Unicast Routing Protocols

Dr Kunwar Pal
Assistant Professor
Dept. of Computer Science and Engineering
National Institute of Technology
Jalandhar-144011, Punjab India

1 INTRODUCTION

An internet is a combination of networks connected by routers. When a datagram goes from a source to a destination, it will probably pass through many routers until it reaches the router attached to the destination network.

- ✔ **Cost or Metric**
- ✔ **Static versus Dynamic Routing Table**
- ✔ **Routing Protocol**

2 INTER- AND INTRA-DOMAIN ROUTING

- An internet can be so large that one routing protocol cannot handle the task of updating the routing tables of all routers.
- For this reason, an internet is divided into autonomous systems.
- An autonomous system (AS) is a group of networks and routers under the authority of a single administration.
- Routing inside an autonomous system is called intra-domain routing.
- Routing between autonomous systems is called inter-domain routing

Figure 1 *Autonomous systems*

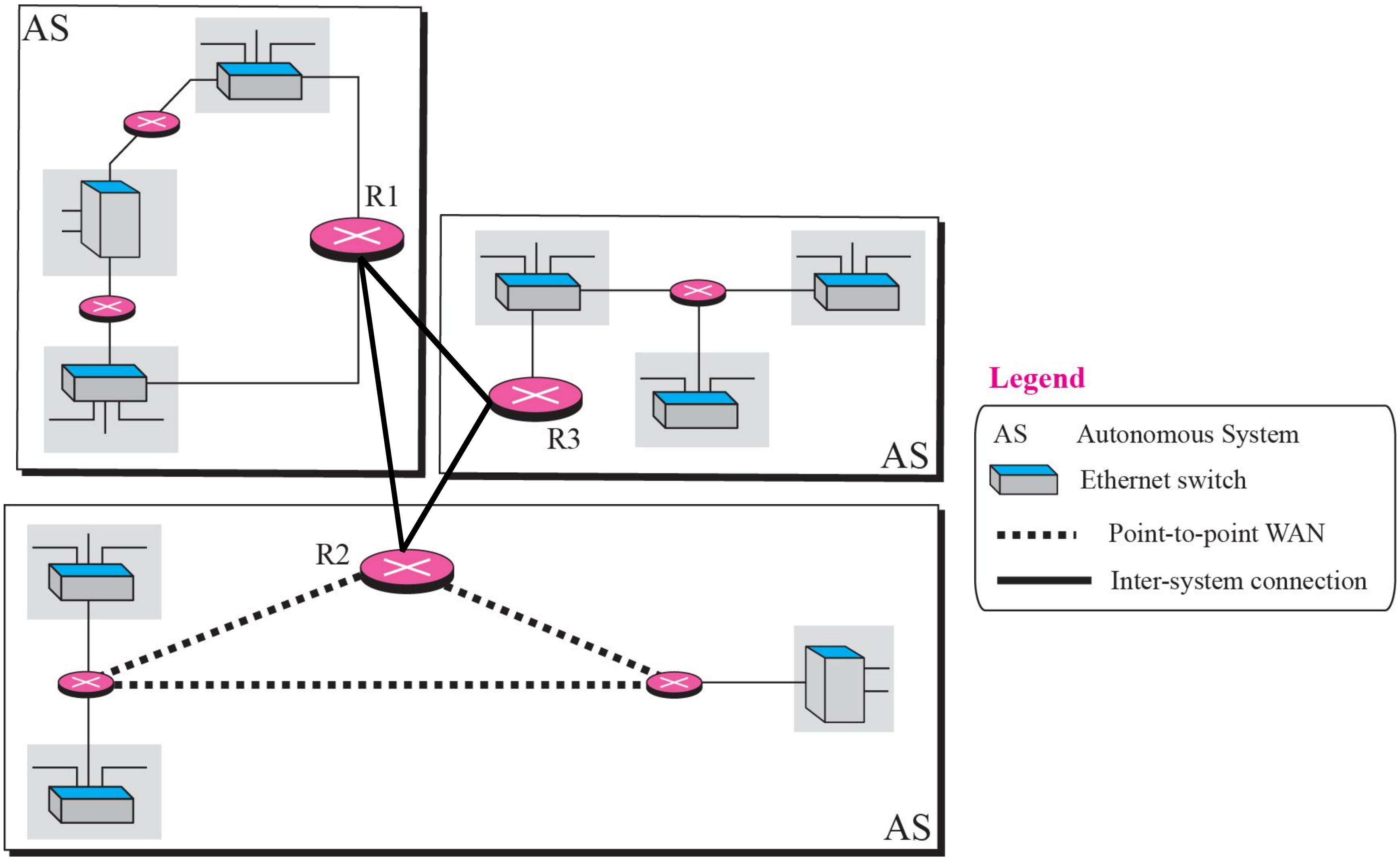
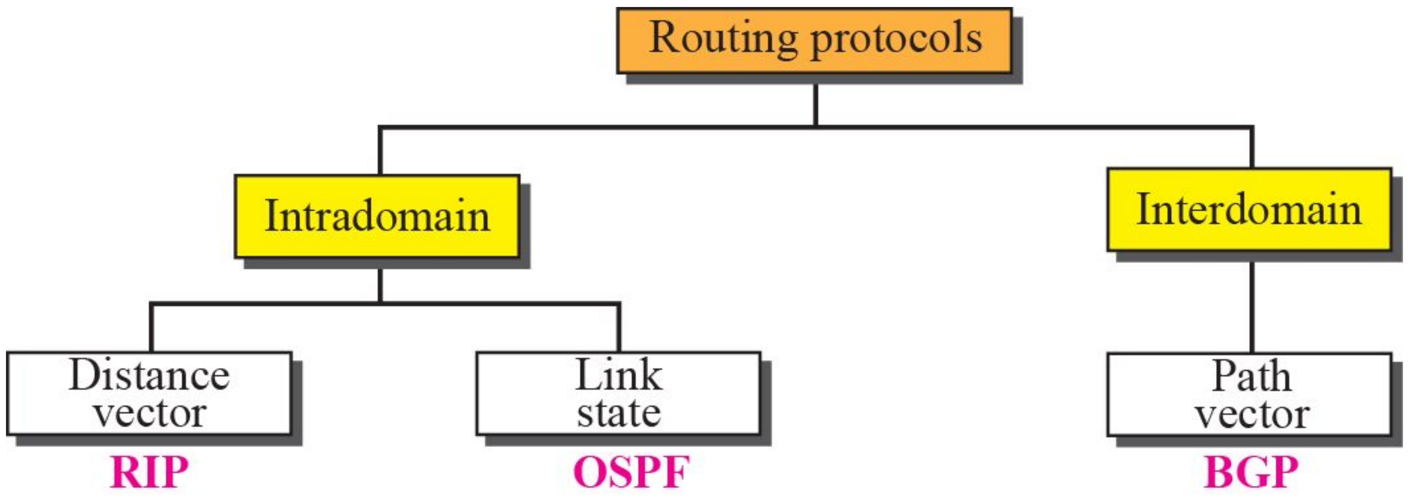


Figure 2 ***Popular routing protocols***



3 DISTANCE VECTOR ROUTING

The distance-vector (DV) routing uses to find the best route.

- DV is simple routing protocol which takes routing decision on the number of hops between source and destination.
- A route with less number of hops is considered as best route.
- Every router advertises its set best routes to other routers.
- Ultimately, all routers build up their network topology based on the advertisements of their peer routers

✔ Bellman-Ford Algorithm

✔ Distance Vector Routing Algorithm

✔ Count to Infinity

Figure 3 *A graph for Bellman-Ford algorithm*

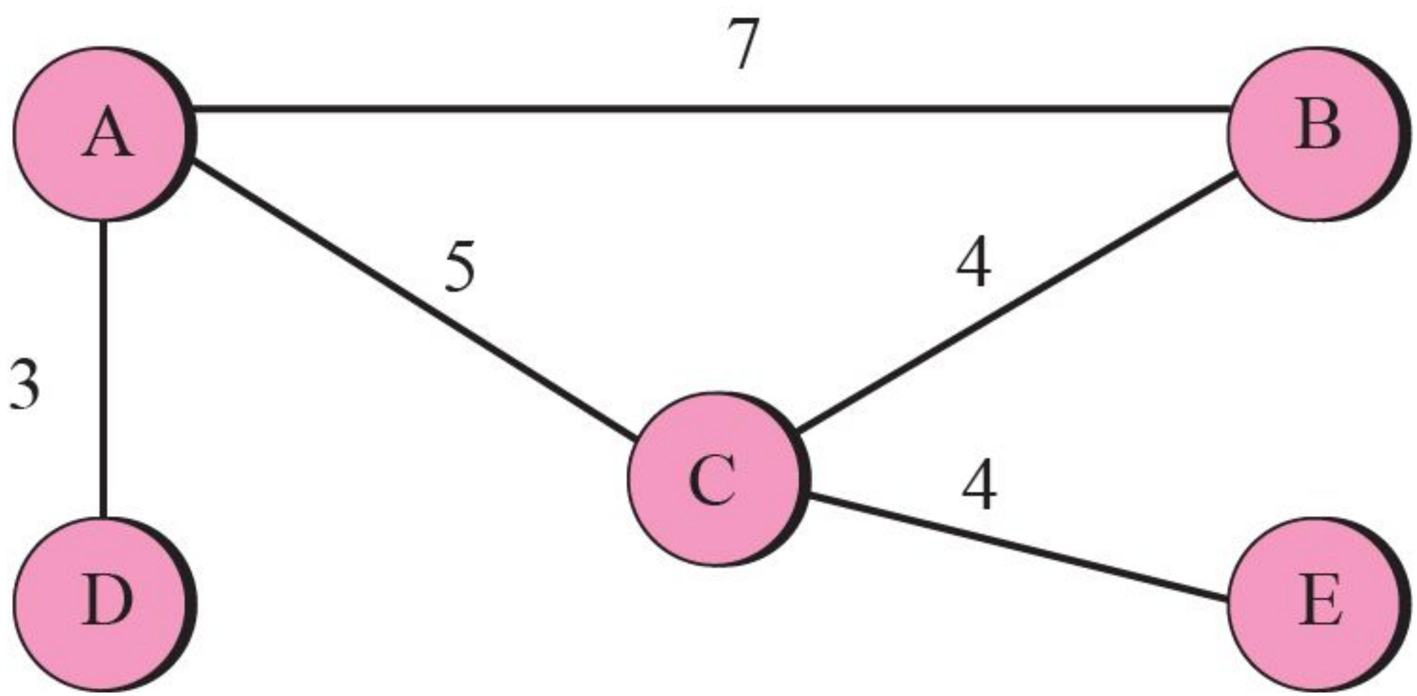
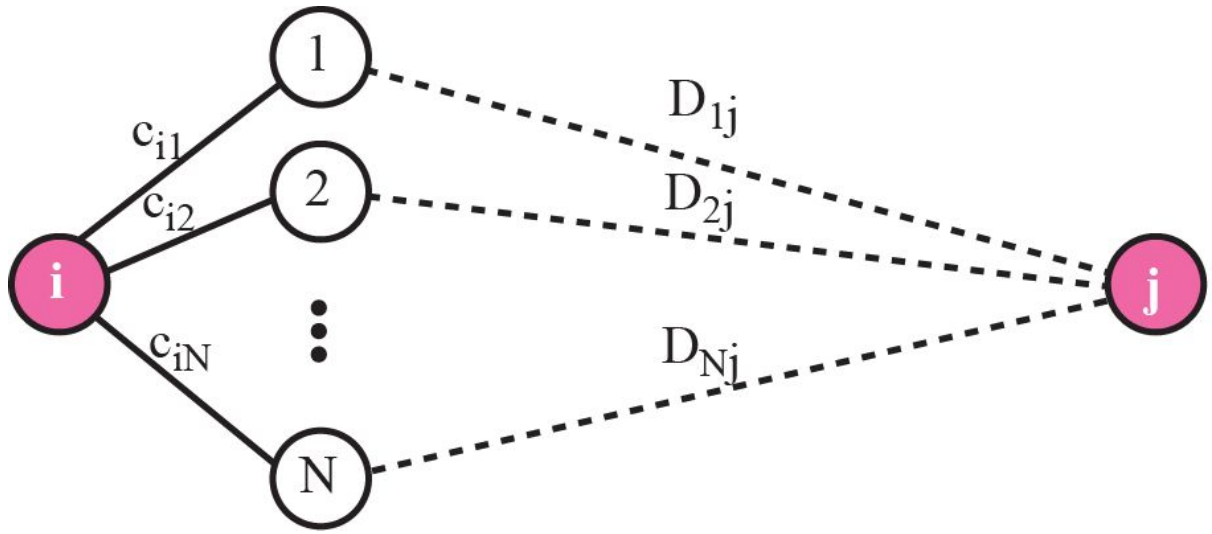


Figure 4 *The fact behind Bellman-Ford algorithm*

$$D_{ij} = \text{minimum } \{(c_{i1} + D_{1j}), (c_{i2} + D_{2j}), \ldots (c_{iN} + D_{Nj})\}$$



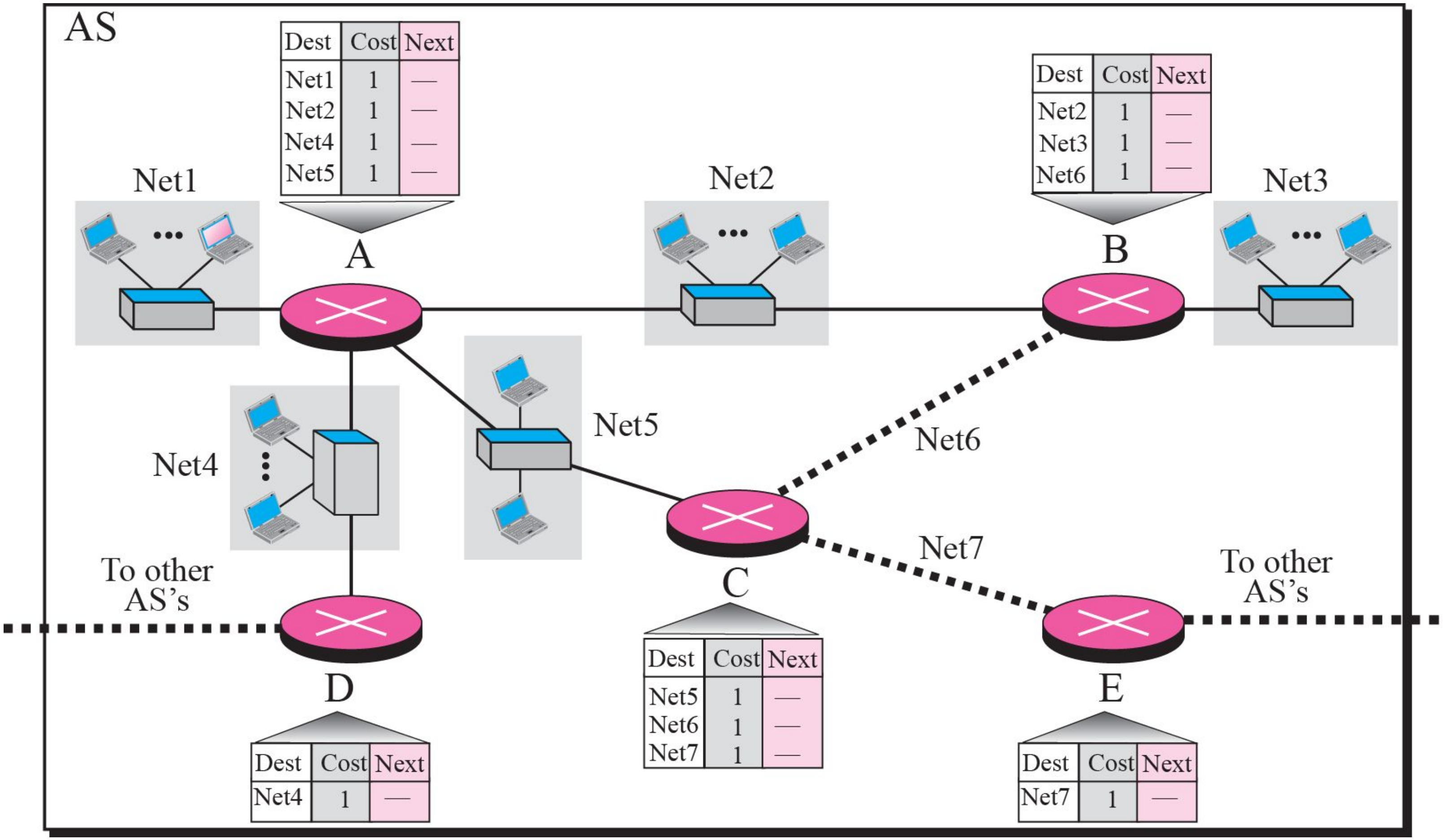
Legend

D_{ij} Shortest distance between i and j
c_{ij} Cost between i and j
N Number of nodes

Example 1

- **Next figure shows the initial routing table for an AS.**
- **Note that the figure does not mean that all routing tables have been created at the same time; each router creates its own routing table when it is booted.**

Figure 5 *Example 1*



Example 2

- **Now assume router A sends four records to its neighbors, routers B, D, and C.**

- **Next figure 6 shows the changes in B's routing table when it receives these records. We leave the changes in the routing tables of other neighbors as exercise.**

Figure 6 *Example 2*

Dest	Cost	Next
Net1	1	—
Net2	1	—
Net4	1	—
Net5	1	—

Dest	Cost	Next
Net2	1	—
Net3	1	—
Net6	1	—

Routing Table B

Dest	Cost	Next
Net1	2	A
Net2	1	—
Net3	1	—
Net6	1	—

After receiving record 1

Routing Table B

Dest	Cost	Next
Net1	2	A
Net2	1	—
Net3	1	—
Net6	1	—

After receiving record 2

Routing Table B

Dest	Cost	Next
Net1	2	A
Net2	1	—
Net3	1	—
Net4	2	A
Net6	1	—

After receiving record 3

Routing Table B

Dest	Cost	Next
Net1	2	A
Net2	1	—
Net3	1	—
Net4	2	A
Net5	2	A
Net6	1	—

After receiving record 4

Example 3

Figure 7 shows the final routing tables for routers in Figure 5.

Figure 7 ***Example 3***

A

Dest	Cost	Next
Net1	1	—
Net2	1	—
Net3	2	B
Net4	1	—
Net5	1	—
Net6	2	C
Net7	2	C

B

Dest	Cost	Next
Net1	2	A
Net2	1	—
Net3	1	—
Net4	2	A
Net5	2	A
Net6	1	—
Net7	2	C

C

Dest	Cost	Next
Net1	2	A
Net2	2	A
Net3	2	B
Net4	2	A
Net5	1	—
Net6	1	—
Net7	1	—

D

Dest	Cost	Next
Net1	2	A
Net2	2	A
Net3	3	A
Net4	1	—
Net5	1	A
Net6	3	A
Net7	3	A

E

Dest	Cost	Next
Net1	3	C
Net2	3	C
Net3	3	C
Net4	3	C
Net5	2	C
Net6	2	C
Net7	1	—

Figure 8 *Two-node instability*

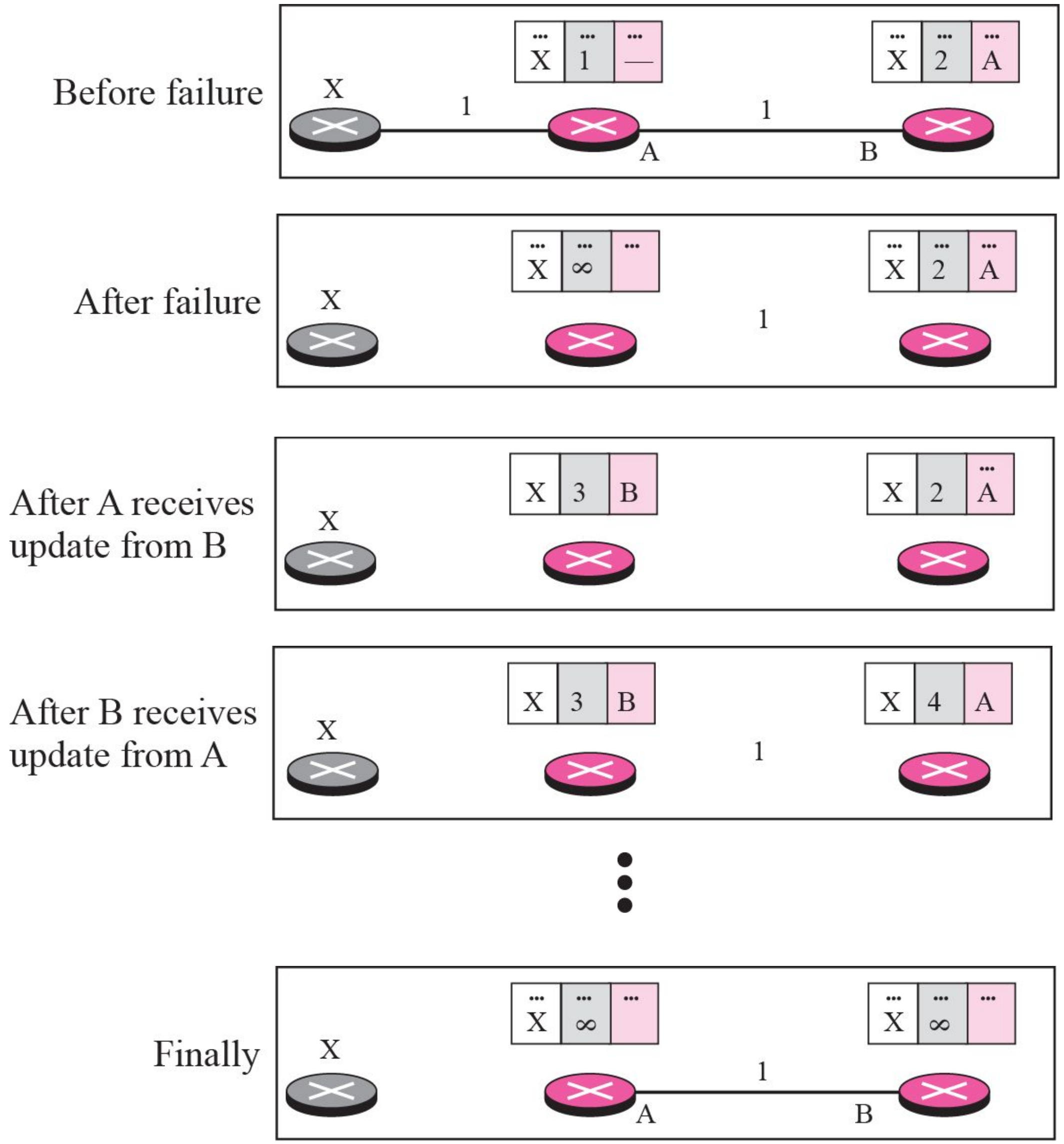
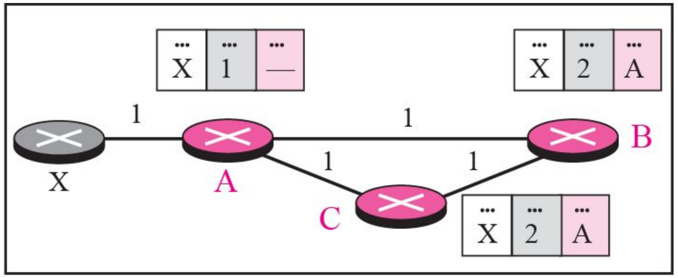
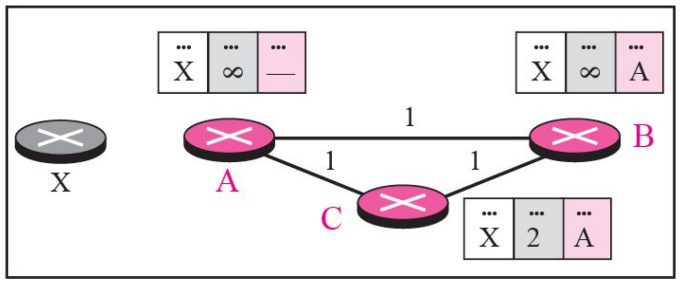


Figure 9 ***Three-node instability***

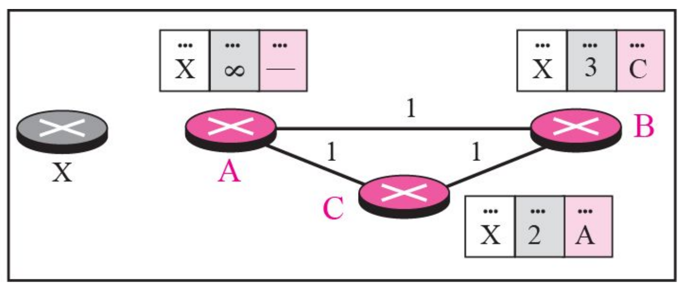
Before failure



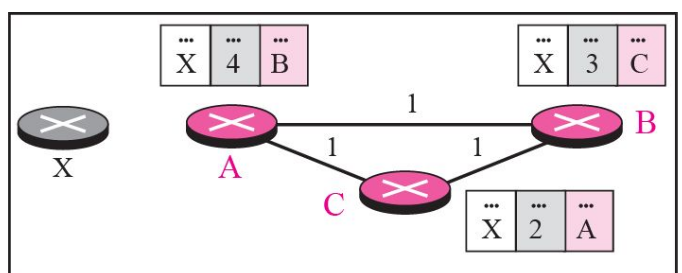
After A sends the route to B and C, but the packet to C is lost



After C sends the route to B



After B sends the route to A



4 RIP

- The Routing Information Protocol (RIP) is an intra-domain (interior) routing protocol used inside an autonomous system.
- It is a very simple protocol based on distance vector routing. RIP implements distance vector routing directly with some considerations.

✔ RIP Message Format

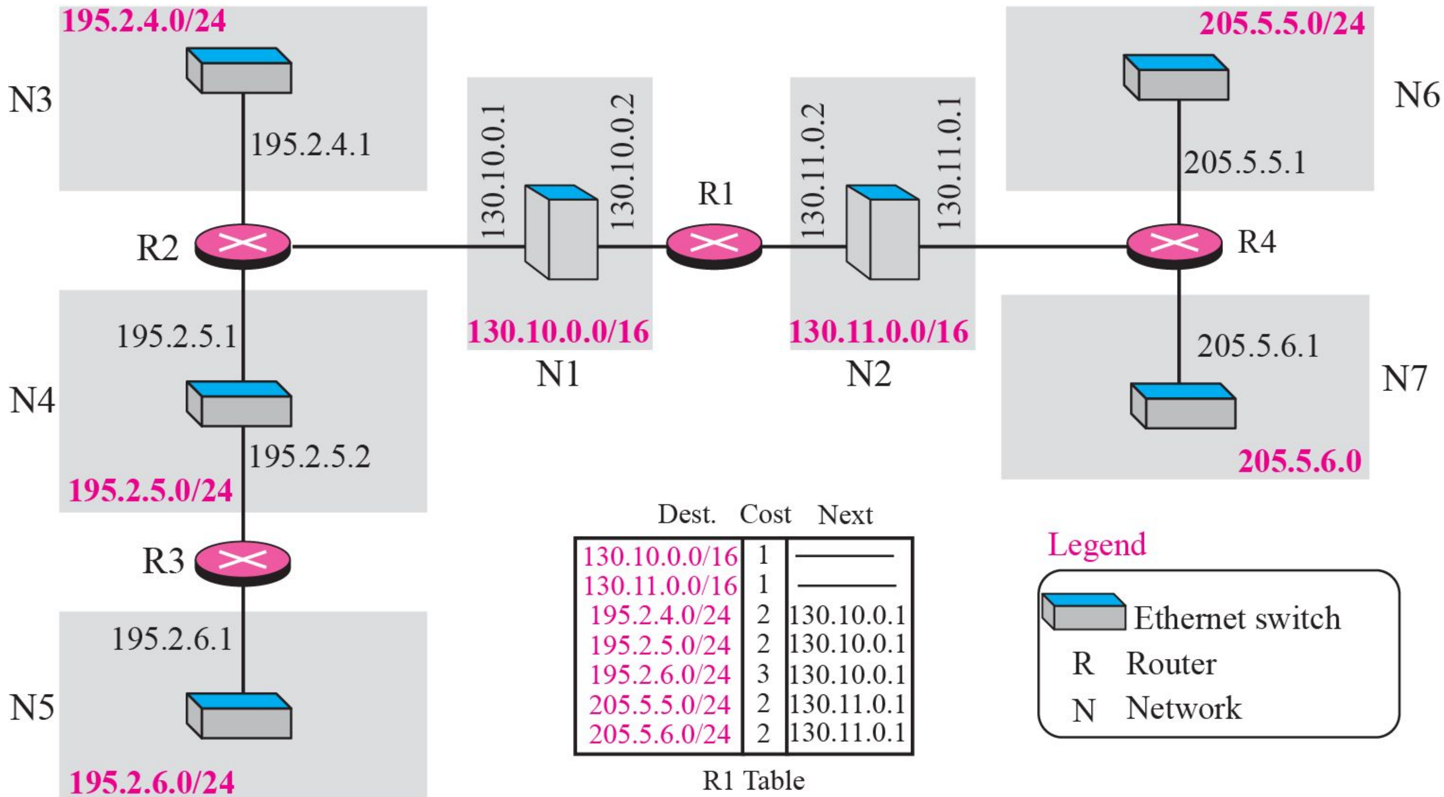
✔ Request and Response

✔ Timers in RIP

✔ RIP Version 2

✔ Encapsulation

Figure 10 *Example of a domain using RIP*



Dest.	Cost	Next
130.10.0.0/16	1	———
130.11.0.0/16	1	———
195.2.4.0/24	2	130.10.0.1
195.2.5.0/24	2	130.10.0.1
195.2.6.0/24	3	130.10.0.1
205.5.5.0/24	2	130.11.0.1
205.5.6.0/24	2	130.11.0.1

R1 Table

Dest.	Cost	Next
130.10.0.0/16	1	———
130.11.0.0/16	2	130.10.0.2
195.2.4.0/24	1	———
195.2.5.0/24	1	———
195.2.6.0/24	2	195.2.5.2
205.5.5.0/24	3	130.10.0.2
205.5.6.0/24	3	130.10.0.2

R2 Table

Dest.	Cost	Next
130.10.0.0/16	2	195.2.5.1
130.11.0.0/16	3	195.2.5.1
195.2.4.0/24	2	195.2.5.1
195.2.5.0/24	1	———
195.2.6.0/24	1	———
205.5.5.0/24	4	195.2.5.1
205.5.6.0/24	4	195.2.5.1

R3 Table

Dest.	Cost	Next
130.10.0.0/16	2	130.11.0.2
130.11.0.0/16	1	———
195.2.4.0/24	3	130.11.0.2
195.2.5.0/24	3	130.11.0.2
195.2.6.0/24	4	130.11.0.2
205.5.5.0/24	1	———
205.5.6.0/24	1	———

R4 Table

Figure 11 *RIP message format*

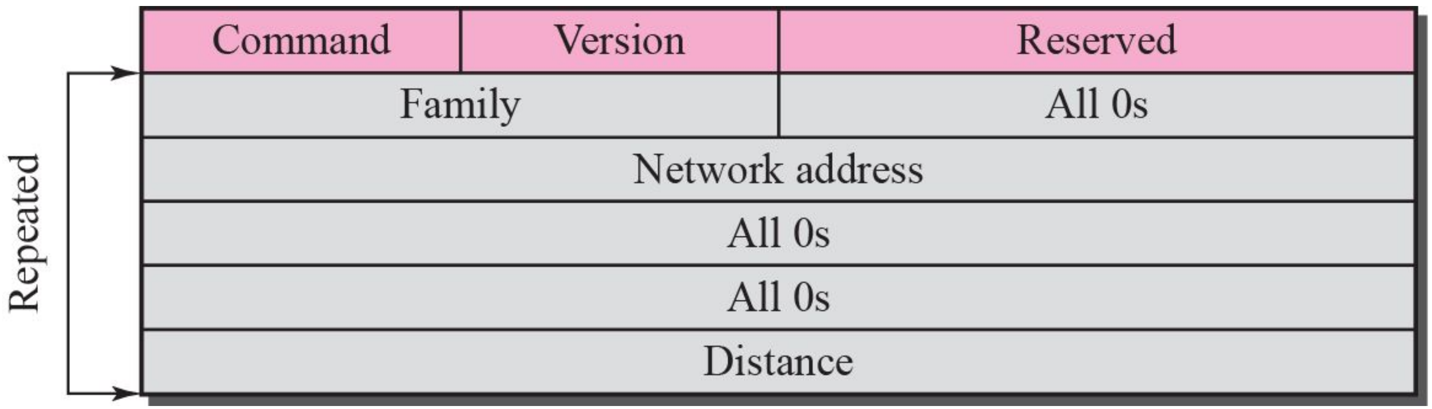


Figure 12 ***Request messages***

Repeated

Com: 1	Version	Reserved
Family		All 0s
Network address		
All 0s		
All 0s		
All 0s		

a. Request for some

Com: 1	Version	Reserved
Family		All 0s
All 0s		
All 0s		
All 0s		
All 0s		

b. Request for all

Example 4

Figure 13 shows the update message sent from router R1 to router R2 in Figure 10. The message is sent out of interface 130.10.0.2.

The message is prepared with the combination of split horizon and poison reverse strategy in mind. Router R1 has obtained information about networks 195.2.4.0, 195.2.5.0, and 195.2.6.0 from router R2. When R1 sends an update message to R2, it replaces the actual value of the hop counts for these three networks with 16 (infinity) to prevent any confusion for R2. The figure also shows the table extracted from the message. Router R2 uses the source address of the IP datagram carrying the RIP message from R1 (130.10.0.2) as the next hop address. Router R2 also increments each hop count by 1 because the values in the message are relative to R1, not R2.

Figure 13 *Solution to Example 4*

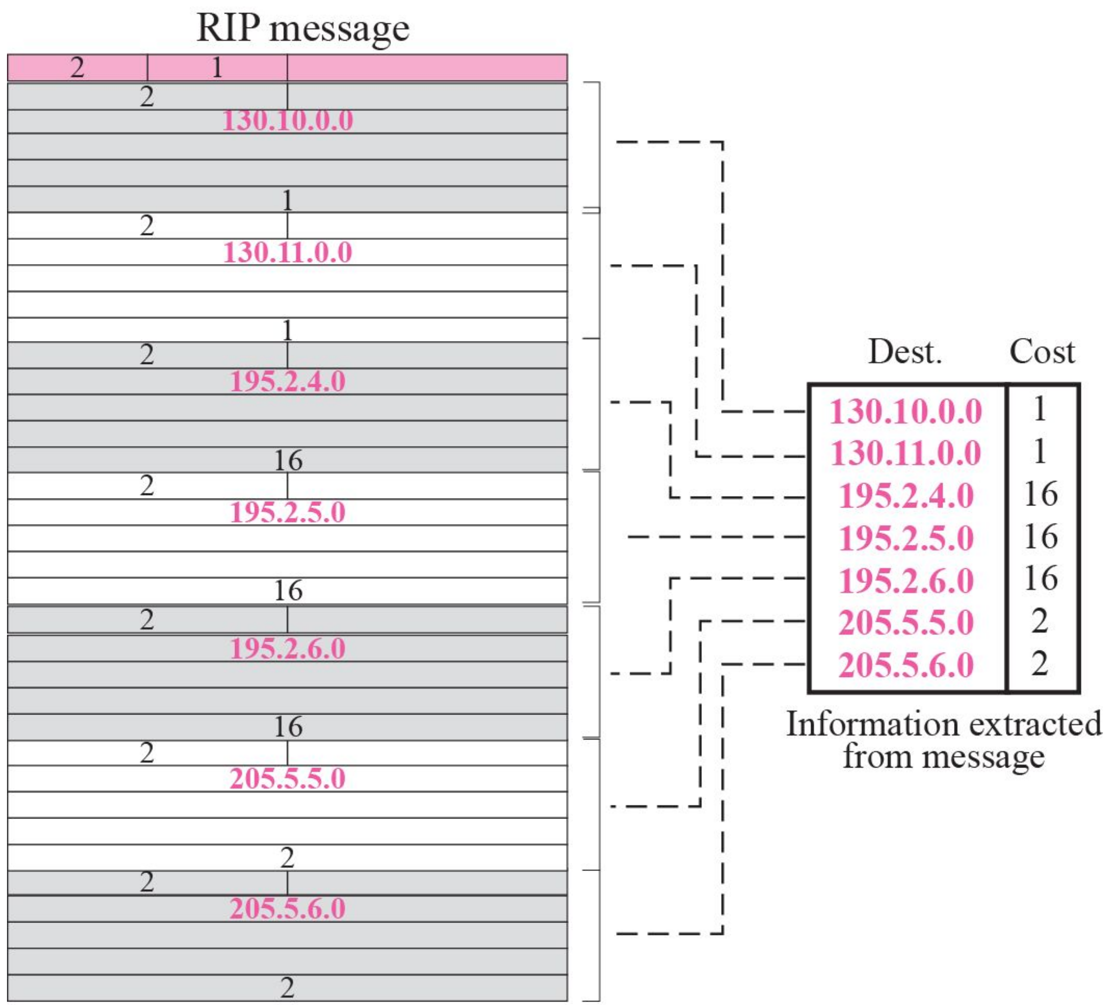


Figure 14 *RIP timers*

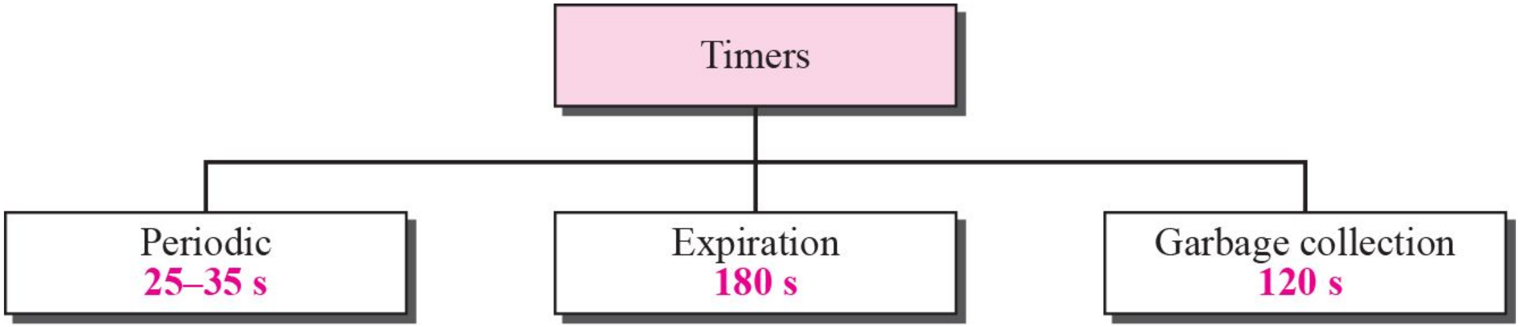


Figure 15 *RIP version 2 format*

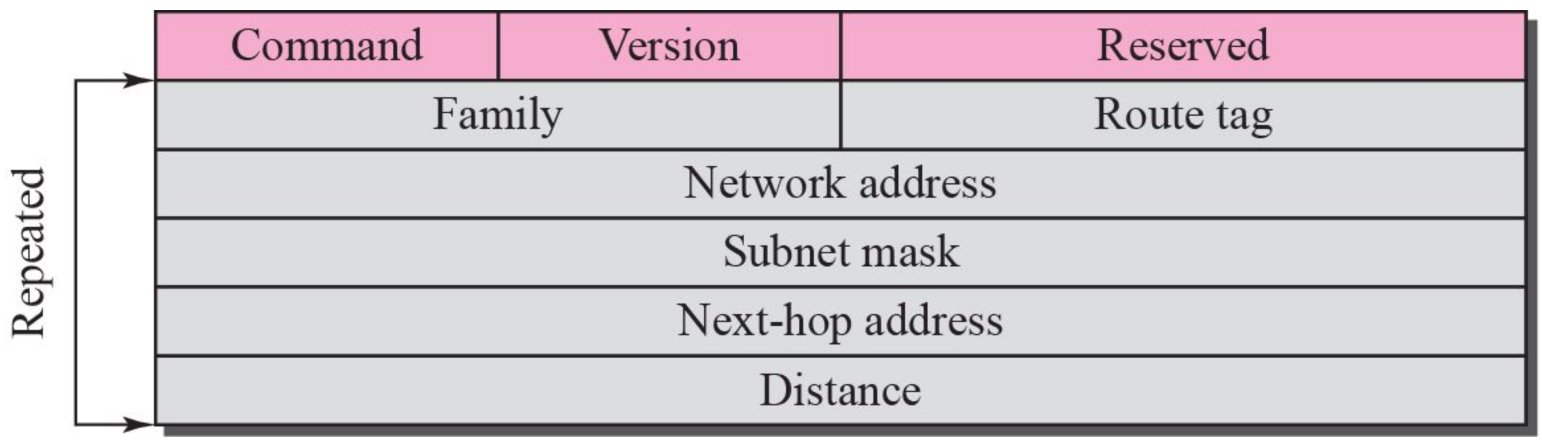
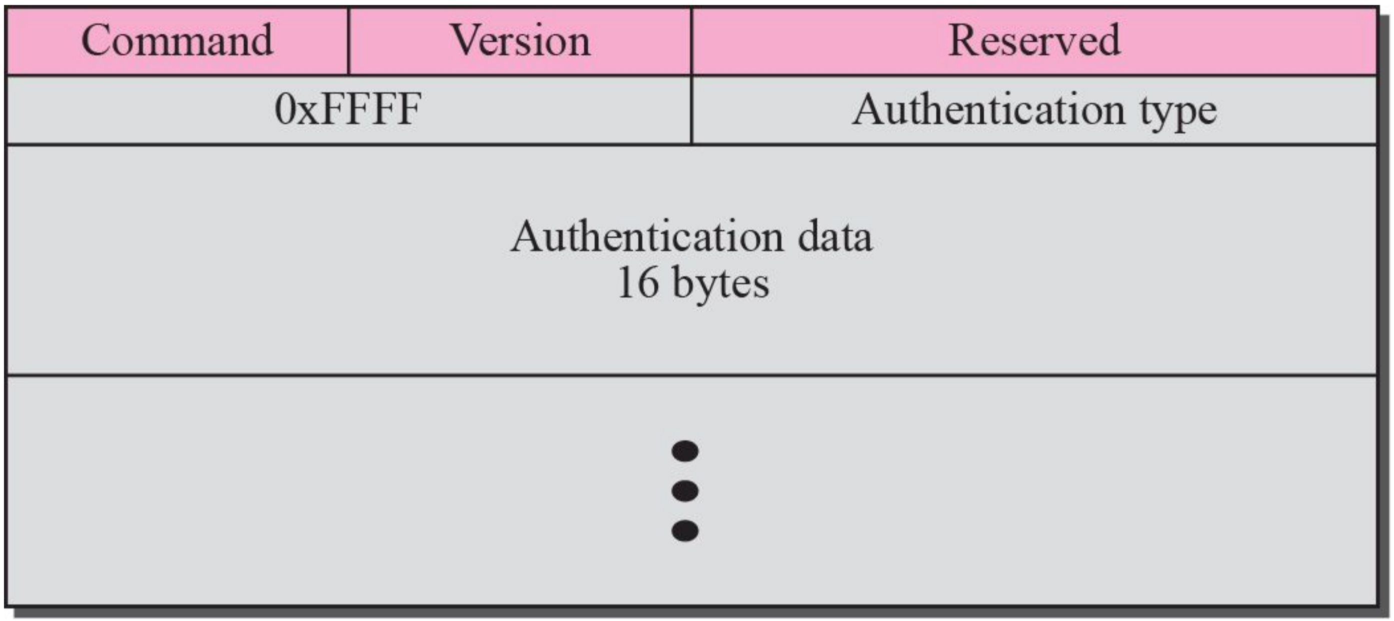


Figure 16 *Authentication*



Note

RIP uses the services of UDP on well-known port 520.

5 LINK STATE ROUTING

Link state routing has a different philosophy from that of distance vector routing.
In link state routing, if each node in the domain has the entire topology of the domain—the list of nodes and links, how they are connected including the type, cost (metric), and the condition of the links (up or down)—the node can use the Dijkstra algorithm to build a routing table.

Figure 17 *Concept of Link state routing*

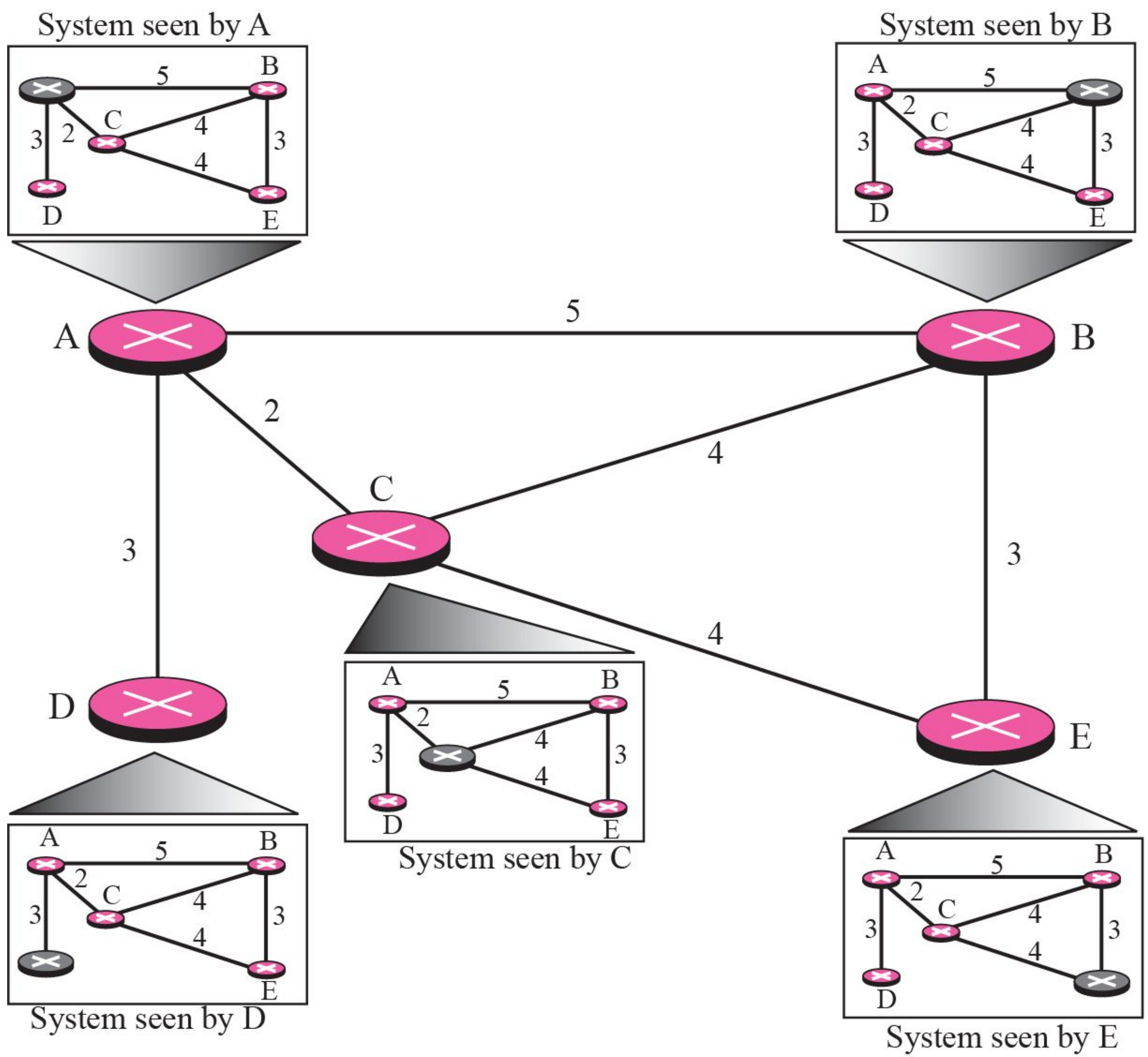
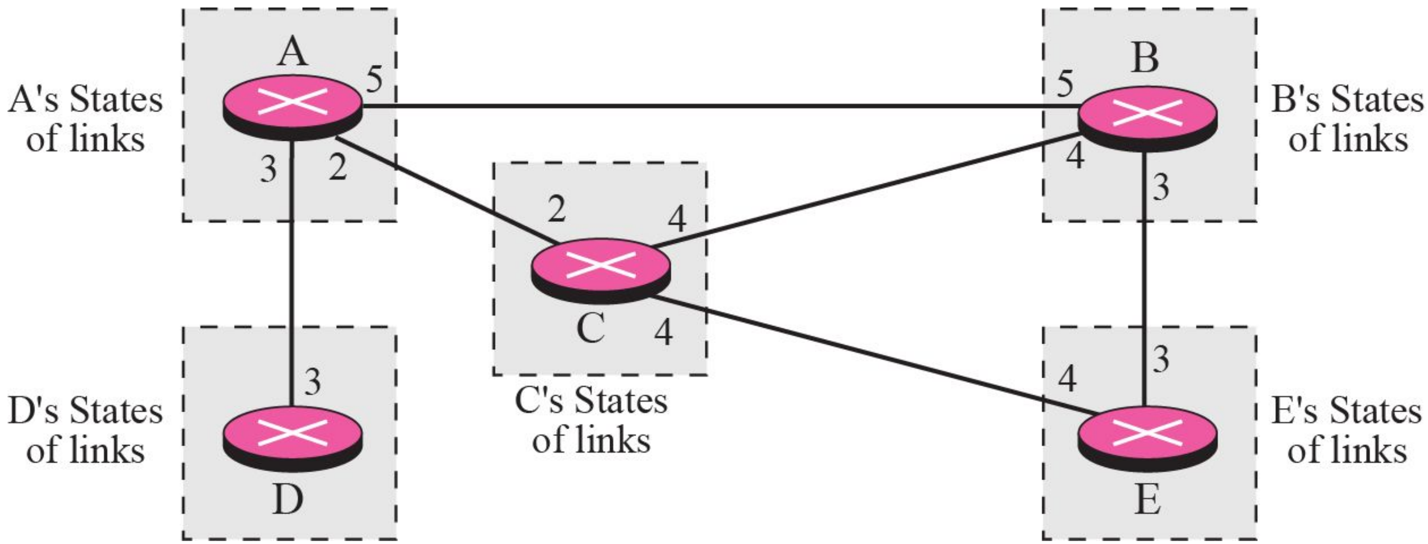


Figure 18 *Link state knowledge*



Building Routing Tables

- Creation of the states of the links by each node, called the link state packet or LSP.
- Dissemination of LSPs to every other router, called **flooding,** in an efficient and reliable way.
- Formation of a shortest path tree for each node.
- Calculation of a routing table based on the shortest path tree.

Creation of Link State Packet (LSP)

- A link state packet (LSP) can carry a large amount of information.
- For the moment, however, we assume that it carries a minimum amount of data:
  - the node identity, the list of links, a sequence number, and age.

- The first two, node identity and the list of links, are needed to make the topology.
- The third, sequence number, facilitates flooding and distinguishes new LSPs from old ones.
- The fourth, age, prevents old LSPs from remaining in the domain for a long time.

LSPs are generated on two occasions

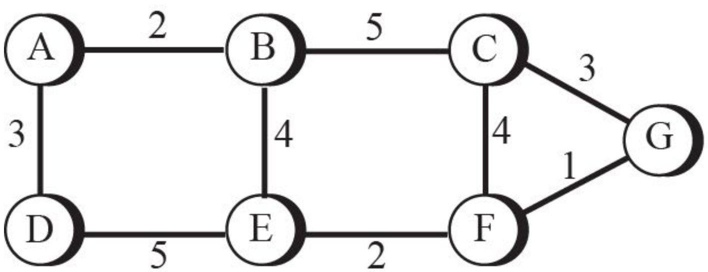
- *When there is a change in the topology of the domain.*
- *On a periodic basis.*
  - The period in this case is much longer compared to distance vector routing. As a matter of fact, there is no actual need for this type of LSP dissemination. It is done to ensure that old information is removed from the domain.
  - The timer set for periodic dissemination is normally in the range of 60 minutes or 2 hours based on the implementation. A longer period ensures that flooding does not create too much traffic on the network.

Flooding of LSPs

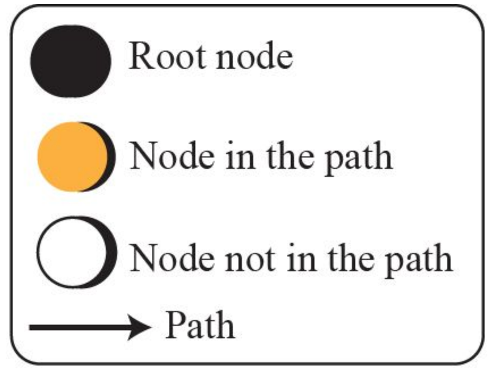
- After a node has prepared an LSP, it must be disseminated to all other nodes, not only to its neighbors.
- The process is called flooding and based on the following:
  - The creating node sends a copy of the LSP out of each interface.
  - A node that receives an LSP compares it with the copy it may already have.
  - If the newly arrived LSP is older than the one it has (found by checking the sequence number), it discards the LSP.
  - If it is newer, the node does the following:
    a. It discards the old LSP and keeps the new one.
    b. It sends a copy of it out of each interface except the one from which the packet arrived. This guarantees that flooding stops somewhere in the domain (where a node has only one interface).

- *Formation of Shortest Path Tree: Dijkstra Algorithm*

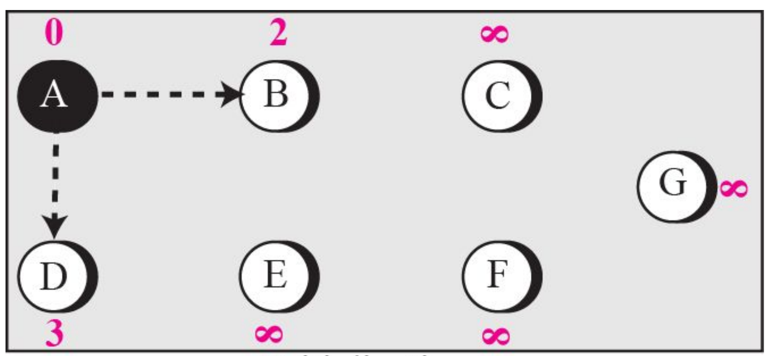
Figure 19 *Forming shortest path three for router A in a graph*



Topology

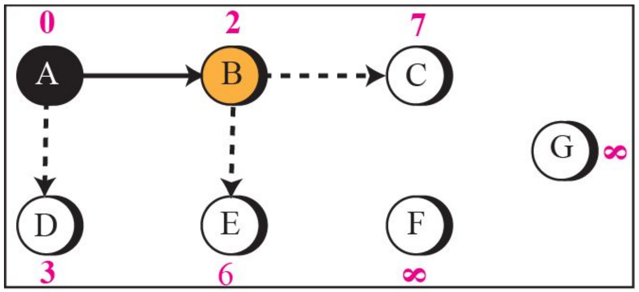


Legend

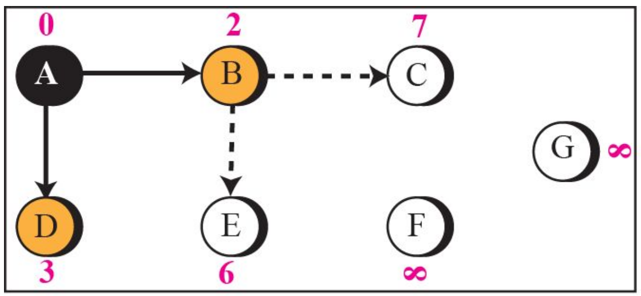


Initialization

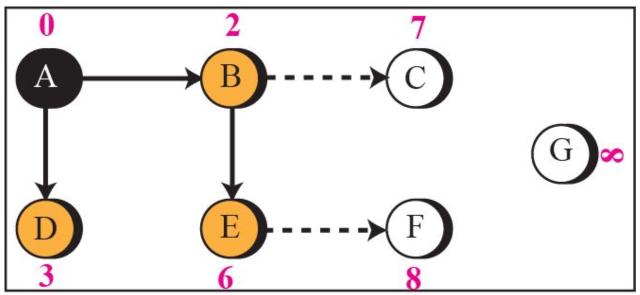
Figure 19 *Continued*



Iteration 1

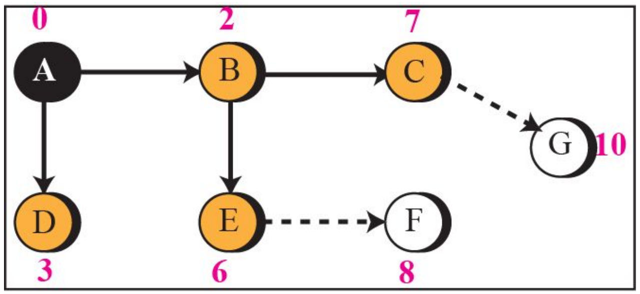


Iteration 2

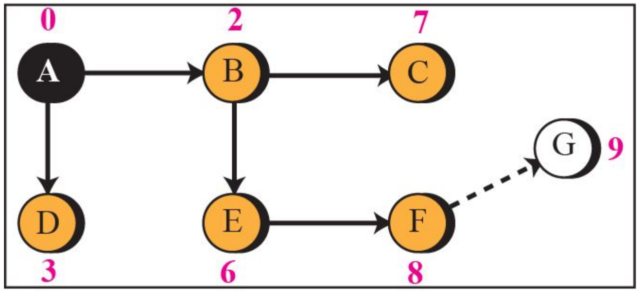


Iteration 3

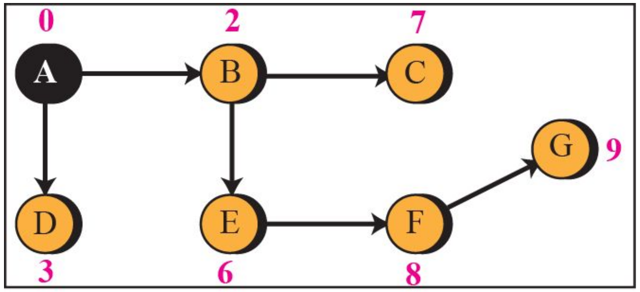
Figure 19 ***Continued***



Iteration 4



Iteration 5



Iteration 6

Example 6

To show that the shortest path tree for each node is different, we found the shortest path tree as seen by node C (Figure 20). We leave the detail as an exercise.

Figure 20 *Example 6*

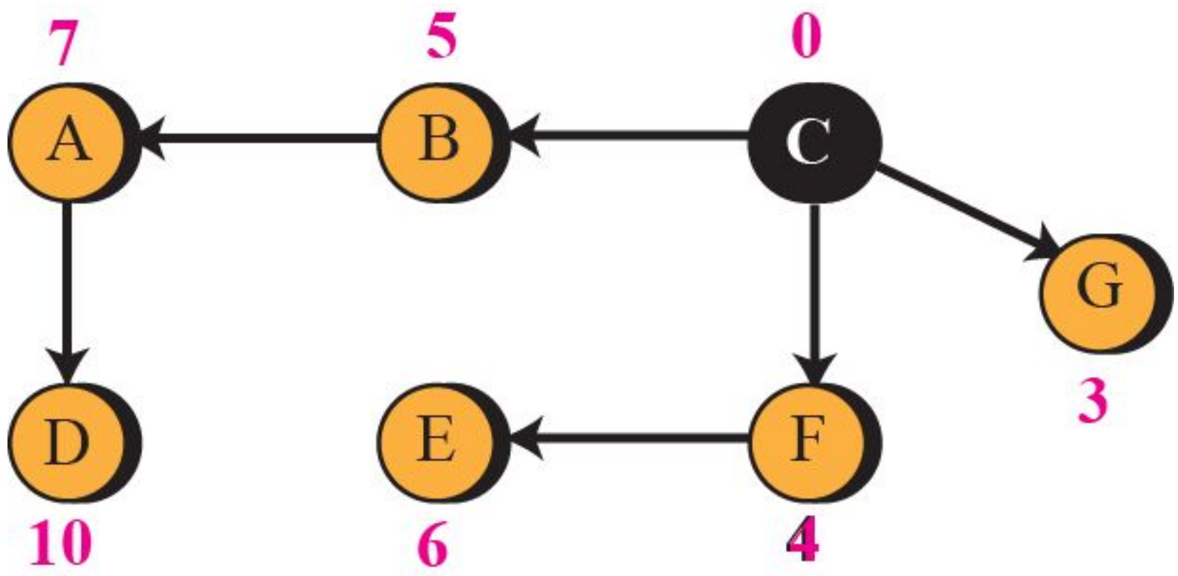


Table 11.4 *Routing Table for Node A*

Destination	*Cost*	*Next Router*
A	0	—
B	2	—
C	7	B
D	3	—
E	6	B
F	8	B
G	9	B

6 OSPF

The Open Shortest Path First (OSPF) protocol is an intra-domain routing protocol based on link state routing. Its domain is also an autonomous system.

- ✔ Area
- ✔ Metric
- ✔ Types of Links
- ✔ Graphical Representation
- ✔ OSPF Packets
- ✔ Link State Update Packet
- ✔ Other Packets
- ✔ Encapsulation

Figure 21 *Areas in an autonomous system*

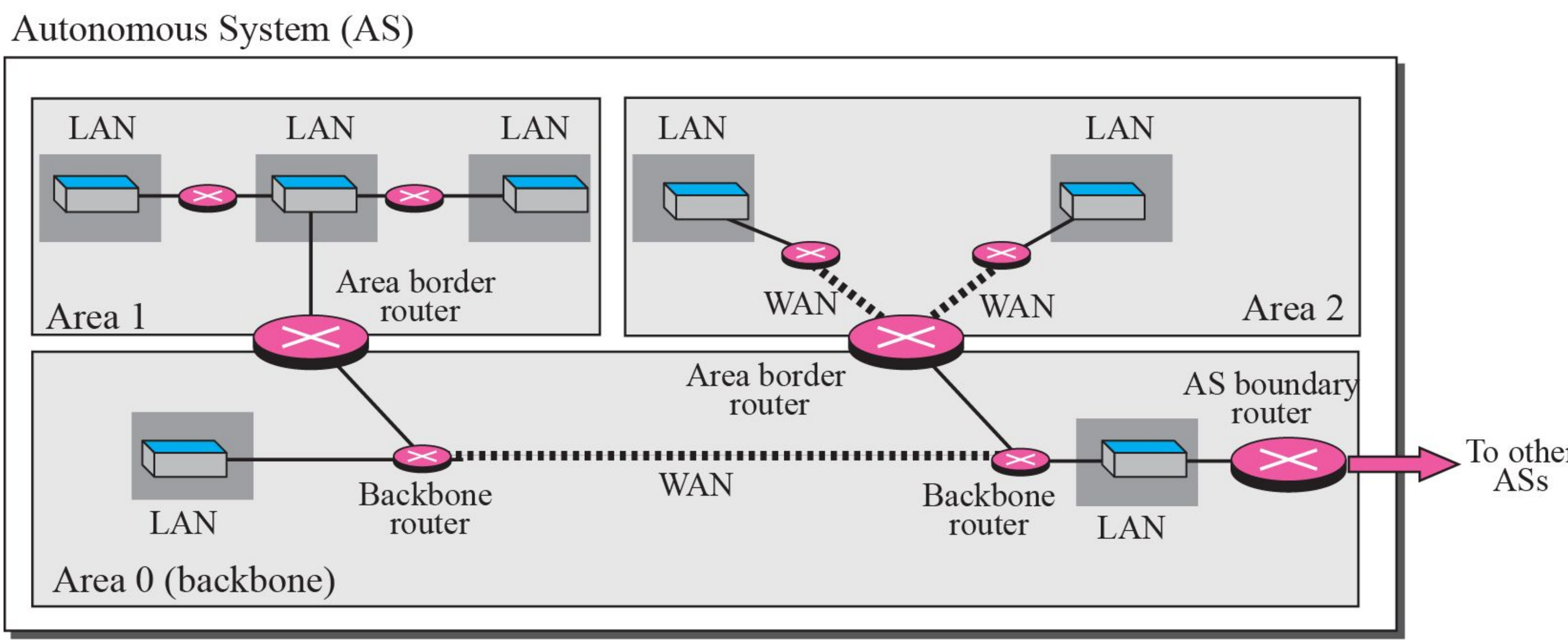


Figure 22 ***Types of links***

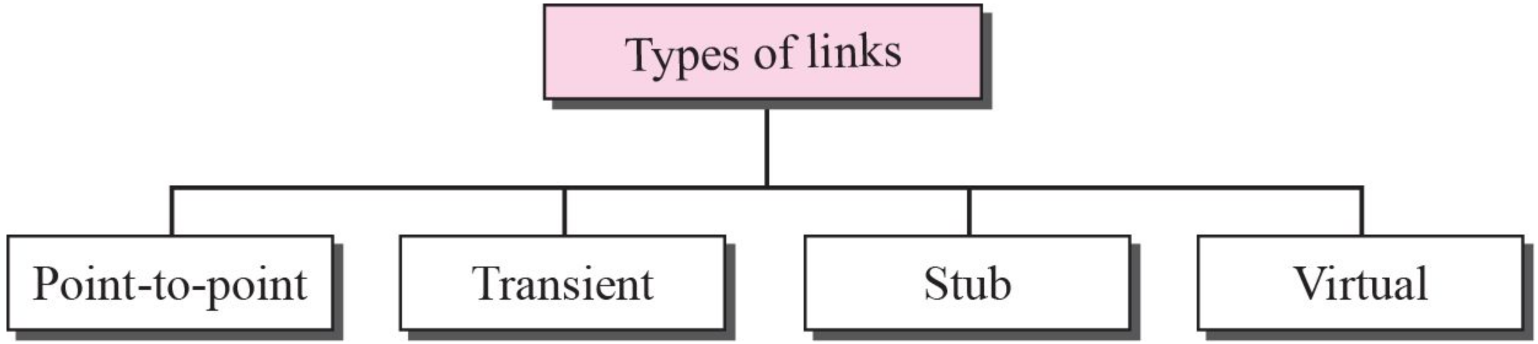


Figure 23 *Point-to-point link*

a. Point-to-point network

b. Representation

Figure 24 *Transient link*

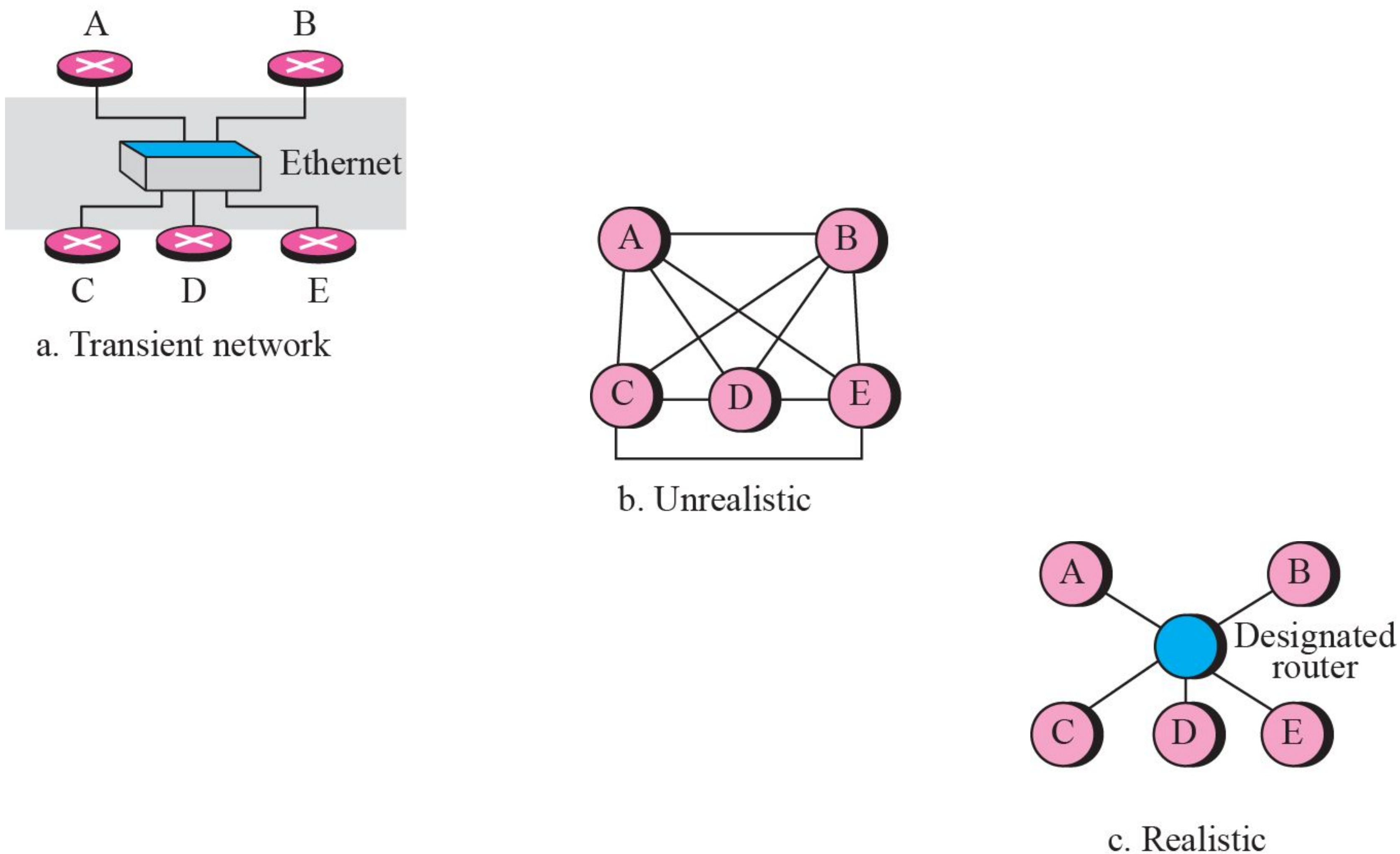
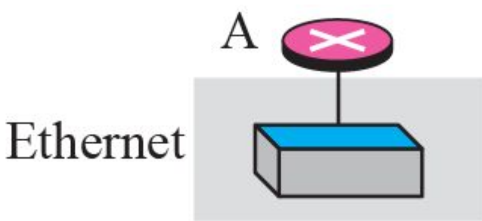
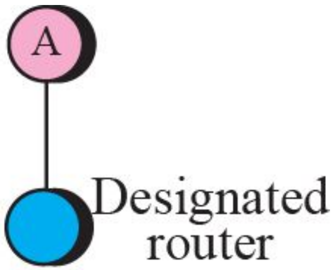


Figure 25 ***Stub link***



a. Stub network



b. Representation

Figure 26 *Example of an AS and its graphical representation in OSPF*

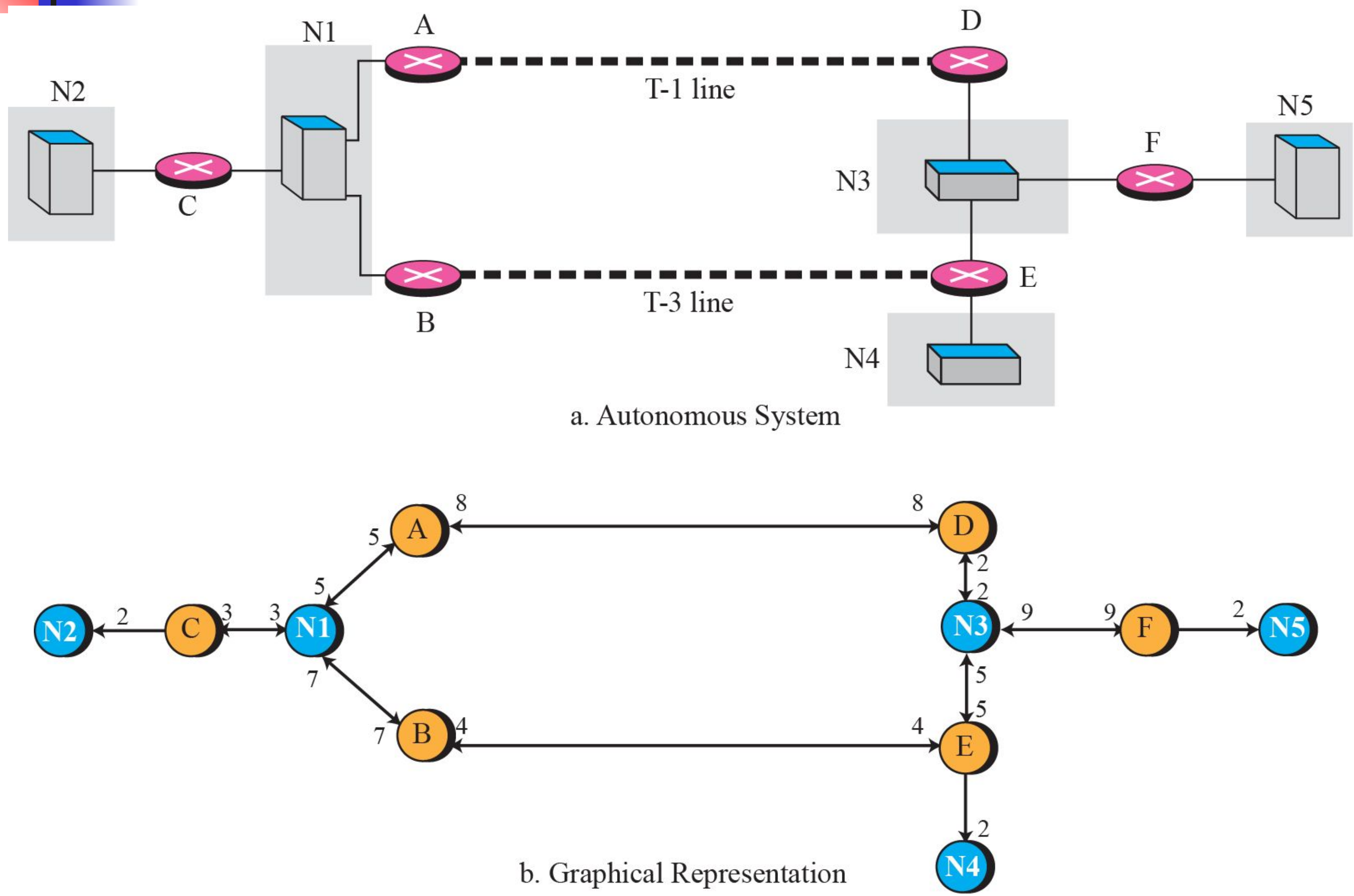


Figure 27 ***Types of OSPF packet***

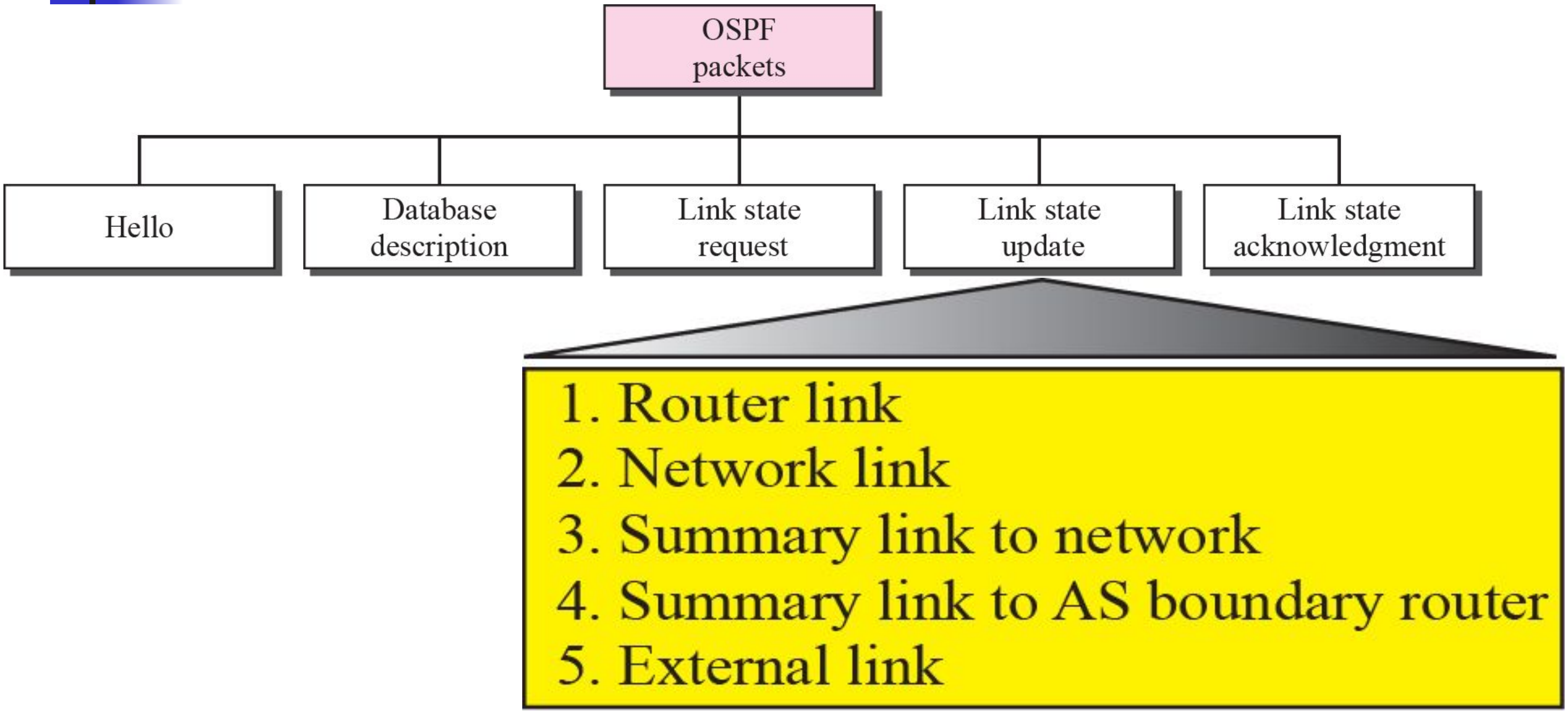


Figure 28 *OSPF common header*

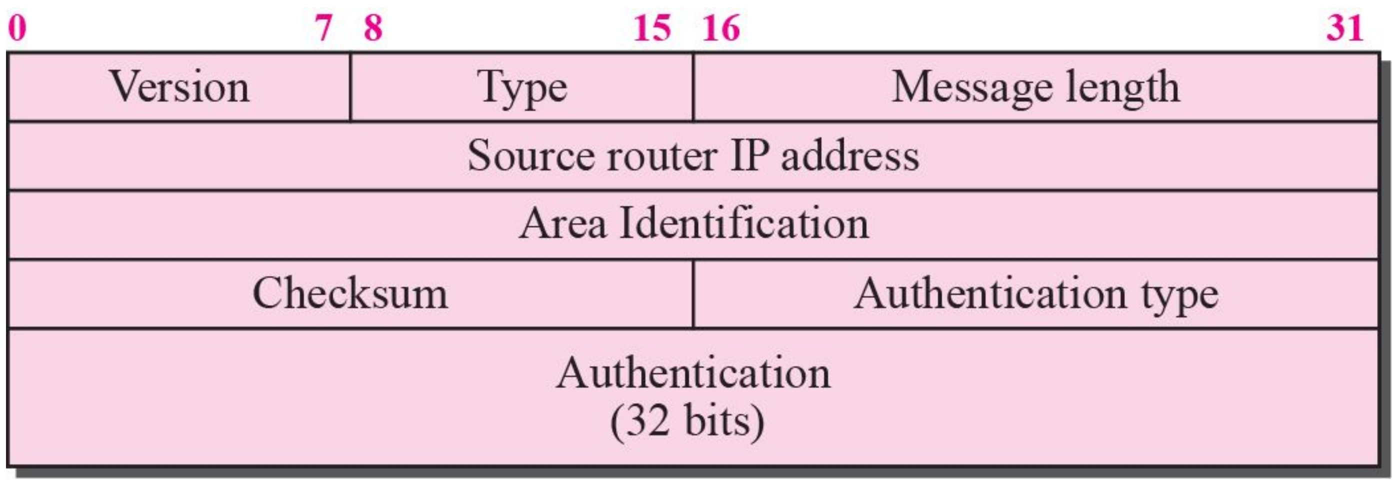


Figure 29 ***Link state update packet***

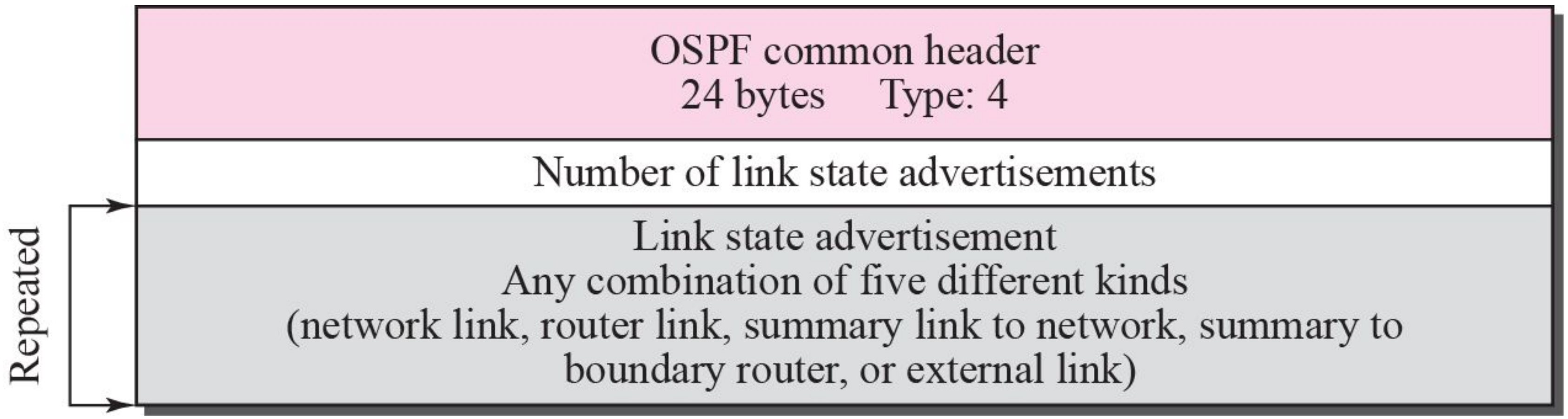


Figure 30 *LSA general header*

<table>
<tr><td>Link state age</td><td>Reserved</td><td>E</td><td>T</td><td>Link state type</td></tr>
<tr><td colspan="5">Link state ID</td></tr>
<tr><td colspan="5">Advertising router</td></tr>
<tr><td colspan="5">Link state sequence number</td></tr>
<tr><td>Link state checksum</td><td colspan="4">Length</td></tr>
</table>

Figure 31 ***Router link***

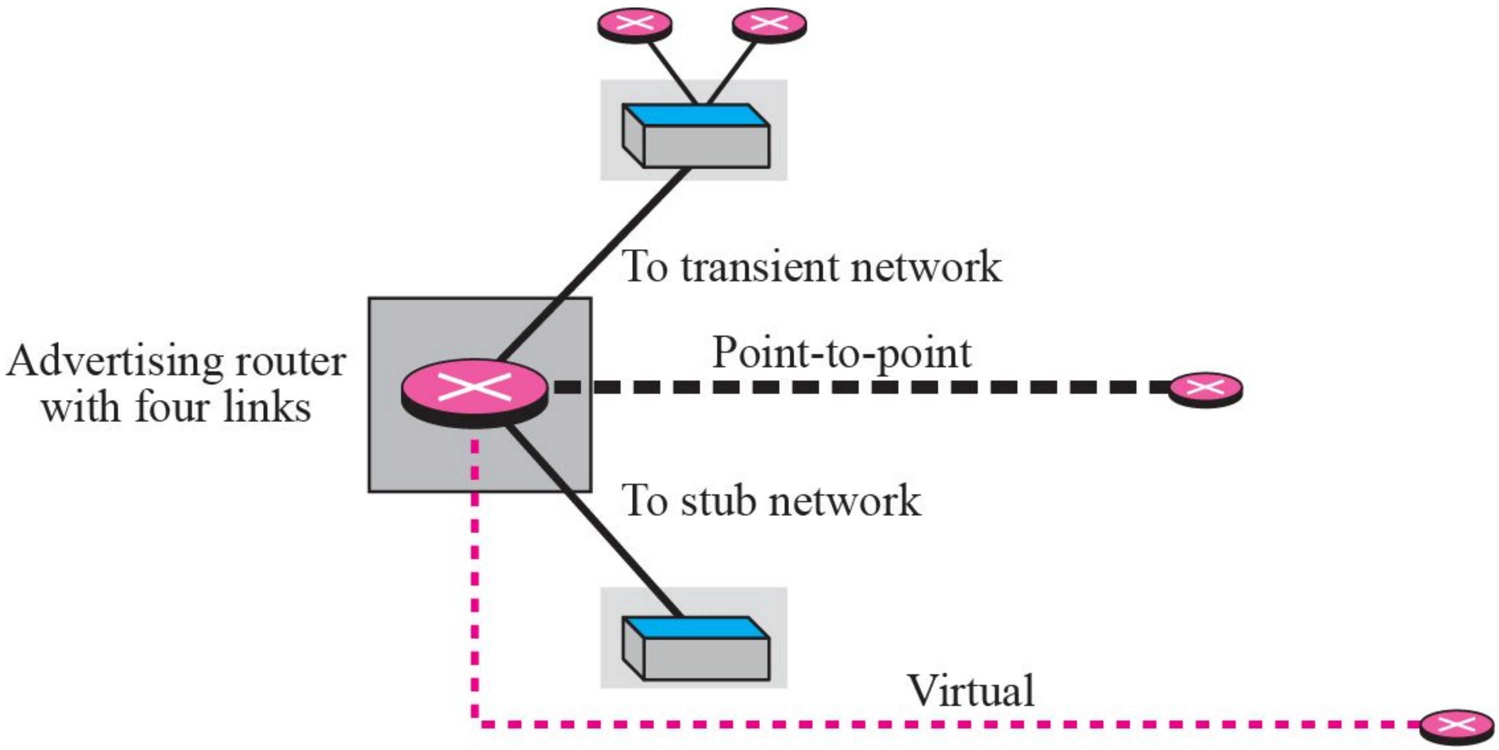


Figure 32 *Router link LSA*

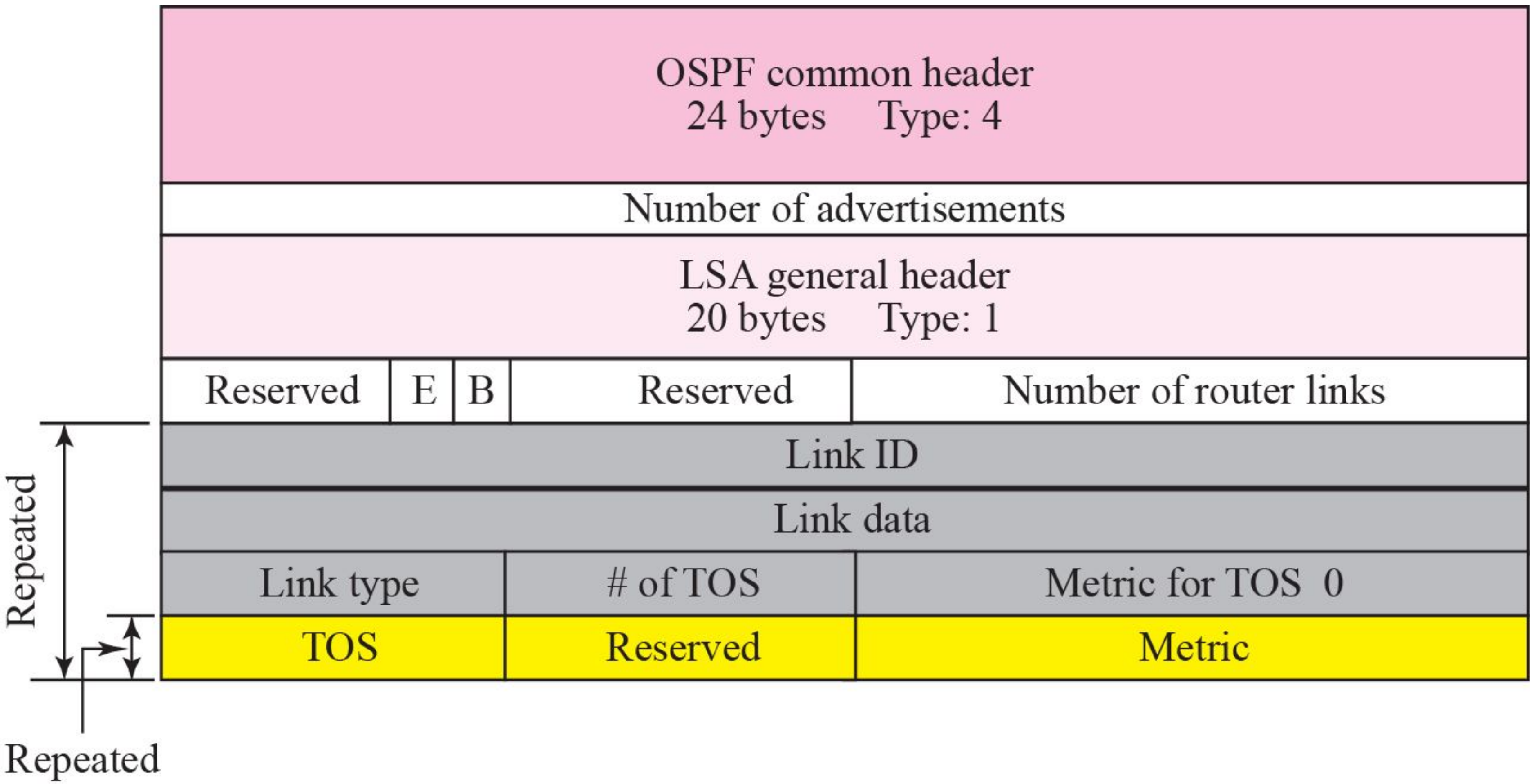


Table 11.5 *Link Types, Link Identification, and Link Data*

Link Type	*Link Identification*	*Link Data*
Type 1: Point-to-point	Address of neighbor router	Interface number
Type 2: Transient	Address of designated router	Router address
Type 3: Stub	Network address	Network mask
Type 4: Virtual	Address of neighbor router	Router address

Example 7

Figure 7 shows the final routing tables for routers in Figure 5.

Solution

This router has three links: two of type 1 (point-to-point) and one of type 3 (stub network). Figure 34 shows the router link LSA.

Figure 33 *Example 7*

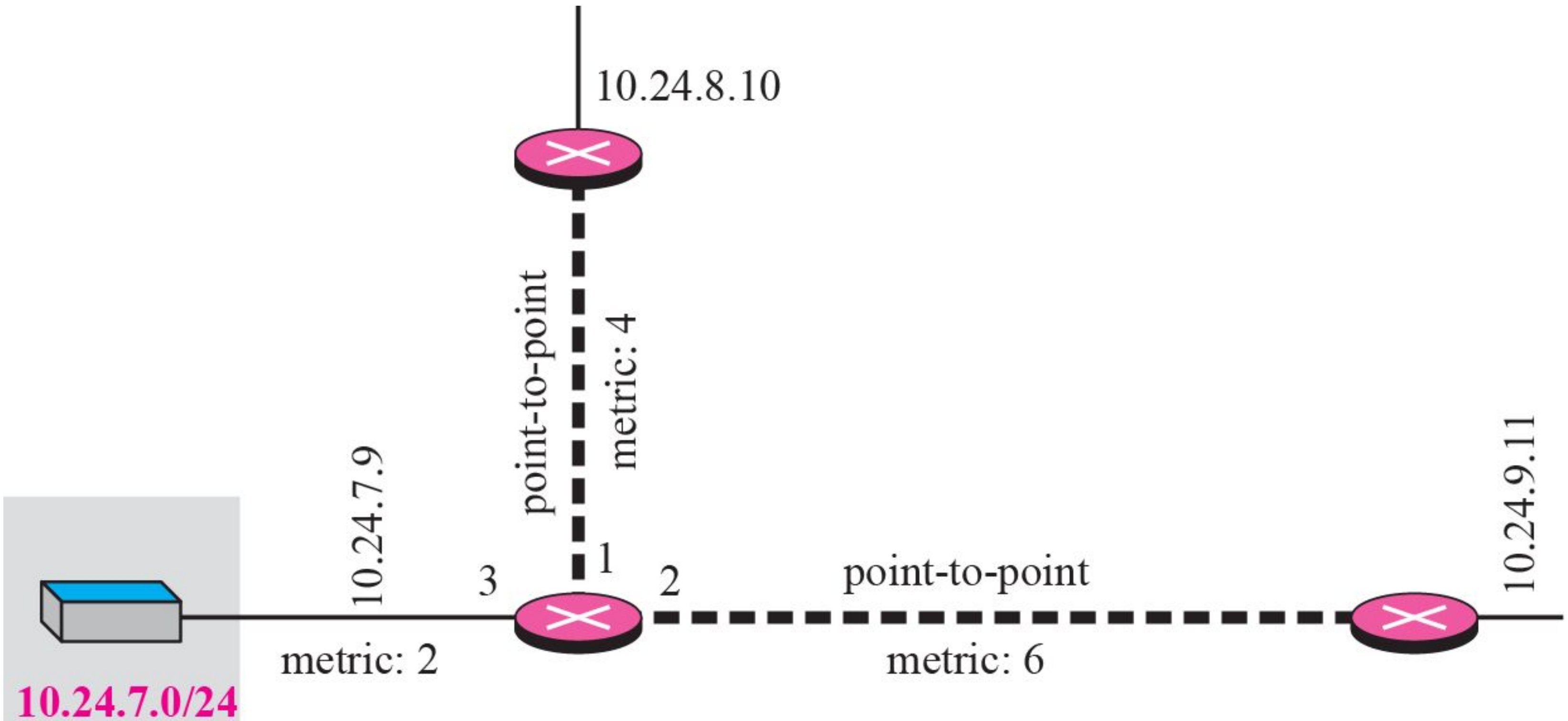


Figure 34 *Solution to Example 7*

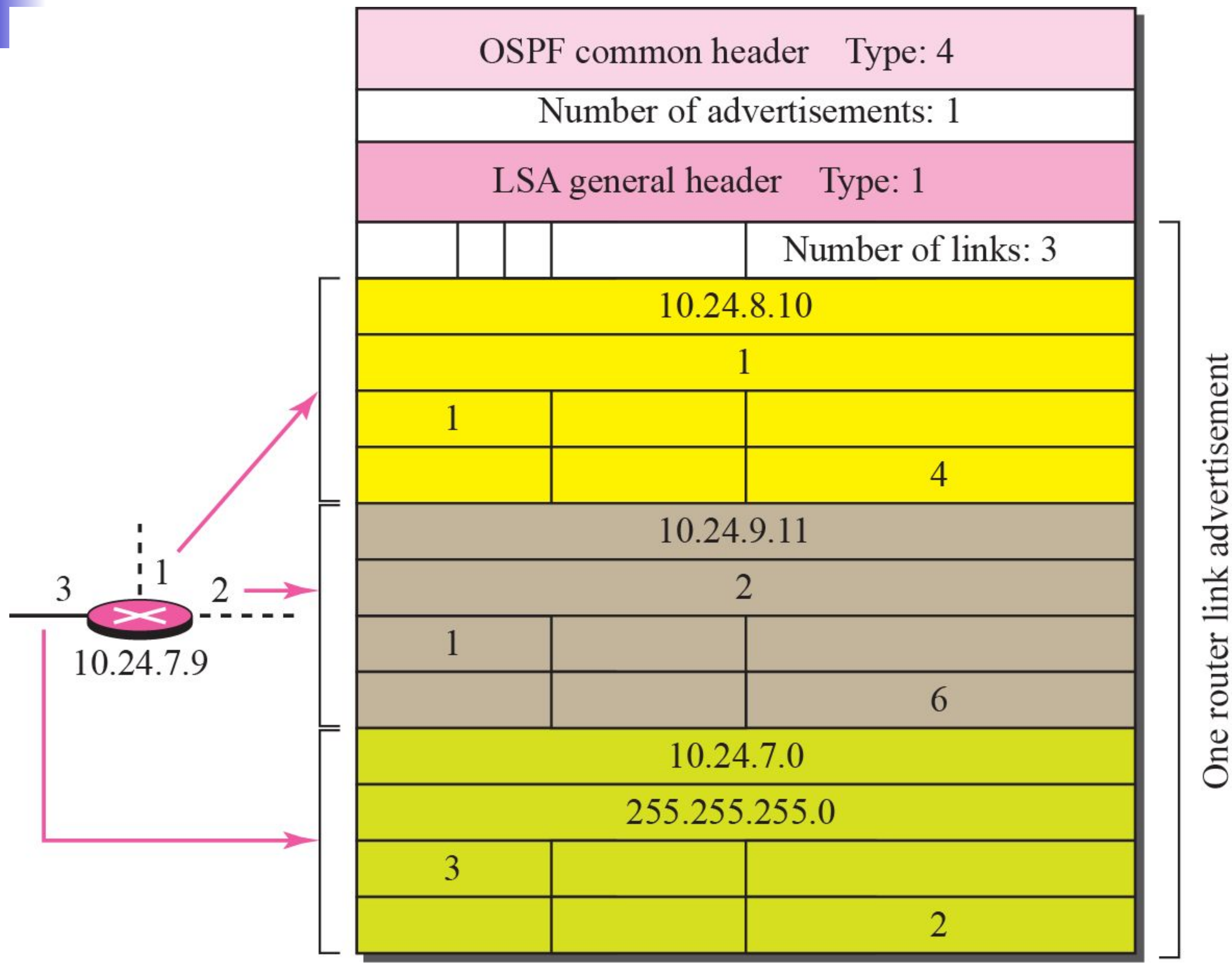


Figure 35 *Network link*

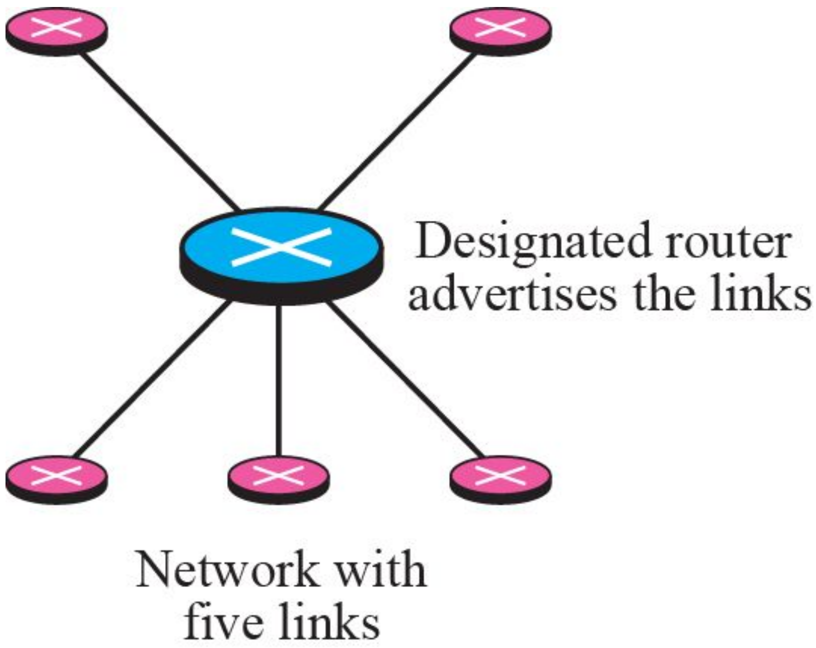
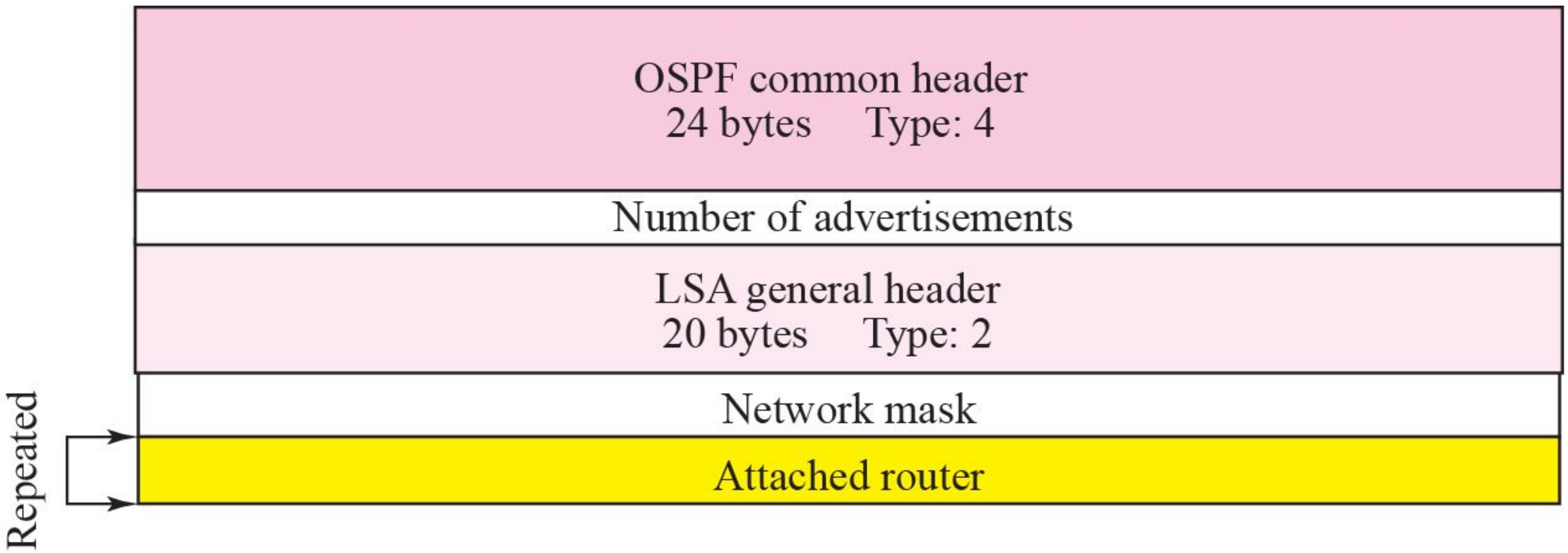


Figure 36 *Network link advertisement format*



Example 8

Give the network link LSA in Figure 37.

Solution

The network for which the network link advertises has three routers attached. The LSA shows the mask and the router addresses. Figure 38 shows the network link LSA.

Figure 37 *Example 8*

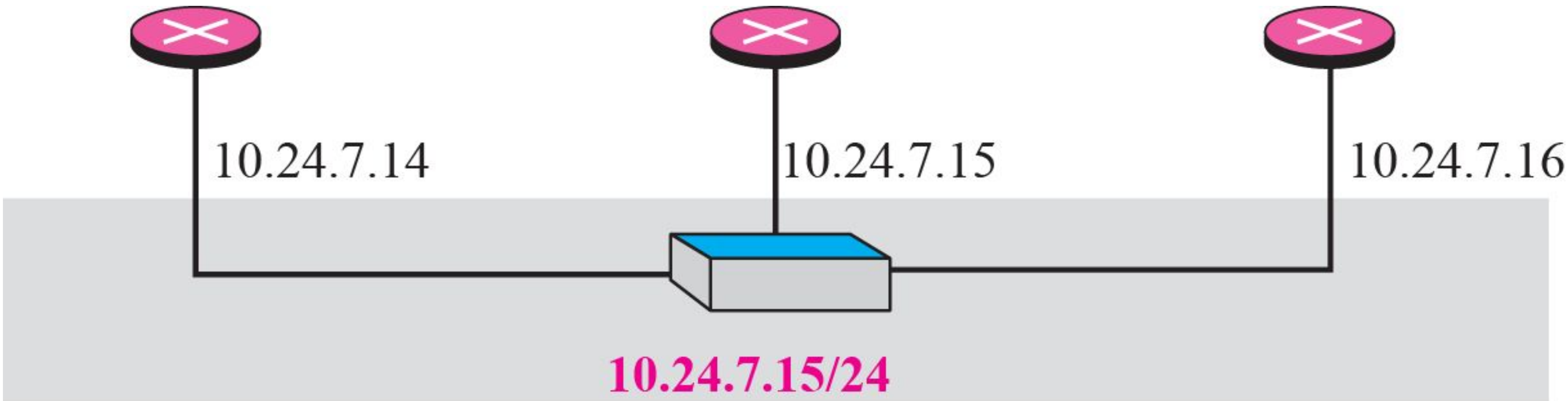


Figure 38 ***Solution to Example 8***

OSPF common header	Type: 4
Number of advertisements: 1	
LSA general header	Type: 2
255.255.255.0	
10.24.7.14	
10.24.7.15	
10.24.7.16	

Example 9

In Figure 39, which router(s) sends out router link LSAs?

Solution

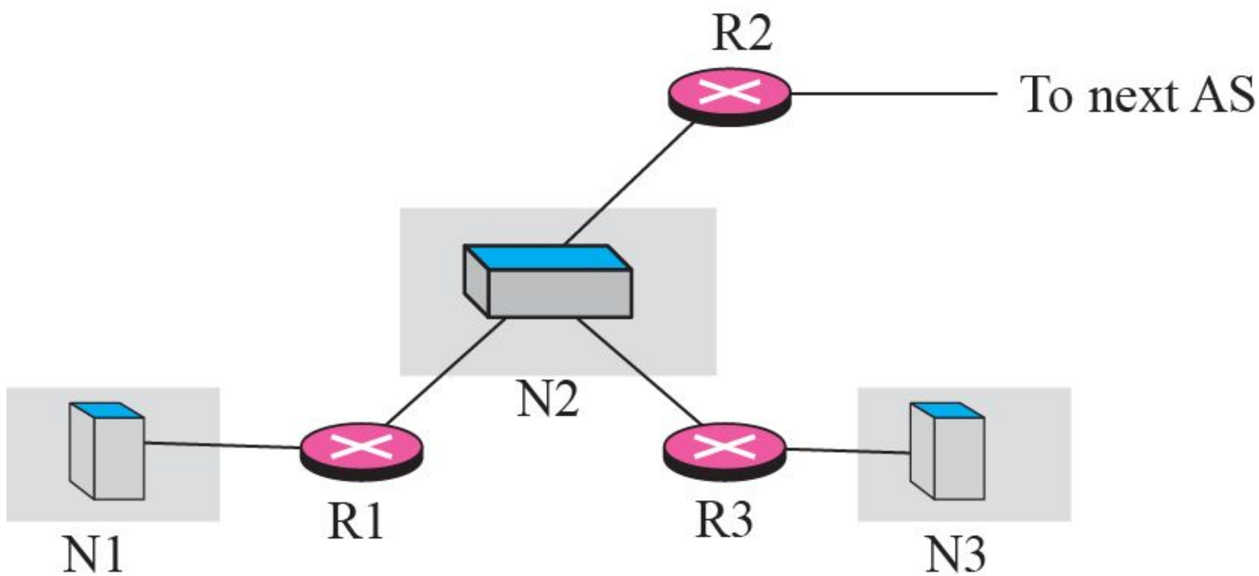
All routers advertise router link LSAs.

a. R1 has two links, N1 and N2.

b. R2 has one link, N2.

c. R3 has two links, N2 and N3.

Figure 39 *Examples 9 and 10*



Example 10

In Figure 39, which router(s) sends out the network link LSAs?

Solution

All three networks must advertise network links:

a. Advertisement for N1 is done by R1 because it is the only attached router and therefore the designated router.

b. Advertisement for N2 can be done by either R1, R2, or R3, depending on which one is chosen as the designated router.

c. Advertisement for N3 is done by R3 because it is the only attached router and therefore the designated router

Figure 40 *Summary link to network*

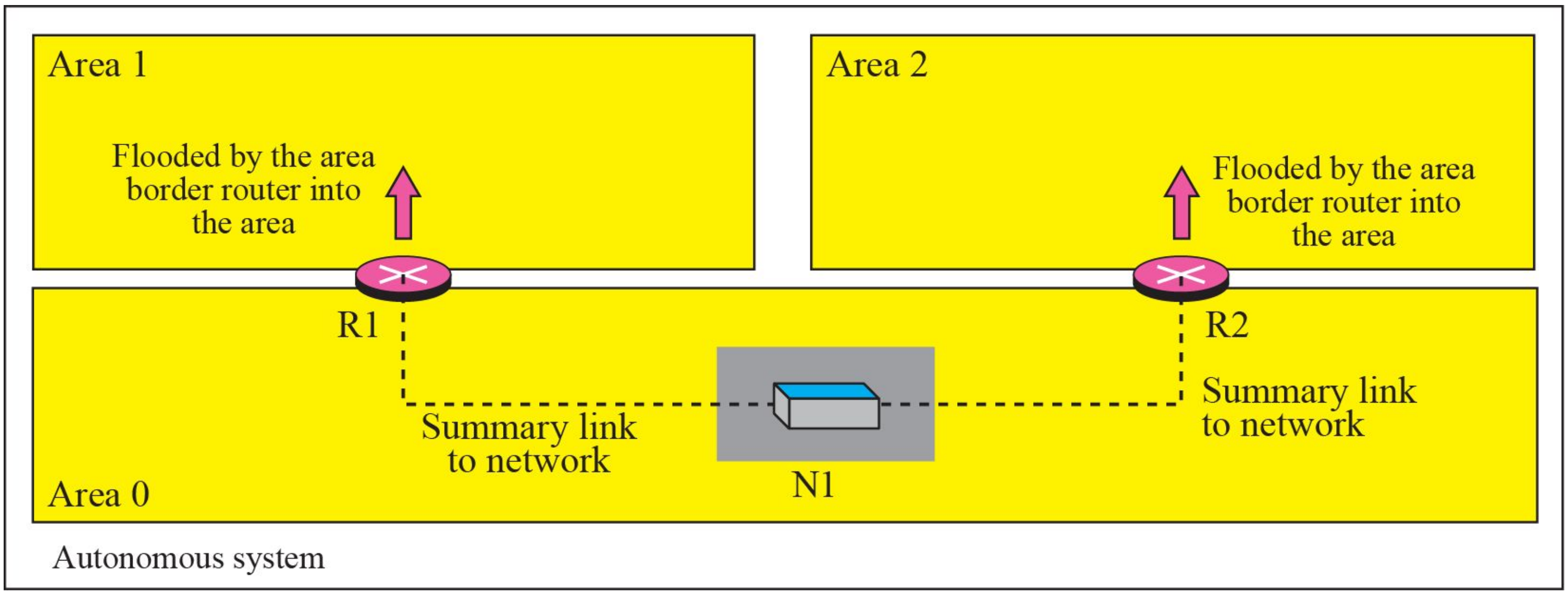


Figure 41 *Summary link to network LSA*

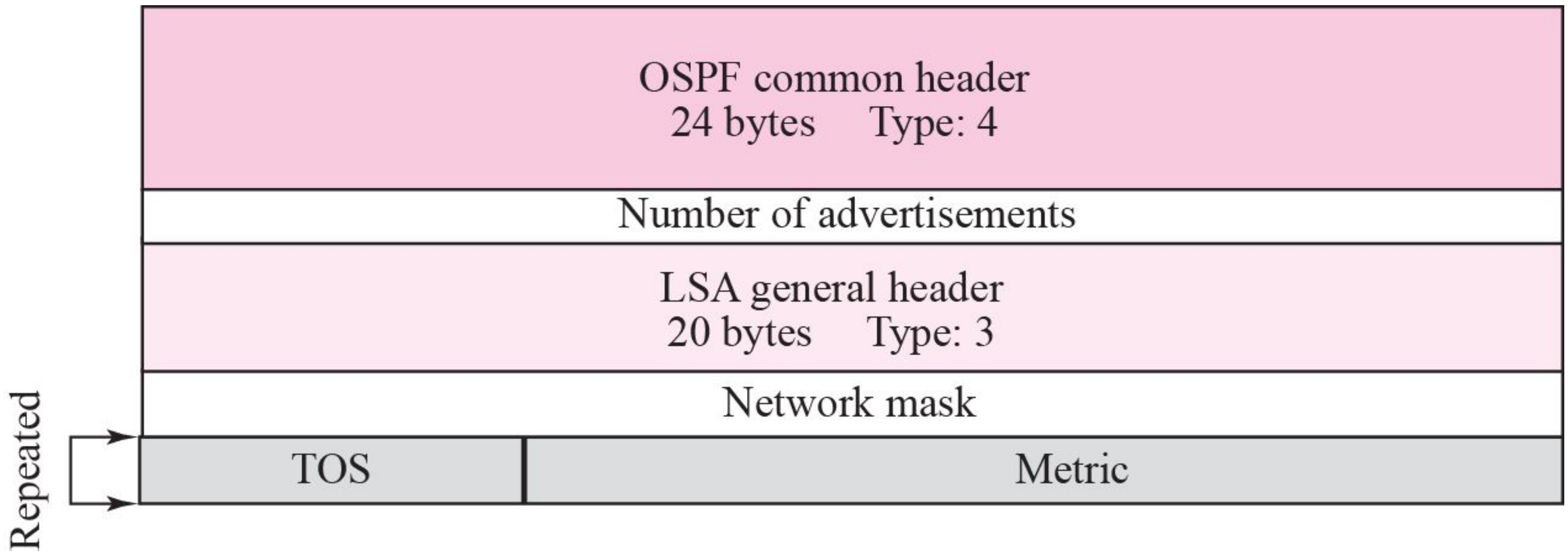


Figure 42 ***Summary link to AS boundary router***

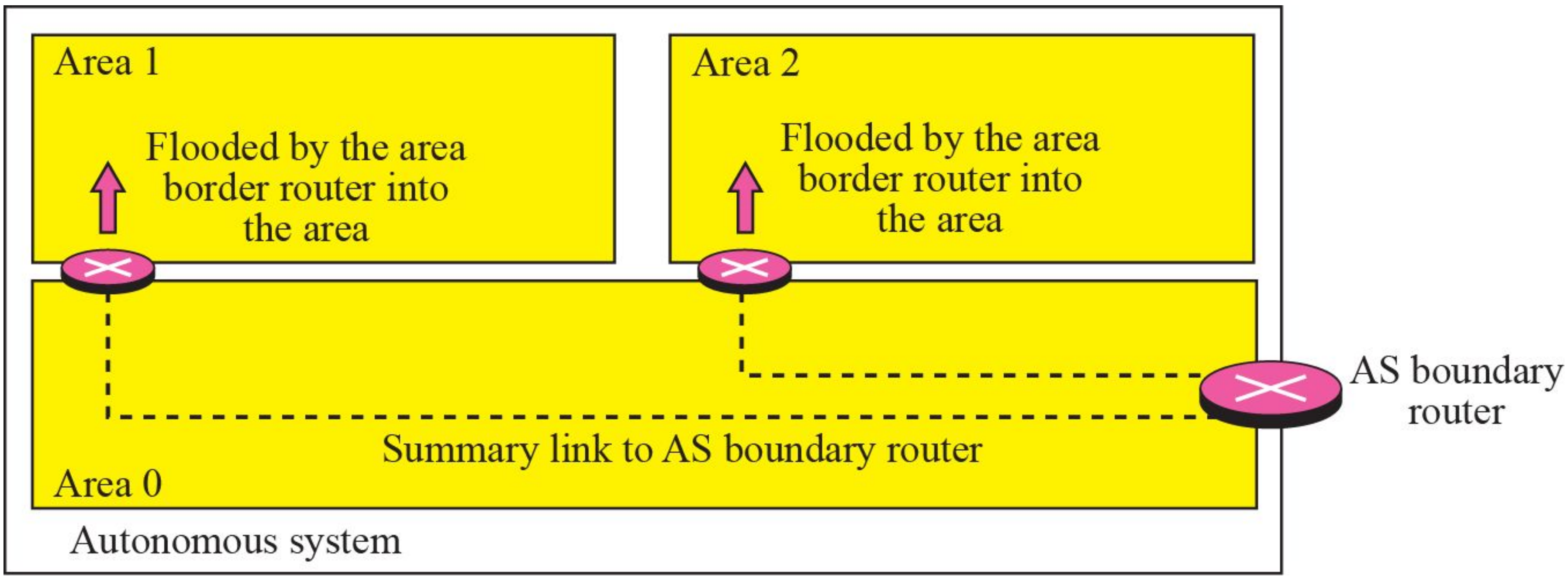


Figure 43 *Summary link to AS boundary router LSA*

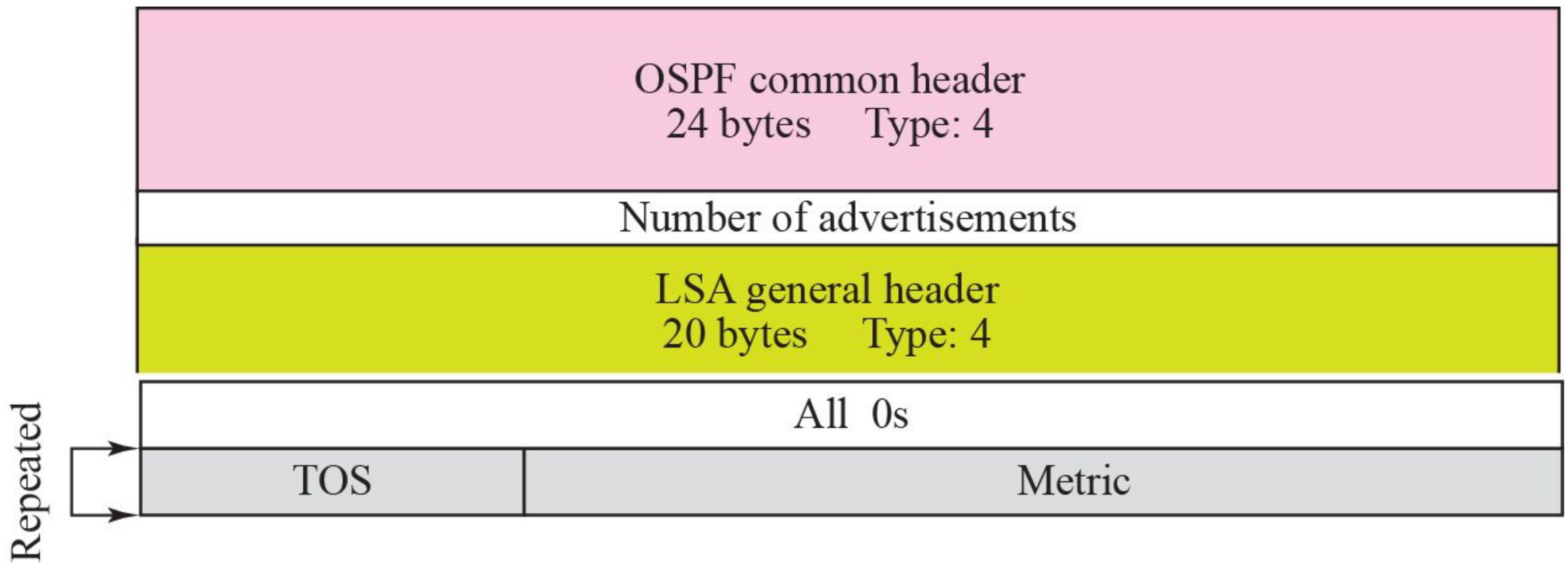


Figure 44 *External link*

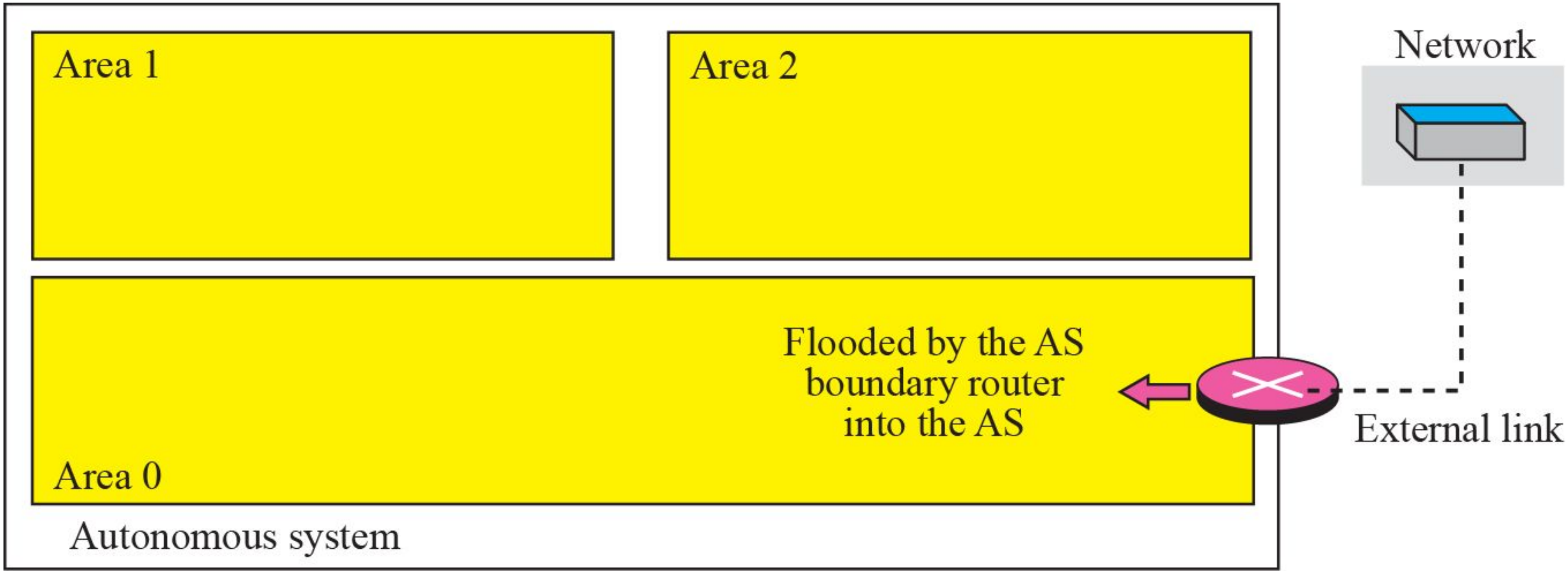


Figure 45 *External link LSA*

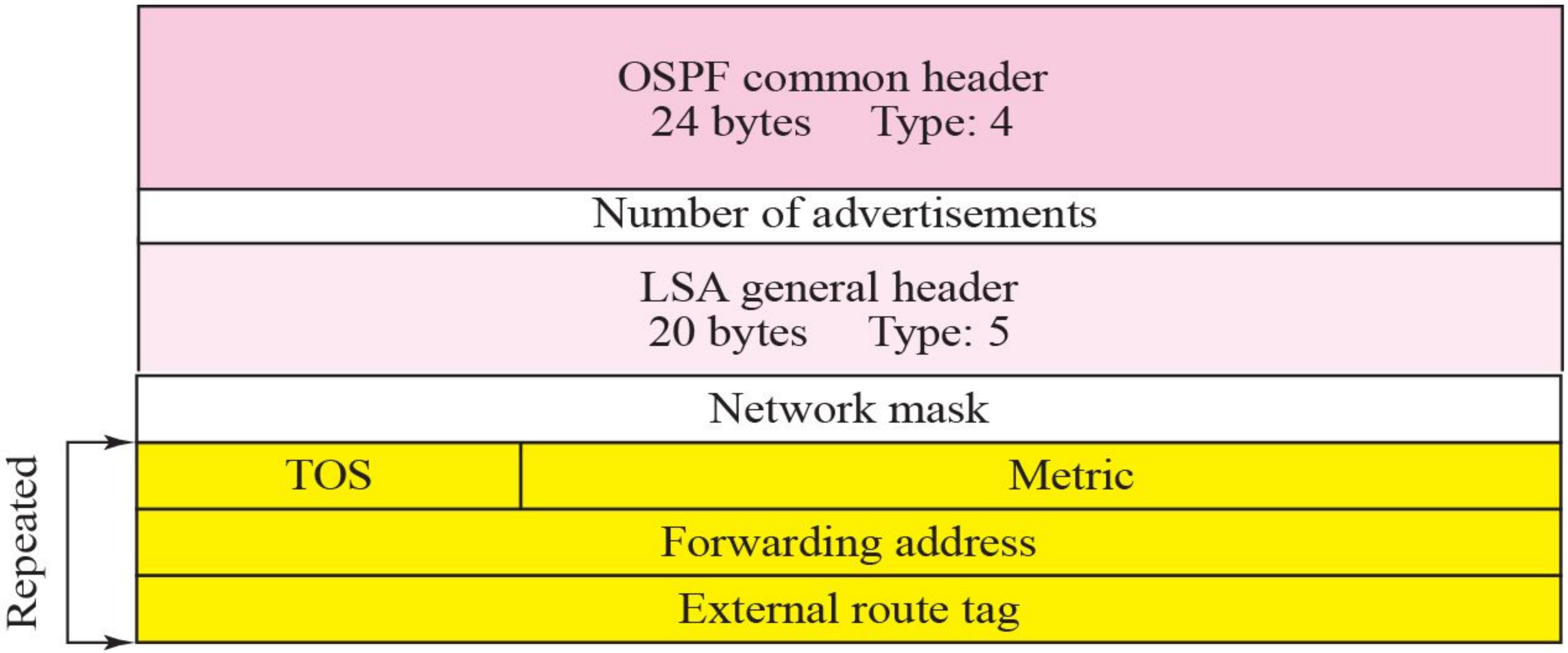


Figure 46 ***Hello packet***

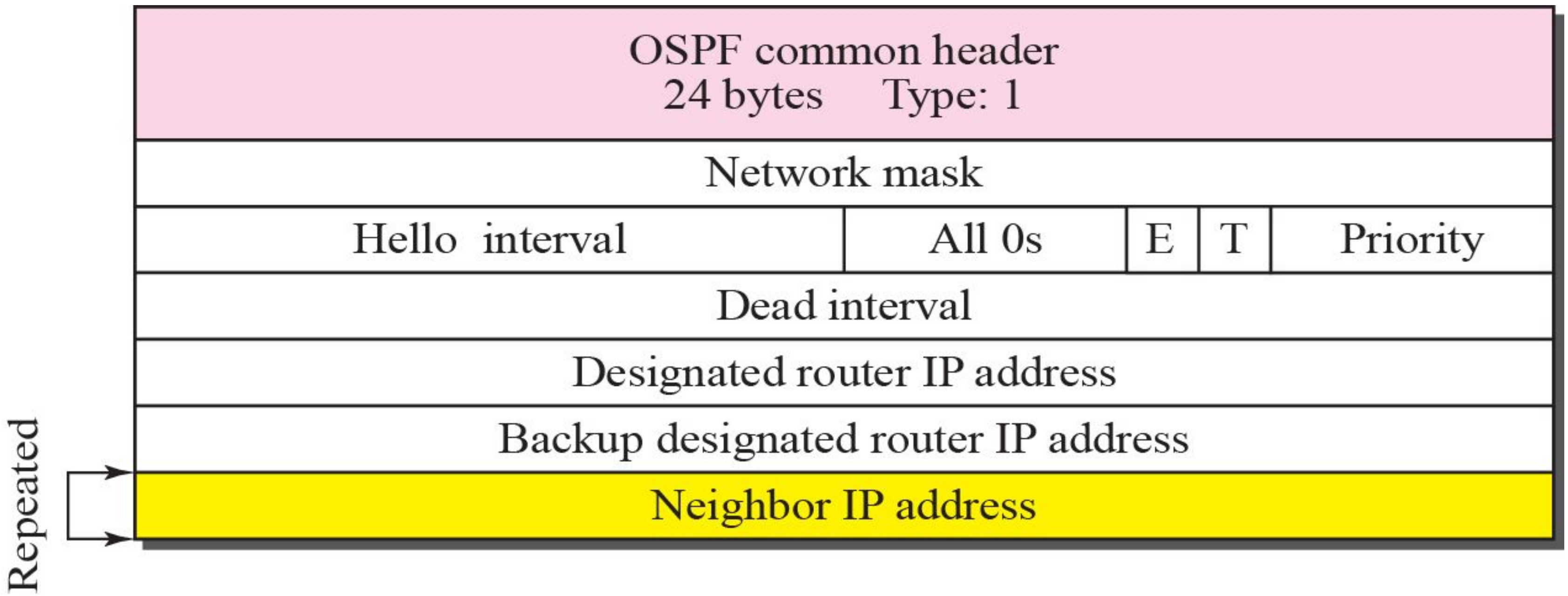


Figure 47 *Database description packet*

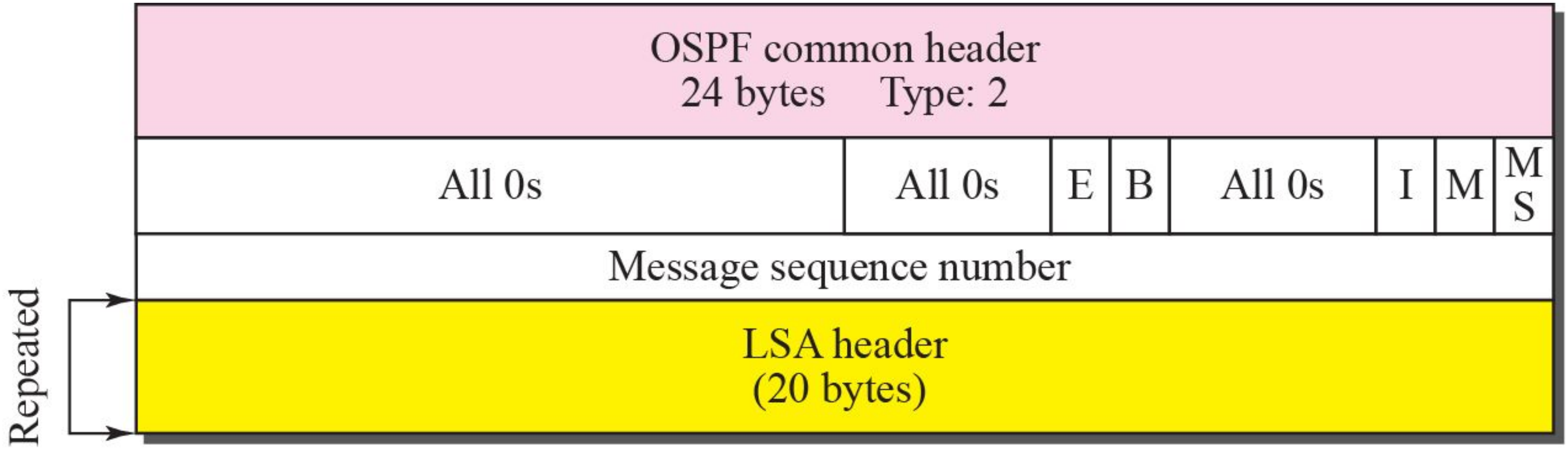


Figure 48 *Link state request packet*

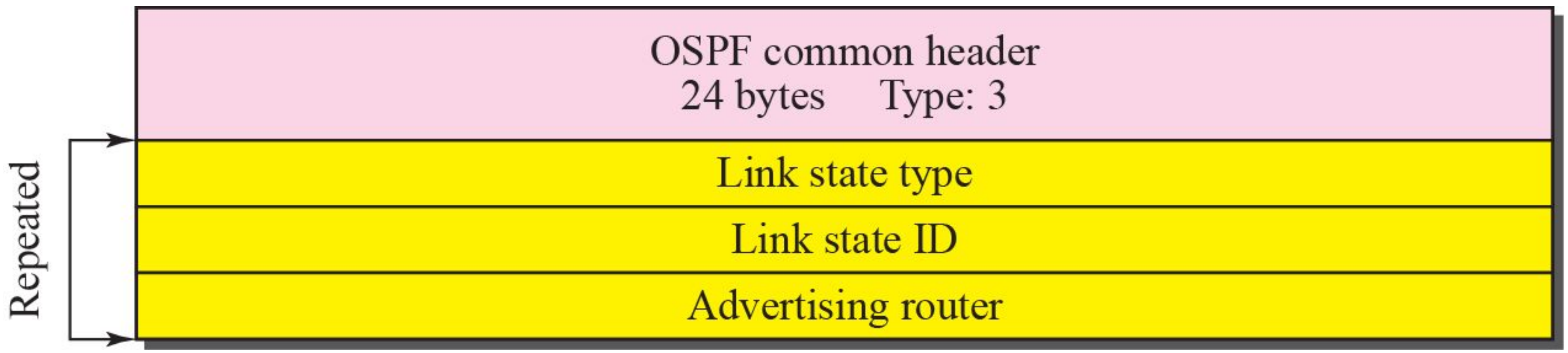
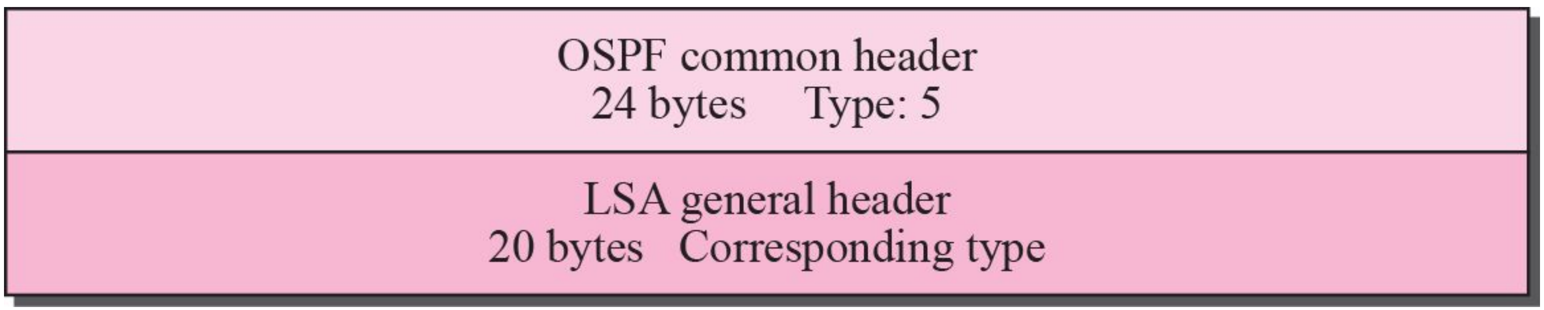


Figure 49 *Link state acknowledgment packet*



Note

OSPF packets are encapsulated in IP datagrams.

7 PATH VECTOR ROUTING

Distance vector and link state routing are both **interior routing protocols**.

They can be used inside an autonomous system. Both of these routing protocols become intractable when the domain of operation becomes large.

Distance vector routing is subject to instability if there is more than a few hops in the domain of operation.

Link state routing needs a huge amount of resources to calculate routing tables. It also creates heavy traffic because of flooding.

There is a need for a third routing protocol which we call path vector routing.

PATH VECTOR ROUTING

Routing takes place between the two autonomous networks.

DV and Link State Routing ignore the internet outside the autonomous system whereas Path Vector routing assumes that internet consists of a collection of interconnected autonomous systems.

Protocol for inter-domain routing are also called as **exterior gateway protocols.**

- ✔ **Reachability**
- ✔ **Routing Table**

Example 10

The difference between the distance vector routing and path vector routing can be compared to the difference between a national map and an international map.

A national map can tell us the road to each city and the distance to be traveled if we choose a particular route;

An international map can tell us which cities exist in each country and which countries should be passed before reaching that city.

Figure 50 *Reachability*

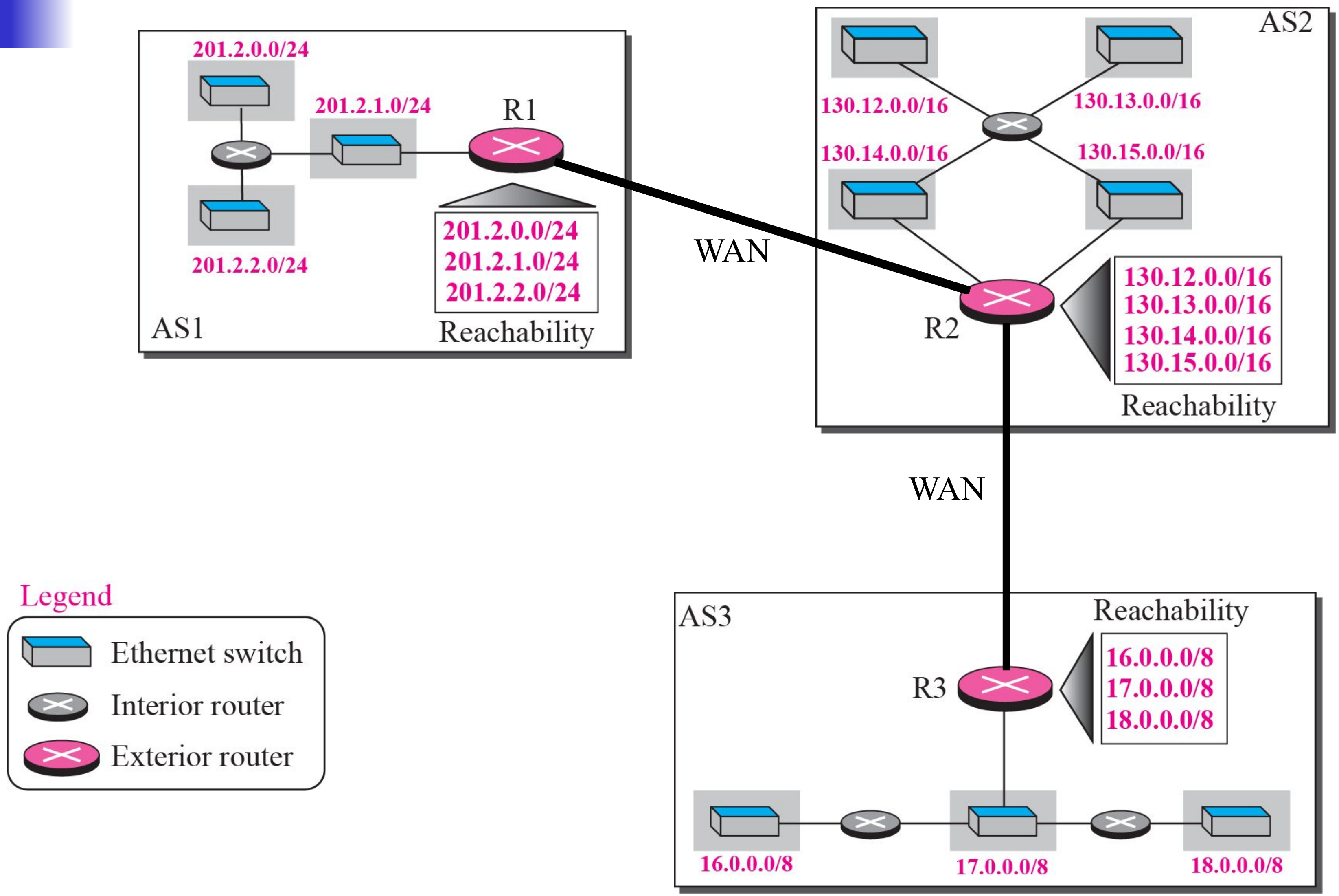


Figure 51 *Stabilized table for three autonomous system*

R1

Network	Path
201.2.0.0/24	**AS1** (This AS)
201.2.1.0/24	**AS1** (This AS)
201.2.2.0/24	**AS1** (This AS)
130.12.0.0/16	AS1, AS2
130.13.0.0/16	AS1, AS2
130.14.0.0/16	AS1, AS2
130.15.0.0/16	AS1, AS2
16.0.0.0/8	AS1, AS2, AS3
17.0.0.0/8	AS1, AS2, AS3
18.0.0.0/8	AS1, AS2, AS3

Path-Vector Routing Table

R2

Network	Path
201.2.0.0/24	AS2, AS1
201.2.1.0/24	AS2, AS1
201.2.2.0/24	AS2, AS1
130.12.0.0/16	**AS2** (This AS)
130.13.0.0/16	**AS2** (This AS)
130.14.0.0/16	**AS2** (This AS)
130.15.0.0/16	**AS2** (This AS)
16.0.0.0/8	AS2, AS3
17.0.0.0/8	AS2, AS3
18.0.0.0/8	AS2, AS3

Path-Vector Routing Table

R3

Network	Path
201.2.0.0/24	AS3, AS2, AS1
201.2.1.0/24	AS3, AS2, AS1
201.2.2.0/24	AS3, AS2, AS1
130.12.0.0/16	AS3, AS2
130.13.0.0/16	AS3, AS2
130.14.0.0/16	AS3, AS2
130.15.0.0/16	AS3, AS2
16.0.0.0/8	**AS3** (This AS)
17.0.0.0/8	**AS3** (This AS)
18.0.0.0/8	**AS3** (This AS)

Path-Vector Routing Table

Figure 52 *Routing tables after aggregation*

R1

Network	Path
201.2.0.0/22	**AS1** (This AS)
130.12.0.0/18	AS1, AS2
16.0.0.0/6	AS1, AS2, AS3

Path-Vector Routing Table

R2

Network	Path
201.2.0.0/22	AS2, AS1
130.12.0.0/18	**AS2** (This AS)
16.0.0.0/6	AS2, AS3

Path-Vector Routing Table

R3

Network	Path
201.2.0.0/22	AS3, AS2, AS1
130.12.0.0/18	AS3, AS2
16.0.0.0/6	**AS3** (This AS)

Path-Vector Routing Table

8 BGP

Border Gateway Protocol (BGP) is an interdomain routing protocol using path vector routing. It first appeared in 1989 and has gone through four versions.

- ✔ **Types of Autonomous Systems**
- ✔ **Path Attributes**
- ✔ **BGP Sessions**
- ✔ **External and Internal BGP**
- ✔ **Types of Packets**
- ✔ **Packet Format**
- ✔ **Encapsulation**

Figure 53 *Internal and external BGP sessions*

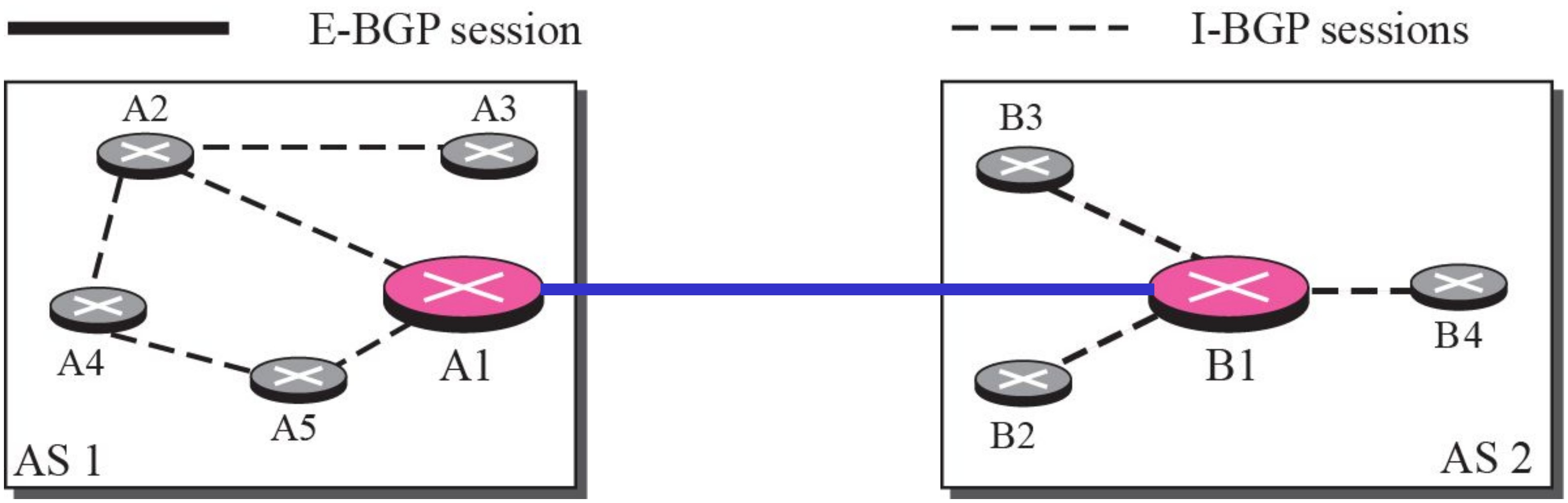


Figure 54 *Types of BGP messages*

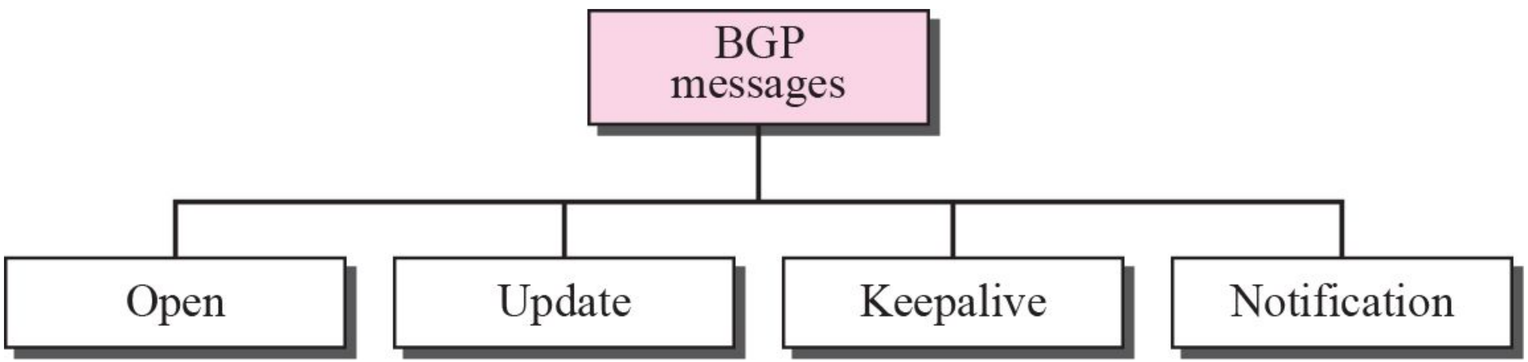


Figure 55 *BGP packet header*

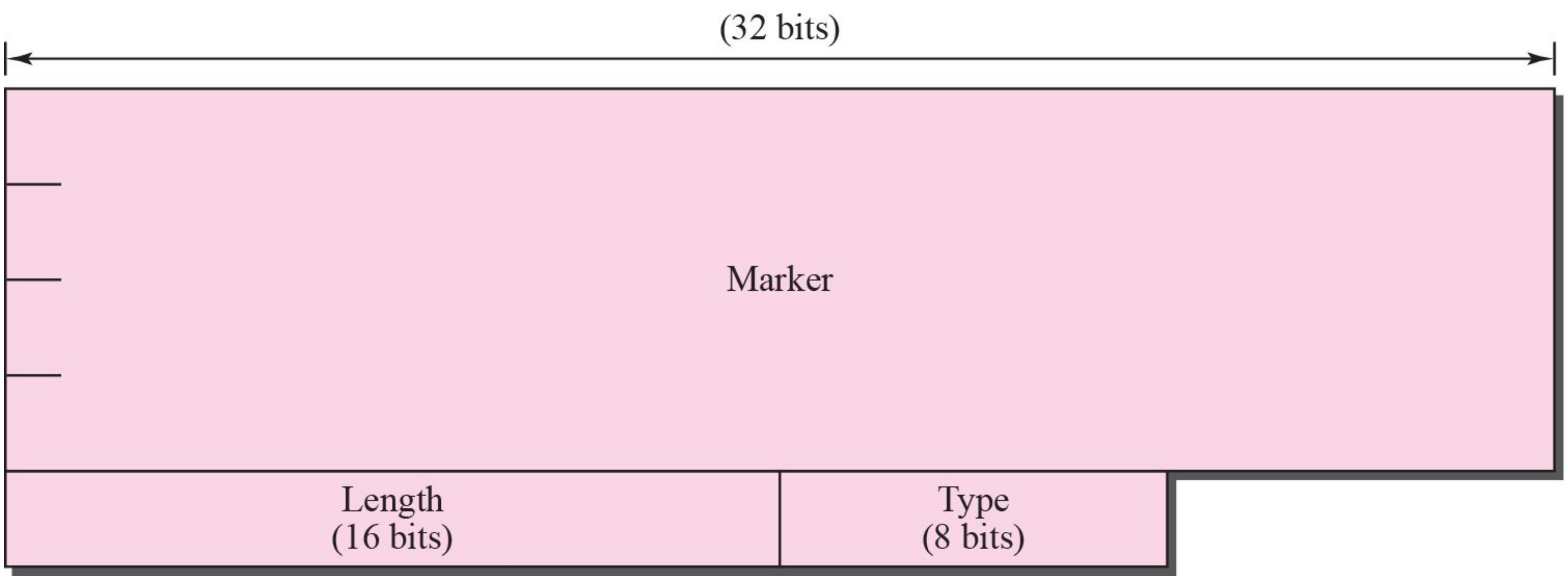


Figure 56 *Open message*

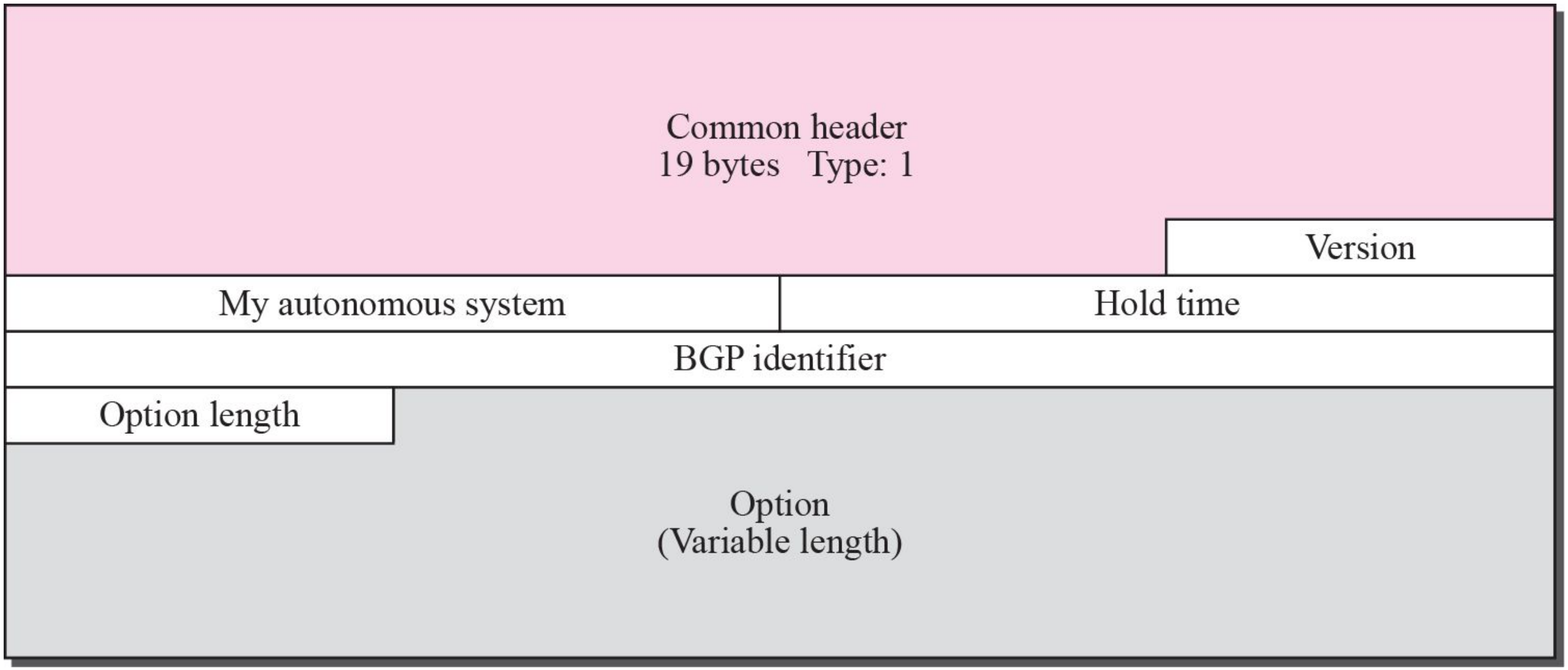
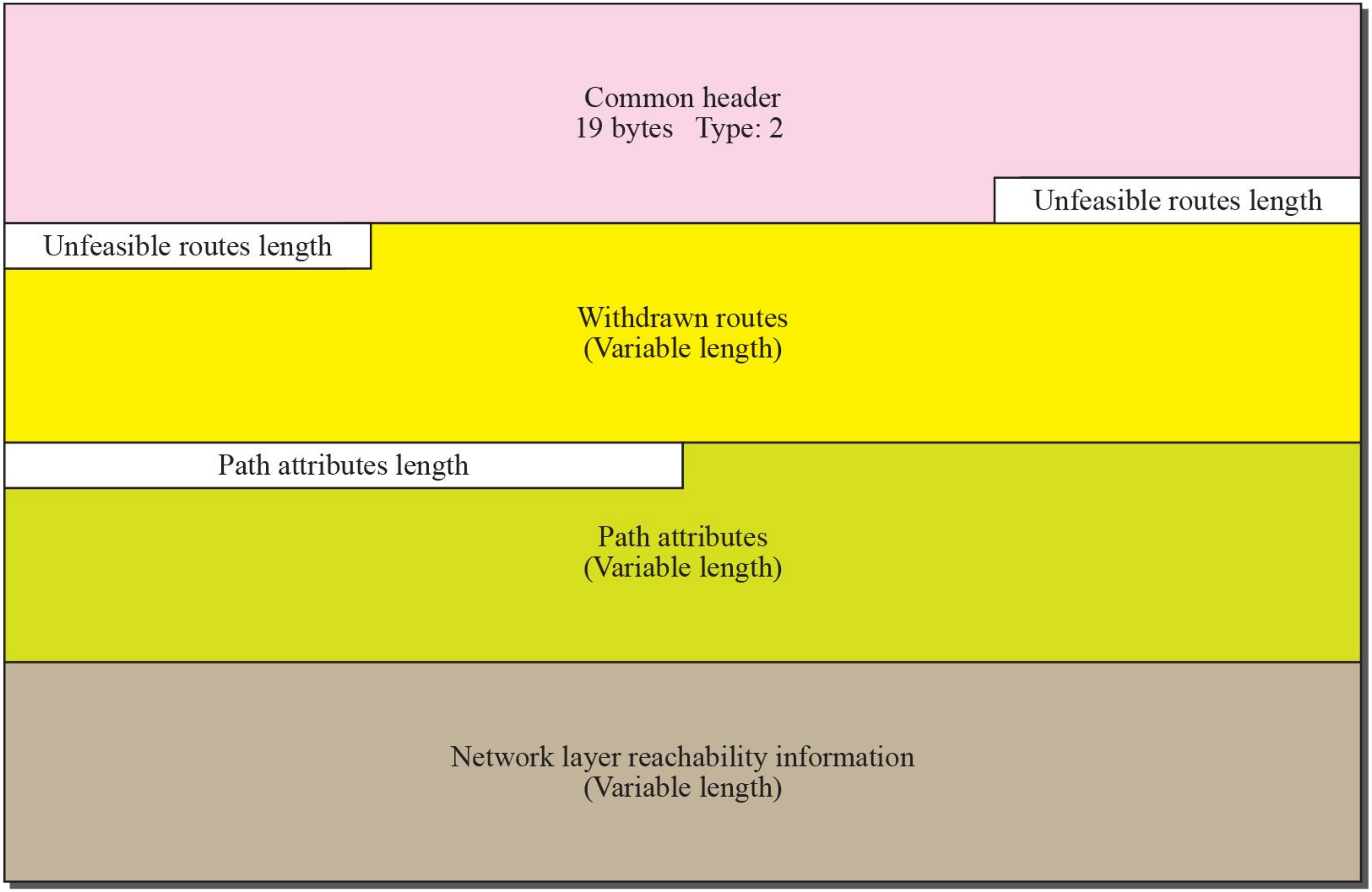


Figure 57 *Update message*



Note

BGP supports classless addressing .

Figure 58 ***Keepalive message***

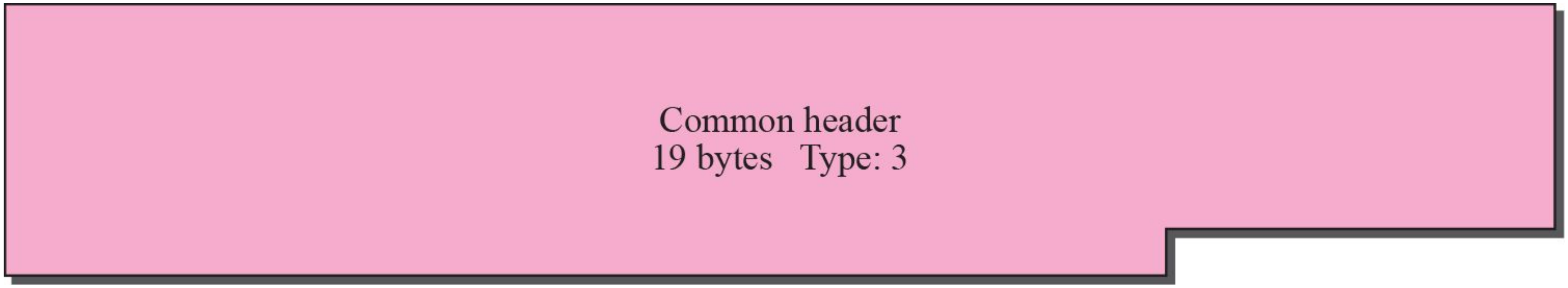


Figure 59 *Notification message*

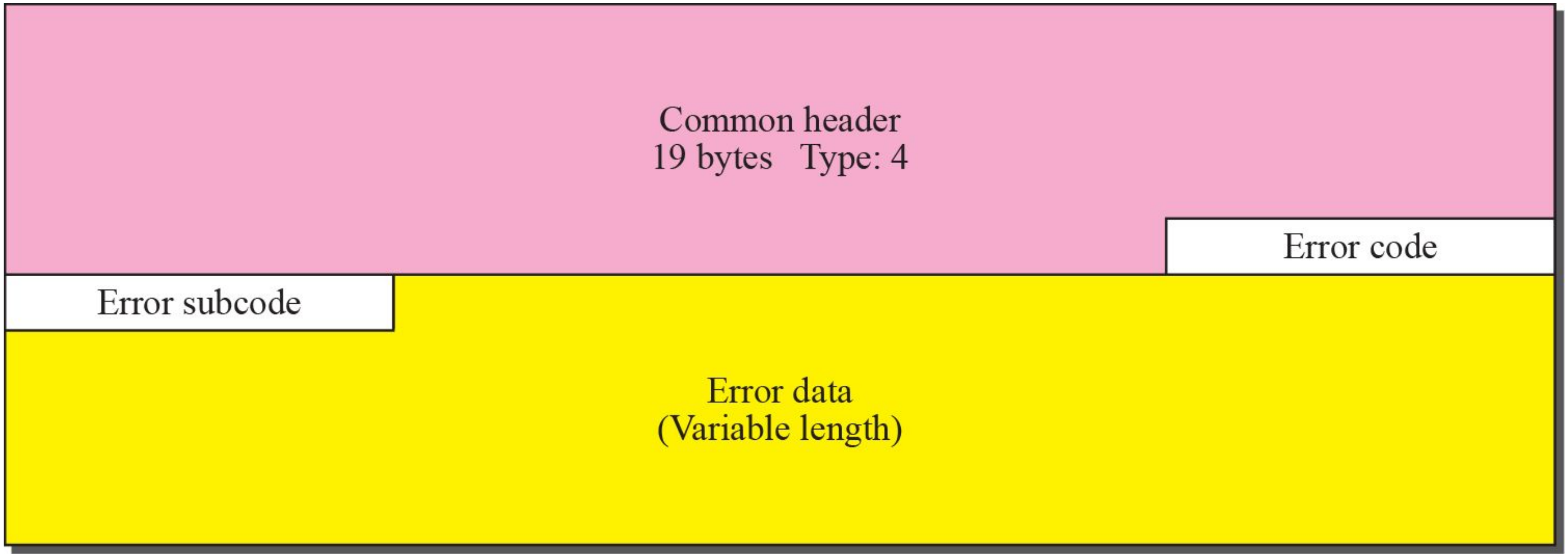


Table 11.6 *Error Codes*

Error Code	*Error Code Description*	*Error Subcode Description*
1	Message header error	Three different subcodes are defined for this type of error: synchronization problem (1), bad message length (2), and bad message type (3).
2	Open message error	Six different subcodes are defined for this type of error: unsupported version number (1), bad peer AS (2), bad BGP identifier (3), unsupported optional parameter (4), authentication failure (5), and unacceptable hold time (6).
3	Update message error	Eleven different subcodes are defined for this type of error: malformed attribute list (1), unrecognized well-known attribute (2), missing well-known attribute (3), attribute flag error (4), attribute length error (5), invalid origin attribute (6), AS routing loop (7), invalid next hop attribute (8), optional attribute error (9), invalid network field (10), malformed AS_PATH (11).
4	Hold timer expired	No subcode defined.
5	Finite state machine error	This defines the procedural error. No subcode defined.
6	Cease	No subcode defined.

Note

BGP uses the services of TCP on port 179.