TCP/IP Protocol & Network Programming Technology

Chapter 4:
Data Link and Network Layer Protocols in TCP/IP

Jungang Xu

Objectives

- Explain the fundamental concepts associated with Data Link layer protocols that operate over different network link types
- Distinguish among the different types of frames used on IP networks
- Describe how hardware address resolution occurs for IPv4 and IPv6 with ARP and NDP protocols
- Explain the essentials of the Internet Protocol, both for IPv4 and IPv6, including route resolution processes, IP datagrams, fragmentation, and the differences between IPv4 and IPv6

Objectives (cont'd.)

- Explain the mechanics of IP routing, such as how entries are placed in a routing table and how the basics of routing operate on an internetwork
- Describe the intricacies of IPv4 and IPv6 routing characteristics, including methods of preventing routing loops, general behaviors of routers in an internetwork, and routing determination
- Provide detailed information about IPv4 and IPv6 routing protocols, including the structure of routing packets and the behavior of each routing protocol

Objectives (cont'd.)

- Describe the various factors involved in choosing how to route between different network environments and infrastructures
- Describe the fundamentals of routing to and from the Internet
- Explain the basics of securing routers
- Describe the tools used to troubleshoot IP routing

Data Link Protocols

- Key functions of the Data Link layer
  - Media Access Control (MAC)
  - Logical Link Control (LLC)
- Point-to-point data transfer
  - Shipping data from one MAC layer address to another

Data Link Protocols (cont'd.)

- Data Encapsulation is used to encapsulate the data packet across the WAN
- WAN encapsulation of frames at Data Link layer involves one or more of the following services
  - Addressing
  - Bit-level integrity check
  - Delimitation (for header, payloads and trailer)
  - Protocol identification (PID)

Point-to-Point Protocol

- Provides
  - Frame delimitation
  - Protocol identification
  - Bit-level integrity check services
- RFC 1661 includes:
  - Encapsulation methods for a variety of protocols, TCP/IP, NetBEUI, IPX/SPX,AppleTalk,SNA,DECNet
  - A special Link Control Protocol (LCP)
  - A collection of negotiation protocols, Network Control Protocol (NCP)

Point-to-Point Protocol (cont'd.)

- Fields in the PPP header and trailer include:
  - Flag: 1 byte, 0x7E(01111110),between 2 frames
  - Protocol identifier: 2 byte, identify upper layer protocol
  - Frame Check Sequence (FCS): 2 byte
- Supports a default MTU of 1,500 bytes
  - Which makes it ideal for interconnecting Ethernet-based networks (or peers)
  - Negotiate Bigger and smaller MTU, 9216 byte for Gigabit Ethernet

Frame Types and Sizes

- At Data Link layer
  - Protocol data units are called frames
- Frame
  - Represents same data that appears in digital form at the Network layer in an IP datagram

Ethernet Frame Types

- Ethernet II frame type
  - De facto standard frame type used for IP datagram transmissions over Ethernet networks
  - Has protocol identification field, 0x0800,IPv6
- Ethernet frame types that TCP/IP can use
  - Ethernet II
  - Ethernet 802.2 Logical Link Control

Ethernet II Frame Structure

- Ethernet II frame type fields and structure
  - Preamble
  - Destination Address Field
  - Source Address Field
  - Type Field
  - Data Field
  - Frame Check Sequence Field

Ethernet II Frame Structure (cont'd.)

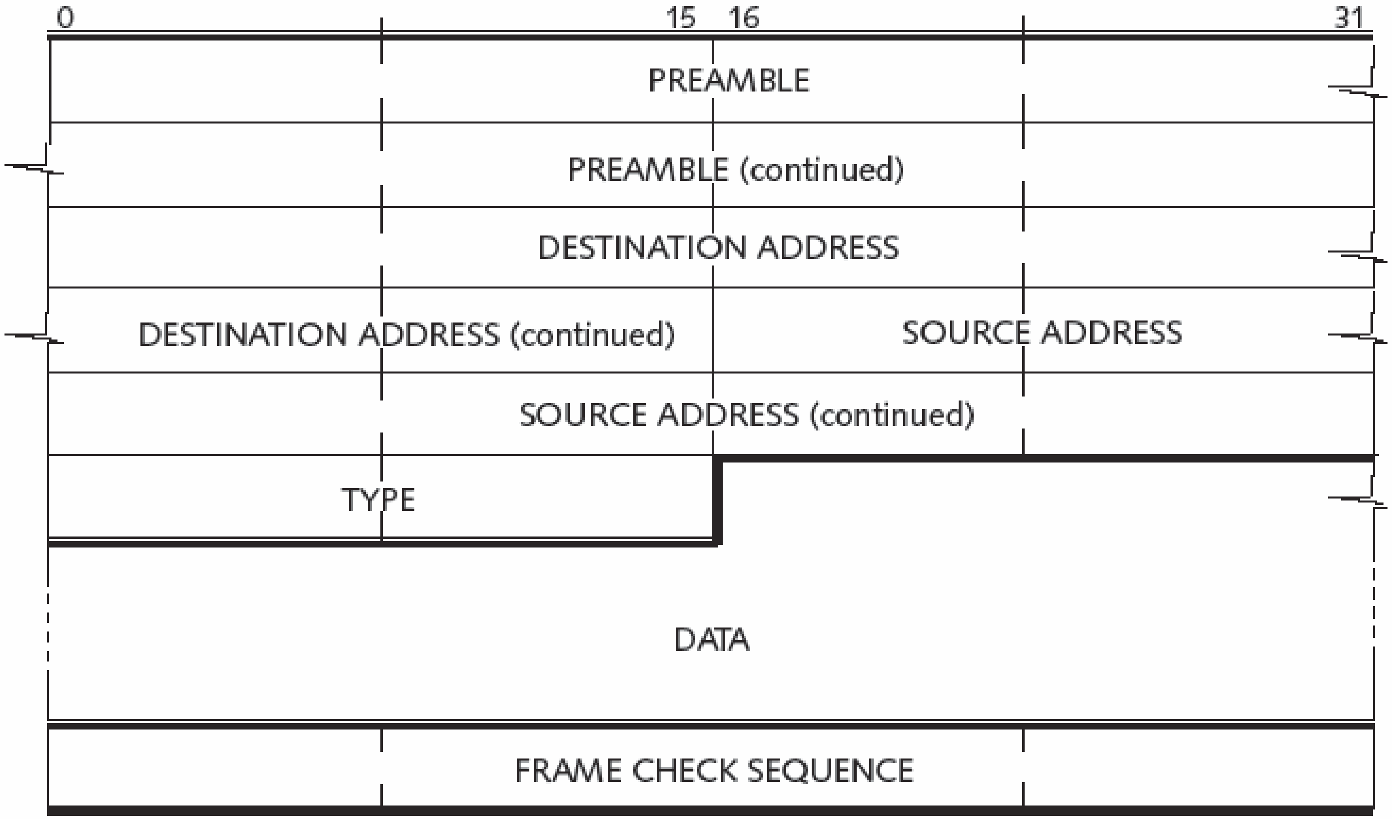


Figure 4-1 Format of an Ethernet II frame

Ethernet II Frame Structure (cont'd.)

Type	Protocol
0x0800	IPv4
0x86dd	IPv6
0x0806	Address Resolution Protocol
0x809B	AppleTalk
0x8137	Novell Internetwork Packet Exchange (IPX)

Table 4-1 Assigned protocol types (by number)

Ethernet 802.2 LLC Frame Structure

- Ethernet 802.2 LLC frame structure
  – Preamble
  – Start Frame Delimiter (SFD)
  – Destination Address Field
  – Source Address Field
  – Length
  – Destination Service Access Point (DSAP)
  – Source Service Access Point (SSAP)
  – Control Field
  – Data Field
  – Frame Check Sequence Field

Ethernet 802.2 LLC Frame Structure(cont'd.)

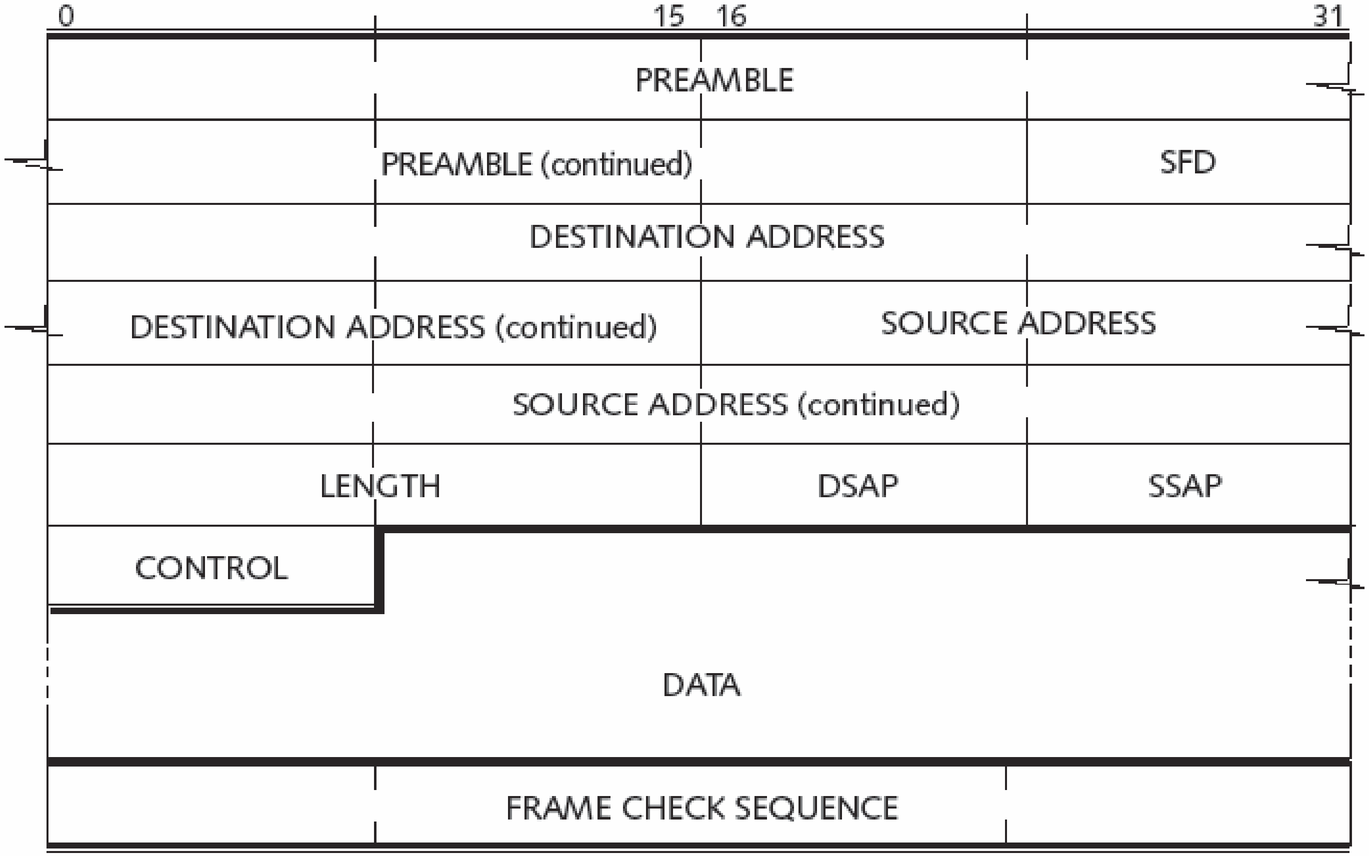


Figure 4-2 Format of an Ethernet 802.2 LLC frame

Hardware Addresses in the IP Environment

- IP addresses
  - Identify individual IP hosts on a TCP/IP internetwork

Address Resolution Protocol and Network Discovery Protocol

- Address Resolution Protocol (ARP)
  - Resolves Network layer (IP) addresses into Data Link layer (physical) addresses
- IPv6 uses Neighbor Discovery Protocol (NDP)
  - Neighbor Solicitation
  - Neighbor Advertisement

ARP Protocol Characteristics and Handling

- ARP cache
  - Table of hardware addresses learned through the ARP process
  - IP host first check ARP cache, if can't find hard address, broadcast the ARP request

ARP Protocol Characteristics and Handling (cont'd.)

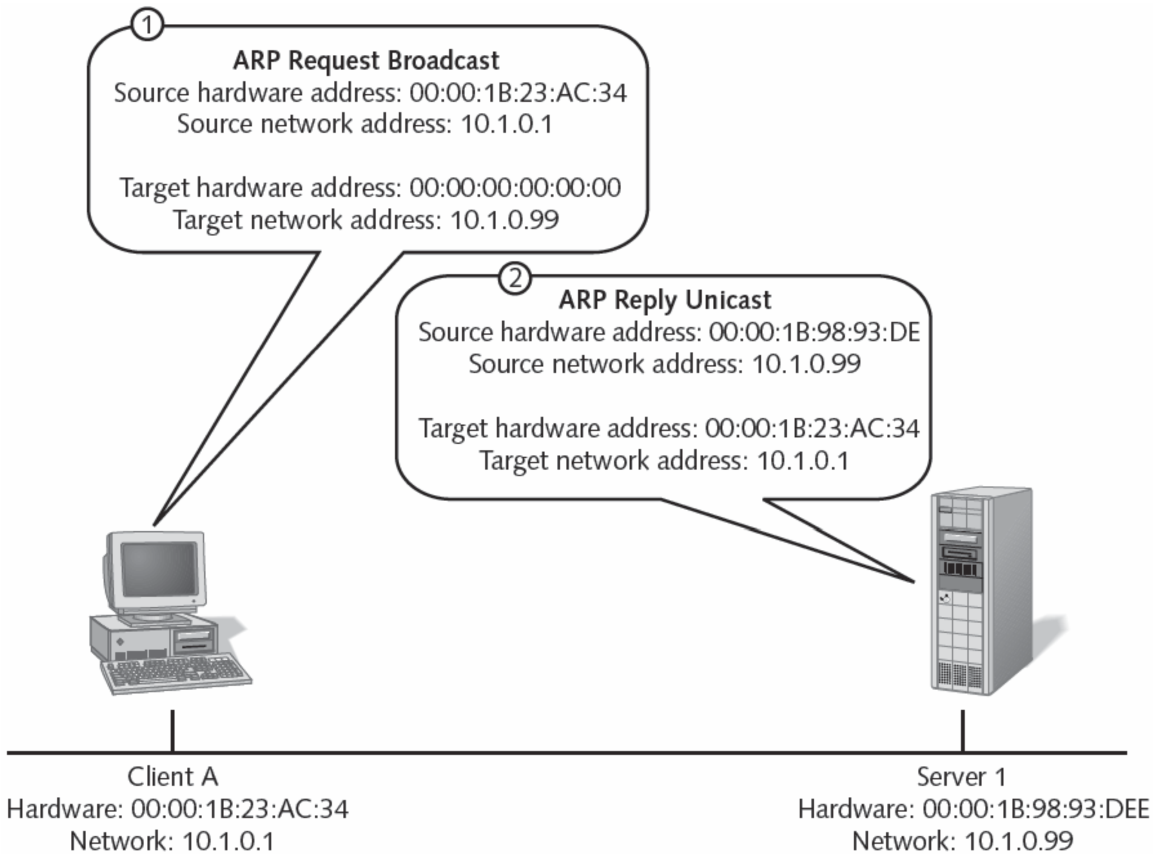
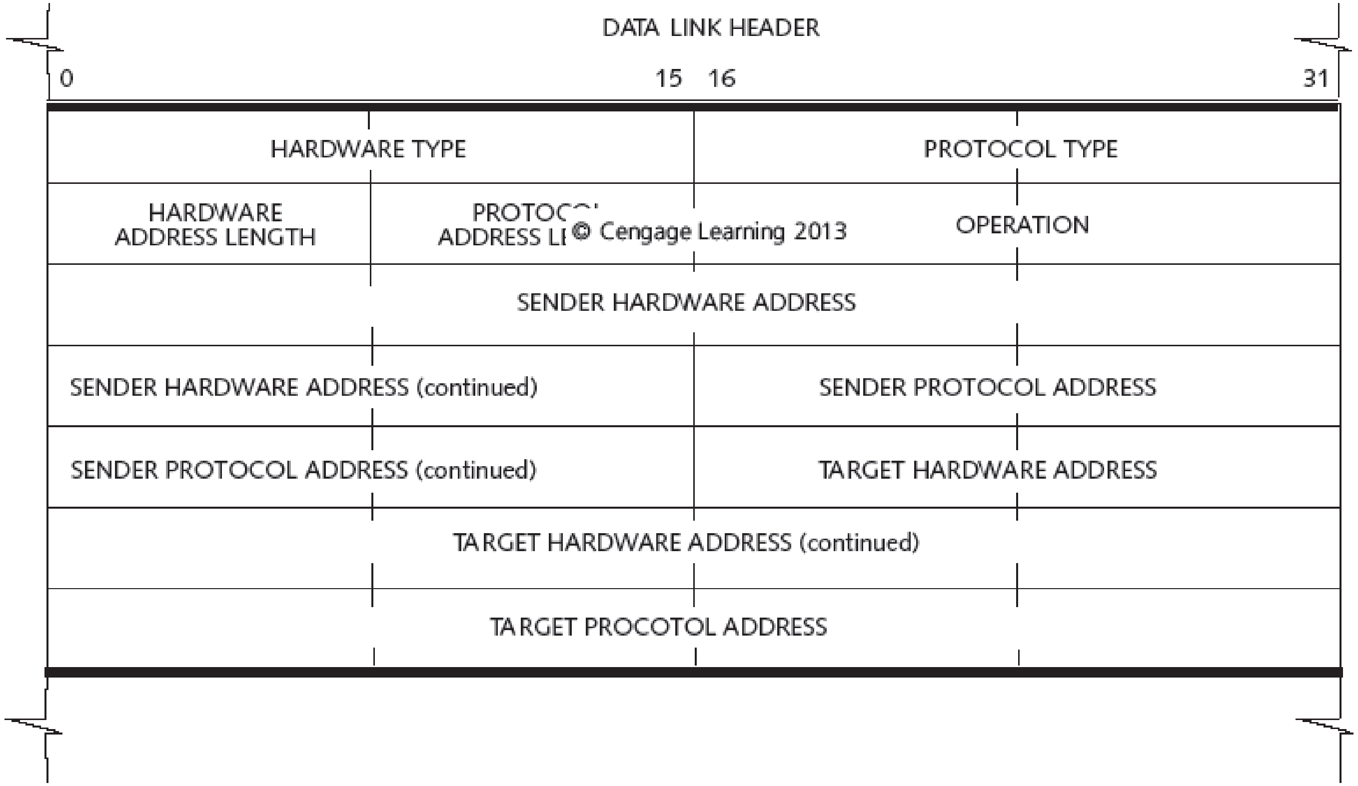


Figure 4-3 ARP broadcasts identifying the source and desired IP address

ARP Protocol Characteristics and Handling (cont'd.)

- The ARP frame structure (Figure 4-4)



ARP Packet Fields and Functions

- Basic ARP packets
  - Broadcast ARP request packet
  - Directed, or unicast, ARP reply packet
- Most confusing part of ARP
  - Interpretation of the sender and target address information
  - Target Internet Address and Target Hardware Address are all set 0

ARP Packet Fields and Functions (cont’d.)

- ARP packet fields:
  - Hardware Type
  - Protocol Type
  - Length of Hardware Address
  - Length of Protocol Address
  - Opcode: 1-ARP request, 2-ARP reply, 3-RARP request, 2-RARP reply
  - Sender’s Hardware Address
  - Sender’s Protocol Address
  - Target Hardware Address
  - Target Protocol Address

ARP Cache

- ARP information
  - Kept in an ARP cache in memory on most operating systems
- Windows-based systems
  - Command *arp -a* is used to view the table contents
  - Have utility to view IP and hardware addresses

Proxy ARP

- Method that allows IP host to use a simplified subnetting design
- Enables a router to “ARP” in response to an IP host’s ARP broadcasts
- Most network configurations
  - May never need to use proxy ARP

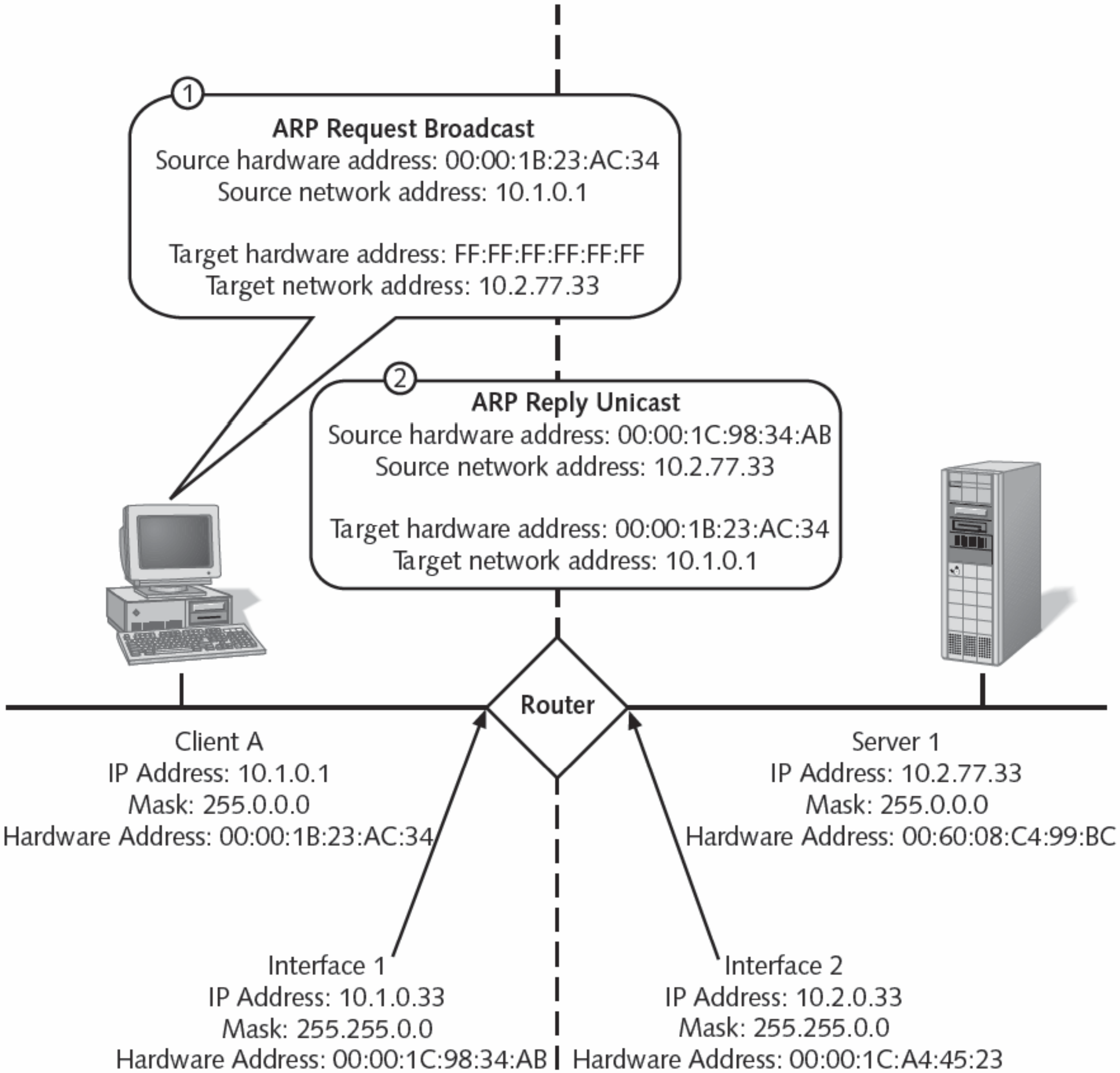


Figure 4-10 ARP proxy network design

Reverse ARP

- Used to obtain an IP address for an associated data link address
- Initially defined to enable diskless workstations to find their own IP addresses upon booting or startup
- BOOTP, and eventually DHCP, replaced RARP

NDP Protocol Characteristics and Handling

- Specified by RFC 4861
- Describes how nodes on the same network link
  - Determine the presence of other nodes
  - Discover the link-layer addresses of another node
  - Find routers
  - Discover network paths to network neighbors

NDP Protocol Characteristics and Handling (cont'd.)

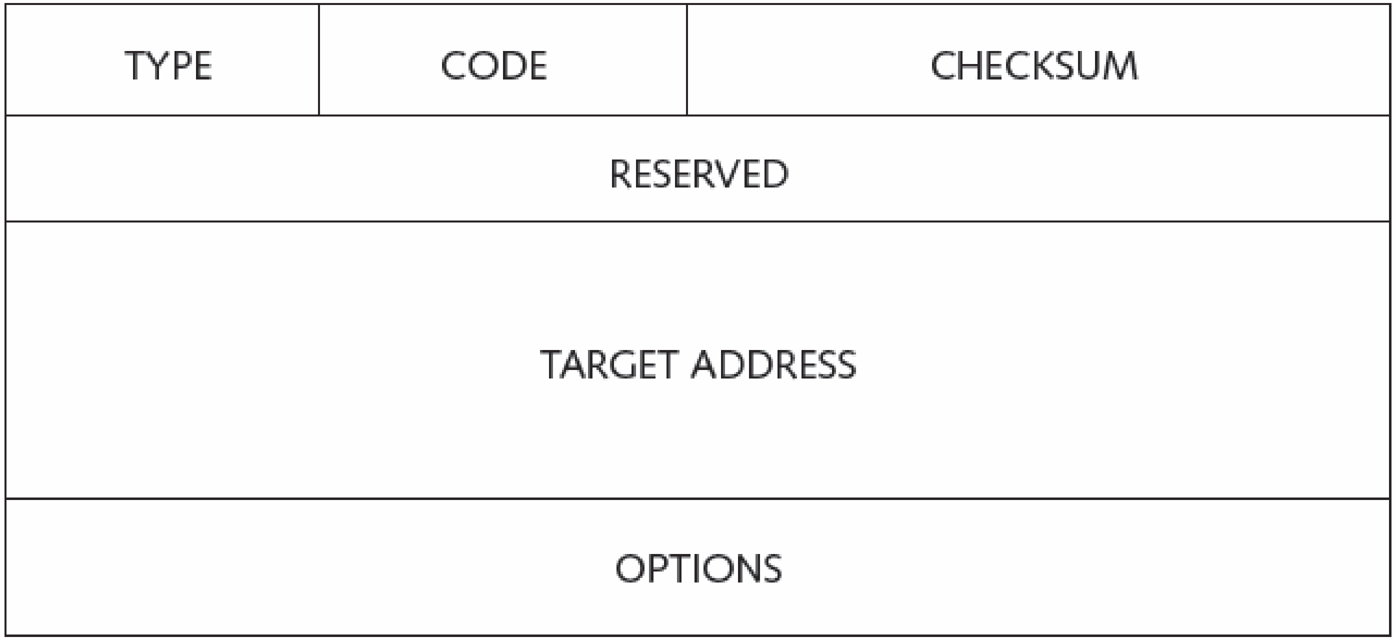


Figure 4-11 Neighbor Solicitation message format

Understanding the Internet Protocol

- Network layer communications
  - End-to-end communications
- Internet Protocol
  - Network layer protocol used in the TCP/IP suite
- IP version 4 (IPv4)
  - Widely implemented
- Internet Protocol version 6 (IPv6)
  - Most used in pilot or experimental implementations

Sending IP Datagrams

- Requirements for building an IP datagram packet to transmit on the wire
  - IP addresses of the source and destination
  - Hardware address of the source and next-hop router
- IP host
  - Can use a manually entered destination IP address or the DNS to obtain a destination's IP address

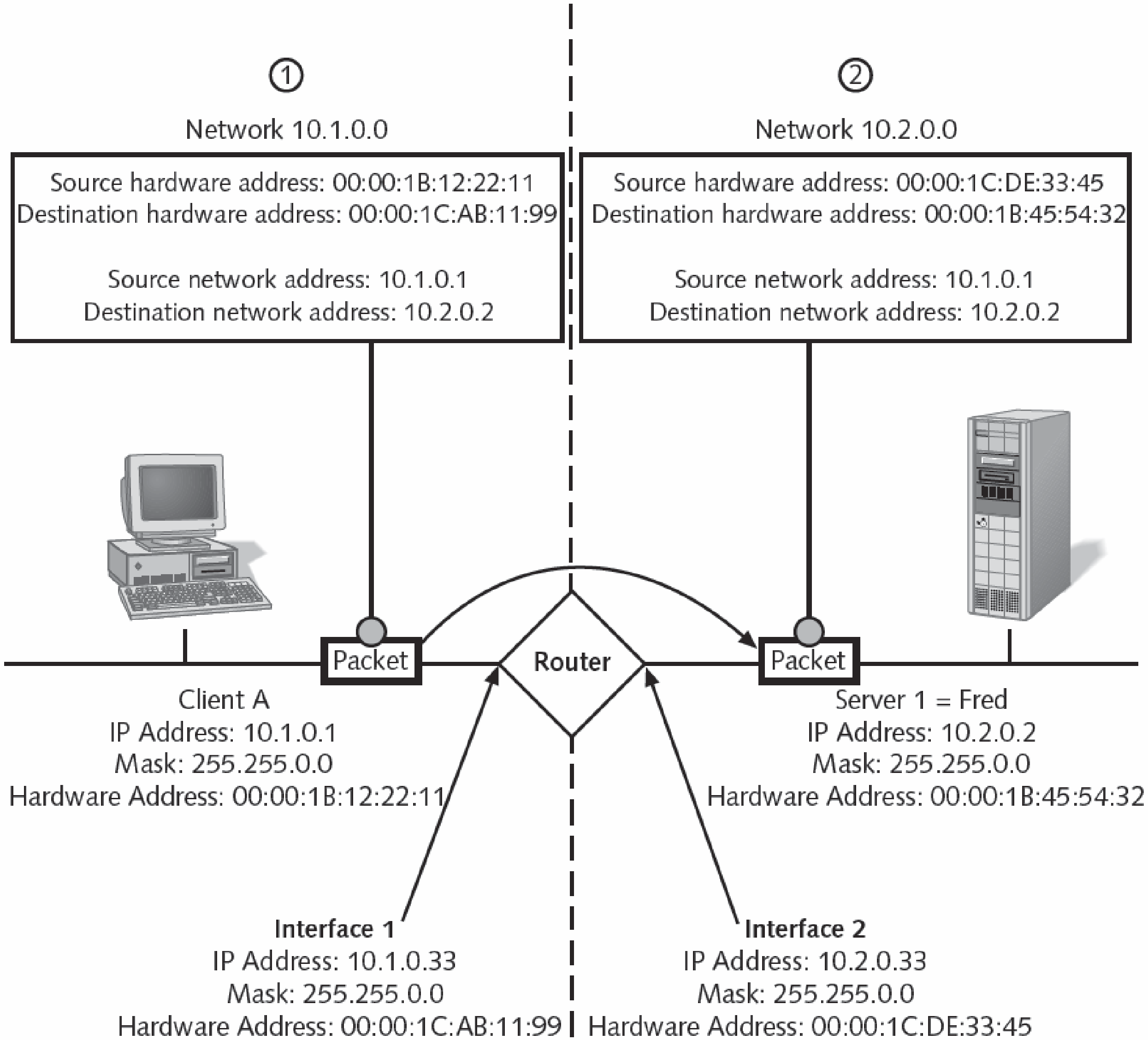


Figure 4-12 Data link header stripped off and reapplied by IP routers when packet is forwarded

The Route Resolution Processes

- Enables IP host to determine if desired destination is local or remote
- Local or Remote Destination?
  - Upon determination of IP address
    - IP host compares network portion of destination address to its own local network address

If Remote, Which Router?

- Types of route table entries
  - Host route entry
  - Network route entry
- Receiving gateway typically does one of the following
  - Forwards packet
  - Sends an ICMP reply
  - Sends an ICMP reply indicating that it is unclear where to send the packet

How IPv4 and IPv6 Differ

- IPv4 uses ARP, while IPv6 uses NDP
- IPv6 packets cannot be fragmented
- For IPv4, the source node must determine if destination node is local or if it must be routed
  - For IPv6, the source node uses PMTU Discovery
- IPv6 uses multicasts rather than broadcasts
- Both need to know DNS servers for name resolution

Lifetime of an IP Datagram

- IP packets
  - Have a pre-defined lifetime indicated in each packet's *Time to Live (TTL)* or *Hop Limit* (IPv6) field
- 64
  - Recommended starting TTL value
- 128
  - Default TTL in Windows 2000, Windows 2003, and Windows XP

Fragmentation and Reassembly

- IP fragmentation
  - Enables a larger packet to be automatically fragmented by a router
- Once fragmented
  - No reassembly occurs until fragments arrive at destination
  - All fragments are given the same TTL value

Service Delivery Options

- IP header support two fields for defining packet priority and route priority
  - Precedence
  - Type of Service
- Recent methods suggested to improve IP-based traffic flows
  - Differentiated Services (Diffserv)
  - Explicit Congestion Notification (ECN)

Precedence

- Precedence
  - Used by routers to determine what packet to send
- There are eight levels of precedence

Binary	Decimal	Meaning
111	7	Network control
110	6	Internetwork control
101	5	CRITIC/ECP (Critical and Emergency Call Processing)
100	4	Flash override
011	3	Flash
010	2	Immediate
001	1	Priority
000	0	Routine

Type of Service

- Type of Service
  - Used to select routing path when *multiple paths* exist
- Routing protocols
  - OSPF
  - *Border Gateway Protocol (BGP)*

Binary	Meaning	Binary	Meaning
0000	Default	0100	Maximum throughput
0001	Minimum cost	1000	Low delay
0010	Maximum reliability	1111	Maximum security

Type of Service (cont'd.)

- RFC 2474, RFC 2475, and RFC 3168
  - Offer a new use of the TOS field bits
  - Suggest that TOS and Precedence field bytes be replaced by a Differentiated Services Code Point (DSCP) field
- Diffserv
  - Uses DSCP value to enable routers to offer varying levels of service to traffic based on marker placed in the DSCP field

Type of Service (cont'd.)

Priority	Class 1	Class 2	Class 3	Class 4	Classless
Low discard priority	DSCP 10	DSCP 18	DSCP 26	DSCP 34	
	AF11	AF21	AF31	AF41	
	001010	010010	001010	100010	
Medium discard priority	DSCP 12	DSCP 20	DSCP 28	DSCP 36	
	AF12	AF22	AF32	AF42	
	00100	010100	011100	100100	
High discard priority	DSCP 14	DSCP 22	DSCP 30	DSCP 38	
	AF13	AF23	AF33	AF43	
	001110	010110	011110	100110	
Expedited forwarding					DSCP 46
					101110

AF: Assured forwarding

Understanding IP Routing

- Routing table
  - Database that lives in the memory of the router
    - Network address
    - Next hop
    - Interface
    - Metrics
    - Vendor information
    - etc
  - Compilation of information about all the networks that the router can reach

Understanding IP Routing (cont'd.)

```
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, * - candidate default
       U - per-user static route, o - ODR

Gateway of last resort is 137.20.25.2 to network 0.0.0.0

       172.168.0.0/24 is subnetted, 2 subnets
B         172.168.70.0 [20/170] via 137.20.10.70, 1:11:35
O E2      172.168.80.0 [110/20] via 137.20.25.2, 01:11:26, Serial0
       200.200.100.0/32 is subnetted, 1 subnets
O IA      200.200.100.1 [100/115] via 137.20.25.2, 01:12:06, Serial0
       137.20.0.0/16 is variably subnetted, 16 subnets, 4 masks
O E1      137.20.200.16/28 [110/164] via 137.20.25.2, 01:12:06, Serial0
O IA      137.20.240.1/32 [110/65] via 137.20.25.2, 01:12:06, Serial0
O IA      137.20.224.0/20 [110/1626] via 137.20.25.2, 01:12:06, Serial0
O IA      137.20.30.0/24 [110/120] via 137.20.25.2, 01:11:51, Serial0
C         137.20.25.0/24 is directly connected, Serial0
C         137.20.10.0/24 is directly connected, Ethernet0
O IA      137.20.60.1/32 [110/75] via 137.20.25.2, 01:12:07, Serial0
O E1      137.20.40.16/28 [110/164] via 137.20.25.2, 01:11:30, Serial0
O IA      137.20.33.0/24 [110/115] via 137.20.25.2, 01:11:33, Serial0
O E2      137.20.86.0/24 [110/20] via 137.20.25.2, 01:12:08, Serial0
B         137.20.81.0/24 [200/0] via 137.20.86.1, 01:11:59
B         137.20.82.0/24 [200/0] via 137.20.86.1, 01:11:59
O IA      137.20.100.33/32 [110/114] via 137.20.25.2, 01:12:08, Serial0
O IA      137.20.100.35/32 [110/114] via 137.20.25.2, 01:11:53, Serial0
O IA      137.20.64.0/20 [110/74] via 137.20.25.2, 01:12:09, Serial0
       170.10.0.0/24 is subnetted, 1 subnets
B         170.10.10.0 [20/170] via 137.20.10.70, 01:11:37
       160.10.0.0/24 is subnetted, 1 subnets
B         160.10.10.0 [20/170] via 137.20.10.70, 01:11:37
       161.10.0.0/24 is subnetted, 1 subnets
B         161.10.10.0 [20/170] via 137.20.10.70, 01:11:37
C         200.200.200.0/24 is directly connected, Loopback0
O*E1   0.0.0.0/0 [110/167] via 137.20.25.2, 01:12:09, Serial0
B         160.0.0.0/4 [20/170] via 137.20.10.70, 01:11:37
```

Tenth entry

Figure 4-18 Routing table from Cisco router

Understanding IP Routing (cont’d.)

```
C:\Documents and Settings\tony.guo>route print
===========================================================================
Interface List
0x1 ........................... MS TCP Loopback interface
0x4 ...00 1e 8c 2b 3b e1 ...... Realtek RTL8139/810x Family Fast Ethernet NIC -
Virtual Machine Network Services Driver
===========================================================================
===========================================================================
Active Routes:
Network Destination        Netmask          Gateway       Interface  Metric
          0.0.0.0          0.0.0.0     192.168.99.1    192.168.99.8       1
        127.0.0.0        255.0.0.0        127.0.0.1       127.0.0.1       1
     192.168.99.0    255.255.255.0     192.168.99.8    192.168.99.8       20
     192.168.99.8  255.255.255.255        127.0.0.1       127.0.0.1       20
   192.168.99.255  255.255.255.255     192.168.99.8    192.168.99.8       20
        224.0.0.0        240.0.0.0     192.168.99.8    192.168.99.8       20
  255.255.255.255  255.255.255.255     192.168.99.8    192.168.99.8       1
Default Gateway:      192.168.99.1
===========================================================================
Persistent Routes:
  None

C:\Documents and Settings\tony.guo>
```

How Entries Are Placed in the Routing Table

- Route entry can be placed in a routing table in three basic ways
  - Through direct connection with netcard
  - Can be manually configured
  - An entry can be placed in a routing table dynamically, by using a *routing protocol*

Routing Protocols and Routed Protocols

- Rout*ing* protocols
  - Used to exchange routing information
  - Routing Information Protocol (RIP)
  - Open Shortest Path First (OSPF)
- Rout*ed* protocols
  - No.3 Layer protocols that are used to get packets through an internetwork
  - IP is the routed protocol for TCP/IP protocol stack

Grouping Routing Protocols

- Group management
  - Interior gateway protocols (IGPs)
    - Routing protocols used inside a routing domain are called interior
  - Exterior gateway protocols (EGPs)
    - Routing protocols used to connect these routing domains
- Communication method
  - Distance vector
  - Link-State

Distance Vector Routing Protocols

- Route Information Protocol (RIP)
- Interior Gateway Routing Protocol (IGRP)
- Border Gateway Protocol (BGP) is also a distance vector routing protocol

Distance Vector Routing Protocols (cont'd.)

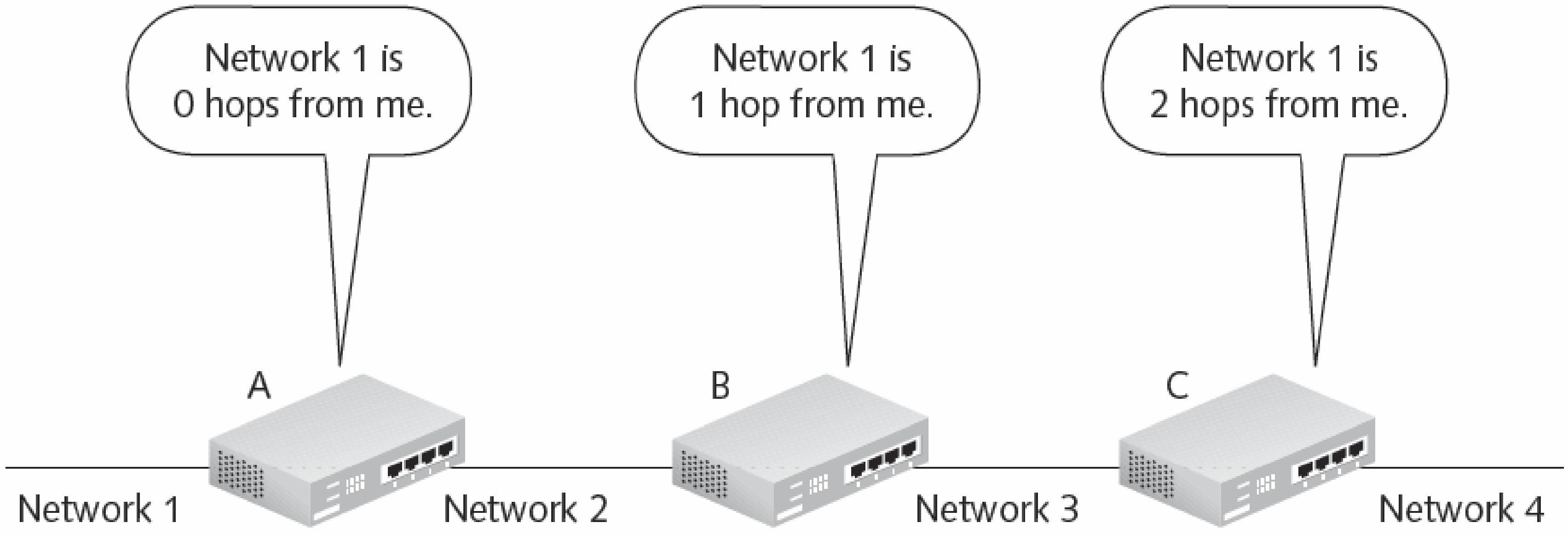
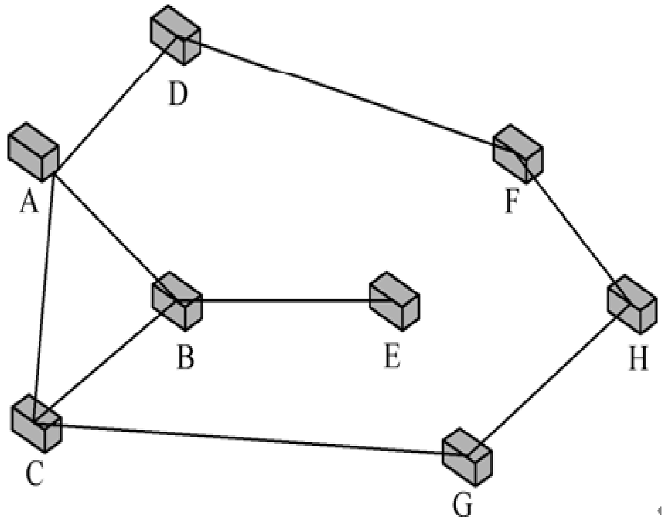


Figure 4-19 Routers on distance vector networks tracking distances

Distance Vector Routing Protocols (cont'd.)

信宿 信源	A	B	C	D	E	F	G	H
A	0	1	1	1	∞	∞	∞	∞
B	1	0	1	∞	1	∞	∞	∞
C	1	1	0	∞	∞	∞	1	∞
D	1	∞	∞	0	∞	1	∞	∞
E	∞	1	∞	∞	0	∞	∞	∞
F	∞	∞	∞	∞	∞	0	1	1
G	∞	∞	1	∞	∞	∞	0	1
H	∞	∞	∞	∞	∞	1	1	0

（a）距离向量初始表



信宿	度量值	下一跳
B	1	B
C	1	C
D	1	D
E	2	B
F	2	D
G	2	C
H	3	C

（b）节点 A 的最终路由表

Routing Loops

- Routing loops
  - One router believes the best path to a network is through a second router
  - At the same time, second router believes the best path to that network is through the first router
  - If C is down, Figure 4-19
- Loop-avoidance schemes
  - Counting to infinity, usually 16
  - Split horizon
    - Often coupled with poison reverse

Link-State Routing Protocols

- Differ from distance vector routing protocols in two ways
  - They do not route by rumor
  - They do not periodically broadcast their entire tables
- Hello process to Neighbor router
- Link- State Advertisement (LSA)
  - Neighbor list
  - Cost to neighbor
- Each router build one network general picture

Link-State Routing Protocols (cont'd.)

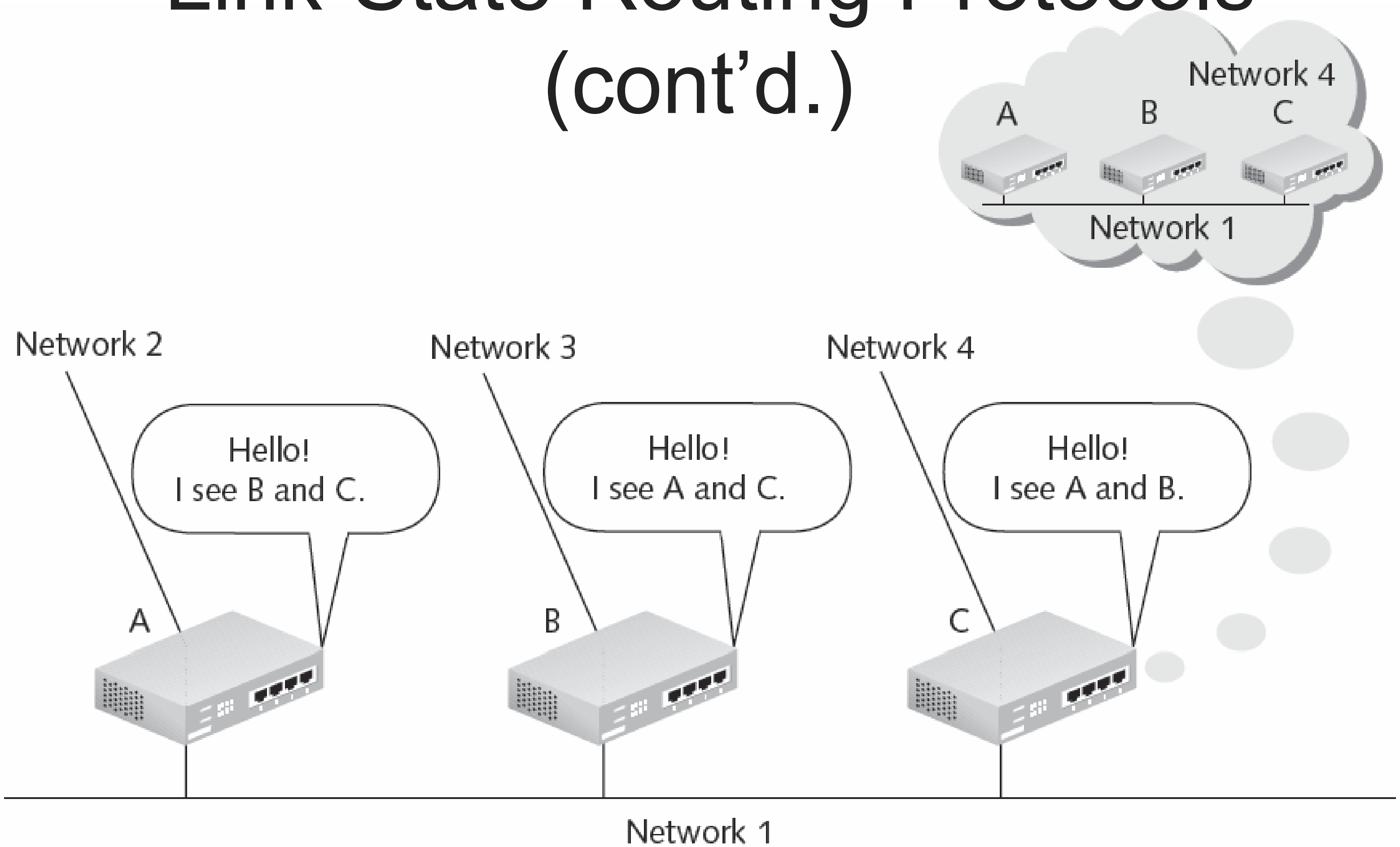
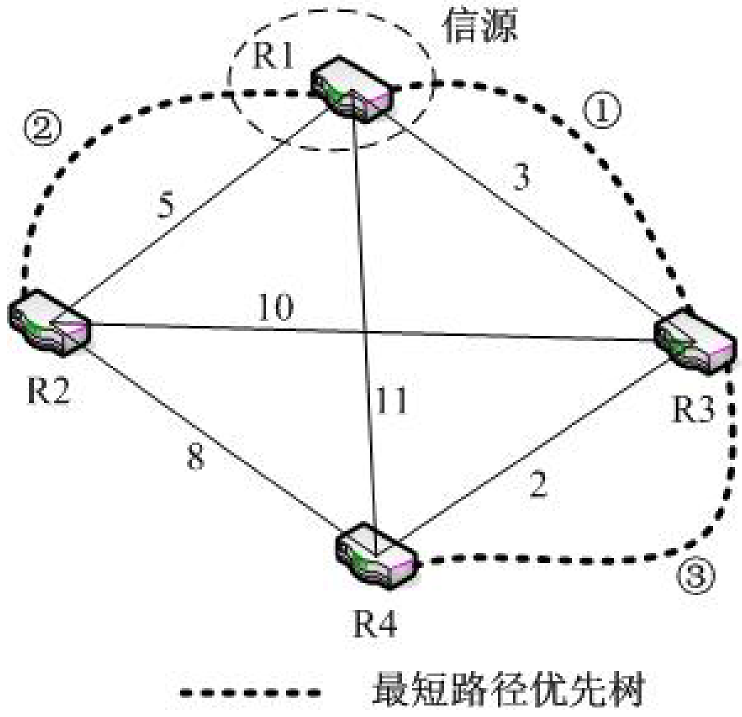
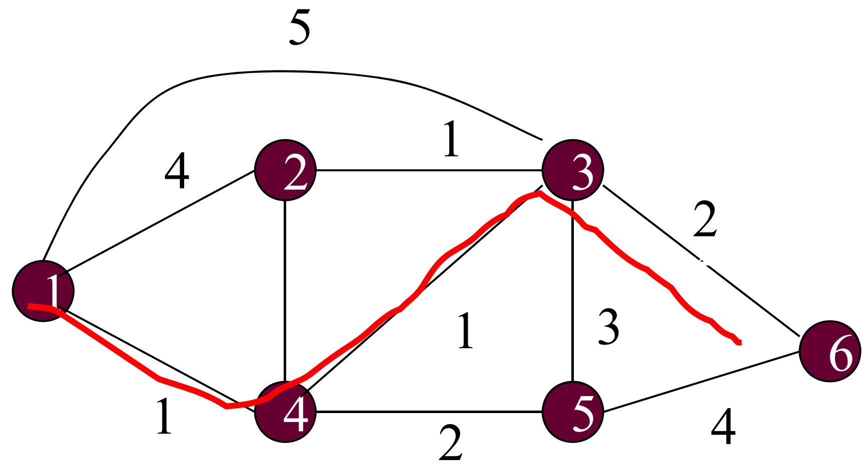


Figure 4-20 Adjacent databases on a link-state network

Link-State Routing Protocols (cont'd.)





Routing Characteristics

- Important to understand
  - Characteristics of the networks involved and how they are interconnected
  - Requirements and limitations of the various routing protocols that may be in use on an internetwork

Route Convergence

- Network
  - Converged when all the routers know a loop-free path to get to all other networks
  - Ideally should be in a converged state

IPv4 Routing Mechanisms

- Help routers learn about paths to destination
- Split horizon
  - Used to speed up the process of convergence and resolve the counting-to-infinity problem
  - When a node sends the update information of the routing table to an adjacent node, it does not return the routing information learned from the adjacent nodes to the node
- Poison reverse
  - Technique for assigning costs to routes designed to prevent routing loops

IPv4 Routing Mechanisms (cont'd.)

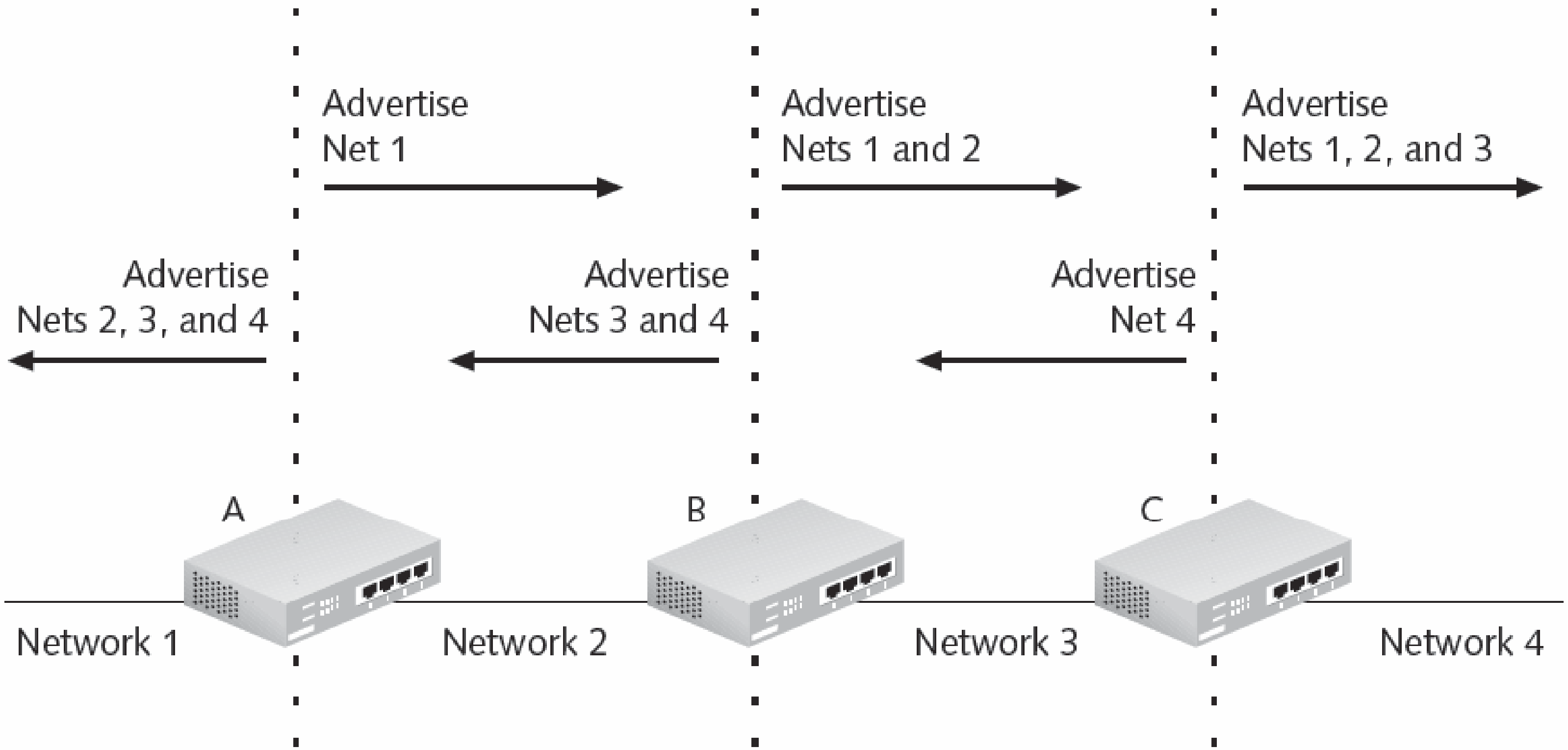


Figure 4-21 The rule of split horizon in operation

IPv4 Routing Mechanisms (cont'd.)

- Time to Live
  - Ensures that packets cannot loop endlessly through a network
  - Defined in the Network layer header
- Multicast Versus Broadcast Update Behavior
  - Routers can be configured to forward multicasts
  - RIPv1 sends broadcast updates
  - RIPv2 can send multicast updates

Areas, Autonomous Systems, and Border Routers

- OSPF utilizes *areas*
  - To reduce the number of entries in the link-state database
- OSPF specification
  - Defines the need for a backbone area, Area 0
- Autonomous systems (ASs)
  - Groups of routers under a single administrative authority

Areas, Autonomous Systems, and Border Routers (cont'd.)

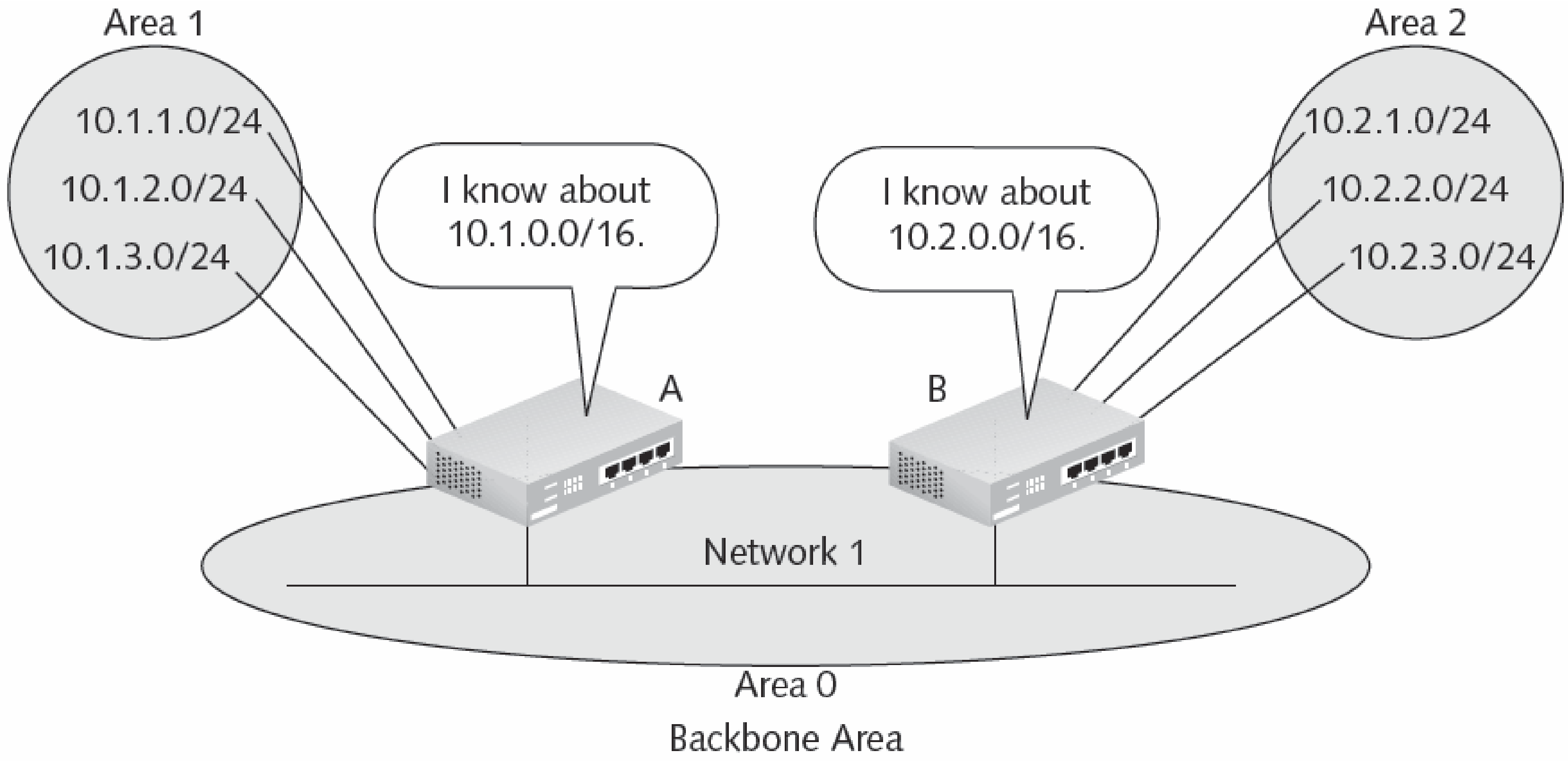


Figure 4-23 Areas connected to backbone area 0

IPv6 Routing Considerations

- IPv6
  - Designed with routing efficiency and throughput in mind
  - Designed to reduce the workload of Internet routers
- Many of the same routing approaches familiar to IPv4
  - Can make the transition to IPv6

IPv6 Routing Mechanisms

- Routing mechanisms
  - Processes by which a router learns about network paths that lead to destinations
- Network nodes must use PMTU Discovery to determine an arbitrary route to its destination
- Dynamic routing protocols in IPv6 include RIPng, EIGRPv6, OSPFv3, IS-IS for IPv6, and BGP for IPv6

IPv6 Routing Mechanisms (cont'd.)

- IPv6 Routing Table Entry Types
  - Nodes and routers maintain routing table
- IPv6 Route Determination Process
  - Device should select one entry as its forwarding decision
  - Router compares the bits in the address prefix for each entry in its routing table
  - Router creates a list of all matching routes
  - Router chooses route with the largest prefix length

IPv6 Routing Mechanisms (cont'd.)

Routing Table Type	Details
Directly attached routes	These are routes with subnet prefixes that are directly attached to the router and that usually have a 64-bit length.
Remote routes	These are routes with prefixes that are either subnet prefixes 64 bits in length or prefixes summarizing an address space that are less than 64 bits in length.
Host routes	This is a route prefix to a specific IPv6 address that is 128 bits in length.
Default route	The IPv6 default route prefix is ::/0.

Table 4-12 **IPv6 routing table types**

IPv6 Routing Mechanisms (cont'd.)

Routing Table Fields	Details
Destination prefix	This is an IPv6 address prefix that can be between 0 and 128 in length.
Next-hop address	This is the IPv6 address to which the packet will next be forwarded.
Interface	This is the network interface on the node used to forward the packet to the next-hop node.
Metric	This is the value indicating the cost of the route allowing the best route, among multiple routes, to be selected.

Table 4-13 **IPv6 routing table fields**

IPv6 Routing Mechanisms (cont'd.)

- How Strong and Weak Hosts Behave
  - RFC 1122 original specification
  - Router has multiple network cards
  - If one interface is not set as source or destination IP address, it can also send or receive data packet, we call it one weak host
  - Network card with specific IP can only send and receive the data packet from the hosts which consider the IP as source or destination address, we call it strong host

IPv6 Routing Mechanisms (cont'd.)

- The IPv6 Delivery Process, End to End
  - If destination is on same local link,source node sends a packet to either a router or to the final destination
    - If sent to a router, the packet is forwarded to either another router or to its final destination node (if node is on a link directly attached to first-hop router)
  - When received by the destination node
    - Packet transfers its data to the desired application on the computer

Multicast Listener Discovery in IPv6

- Understanding MLD and MLDv2
  - MLD enables each IPv6 router to discover nodes that want to receive multicast messages
  - MLDv2 updates the protocol to support source filtering
- IPv6 Multicast Behaviors
  - Multicast is a method of sending a “one-to-many” message in real time
    - Usually using ICMPv6 as the transport mechanism

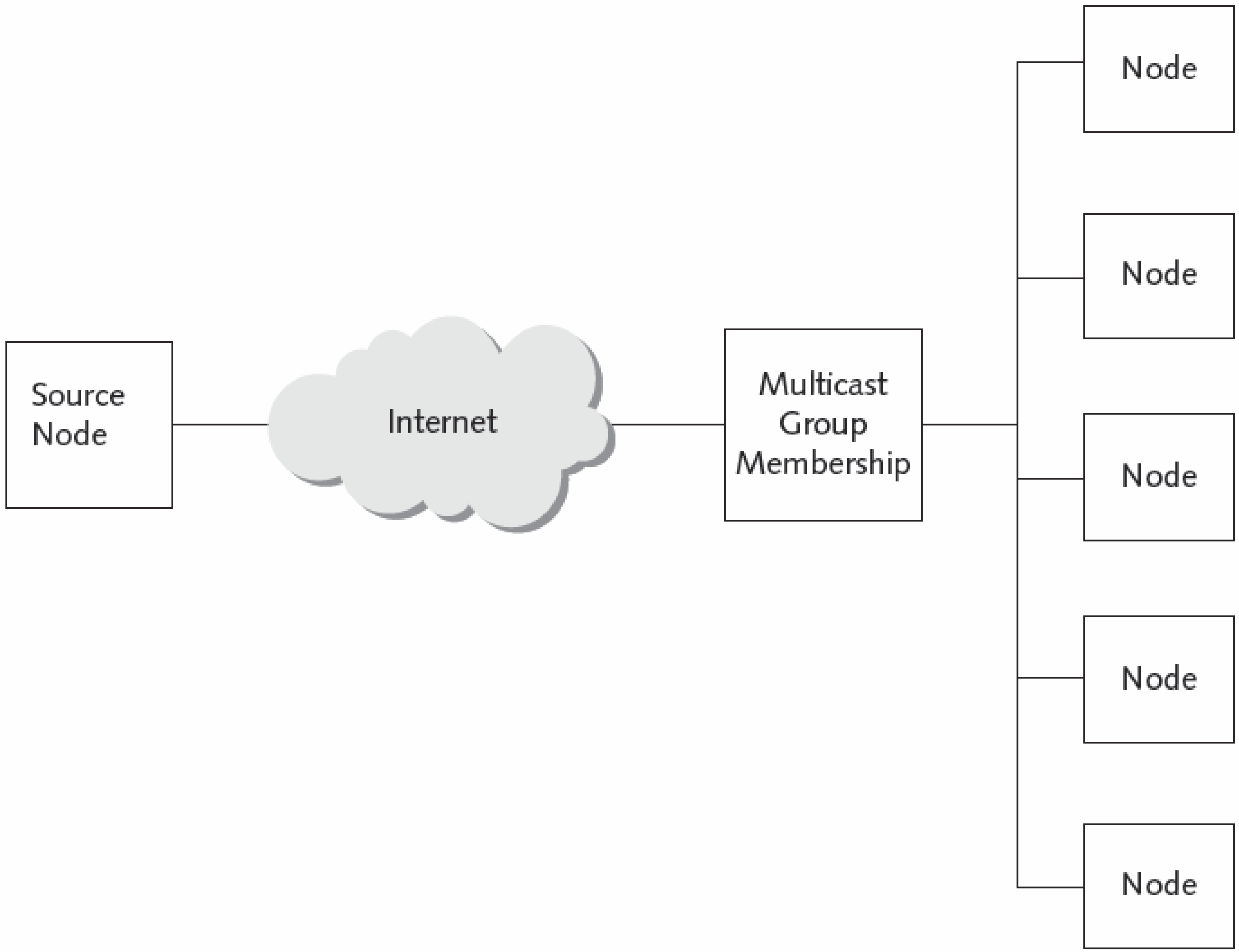


Figure 4-25 Logical diagram of a multicast message being sent to a multicast group

MLD and MLDv2 Packet Structures and Messages

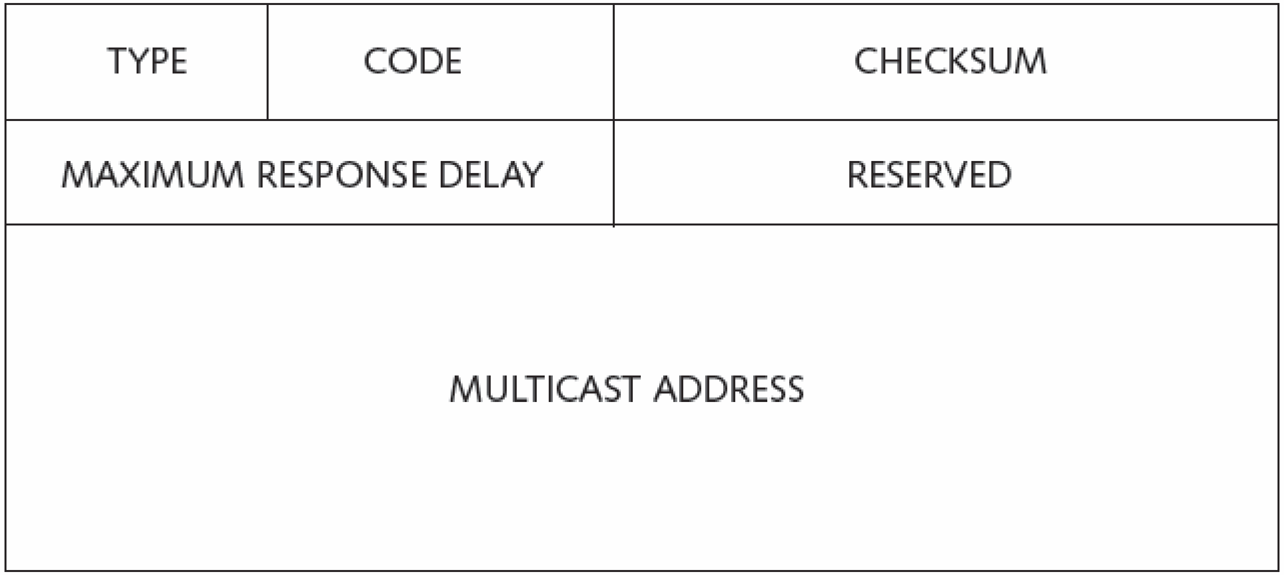


Figure 4-26 MLDv1 packet structure

MLD and MLDv2 Packet Structures and Messages (cont'd.)

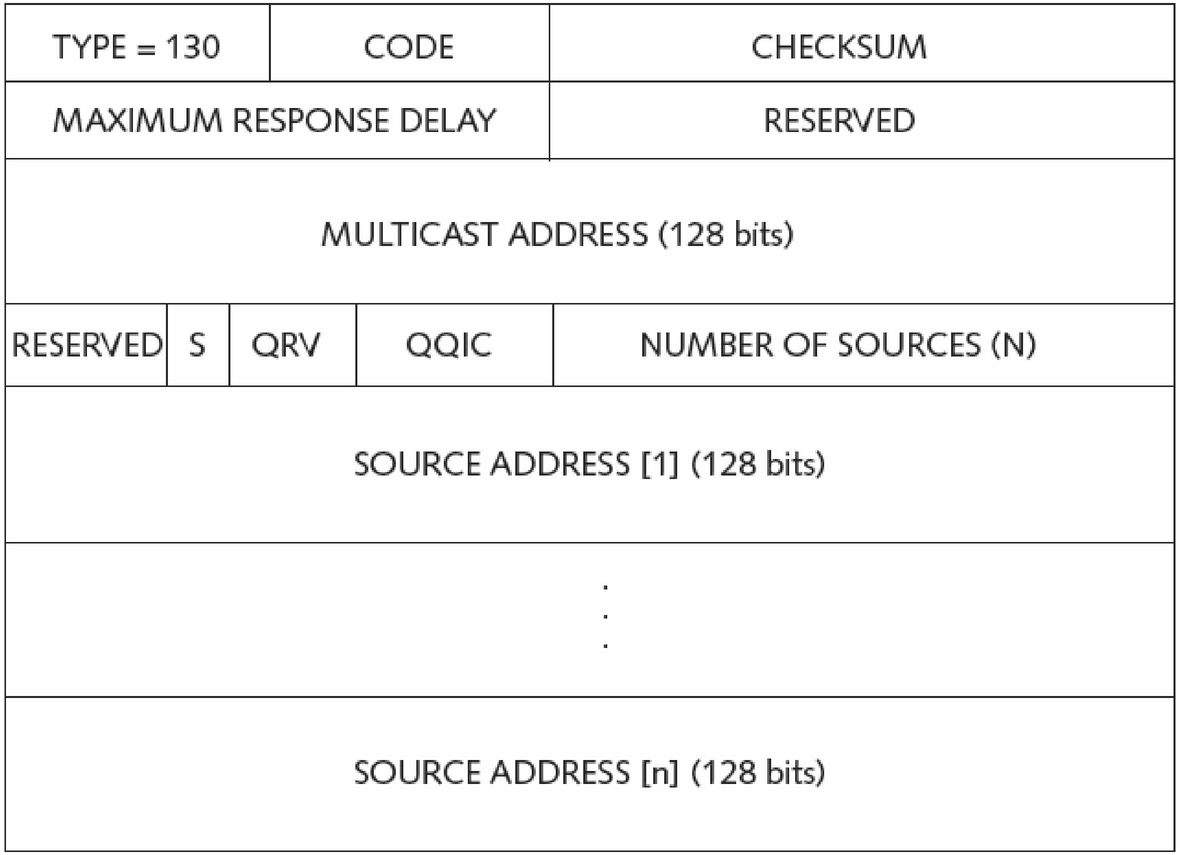


Figure 4-27 MLDv2 packet structure for Multicast Listener Query message

Routing Protocols

- Routing protocols
  - Define how routers communicate with each other
    - For the purpose of sharing and updating routing information
  - Implementation of routing algorithm that uses metrics to discover paths
- Information about each path to destinations is stored in a routing table
  - Table can be updated manually or dynamically

IPv4

- RIP
  - Basic distance vector routing protocol
  - Two versions: RIPv1 and RIPv2
  - Communications are UDP based
  - RIP-based routers send and receive datagrams on UDP port number 520

Open Shortest Path First

- Defined in RFC 2328
- The premier link-state routing protocol used on TCP/IP networks
- Based on
  - Configurable values (metrics) that may be based on network bandwidth, delay, or monetary cost

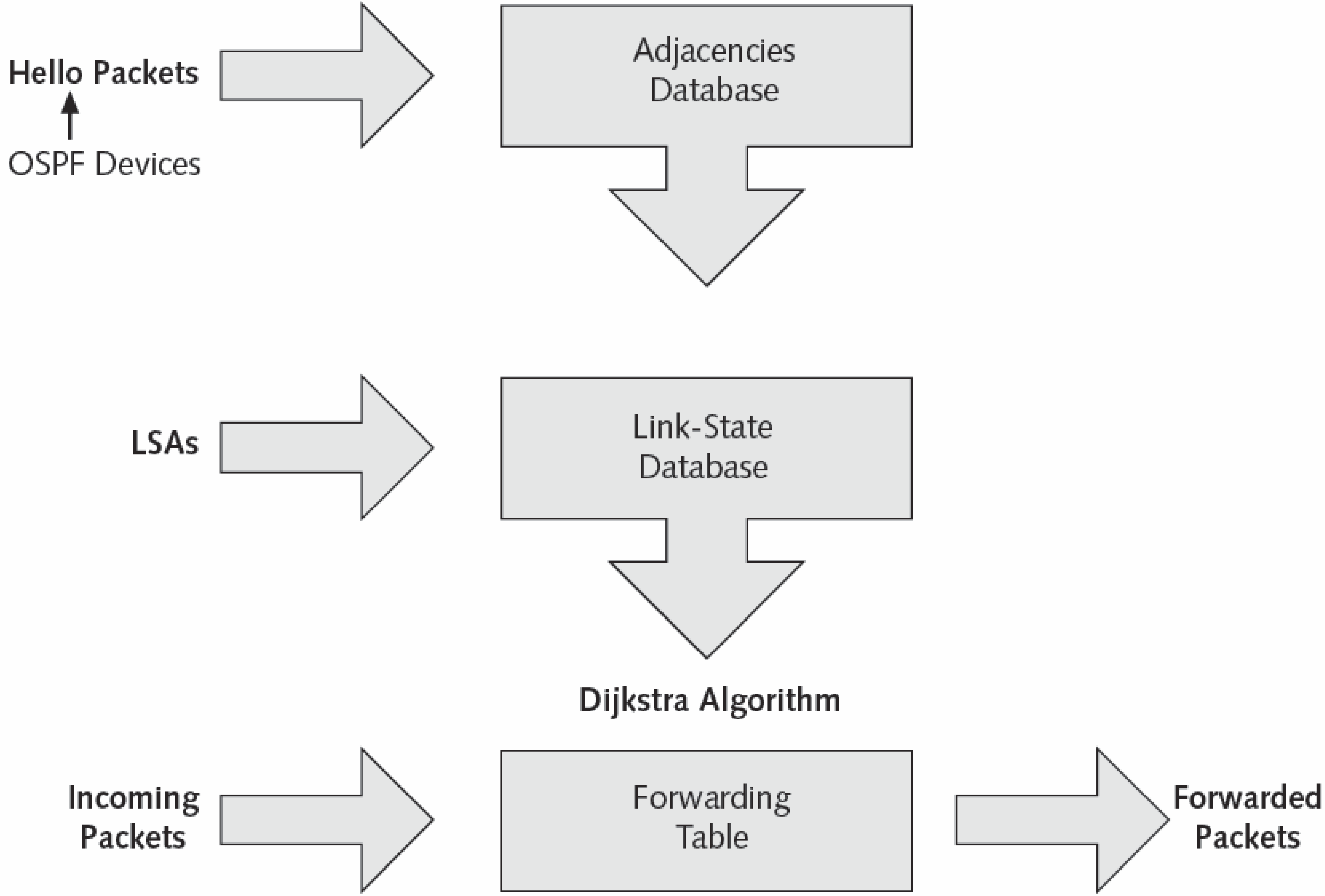


Figure 4-33 OSPF architecture

Open Shortest Path First (cont’d.)

- Six basic types of LSAs:
  - Type 1 (Router Links Advertisement)
  - Type 2 (Network Links Advertisement)
  - Type 3 (Network Summary Link Advertisement)
  - Type 4 (AS Boundary Router Summary Link Advertisement)
  - Type 5 (AS External Link Advertisement)
  - Type 7 (Not So Stubby Area Networks Advertisement)

Open Shortest Path First (cont'd.)

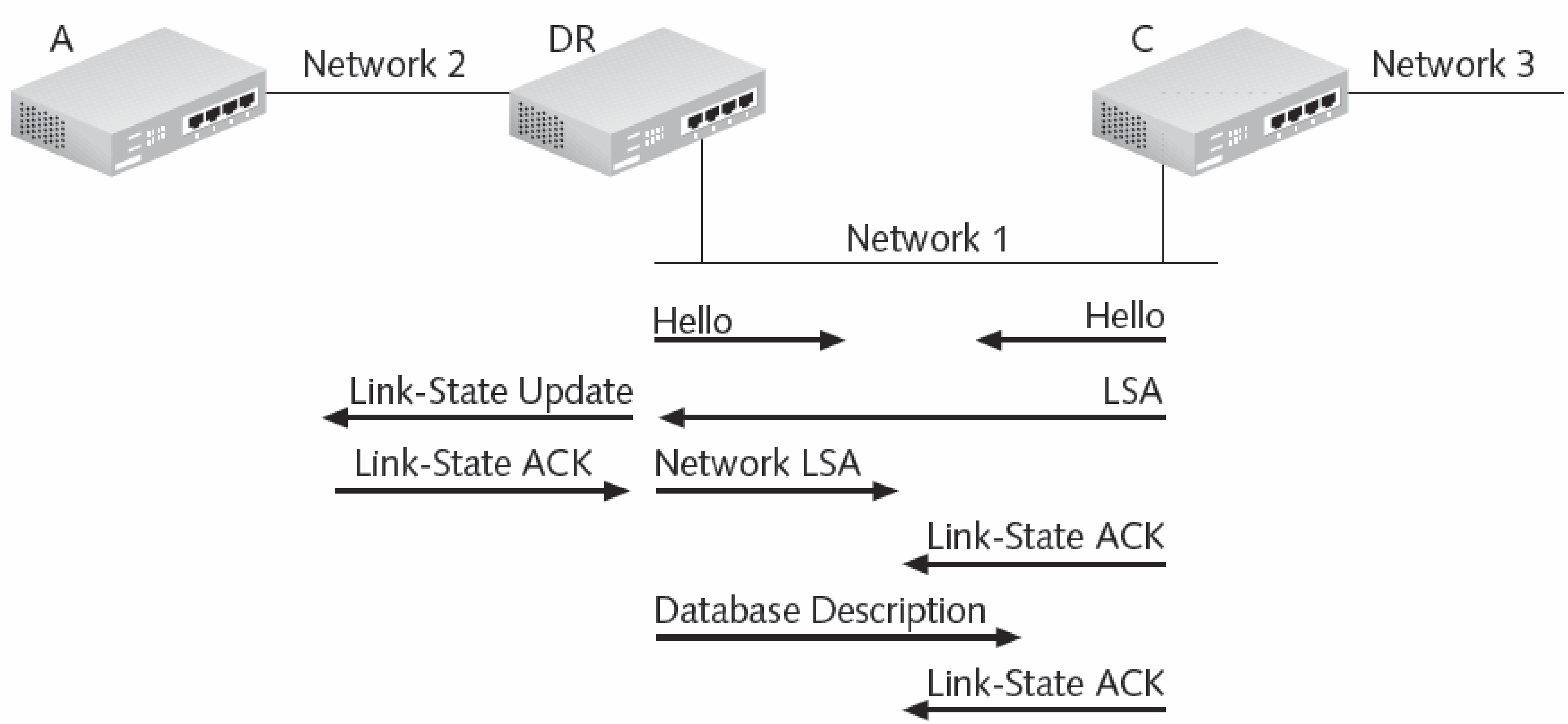


Figure 4-34 OSPF Hello, LSA, and link-state update process

Open Shortest Path First (cont'd.)

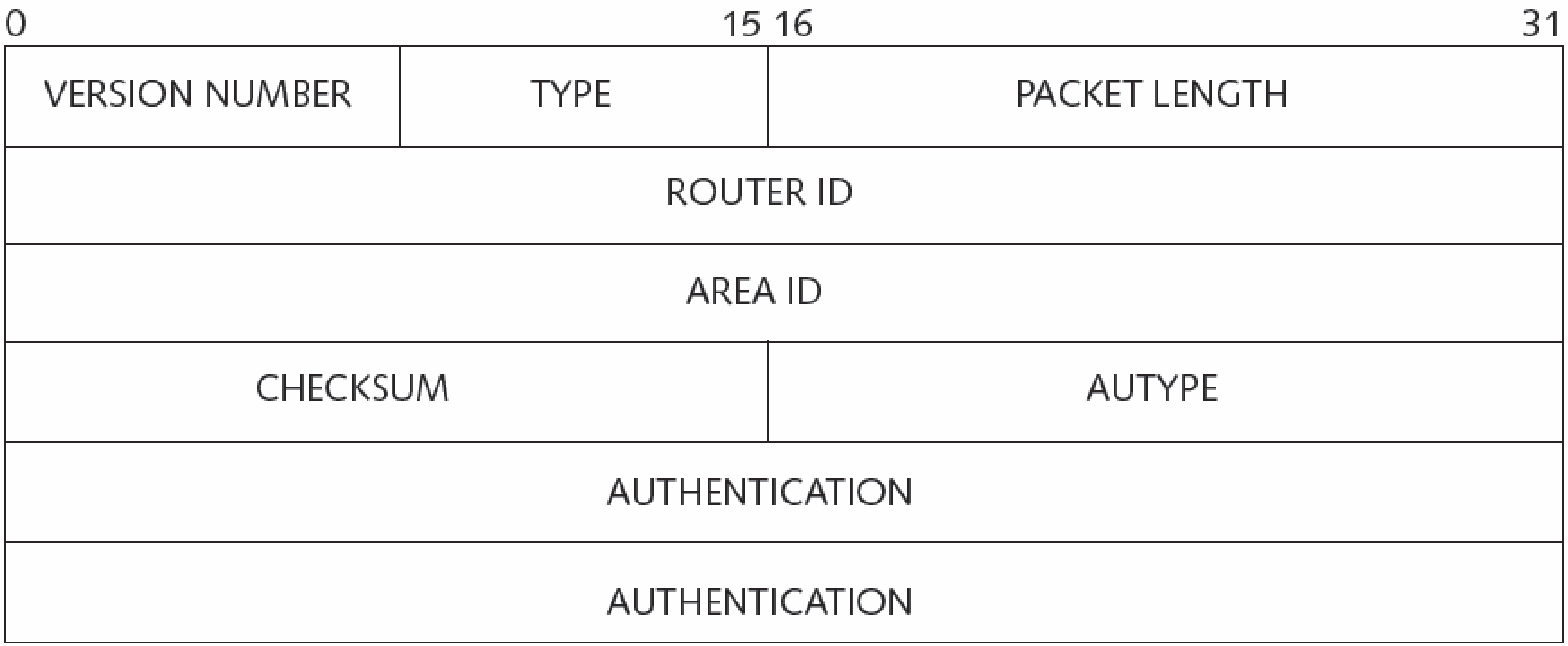


Figure 4-35 Standard OSPF header structure

Enhanced Interior Gateway Routing Protocol

- Interior Gateway Routing Protocol (IGRP)
  - Developed in the 1980s by Cisco Systems
  - Updated in the early 1990s (Enhanced Interior Gateway Routing Protocol)
- Enhanced Interior Gateway Routing Protocol
  - Integrates the capabilities of link-state routing into a distance vector routing protocol

Border Gateway Protocol

- Exterior gateway protocols (EGP)
  - Used to exchange routing information between separate autonomous systems
  - Defined in RFC 904
  - Border Gateway Protocol (BGP) replaces EGP routing

Border Gateway Protocol (cont'd.)

- Offers three types of routing operations
  - Inter-autonomous system routing
  - Intra-autonomous system routing
  - Pass-through autonomous system routing
- When configured for intra-autonomous system routing
  - BGP routers are located within the same AS
- Pass-through autonomous system routing
  - Enables BGP peer routers to exchange routing information across an AS that does not support BGP

Border Gateway Protocol (cont'd.)

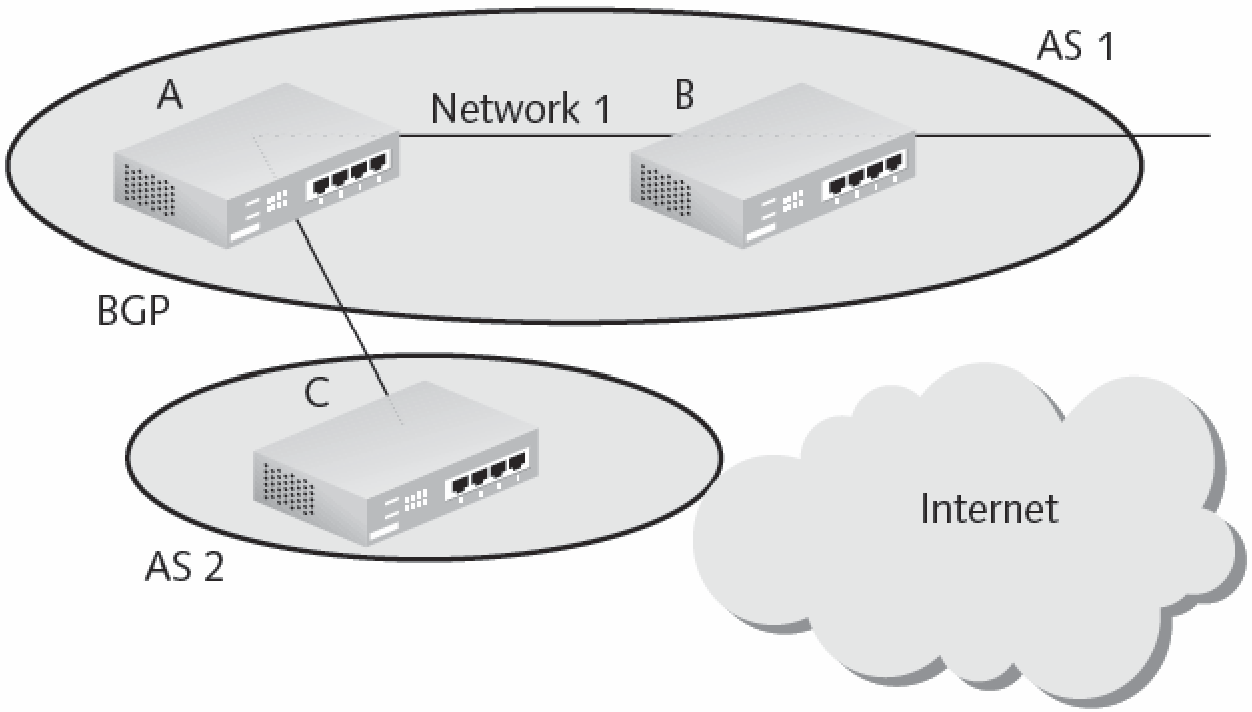


Figure 4-37 Typical BGP design

IPv6 routing protocol

- RIPng for IPv6
  - Based on protocols and algorithm currently being used by RIP IPv4
  - RFC 2080 proposes standard for RIPng for IPv6
- OSPFv3 for IPv6
  - Specified by RFC 5340
  - Based on OSPFv2 with enhancements added
- EIGRP for IPv6
  - Essentially the same protocol

IPv6 routing protocol (cont'd.)

COMMAND (1)	VERSION (1)	MUST BE ZERO (2)
ROUTE TABLE ENTRY 1 (20)		
...		
ROUTE TABLE ENTRY N (20)		

Figure 4-38 RIPng packet format

IPv6 routing protocol (cont'd.)

<table>
<tr><td>VERSION</td><td>TYPE</td><td colspan="3">PACKET LENGTH</td></tr>
<tr><td colspan="5">ROUTER ID</td></tr>
<tr><td colspan="5">AREA ID</td></tr>
<tr><td colspan="3">CHECKSUM</td><td>INSTANCE ID</td><td>0</td></tr>
</table>

Figure 4-39 OSPFv3 header format

IPv6 routing protocol (cont'd.)

- IS-IS for IPv6
  - Intermediate System-to-Intermediate System
  - Intradomain routing information exchange protocol
  - Described by RFC 5308
- MP-BGP
  - RFC 4760 defines the multiprotocol extensions
  - Make BGP-4 available to other network layer protocols

Routing on and off a Wide Area Network

- Consider many factors when choosing a routing protocol for your enterprise

Several Small Offices

- If the network in question is relatively small
  - Consider using no routing protocol
- If a routing protocol is required, however, then RIP may be the simplest solution
  - When connecting to a server that supports routing
  - When routers that do not support any other protocols

Hub and Spoke

- On-Demand Routing (ODR) protocol
  - A good choice
- Frame relay and ATM
  - Use virtual circuits
    - So one physical interface can actually be logically partitioned into several logical interfaces

Multiprotocol

- Cisco's EIGRP protocol
  - Support routing of both IPv4 and IPv6 traffic at the same time

Mobile Users

- Some of the most difficult types of networks to implement and manage
- Last few years witnessed a remarkable maturation of technologies

Mobile IP

- Mobile IP
  - Defined by the IETF in RFCs 2003 through 2006 and RFC 3220 (which obsoletes RFC 2002)

Local Area Mobility

- Local Area Mobility
  - Cisco proprietary feature that is similar to Mobile IP
  - Operates by using the routing table
- When a router is configured with Local Area Mobility (LAM):
  - It watches for traffic on its LAN that does not match its own IP address
  - When it finds this traffic, it installs an ARP entry in its cache and a host route in its routing table

Routing to and from the Internet

- BGPv4
  - The exterior routing protocol in use on the Internet
  - Should only be used by networks that connect to multiple Internet providers
  - Tracks hops between pairs of autonomous systems instead of tracking hop counts for actual routers

Securing Routers and Routing Behavior

- Securing routers
  - Turn off unnecessary services
  - Shut down unnecessary listening ports
  - Configure strong access security to prevent tampering
  - Secure physical access to the boxes
- Securing routing protocols
  - Requires cooperation from the protocols themselves
- Details in Chapter 11

Troubleshooting IP Routing

- Traceroute/Tracert
  - Used to trace an IP packet from a source computer to its destination
- Ping
  - Sends ICMP Echo messages and test connectivity
- Pathping
  - Used to discover path from host to destination

Summary

- Data link protocols
  - Manage transfer of datagrams across the network
- At Data Link layer
  - Protocols must deliver services, such as delimitation, bit-level integrity checks, addressing, and protocol identification
- Ethernet II frames
  - Most common frame type on LANs

Summary (cont'd.)

- Understanding frame layouts
  - Crucial for proper handling of contents
- At the lowest level of detail
  - Important to understand the differences in field layouts and meanings of different Ethernet II frames
- Imperative to understand how TCP/IP manages the translation between MAC layer addresses and numeric IP addresses

Summary (cont'd.)

- Proxy ARP
  - Permits router to interconnect multiple network segments
- Network layer protocols
  - Make their way into the Data Link layer through a process known as *data encapsulation*
- Important characteristics of IP datagrams
  - Time to Live (TTL) values
  - Fragmentation of incoming frames
  - Service delivery options

Summary (cont'd.)

- Routing protocols and routers
  - Provide a mechanism that can forward traffic from a sender's subnet to an intended receiver's subnet
- Routers
  - Depend on various routing protocols to manage the packet forwarding process
- Distance vector routing protocols such as RIP
  - Provides a crude metric of routing cost

Summary (cont'd.)

- The OSPF protocol
  - Supports much more sophisticated routing structures that break up a network into routing areas
- Routing characteristics
  - Help to determine what kinds of routing protocols to use in specific applications
- Managing routing on a complex network involves:
  - Understanding how and when to use exterior and interior routing protocols

Summary (cont'd.)

- Router tables define the topology and behavior of IP networks
  - Essential to manage router security and updates as safely as possible
- IPv6 Routing Mechanisms perform the same function as their IPv4 counterparts
- IPv6 Multicast Listener Discovery specifies multicast behaviors as well as MLD packets and message types
- IPv6 routing protocols are RIPng, OSPFv3, EIGRP for IPv6, IS-IS for IPv6, and MPBGP

Homework

- Developing a glossary of this chapter
  - English full name (Acronym): Chinese full name
  - E.g. Transmission Control Protocol (TCP): 传输控制协议
- Exercises
  - 1, 3, 4, 7, 8, 10, 12, 15, 17, 19, 20, 22
- Hands-on projects
  - 4-1~4-4
  - Write only one project report for all projects