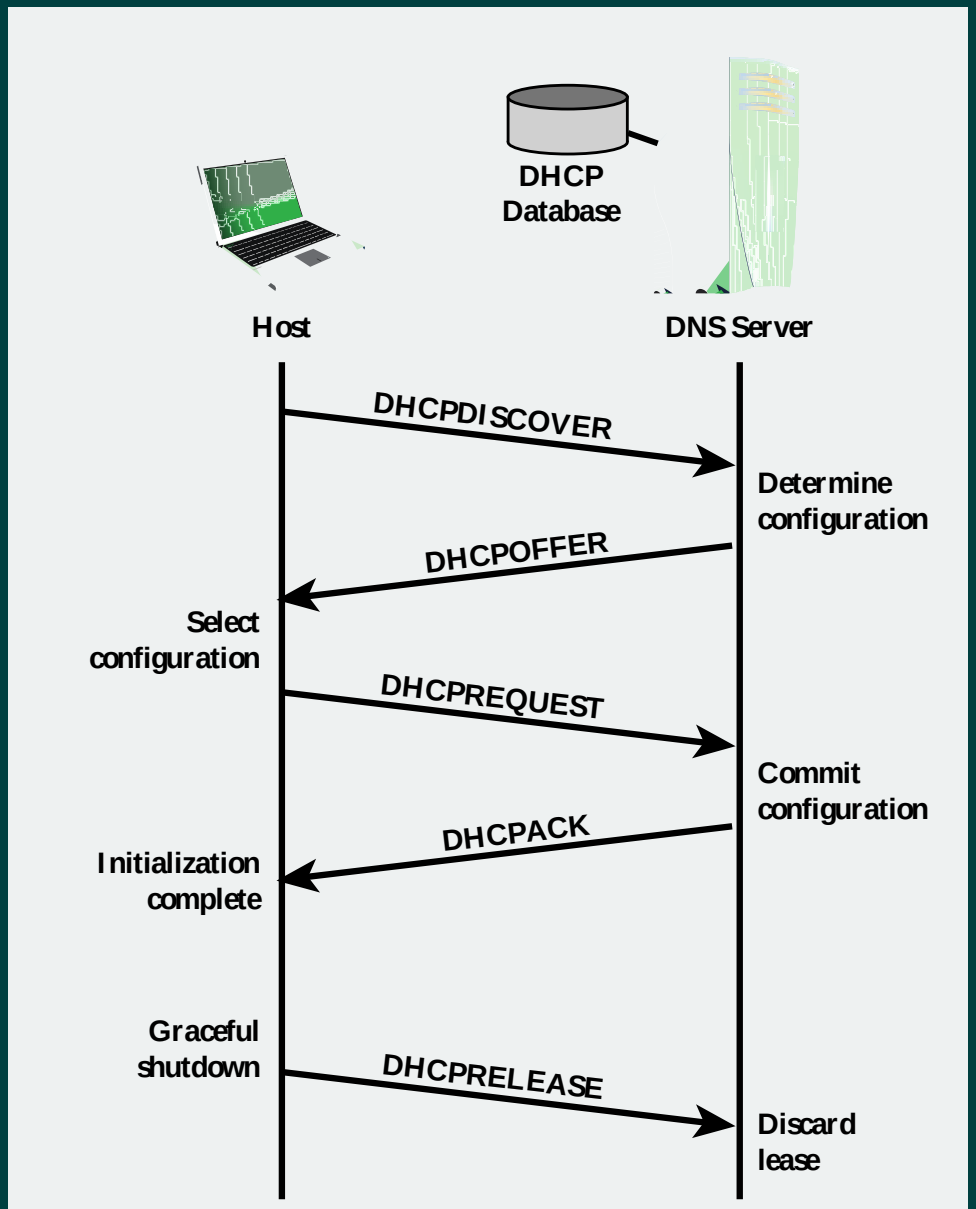


Figure 21.15 DHCP Message Exchange