Routing Protocols

(RIP, OSPF, and BGP)

Introduction

- An internet is a combination of networks connected by routers
- A ***metric*** is a cost assigned for passing through a network.
  - the total metric of a particular route is equal to the sum of the metrics of networks that comprise the route.
  - the router chooses the route with the shortest (smallest) metric
- RIP (Routing Information Protocol) : treating each network equals.
  - The cost of passing through each network is the same.
    - so if a packet passes through 10 networks to reach the destination, the total cost is hop counts.

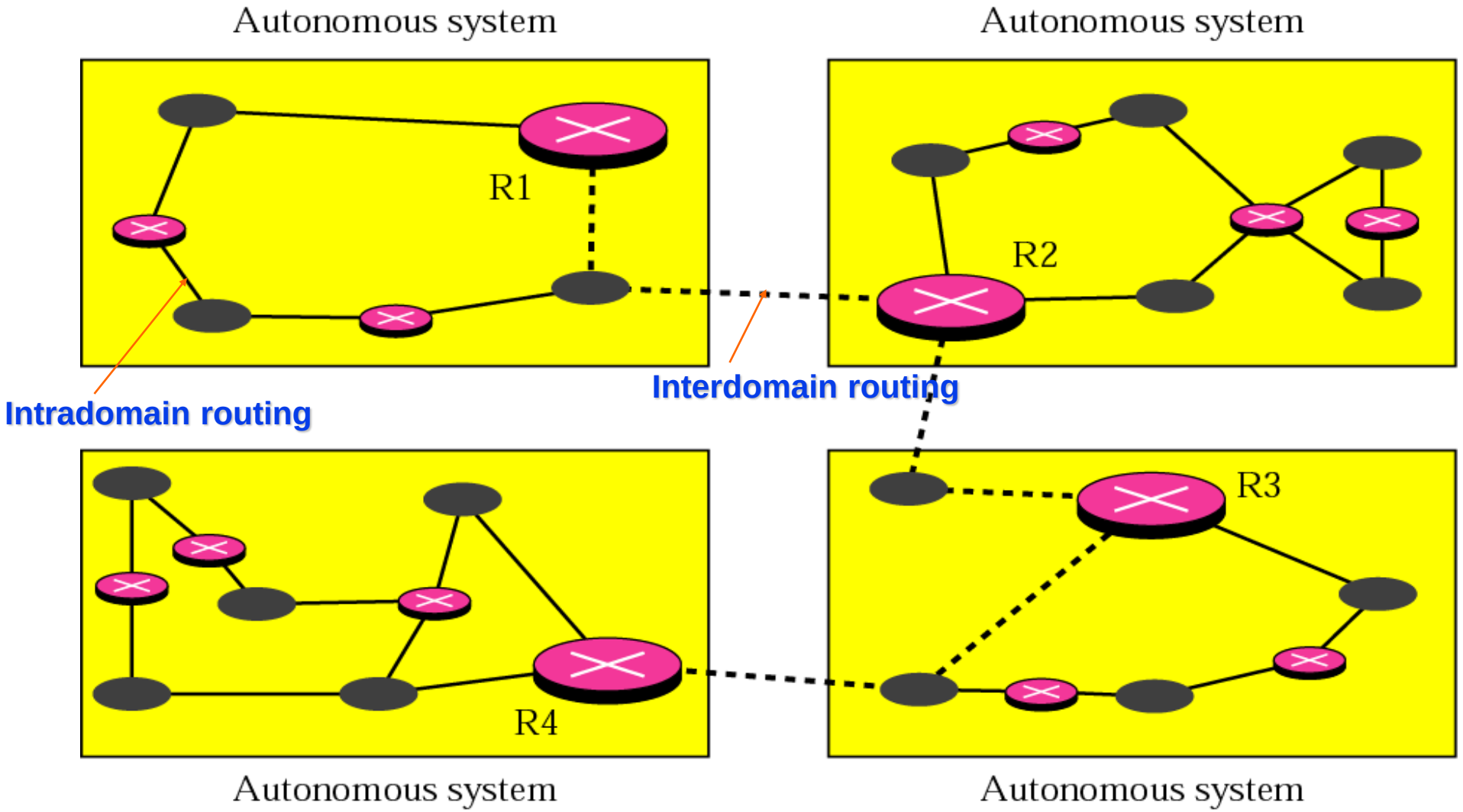
Introduction

- OSPF(Open Shortest Path First)
  - allowing the administrator to assign a cost for passing through a network based on the type of service required.
  - A route through a network can have different costs (metrics)
- BGP (Border Router Protocol)
  - Criterion is the policy, which can be set by the administrator.
  - Policy defines what paths should be chosen.
- Static and Dynamic tables
- Unicast Routing and Multicast Routing

14.1 Intra and Extra Domain Routing

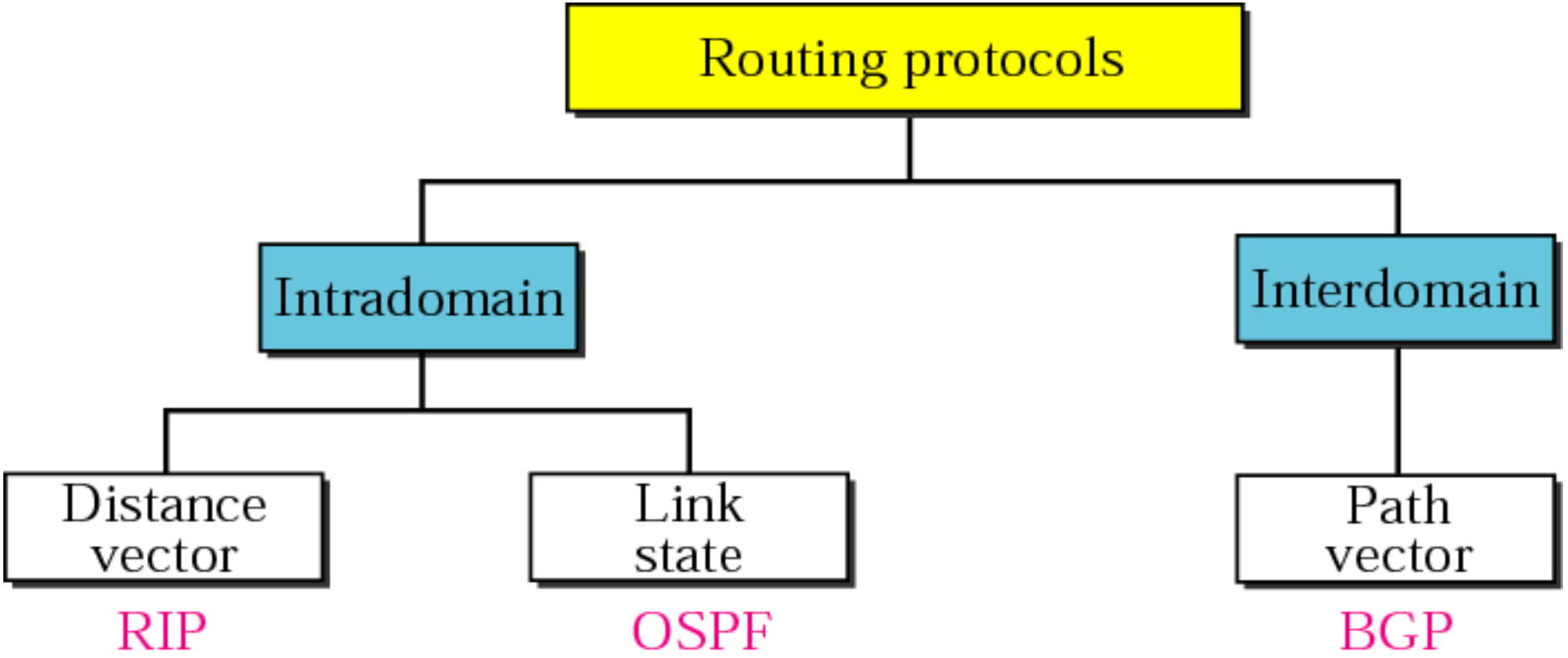
- Because an internet can be so large, one routing protocol cannot handle the task of updating routing tables of all routers.
- So, an internet is divided into autonomous systems.
- An *autonomous system (AS)* is a group of networks and routers under the authority of a single administration.
- Intradomain routing
  - used for the routing inside an autonomous system
- Interdomain routing
  - used for the routing between autonomous systems

Autonomous Systems



Intra and Extra Domain Routing (Cont'd)

- **Popular routing protocols**

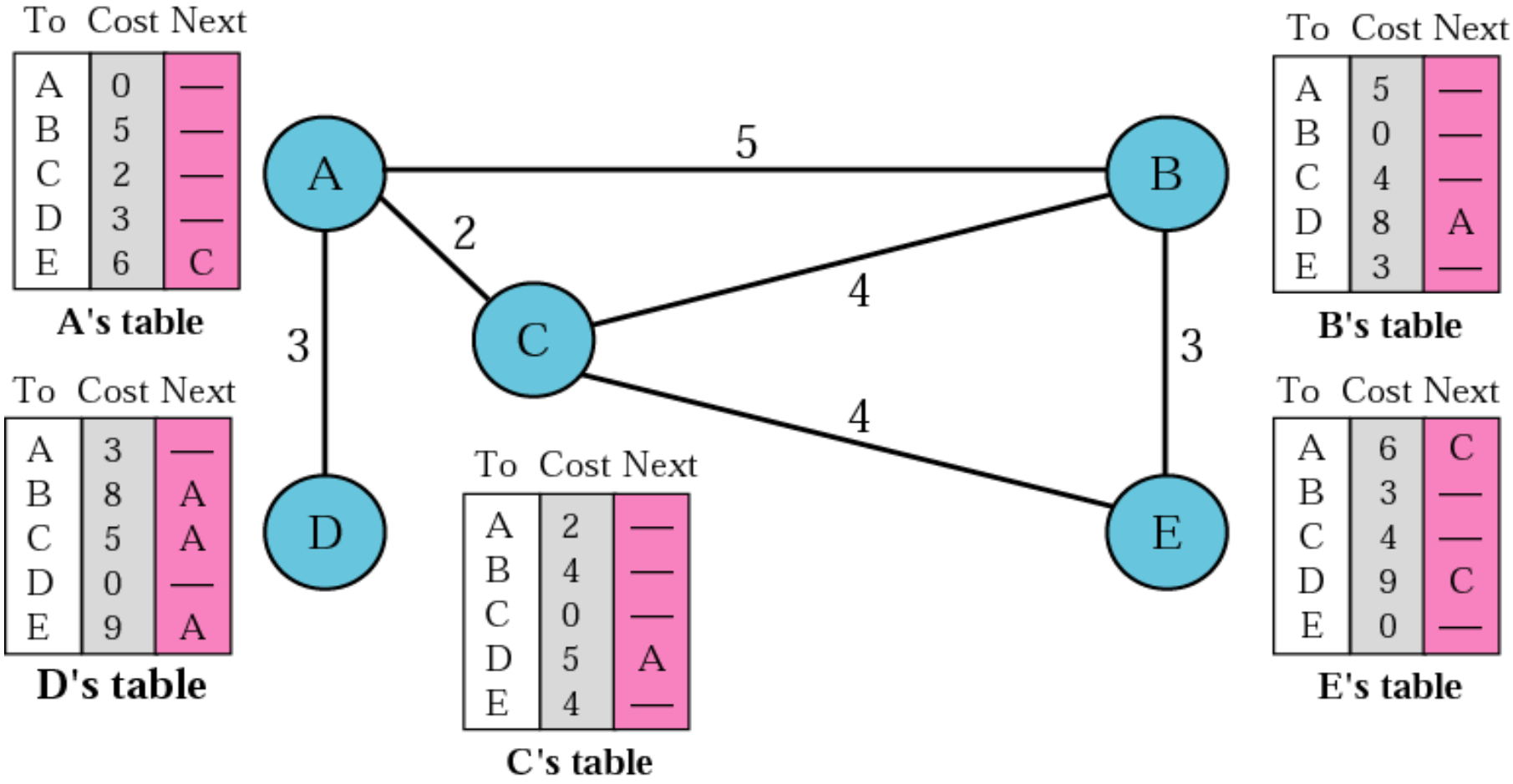


14.2 Distance Vector Routing

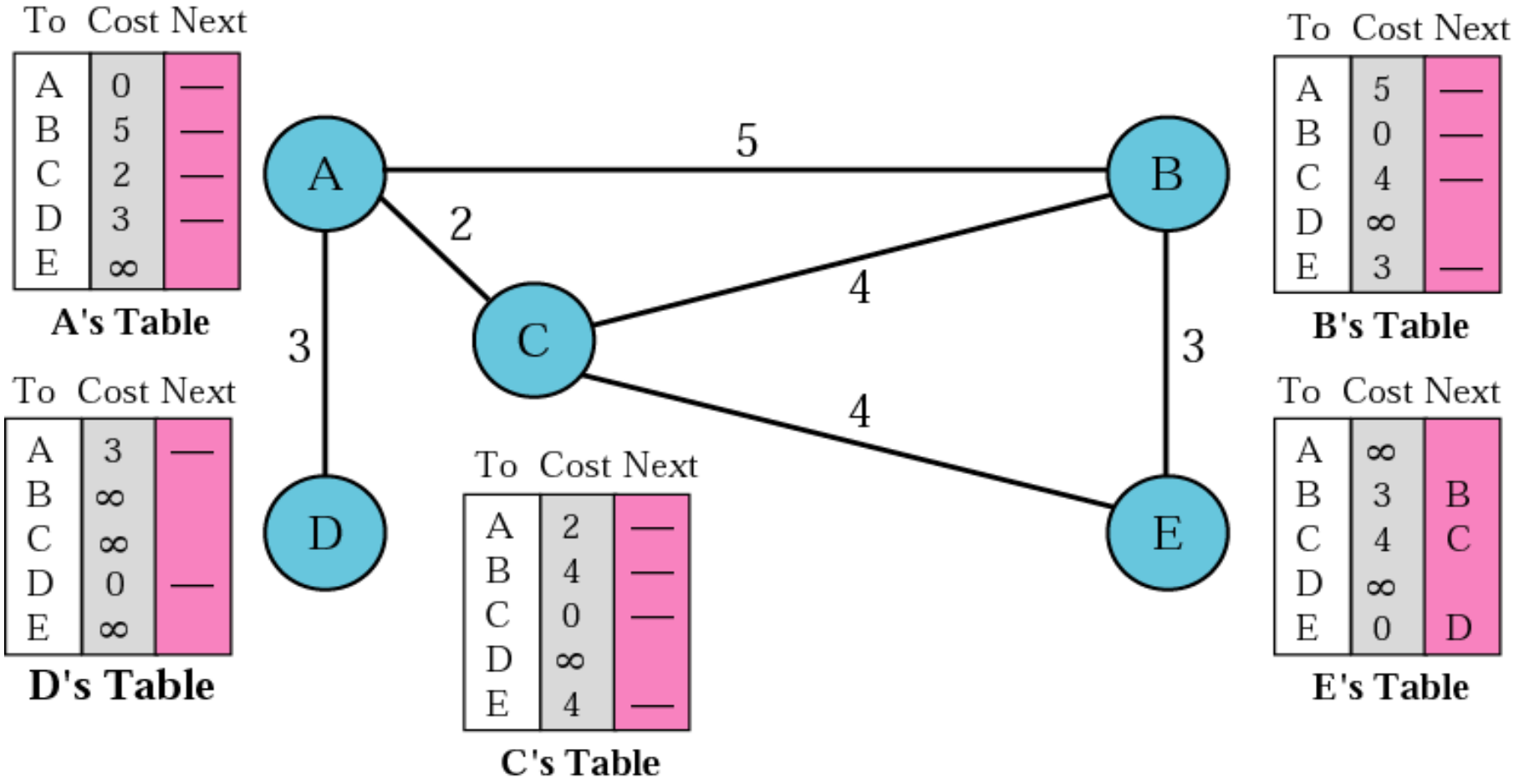
- In distance vector routing, the least cost route between any two nodes is the route with minimum distance. In this protocol each node maintains a vector (table) of minimum distances to every node
- The table at the each node also guides the packet to the desired node by showing the next stop in the route (next-hop routing)
- Distance Vector Routing
  - each router periodically shares its knowledge about the entire internet with neighbors
  - the operational principles of this algorithm

1. Sharing knowledge about the entire autonomous system
2. Sharing only with neighbors
3. Sharing at regular intervals (ex, every 30 seconds) and when there is a change

Distance Vector Routing Tables

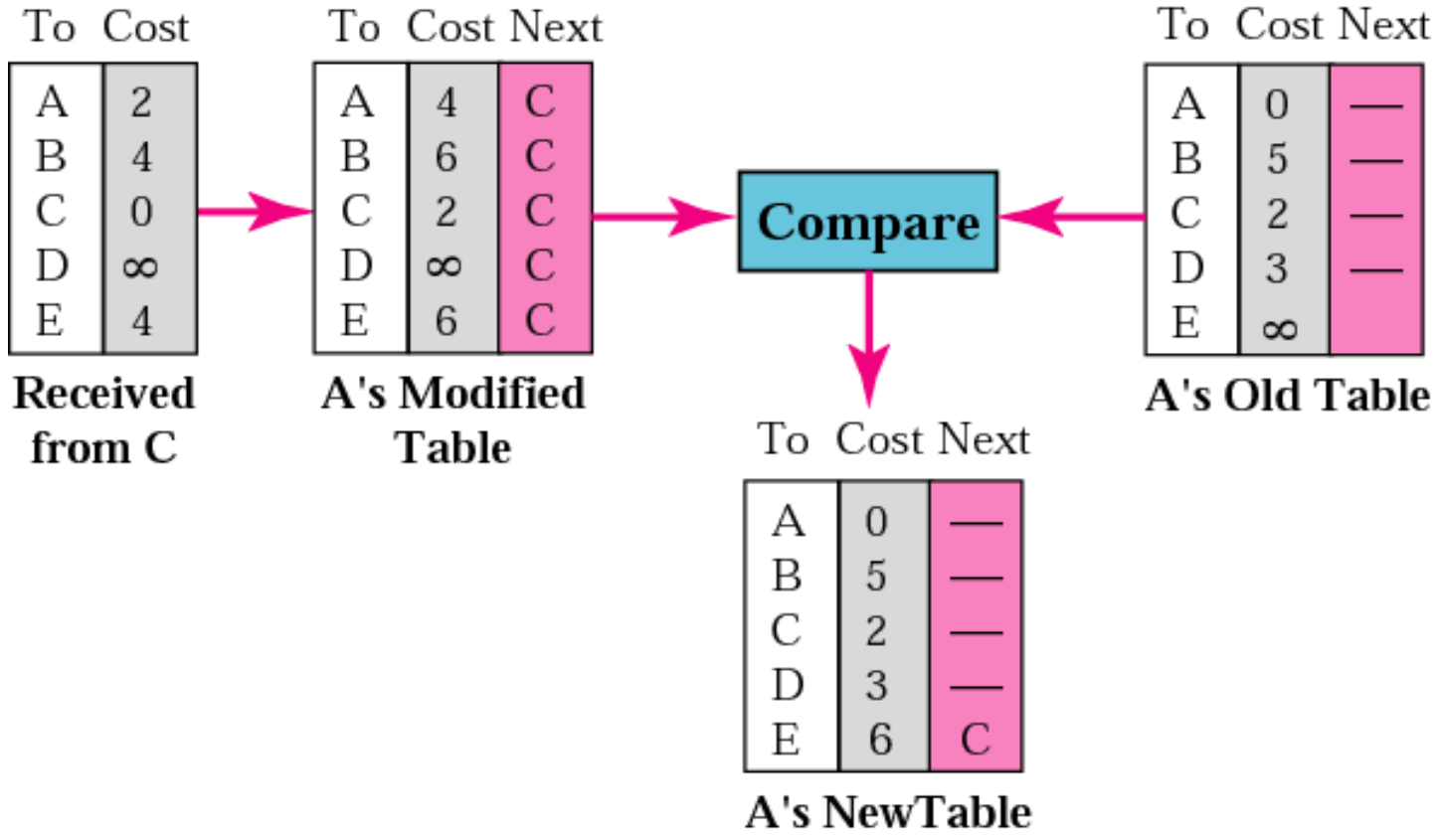


Initialization of Tables in Distance Vector Routing

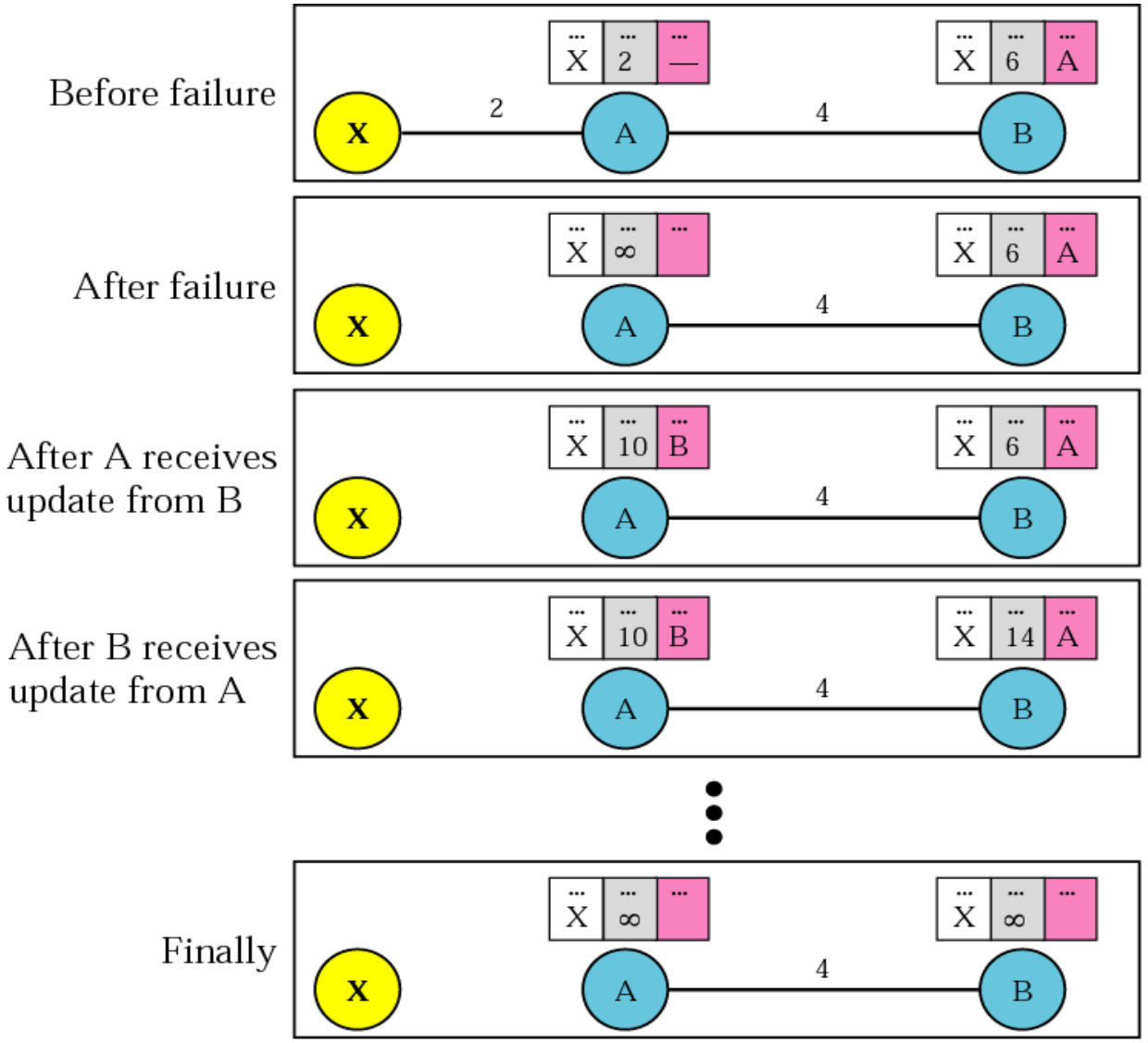


Updating in Distance Vector Routing

- In distance vector routing, each node shares its routing table with its immediate neighbors periodically and when there is a change.

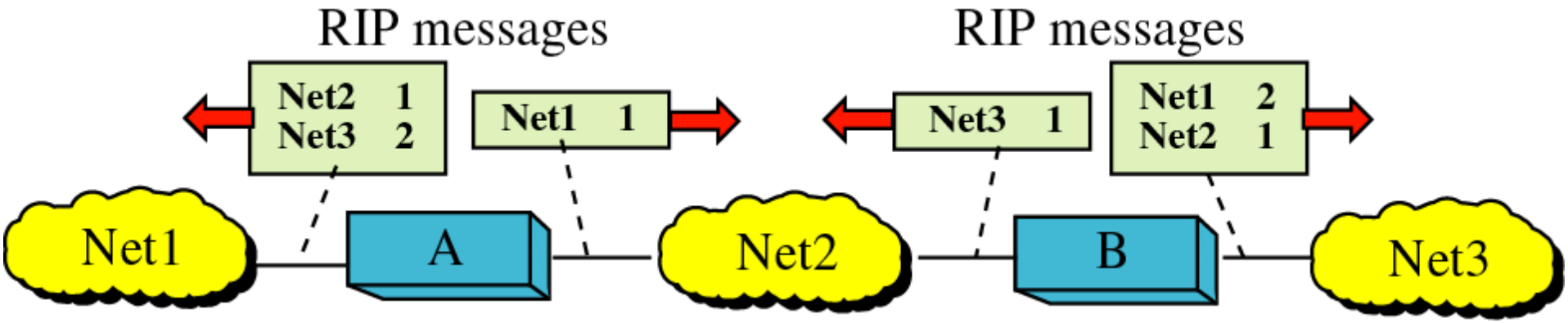


Two-Node Loop Instability



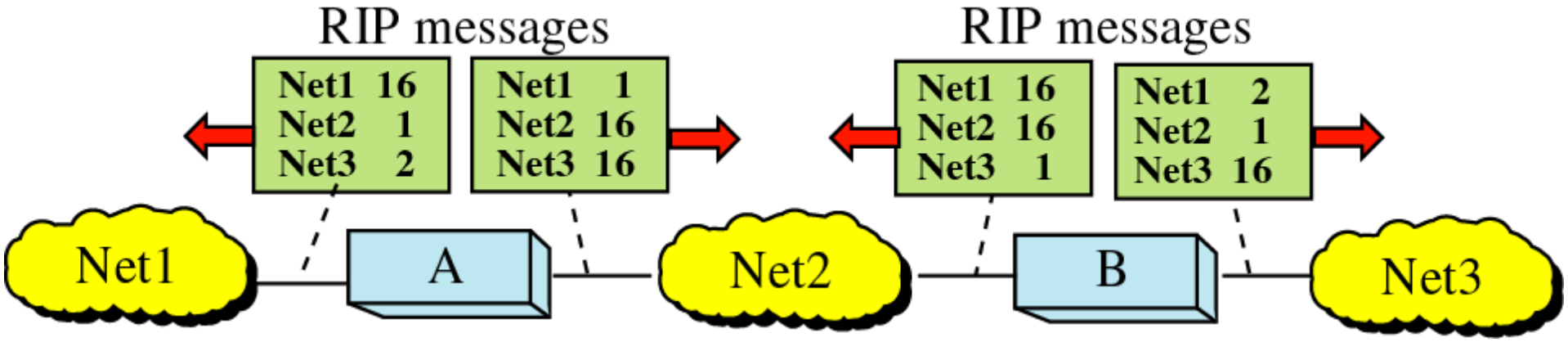
Distance Vector Routing

- **Some Remedies for Instability**
  - **Split Horizons**

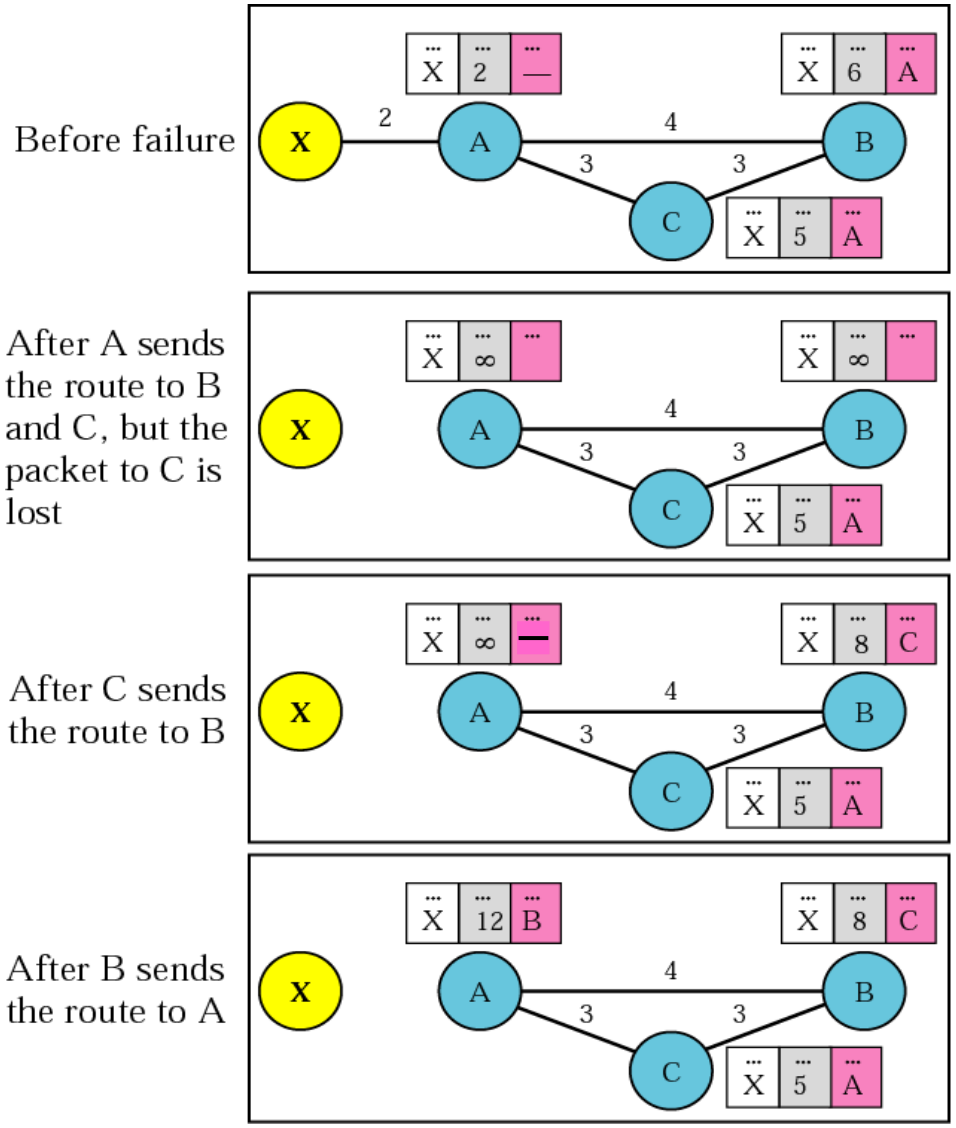


Distance Vector Routing

- **Poison Reverse : a variation of split horizons**



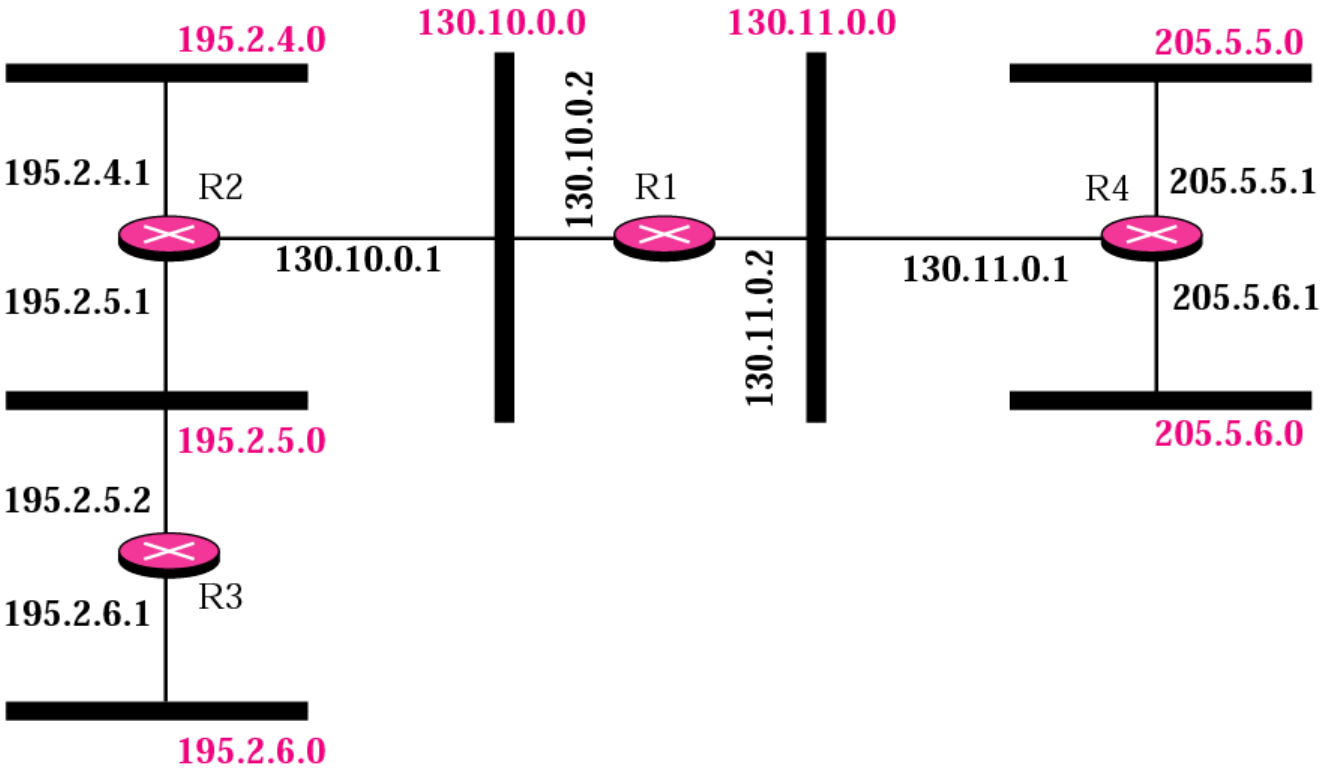
Three-Node Instability



14.3 RIP

- The Routing Information Protocol (RIP) is an intradomain routing protocol used inside an autonomous system. It is a very simple protocol based on distance vector routing.
- The destination in a routing table is a network, which means the first column defines a network address.
- A metric in RIP is called a hop count; distance; defined as the number of links (networks) that have to be used to reach the destination.

Example of a Domain Using RIP



Dest.	Hop	Next
130.10.0.0	1	———
130.11.0.0	1	———
195.2.4.0	2	130.10.0.1
195.2.5.0	2	130.10.0.1
195.2.6.0	3	130.10.0.1
205.5.5.0	2	130.11.0.1
205.5.6.0	2	130.11.0.1

R1 Table

Dest.	Hop	Next
130.10.0.0	1	———
130.11.0.0	2	130.10.0.2
195.2.4.0	1	———
195.2.5.0	1	———
195.2.6.0	2	195.2.5.2
205.5.5.0	3	130.10.0.2
205.5.6.0	3	130.10.0.2

R2 Table

Dest.	Hop	Next
130.10.0.0	2	195.2.5.1
130.11.0.0	3	195.2.5.1
195.2.4.0	2	195.2.5.1
195.2.5.0	1	———
195.2.6.0	1	———
205.5.5.0	4	195.2.5.1
205.5.6.0	4	195.2.5.1

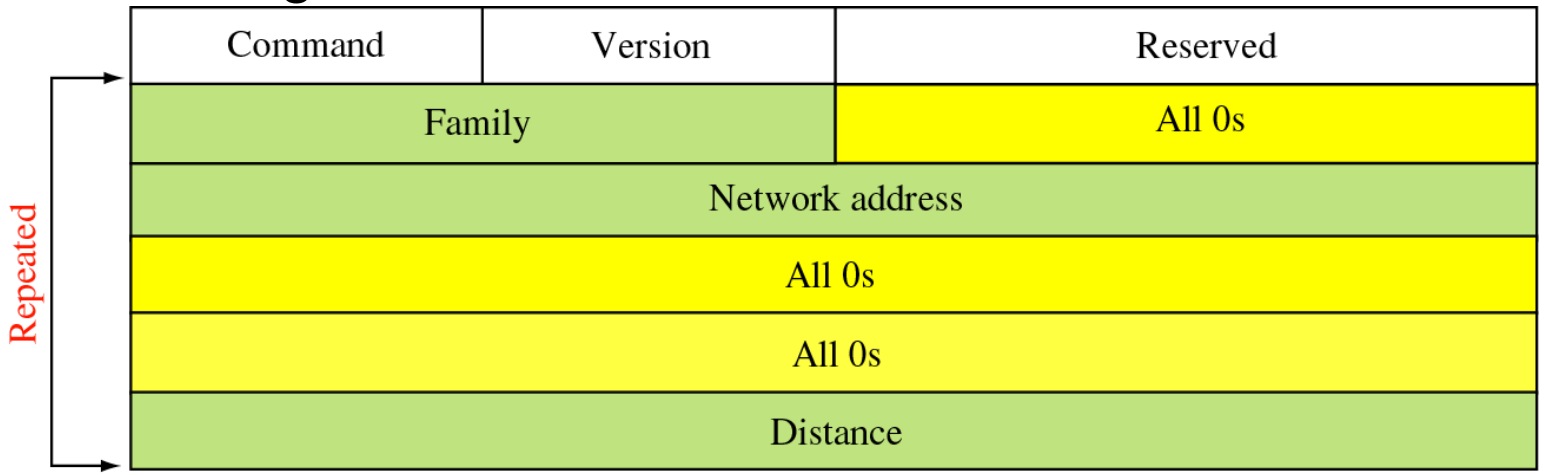
R3 Table

Dest.	Hop	Next
130.10.0.0	2	130.11.0.2
130.11.0.0	1	———
195.2.4.0	3	130.11.0.2
195.2.5.0	3	130.11.0.2
195.2.6.0	4	130.11.0.2
205.5.5.0	1	———
205.5.6.0	1	———

R4 Table

RIP (cont'd)

- **RIP Message Format**

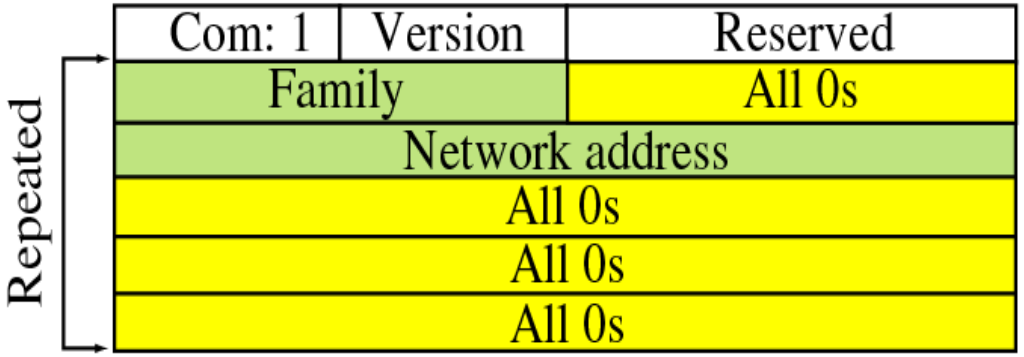


- **Command : request (1) or response (2)**
- **Version**
- **Family : For TCP/IP the value is 2**
- **Address : destination network address**
- **Distance : defining the hop count from the advertising router to the destination network**

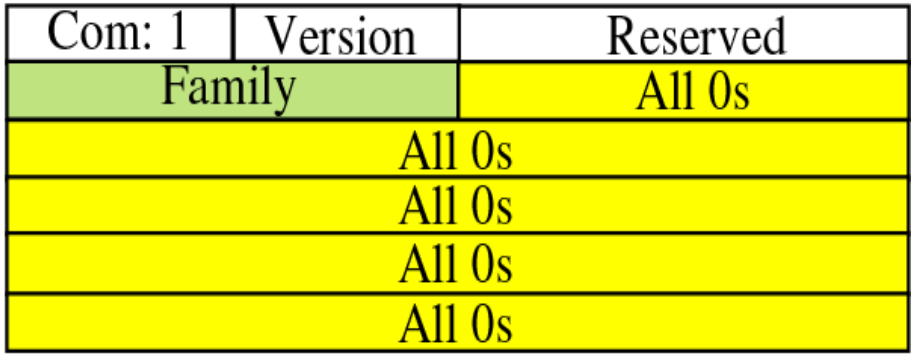
*** Part of the message (entry) is repeated for each destination network.**

RIP (cont'd)

- **Requests and Response**
  - **Request messages : sent by a router that has just come up or by a router that has some time-out entries.**



a. Request for some



b. Request for all

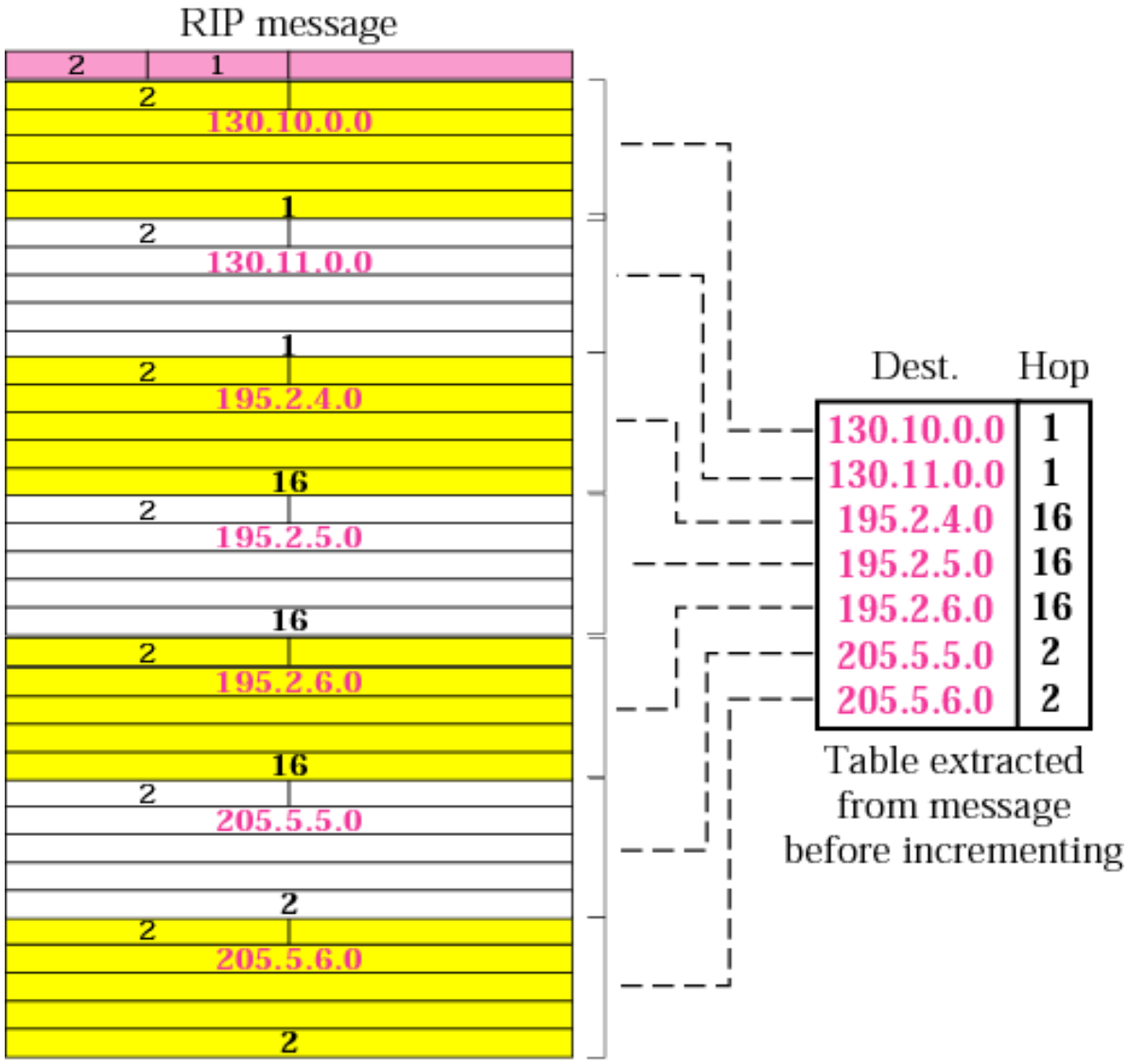
RIP (cont'd)

- **Response**
  - **solicited response**
    - **is sent only in answer to a request**
    - **containing information about the destination specified in the corresponding request**
  - **unsolicited response**
    - **is sent periodically, every 30 seconds**
    - **containing information covering the whole routing table**

Example 1

- Figure 14.11 shows the update message sent from router R1 to router R2 in Figure 14.8. The message is sent out of interface 130.10.0.2.

Solution to Example 1



RIP (cont'd)

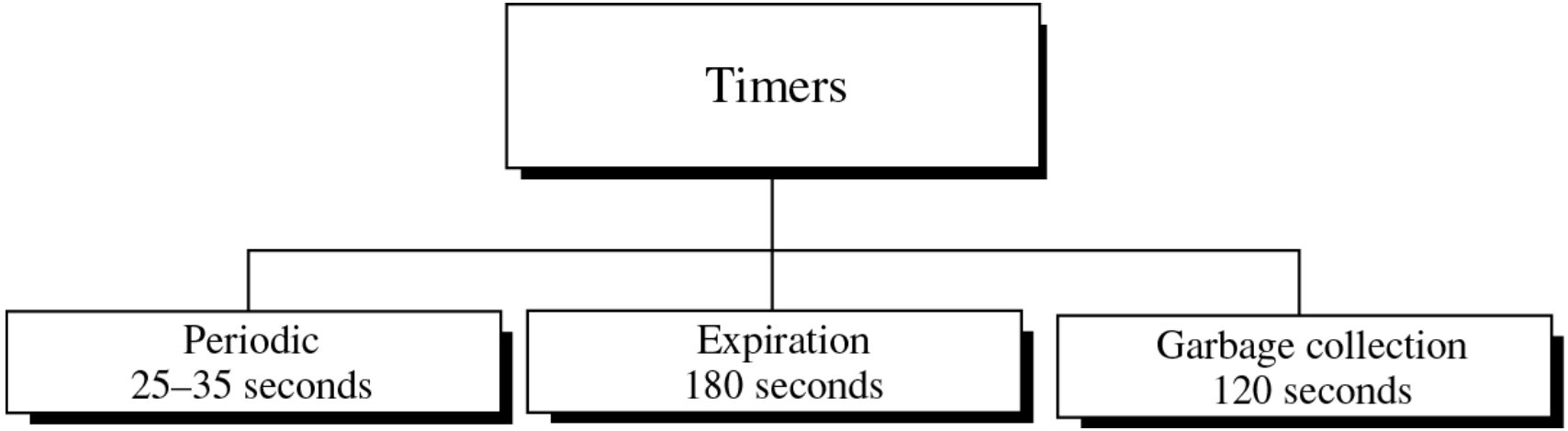
- Timers in RIP
  - Periodic timer : controlling the advertisements of regular update messages
  - expiration timer : governing the validity of a route
  - the garbage collection timer : advertising the failure of a route
- Periodic timer
  - controlling the advertising of regular update messages
  - using random number between 25 to 35 seconds

RIP (cont'd)

- Expiration timer
  - In normal situation, the new update for a route occurs every 30 seconds
  - But, if there is a problem on an Internet and no update is received within the allotted 180 seconds, the route is considered expired and the hop count of the route is set to 16.
  - Each router has its own expiration timer.
- Garbage Collection Timer
  - When the information about a route becomes invalid, the router continues to advertise the route with a metric value of 16 and the garbage collection timer is set to 120 sec for that route
  - When the count reaches zero, the route is purged from the table.

RIP (cont'd)

- **RIP timers**



RIP (cont'd)

Example 2

A routing table has 20 entries. It does not receive information about five routes for 200 seconds. How many timers are running at this time?

The timers are listed below:

Periodic timer: 1

Expiration timer: 20 - 5 = 15

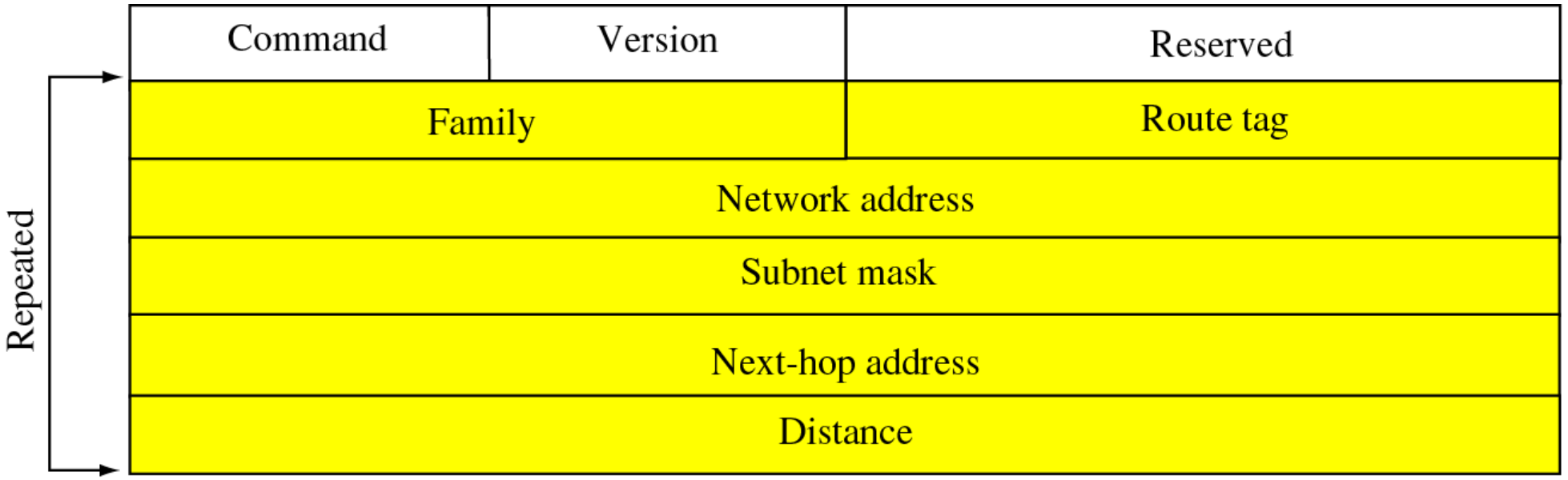
Garbage collection timer: 5

RIP Version 2

- Designed for overcoming some of the shortcomings of version 1
- Replaced fields in version 1 that were filled with 0s for the TCP/IP protocols with some new fields
- Can use classless addressing

RIP Version 2 (cont'd)

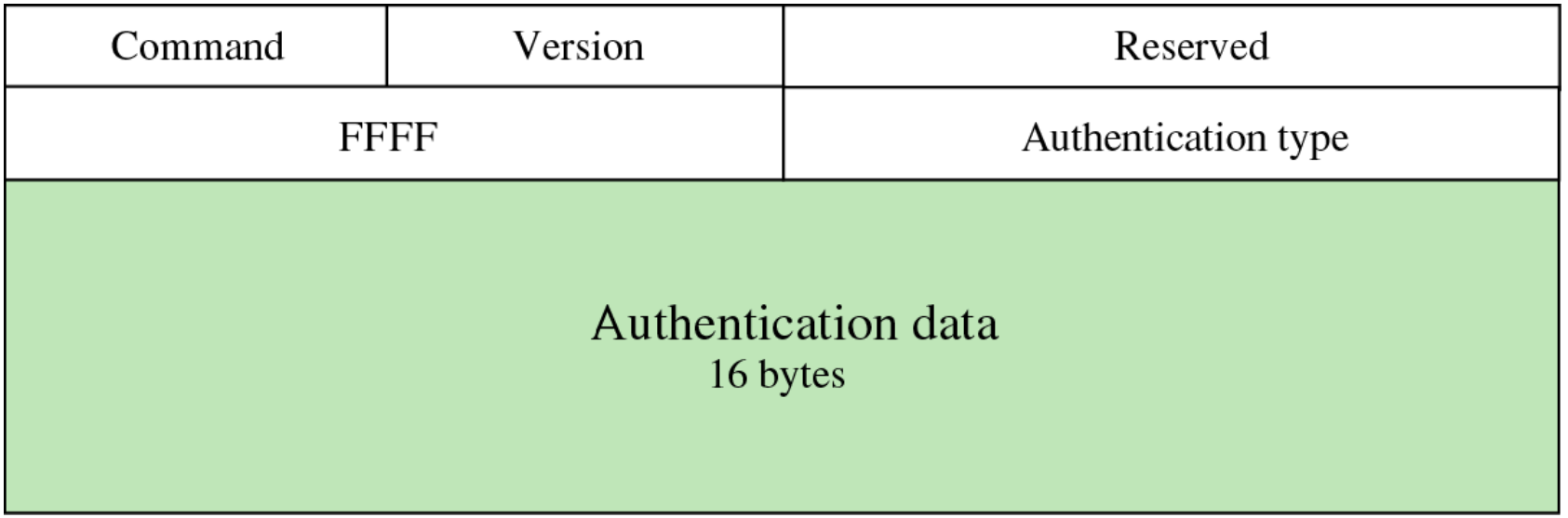
- **RIP version 2 format**



- Route Tag : carrying information such as the autonomous system number
- Subnet mask : carrying the subnet mask
- Next-hop address : showing the next hop
  - In case that shares a network backbone by two ASs, the message can define the router to which the packet should go next

RIP Version 2 (cont'd)

- **Authentication**
  - added to protect the message against unauthorized advertisement
  - Value of $FFFF_{16}$ is entered in the family field
  - Authentication type : protocol used for authentication



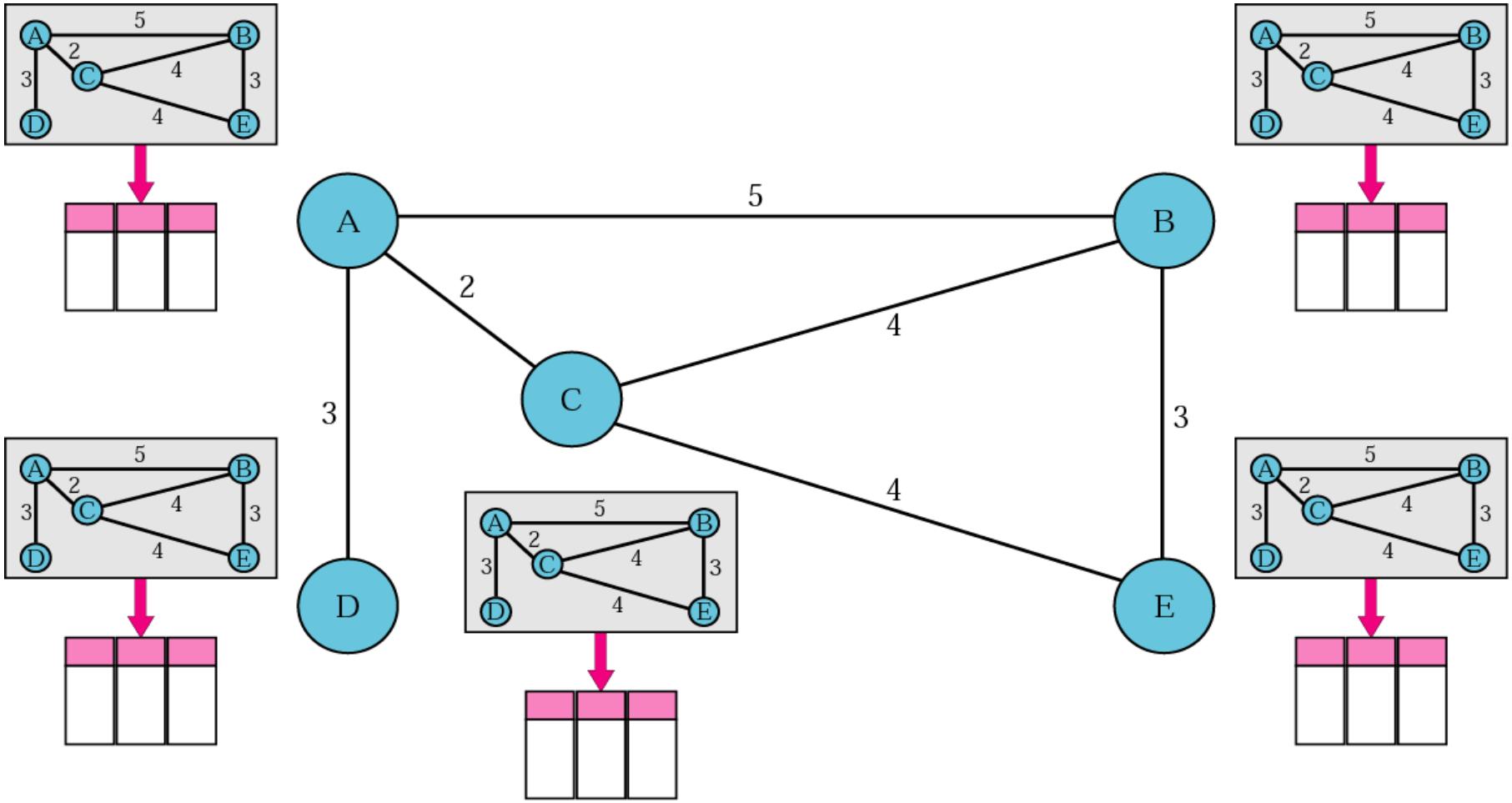
RIP Version 2 (cont'd)

- **Multicasting**
  - Using the multicast address 224.0.0.9 to multicast RIP messages only to RIPv2 routers in the network
- **Encapsulation of RIP messages**
  - encapsulated in UDP user datagram
  - not included a field that indicates the length of the message
  - Well-known port assigned to RIP in UDP is port 520

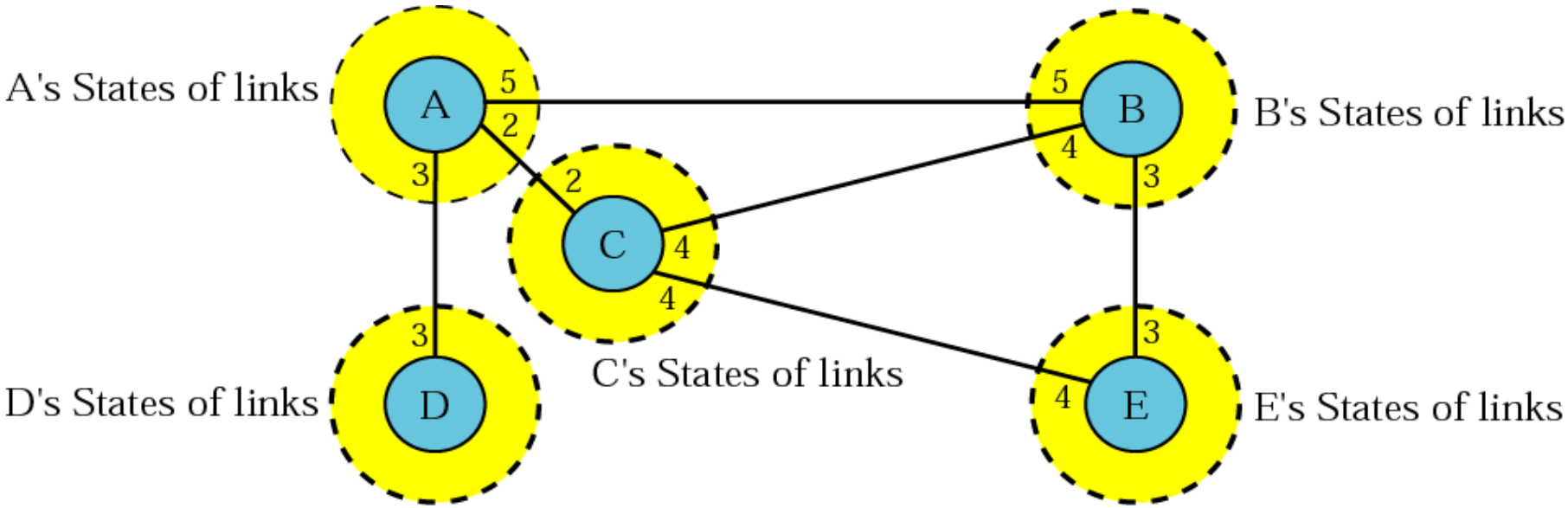
14.4 Link State Routing

- In link state routing, if each node in the domain has the entire topology of the domain, the node can use Dijkstra's algorithm to build a routing table.

Concept of Link State Routing



Link State Knowledge



Building Routing Tables

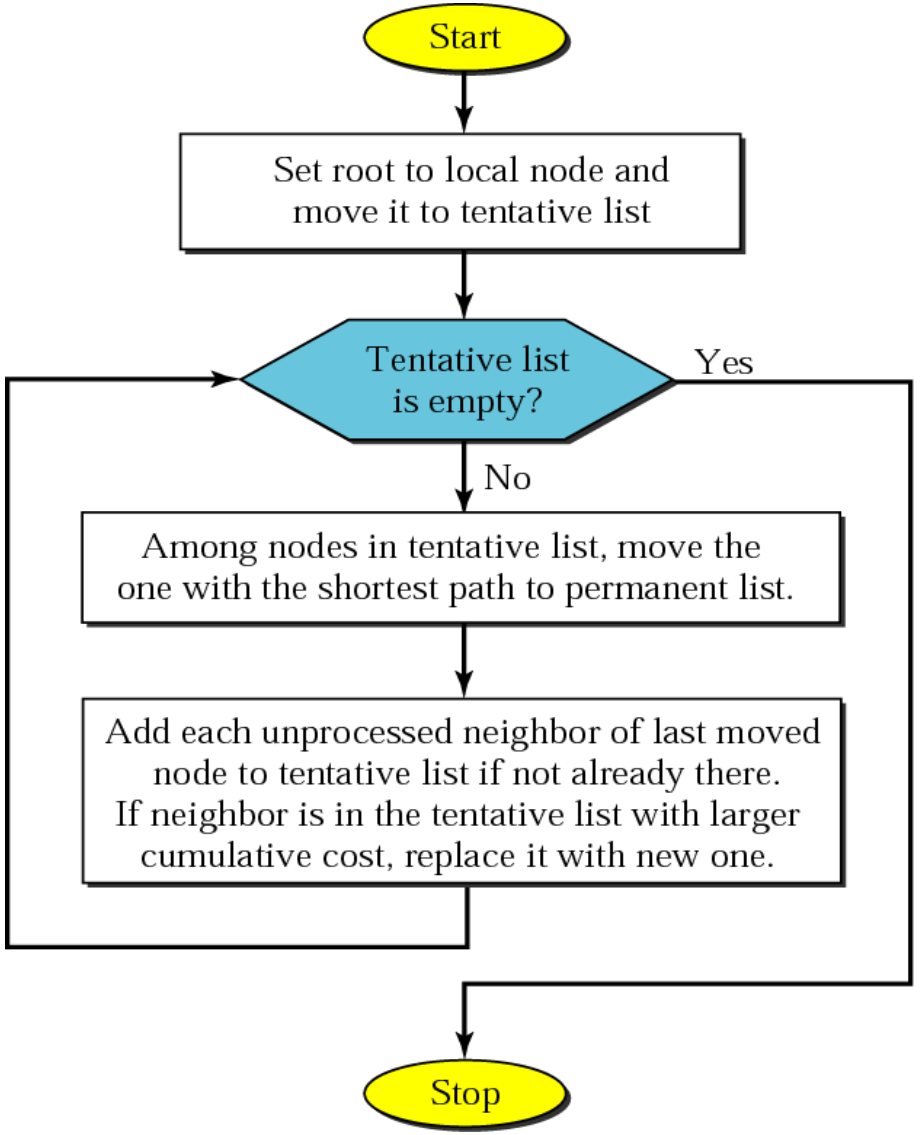
1. Creation of the states of the links by each node, called the link state packet or LSP.
2. Dissemination of LSPs to every other router, called flooding, in an efficient and reliable way
3. Formation of a shortest path tree for each node
4. Calculation of a routing table based on the shortest path tree

Creation of LSP

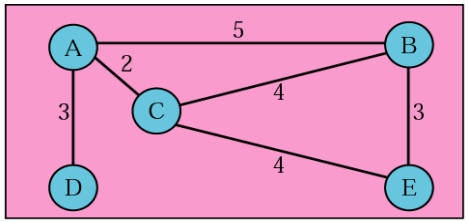
- When there is a change in the topology of the domain
- On a periodic basis
  - 60 minutes or 2 hours

Formation of Shortest Path Tree

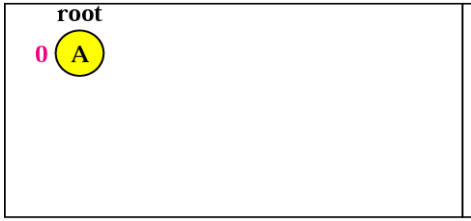
- **Dijkstra Algorithm**



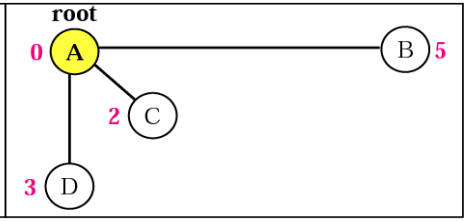
Example of formation of Shortest Path Tree



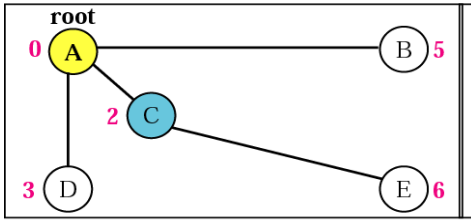
Topology



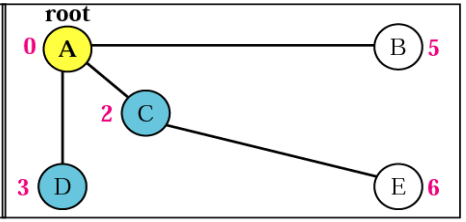
1. Set root to A and move A to tentative list



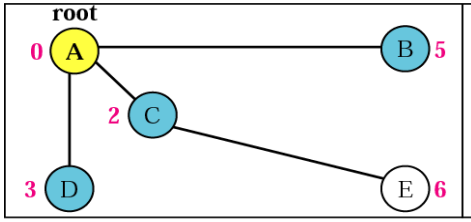
2. Move A to permanent list and add B, C, and D to tentative list



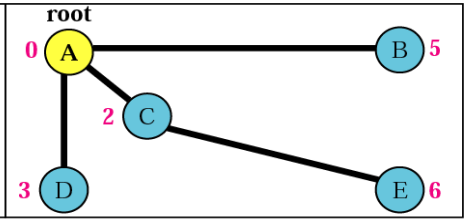
3. Move C to permanent and add E to tentative list



4. Move D to permanent list.



5. Move B to permanent list



6. Move E to permanent list (tentative list is empty)

Calculating of Routing Table from Shortest Path Tree

Table 14.1 Routing table for node A

Node	*Cost*	*Next Router*
A	0	—
B	5	—
C	2	—
D	3	—
E	6	C

14.5 OSPF (Open Shortest Path First)

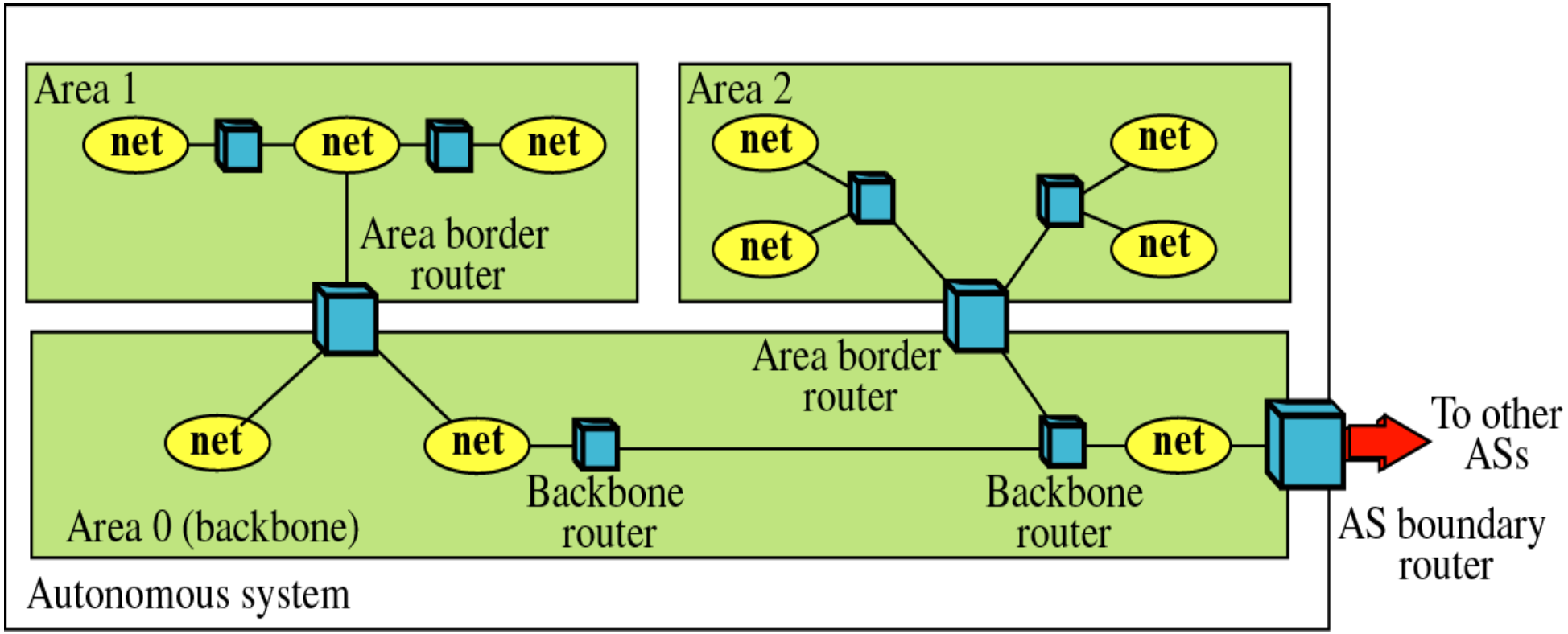
- The Open Shortest Path First (OSPF) protocol is an intradomain routing protocol based on link state routing. Its domain is also an autonomous system
- Dividing an AS into areas
  - to handle routing efficiently and in a timely manner

OSPF (cont'd)

- Areas
  - Is a collection of networks, hosts, and routers in AS
  - AS can be divided into many different areas.
  - All networks inside an area must be connected.
  - Routers inside an area flood the area with routing information.
- Area Border Router
  - Summarizes the information about the area and sends it to other areas
- Backbone
  - All of the areas inside an AS must be connected to the backbone
  - Serving as a primary area (Area 0)
  - Consisting of backbone routers
  - Back bone routers can be an area border router

OSPF (cont'd)

- **Areas in an AS**

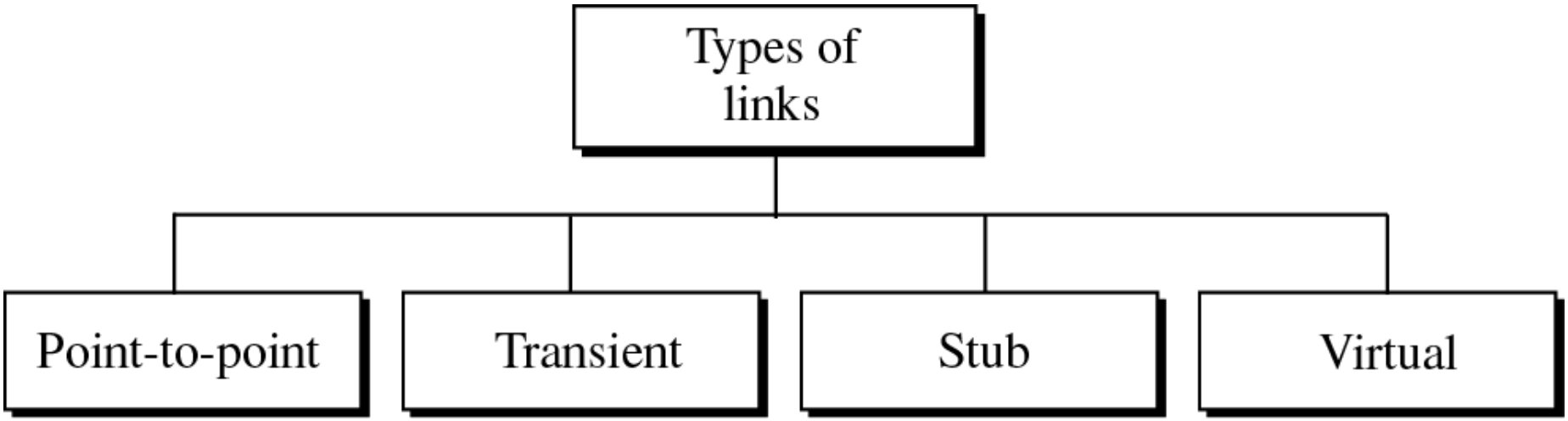


OSPF (cont'd)

- **Metric**
  - OSPF protocol allows the administrator to assign a cost, called the ***metric***, to each route.
  - Based on a type of service (minimum delay, maximum throughput, bandwidth and so on).
  - A router can have multiple routing tables, each based on a different type of service.
- **Link State Routing**
  - OSPF uses Link State Routing to update the routing tables in an area
  - Each router shares its knowledge about its neighborhood with every router in the area.

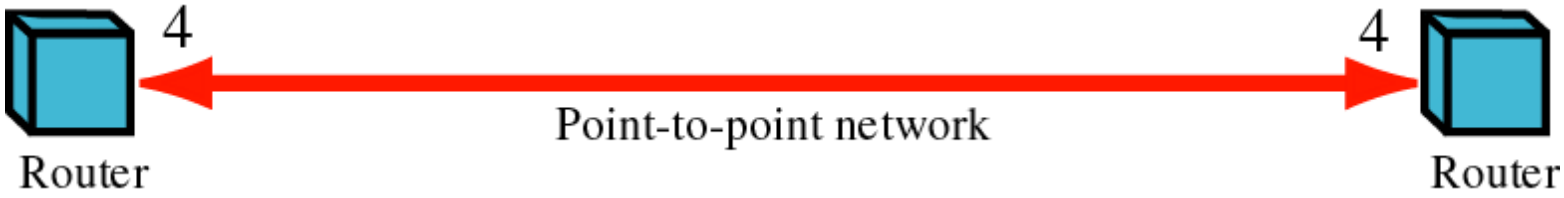
OSPF (cont'd)

- **Types of Links**
  - **In OSPF terminology, a connection is called a *link*.**



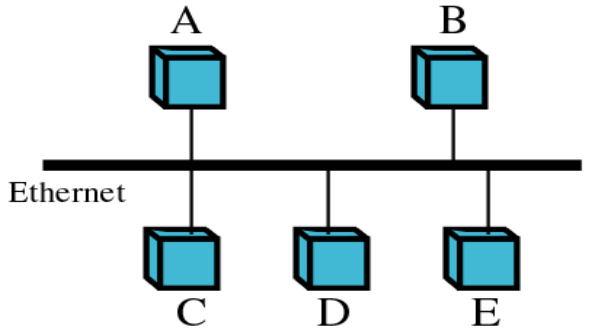
OSPF (cont'd)

- **Point-to-point Link**
  - **Routers are represented by nodes and the link is represented by a bidirectional edge connecting the nodes.**
  - **Each router has only one neighbor at the other side of the link.**

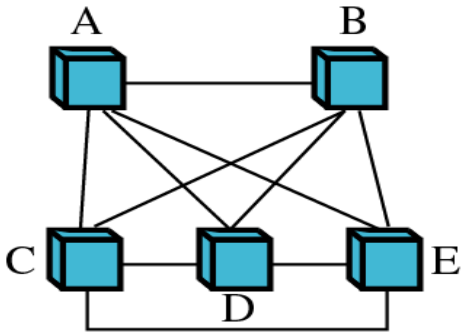


OSPF (cont'd)

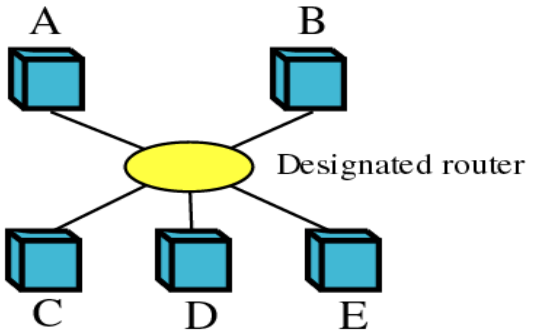
- **Transient Link**
  - **is a network with several routers attached to transient Link**



a. Transient network



b. Unrealistic representation



c. Realistic representation

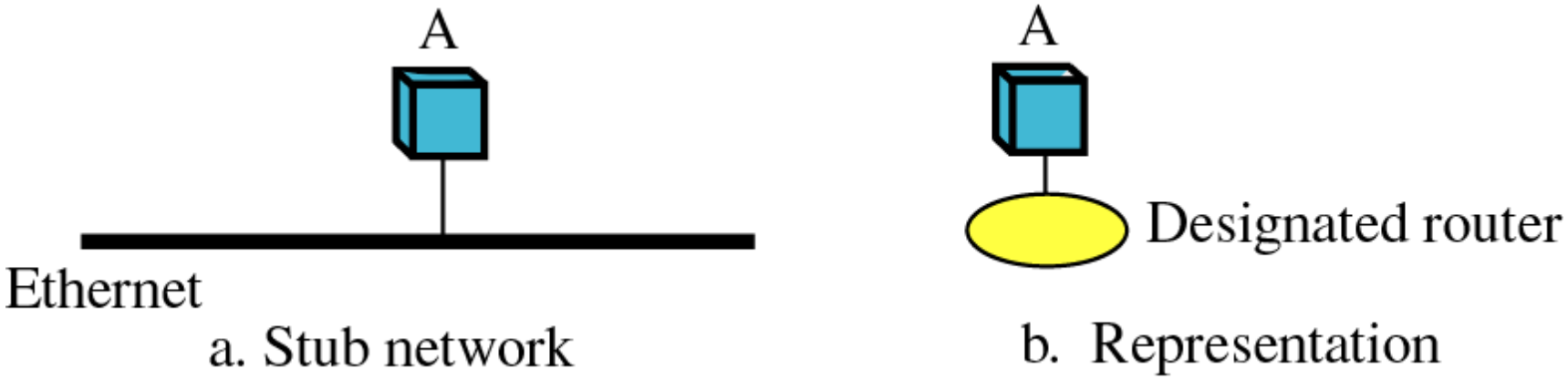
(network)

OSPF (cont'd)

- In "c", each router has only one neighbor, the designated router (network)
  - The designated router has five neighbors.
  - Number of neighbor announcements is reduced from 20 to 10
  - There is no metric from the designated router to any other node.
    - Because the designated router represents the network.

OSPF (cont'd)

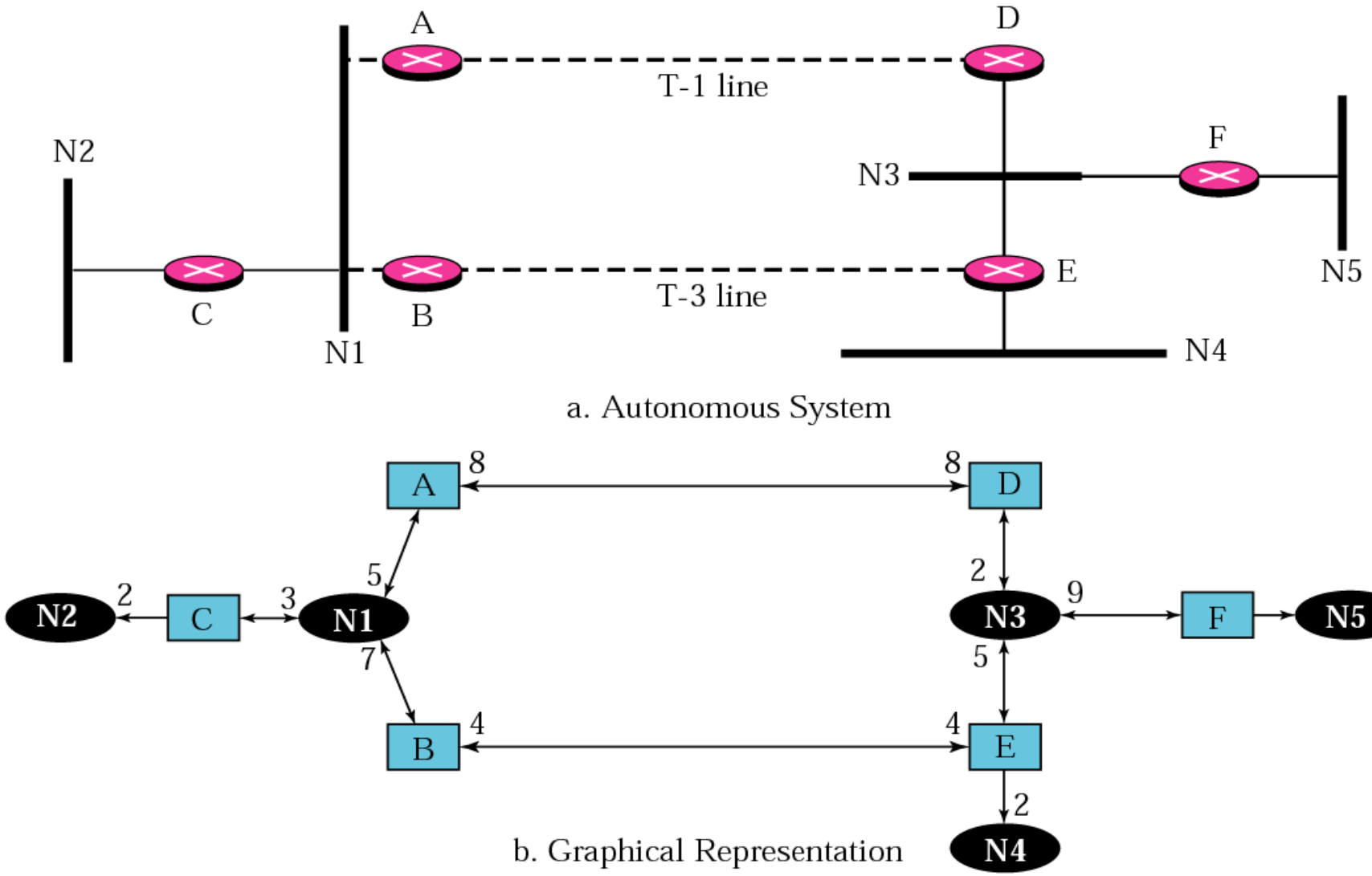
- **Stub Link**
  - **is a network that is connected to only one router**
  - **is a special case of transient network**



OSPF (cont'd)

- **Virtual Link**
  - **When the link between two routers is broken, the administration may create a virtual link between them using a longer path**

AS and its Graphical Representation in OSPF

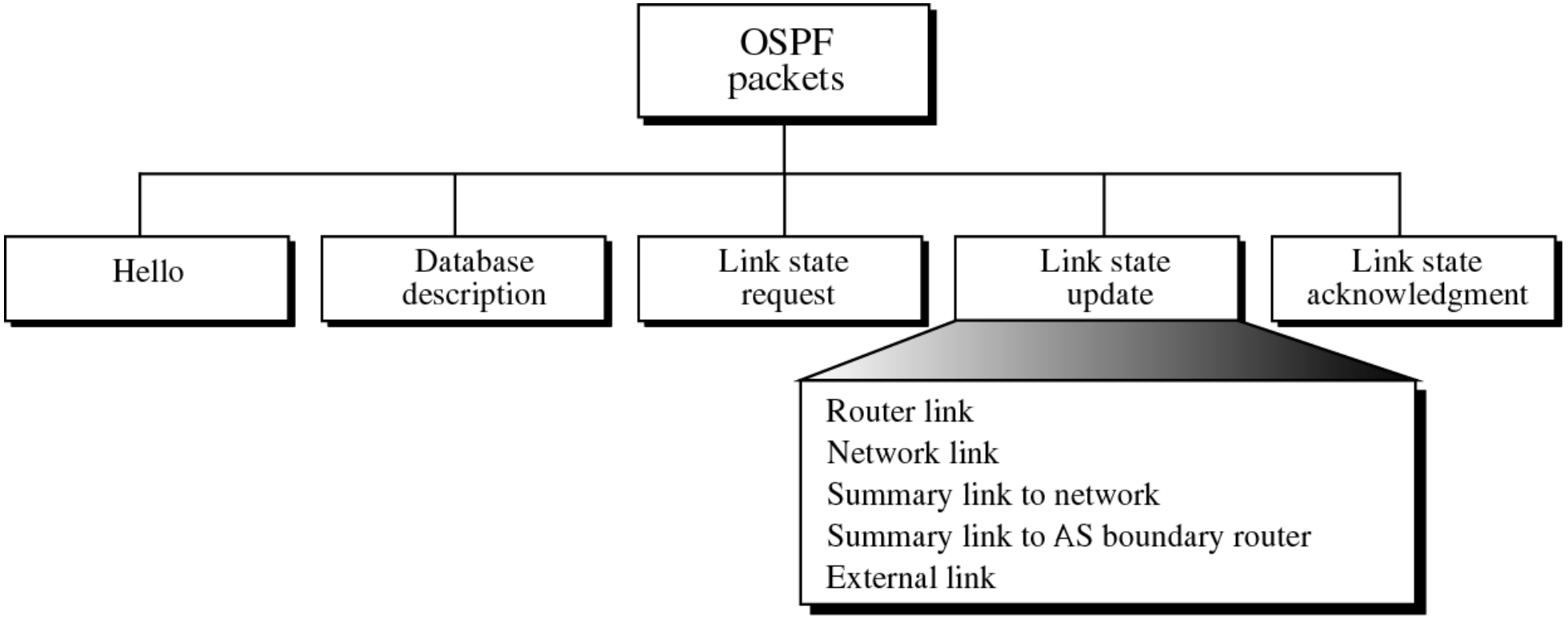


AS and its Graphical Representation in OSPF

- **Graphical Representation (cont'd)**
  - **N1, N3 : transient, N2, N4, N5 : Stub**
  - **using square nodes for the routers and ovals for the networks**

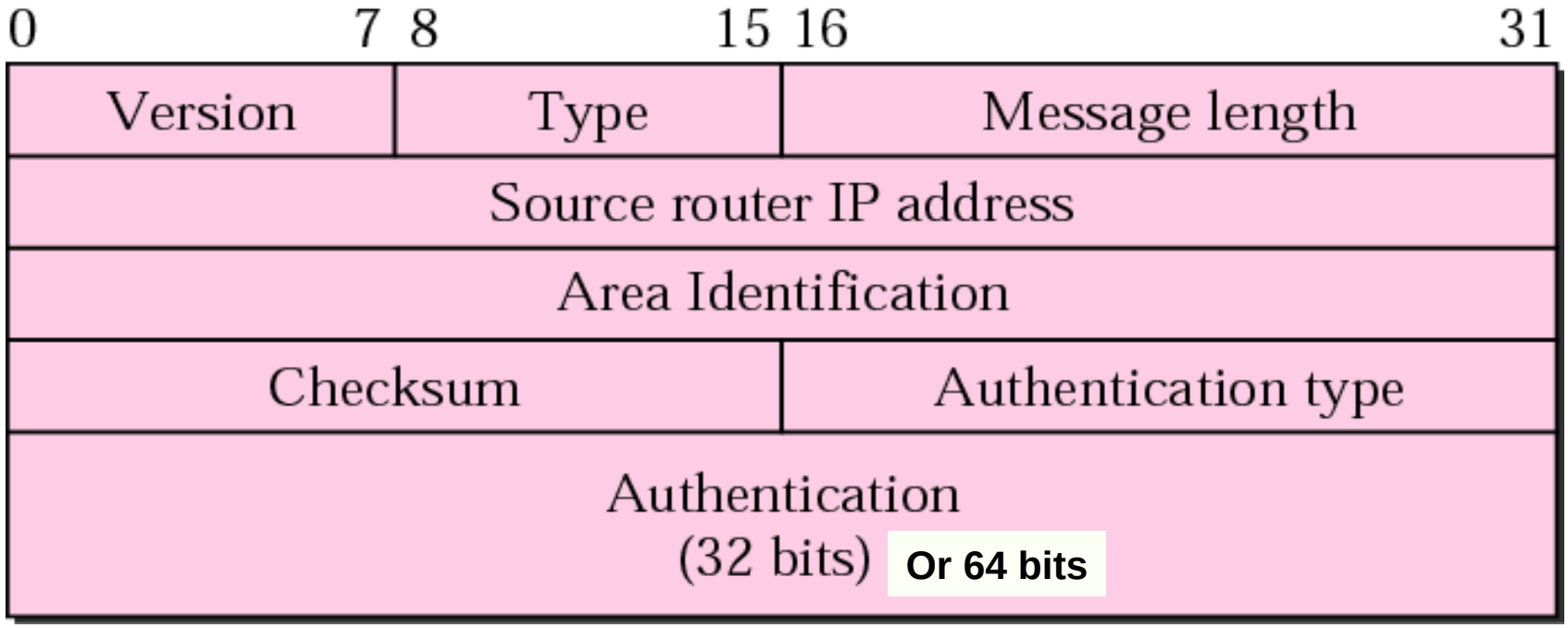
OSPF Packets

- **Types of OSPF Packets (Messages)**



OSPF (cont'd)

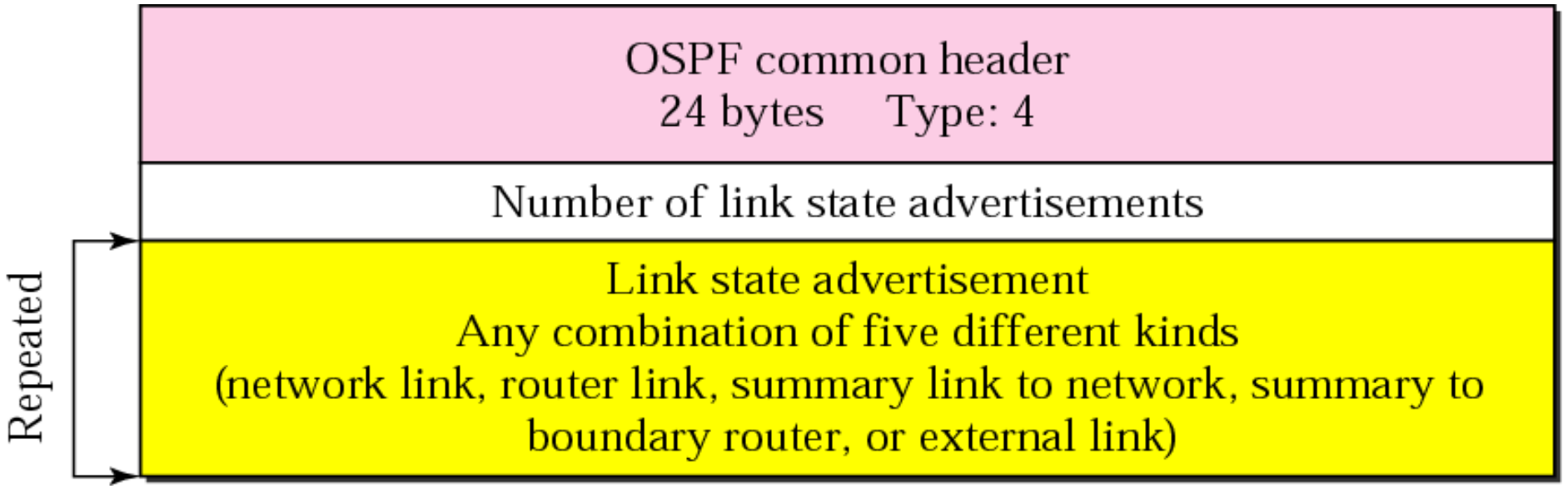
- **OSPF Common Header**



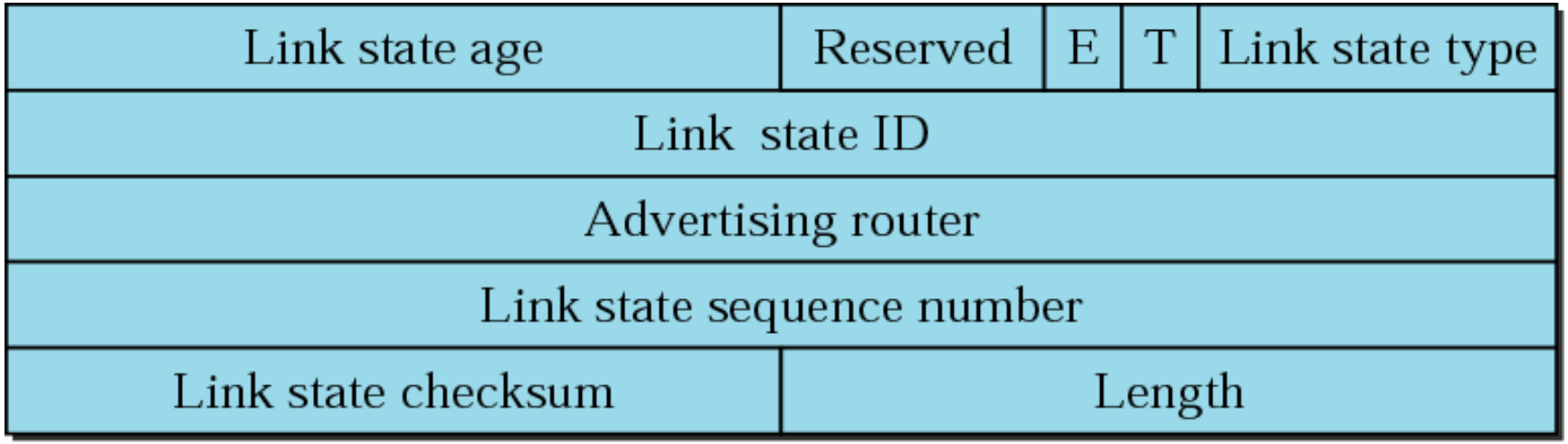
- authentication type : 0 for none, 1 for password
- packet type : five types

Link State Update Packet

- Used by a router to advertise the states of its links

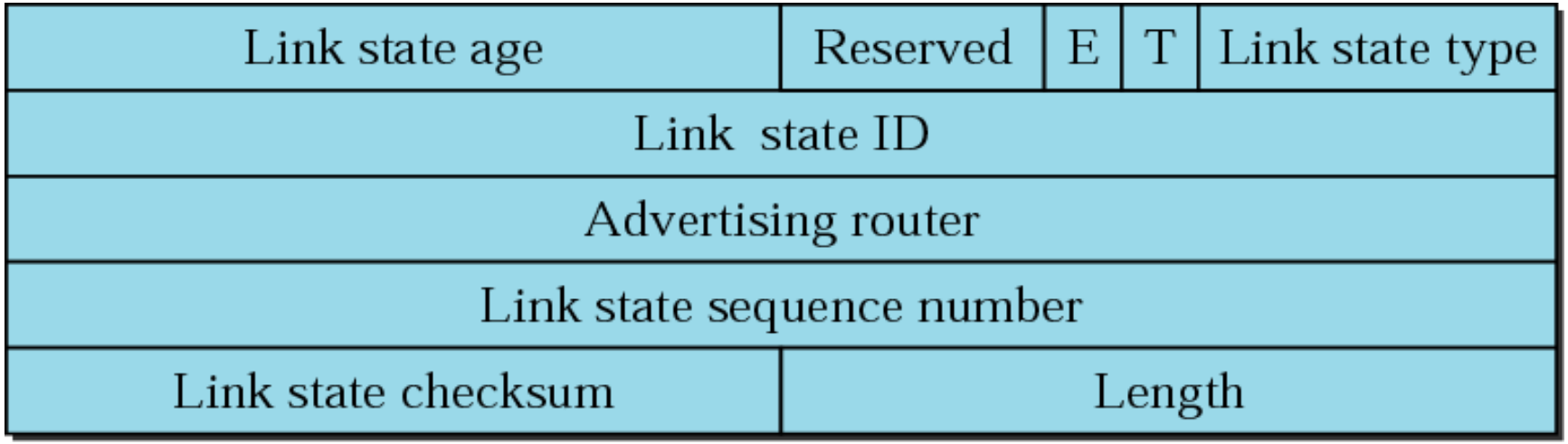


LSA General Header



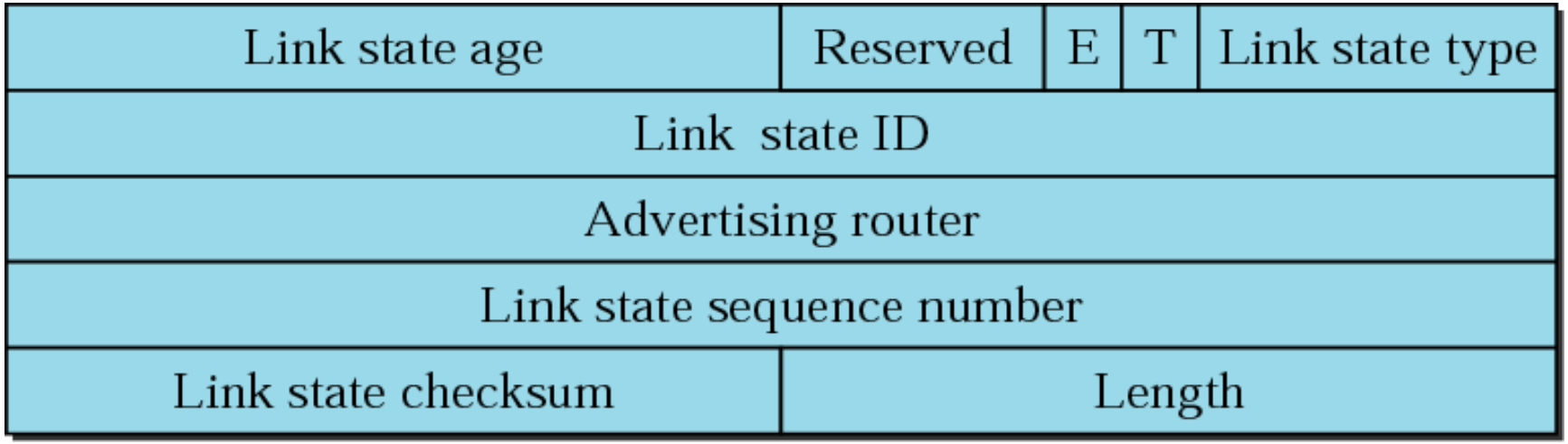
- **Link state age**
  - **When a router creates the message, the value of this field is 0**
  - **When each successive router forwards this message, it estimates the transit time and adds it to the cumulative value of this field.**

LSA General Header



- **E flag : 1 means that the area is a stub area**
- **T flag : 1 means that the router can handle multiple types of service**
- **Link state type : 1) router link, 2) network link, 3) summary link to network, 4) summary link to AS boundary router**

LSA General Header



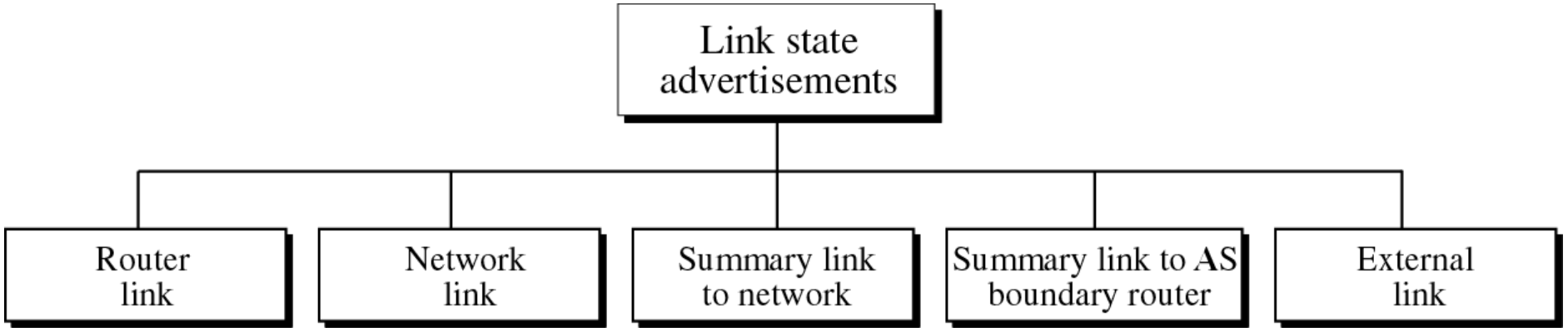
- **Advertising router**
  - The IP address of the router advertising this message
- **Link state sequence number**
  - A sequence number assigned to each link state update message

LS Type & LS ID

Link state type	Link state ID
Router link	IP address of the router
Network link	IP address of the designated router
Summary link to network	Address of the network
Summary link to AS boundary	IP address of the boundary router
External link	Address of the network

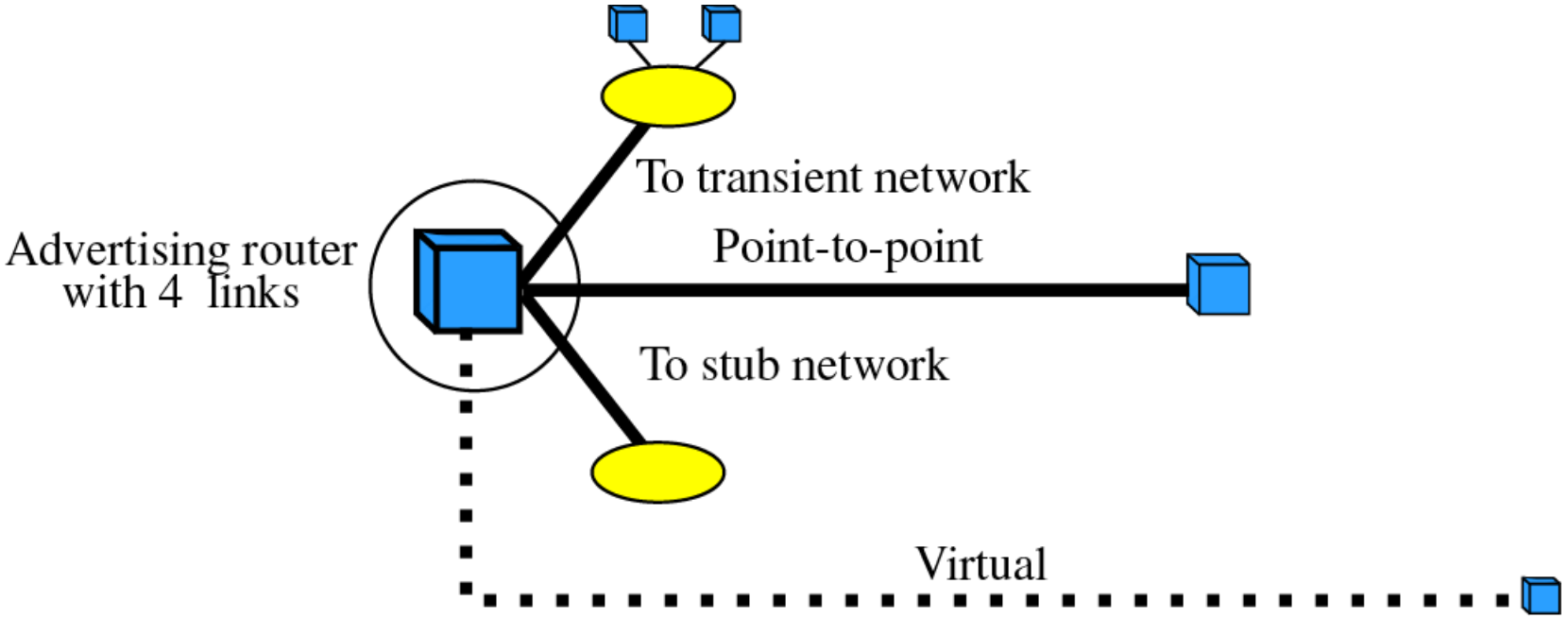
LSA

- **Link State Advertisements**
  - **to share information about neighbors, each router distributes link state advertisements (LSAs)**

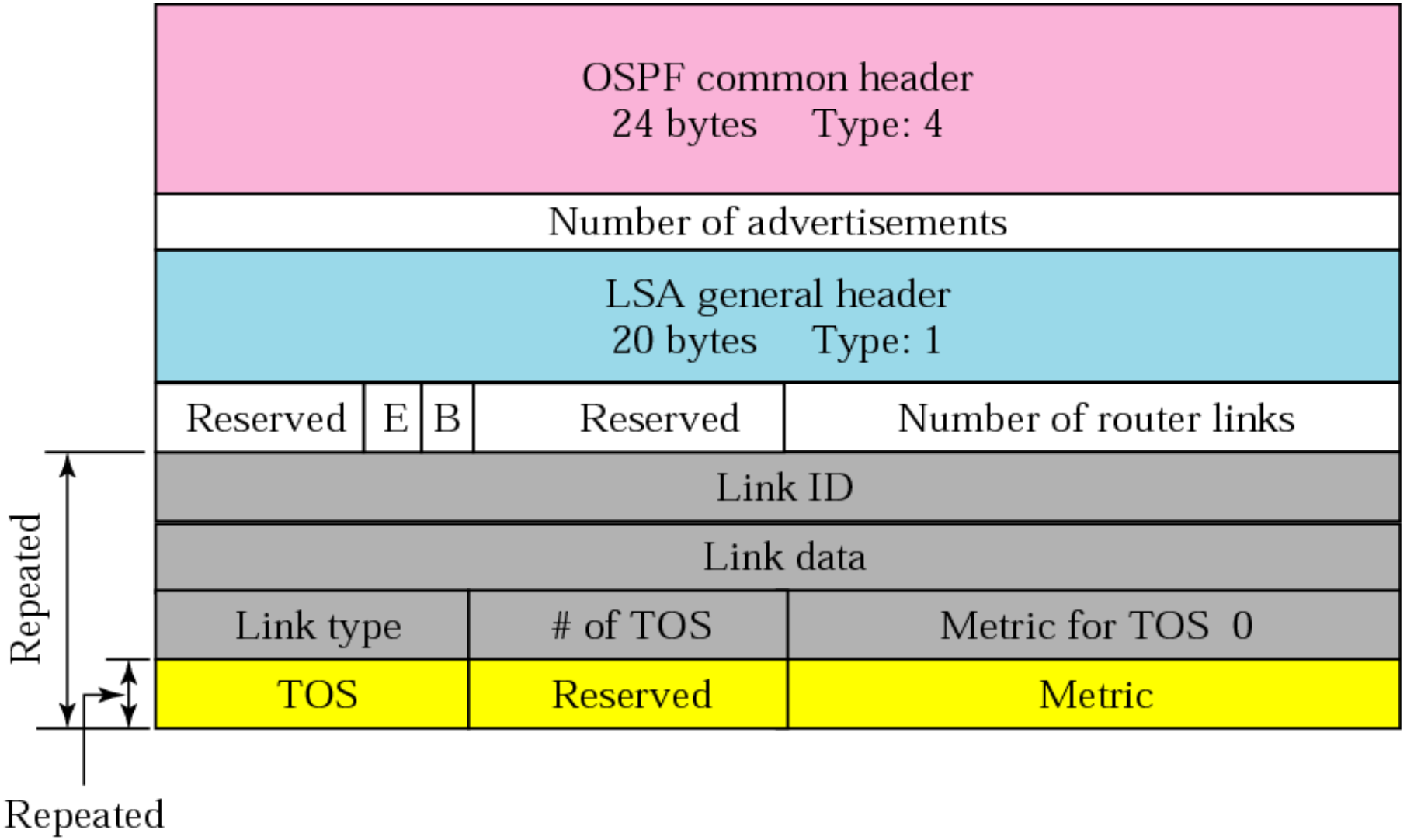


Router Link LSA

- **Router Link**
  - **defining the links of a true router**
  - **A true router uses the advertisement to announce information about all of its links and what is at the other side of the link (neighbors)**



Router Link LSA (cont'd)



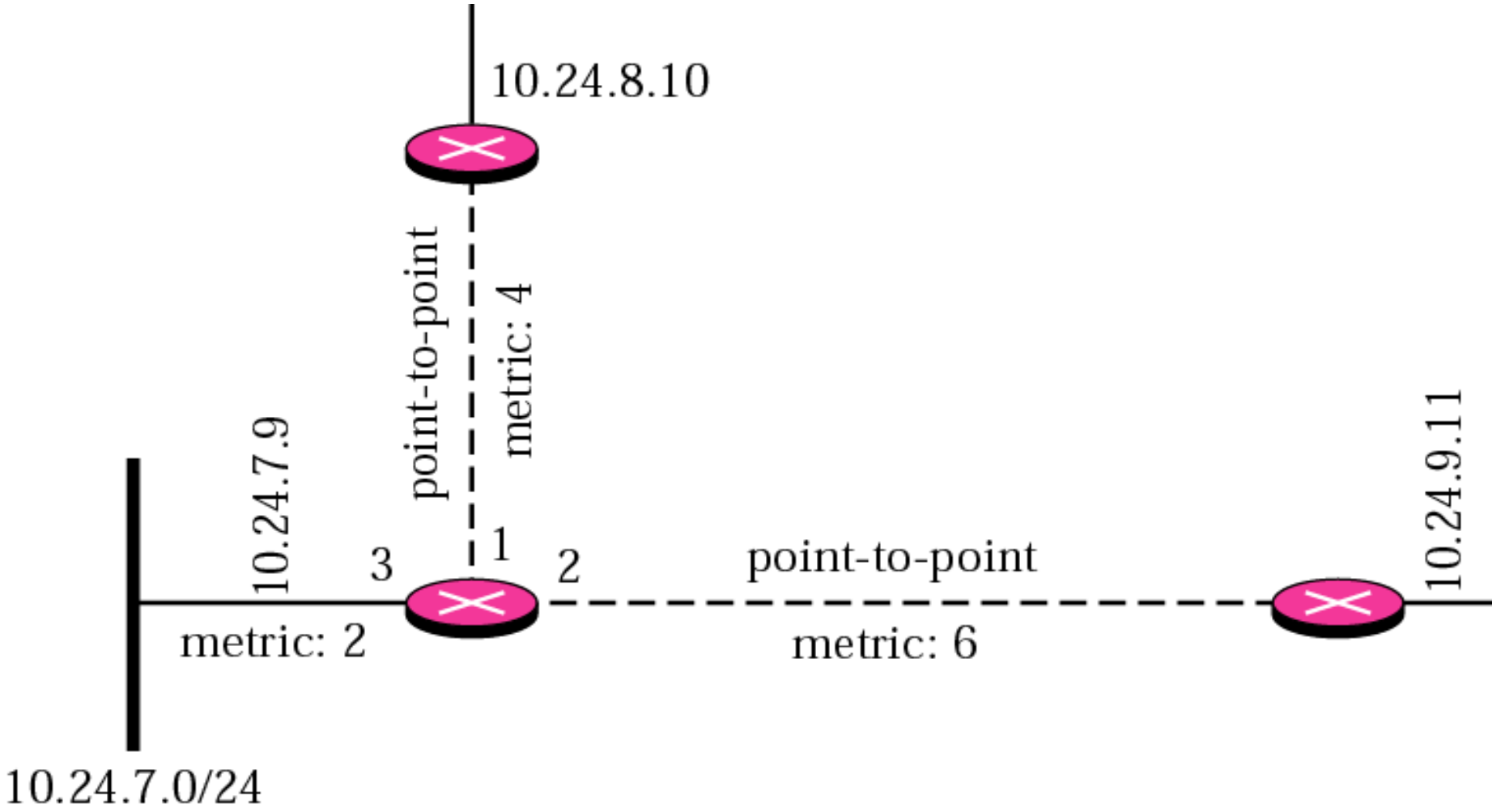
Router Link LSA (cont'd)

Table 14.2 Link types, link identification, and link data

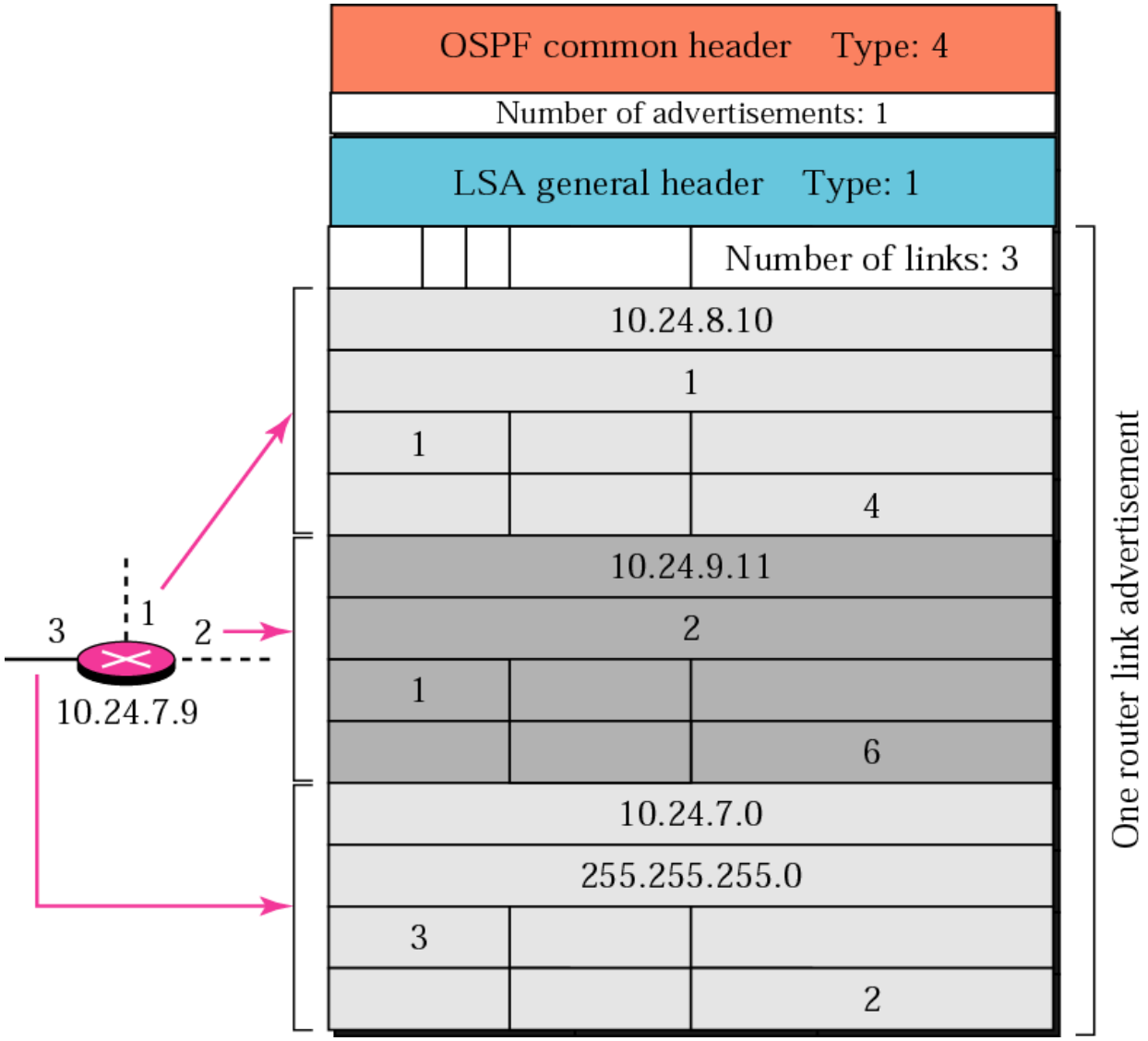
Link Type	*Link Identification*	*Link Data*
Type 1: Point-to-point	Address of neighbor router	Interface number
Type 2: Transient	Address of designated router	Router address
Type 3: Stub	Network address	Network mask
Type 4: Virtual	Address of neighbor router	Router address

Example 3

- Give the router link LSA sent by router 10.24.7.9 in Figure 14.31.

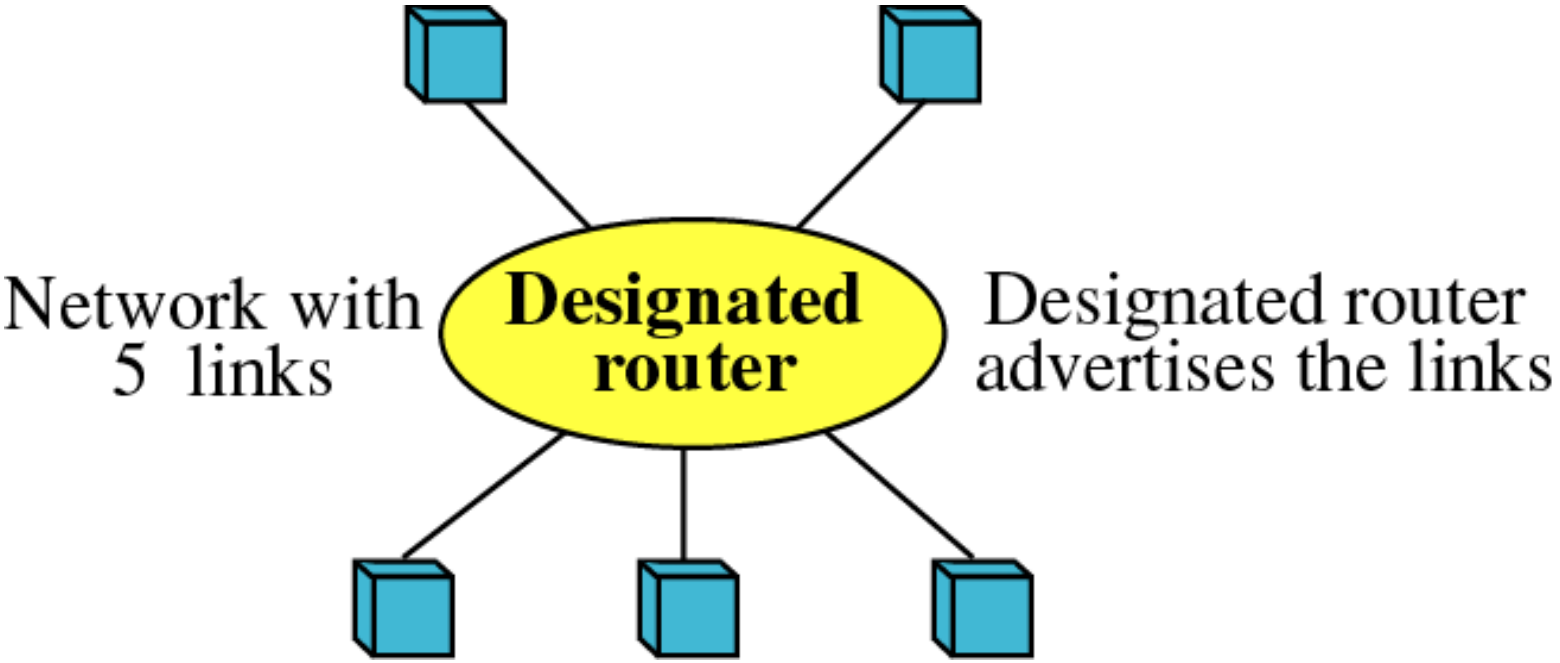


Example 3 : solution



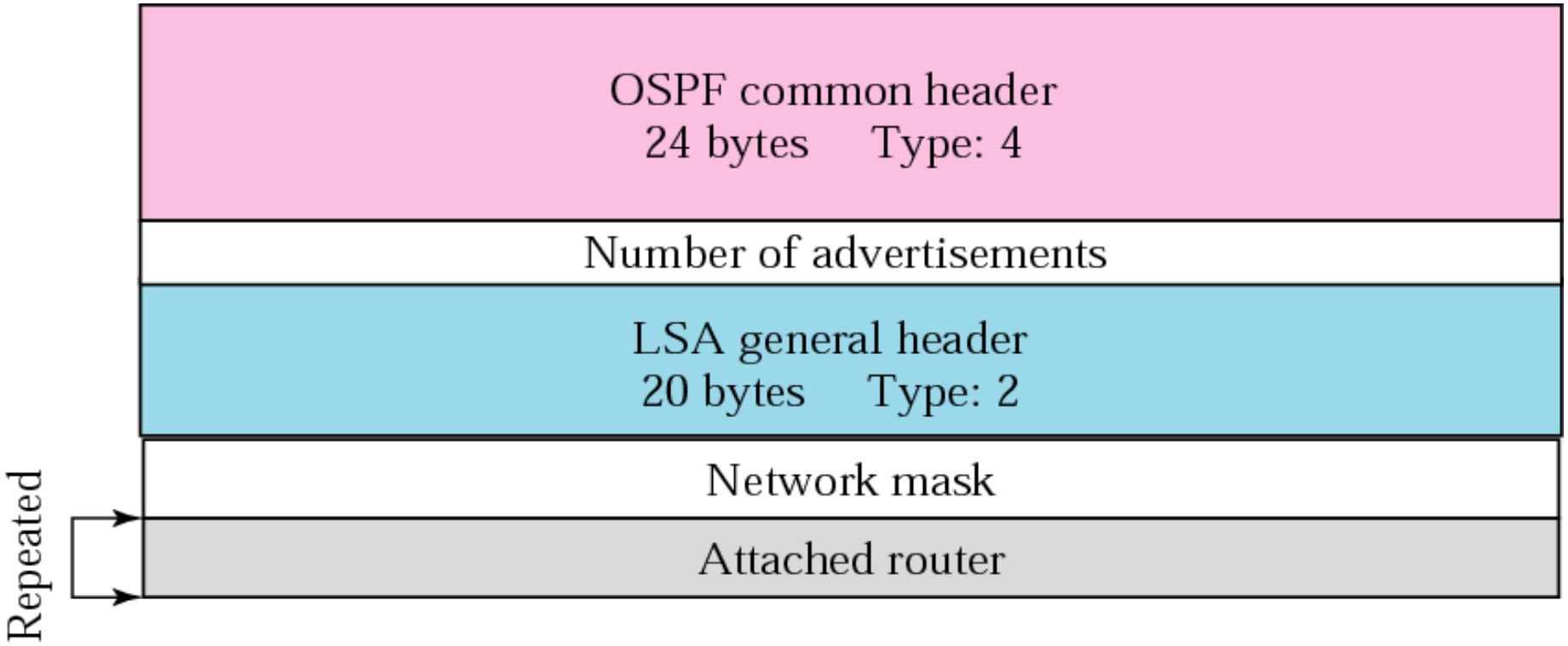
Network LINK LSA

- **Network Link**
  - **defines the links of a network**
  - **A designated router distributes this type of LSA packet.**
  - **The packet announces the existence of all of the routers connected to the network.**



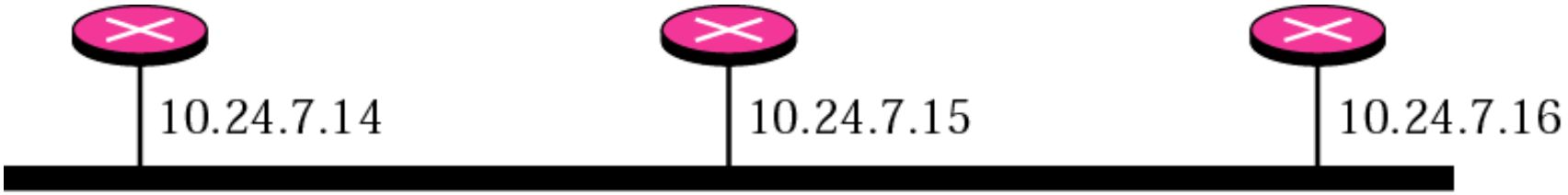
Network LINK LSA

- **Network Link Advertisement Format**

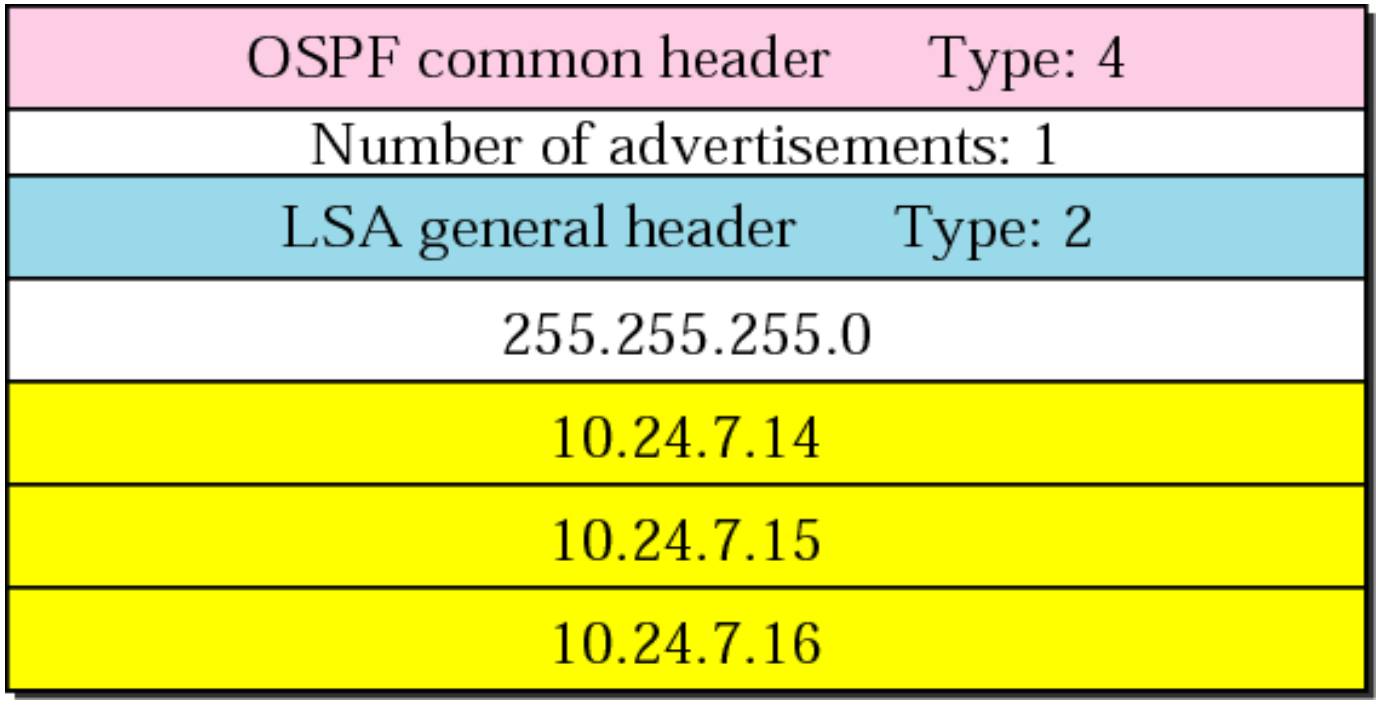


Example 4

- Give the network link LSA in Figure 14.35.

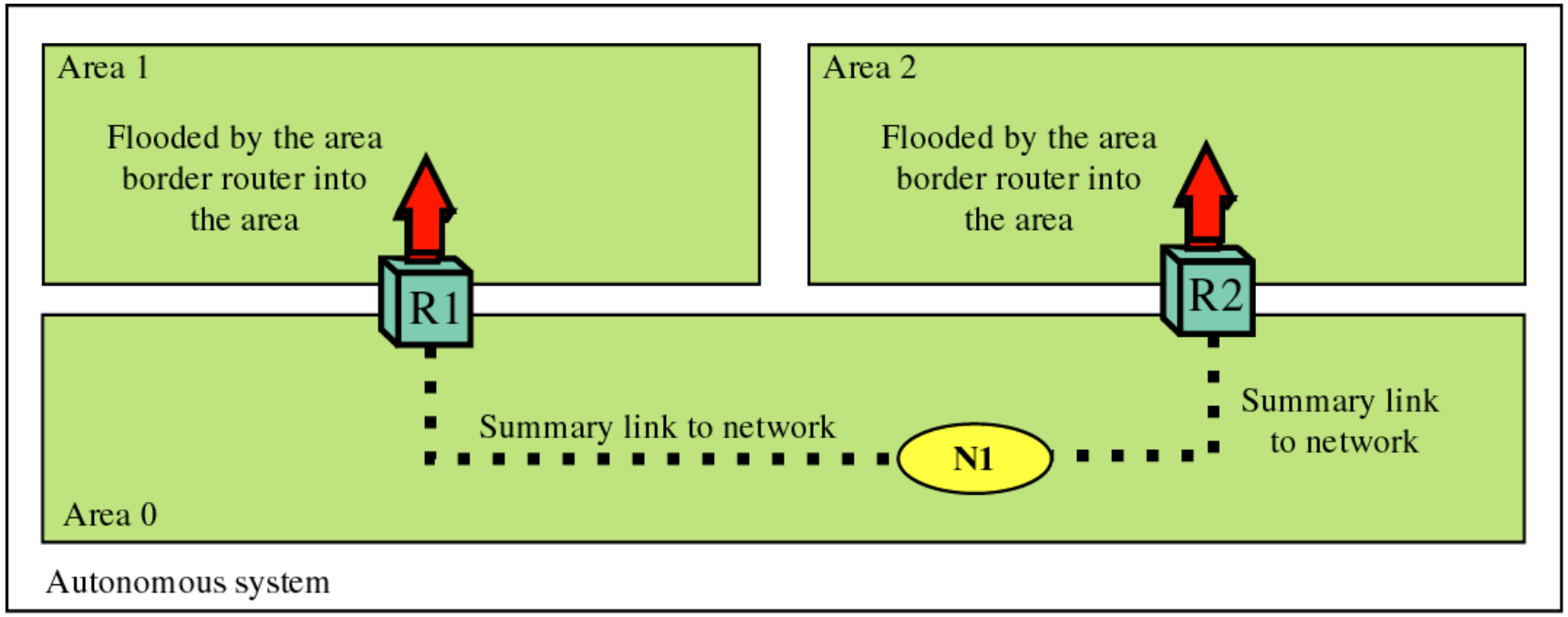


Example 4 : solution



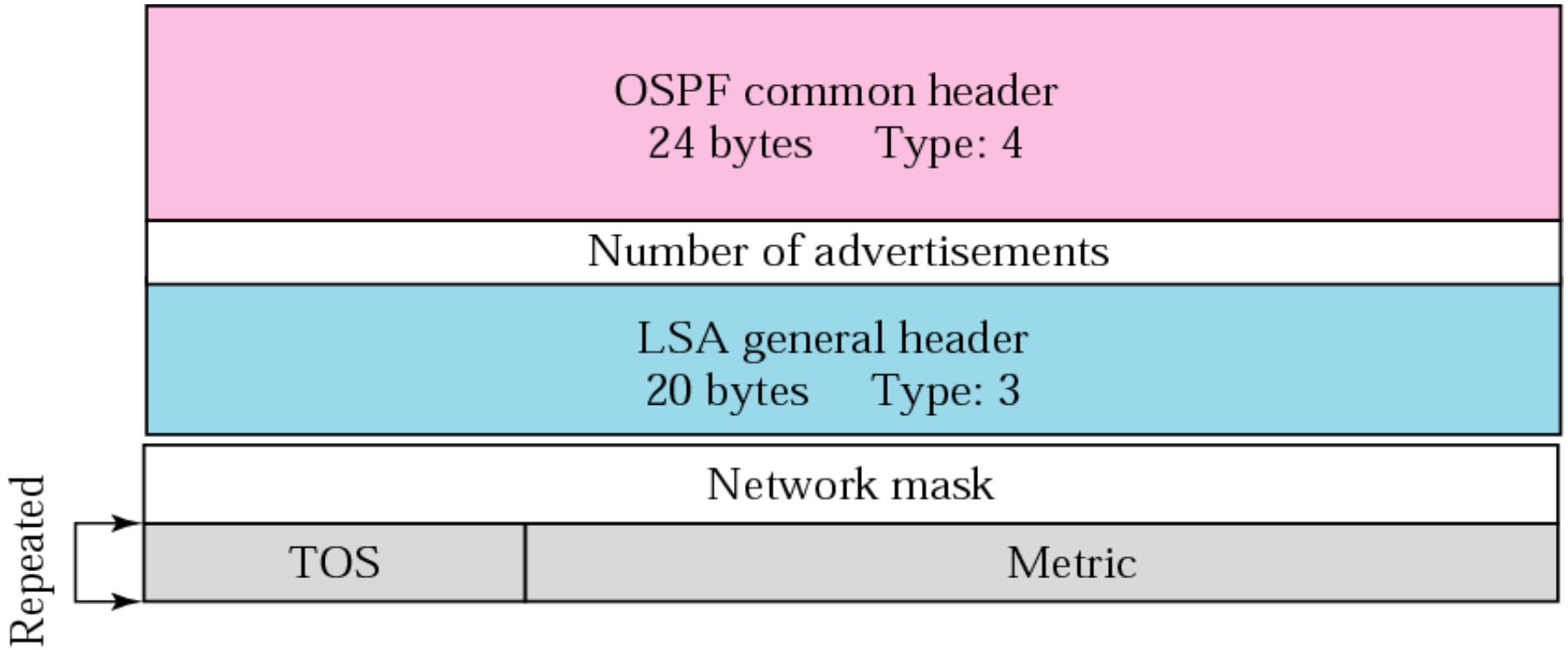
Summary Link to Network LSA

- An border router is active in more than one area and creates routing table for each area.



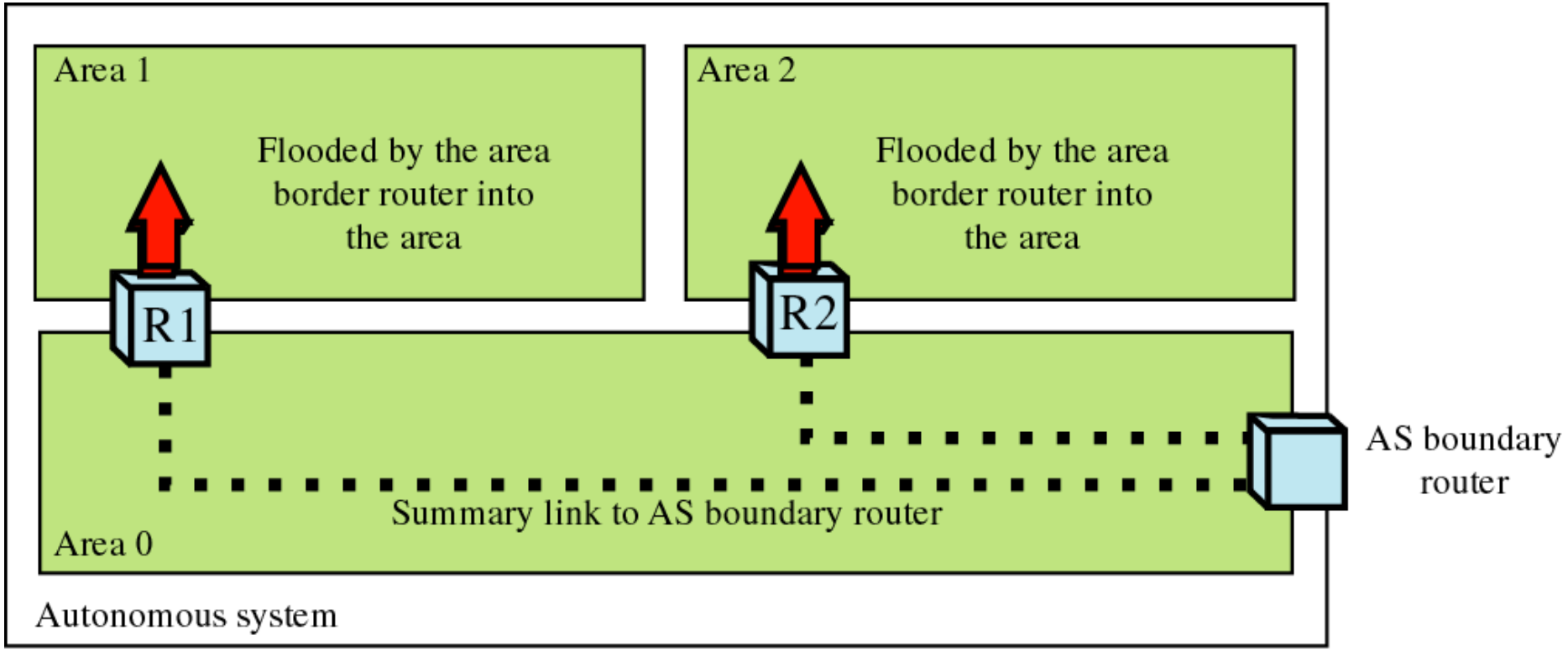
- Router R1 floods area 1 with information about how to reach a network located in area 0.

Summary Link to Network LSA (cont'd)

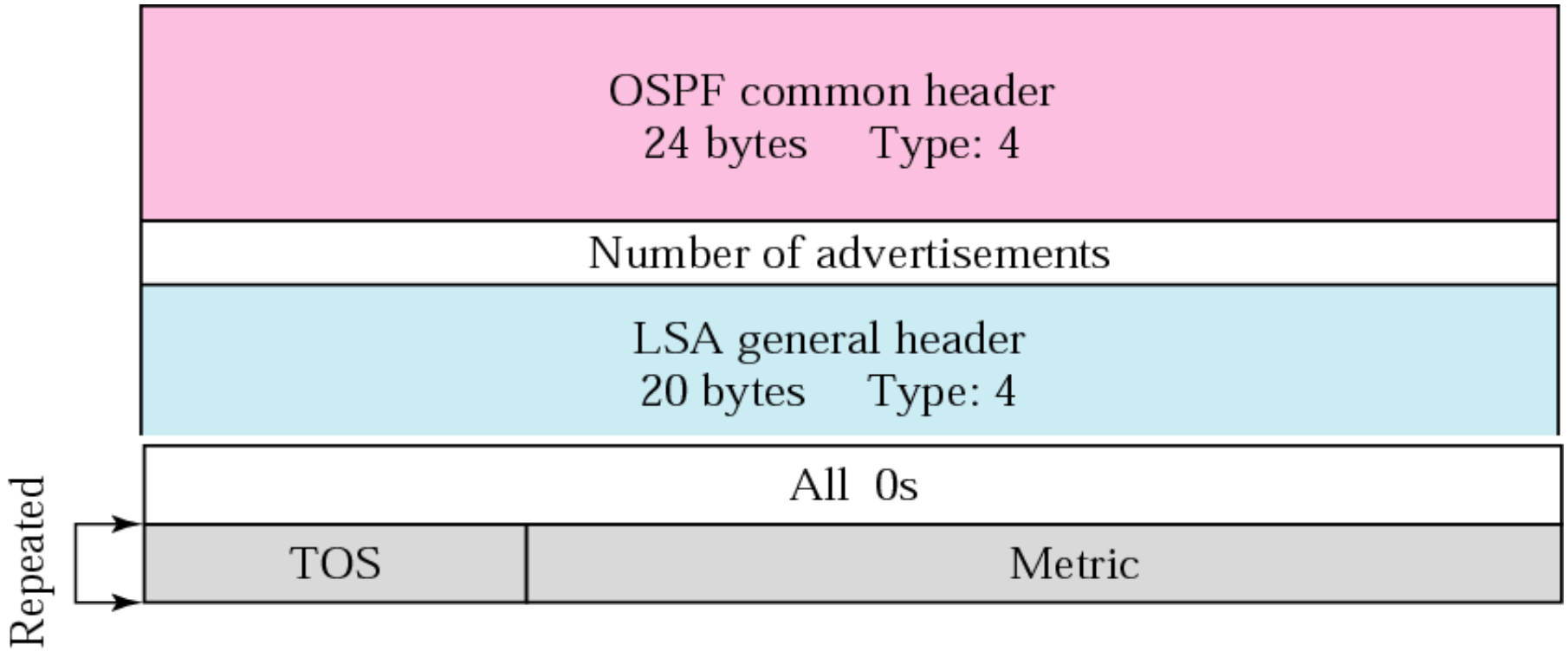


Summary Link to AS Boundary Router

- providing the information of the route to an autonomous system boundary router
  - used for a router that sends a packet outside the autonomous system

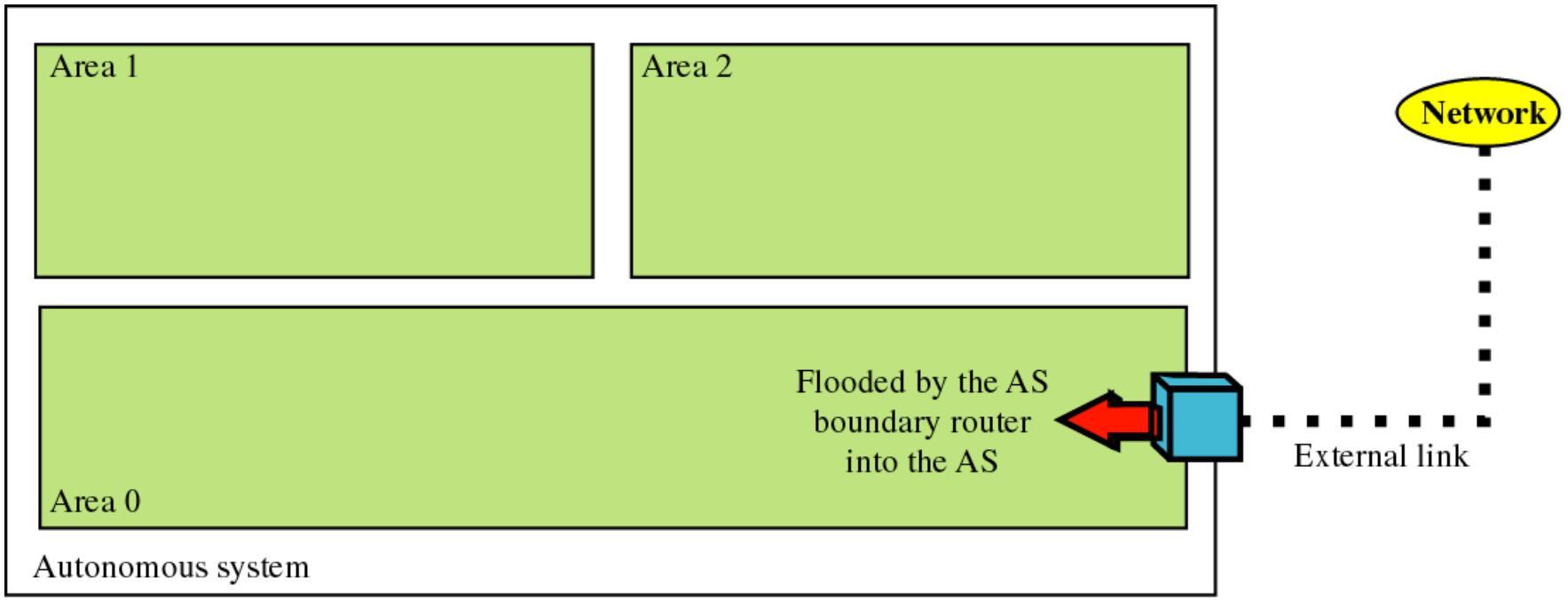


Summary Link to AS Boundary Router (cont'd)

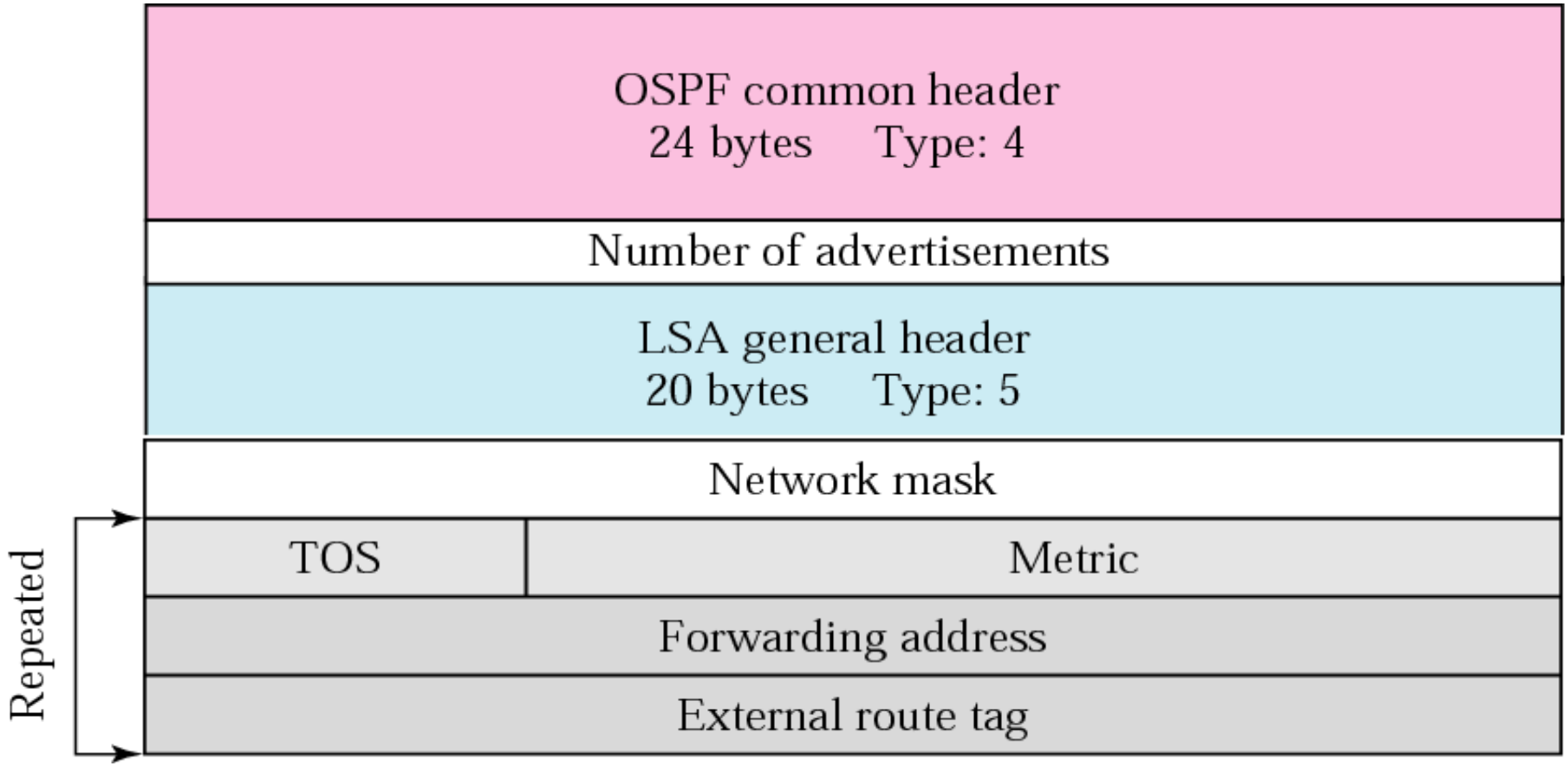


External Link

- used to know which networks are available outside the autonomous system

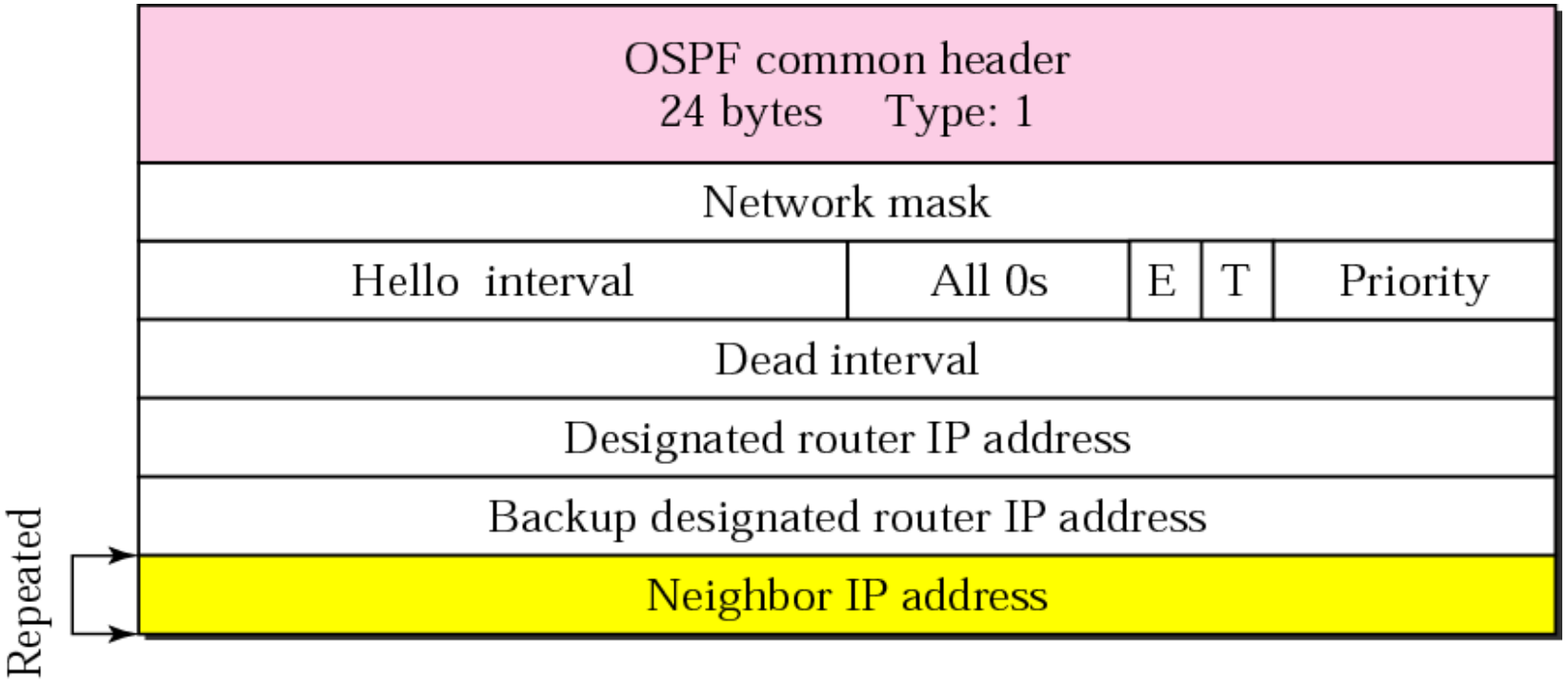


External Link (cont'd)



Other Packets

- **Hello message**
  - **uses to create neighborhood relationships and to test the reachability of neighbors**
  - **is the first step in link state routing**



Other Packets (cont'd)

- **Database description message**
  - **When router is connected to the system for the first time or after a failure, it needs the complete link state database immediately**
  - **used when a router is connected to the system for the first time or after a failure**
    - **After a router is connected to the system, the router sends hello packets to greet its neighbor.**
    - **If it is first time that neighbors hear from the router, they send a** *database description packet.*
      - *The packet does not contain complete database information*
    - **Then, the router sends one or more link state request packets to get full information about that particular link**

- **Link State Request Packet**
  - **is sent by a router that needs information about a specific route or routes**
  - **It is answered with a link state update packet.**

Other Packets (cont'd)

- Link state acknowledgment packet
  - OSPF makes routing more reliable by forcing every router to acknowledge the receipt of every link state update packet.
- Link State Update Packet
  - used by a router to advertise the states of its links

Encapsulation of OSPF Packets

- **Encapsulation**
  - **OSPF packets are encapsulated in IP datagram**
    - **These packets contain the acknowledgment mechanism for flow and error control**
    - **Do not need a transport layer protocol to provide these services**

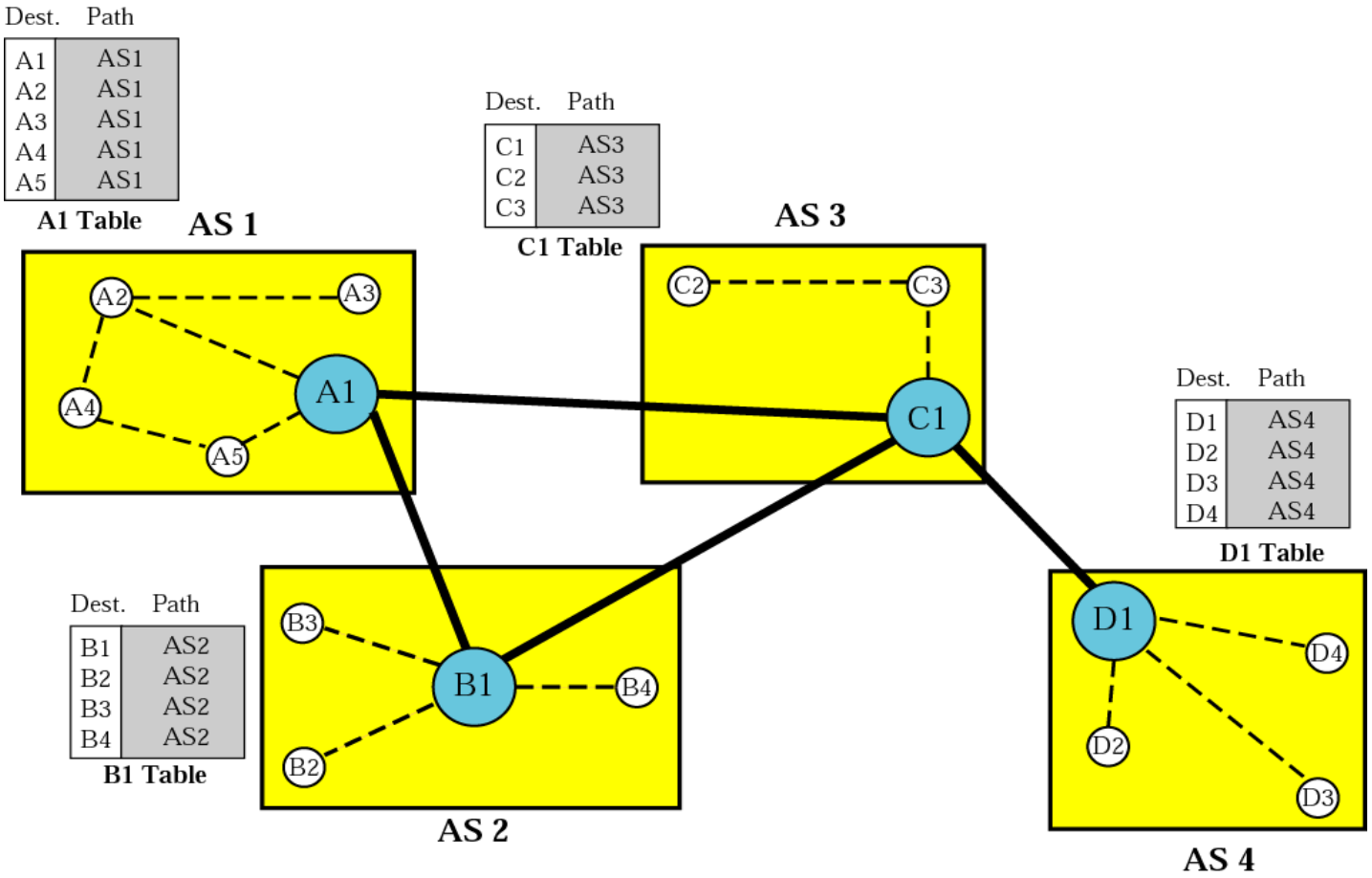
14.6 Path Vector Routing

- is similar to distance vector routing
- Assuming that there is one node in each AS that acts as on behalf of the entire AS : Speaker Node
- Speaker node creates a routing table and advertises it speaker nodes in the neighboring ASs
  - advertising the path, not the metric of the nodes

Path Vector Routing (cont'd)

Initialization

- Each speaker node can know only the reachability of nodes inside its AS



14.7 BGP

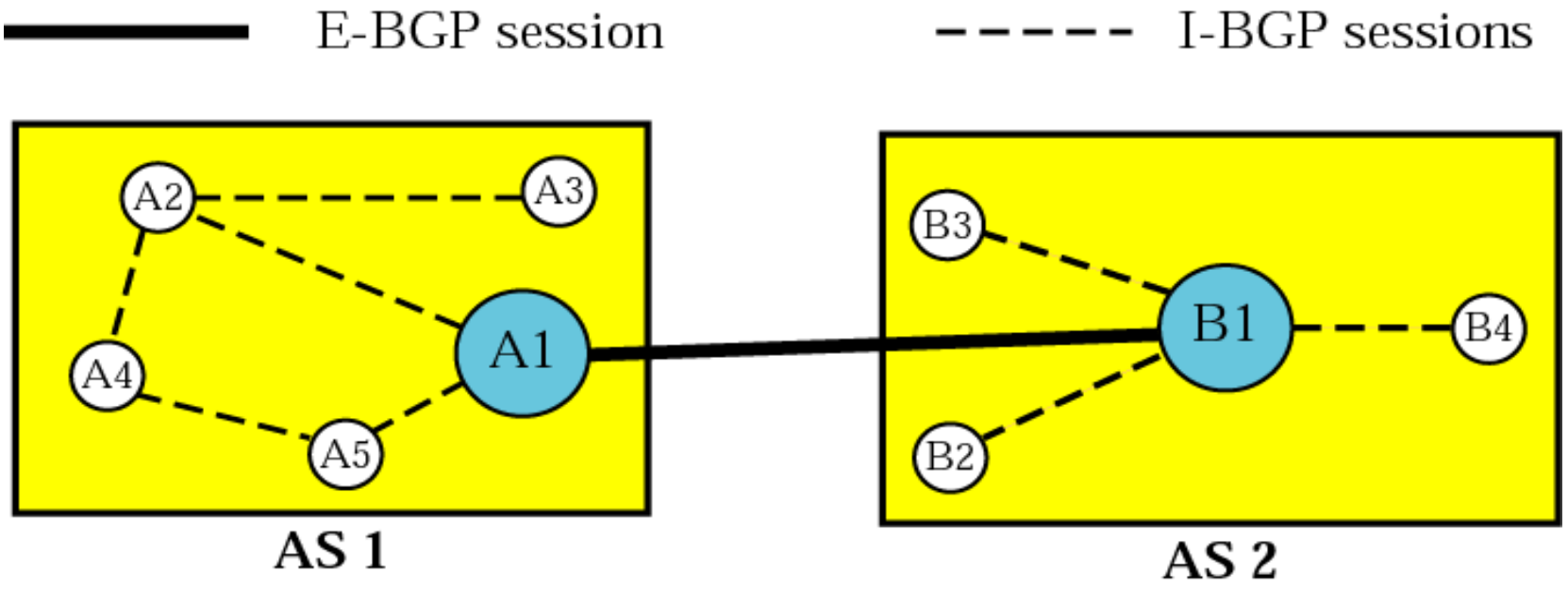
- Border Gateway Protocol is an interdomain routing protocol using path vector routing
- Distance vector routing and link state routing
  - distance vector routing : just considering the number of hops
  - link state routing : requiring each router to have a huge link state database
- Path Vector Routing
  - Each entry in the routing table contains the destination network, the next router, and the path to reach the destination
  - The path is usually defined as an ordered list of autonomous systems that a packet should travel through to reach the destination

Types of Autonomous Systems

- **Stub AS**
  - **has only one connection to another AS**
- **Multihomed AS**
  - **has more than one connection to other ASs**
- **Transit AS**
  - **is a multihomed AS that also allows transient traffic.**
    - **ex. national and international ISPs**

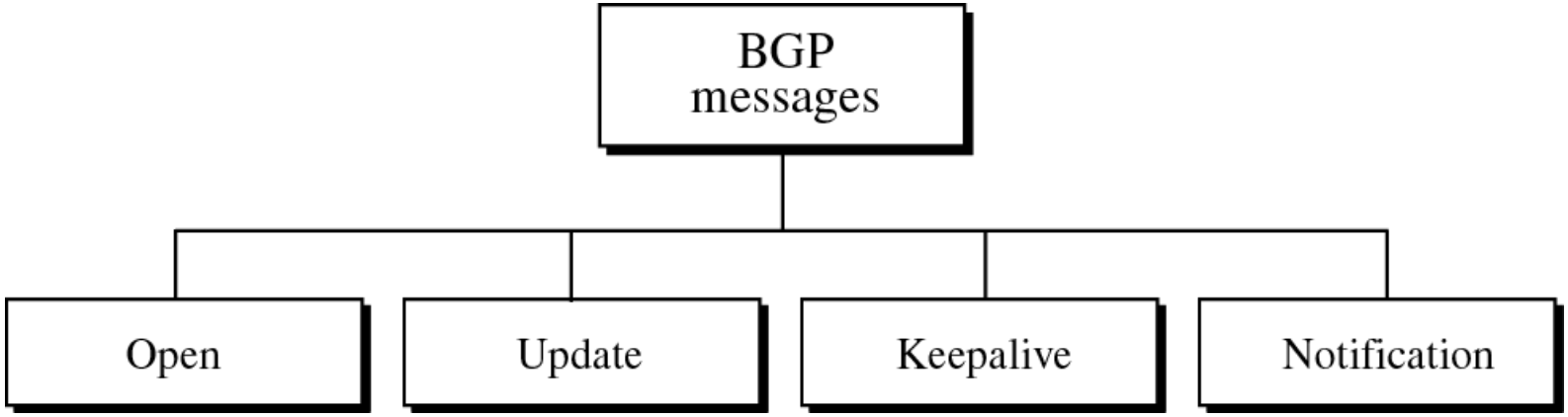
Types of BGP Connections

- External and Internal BGP

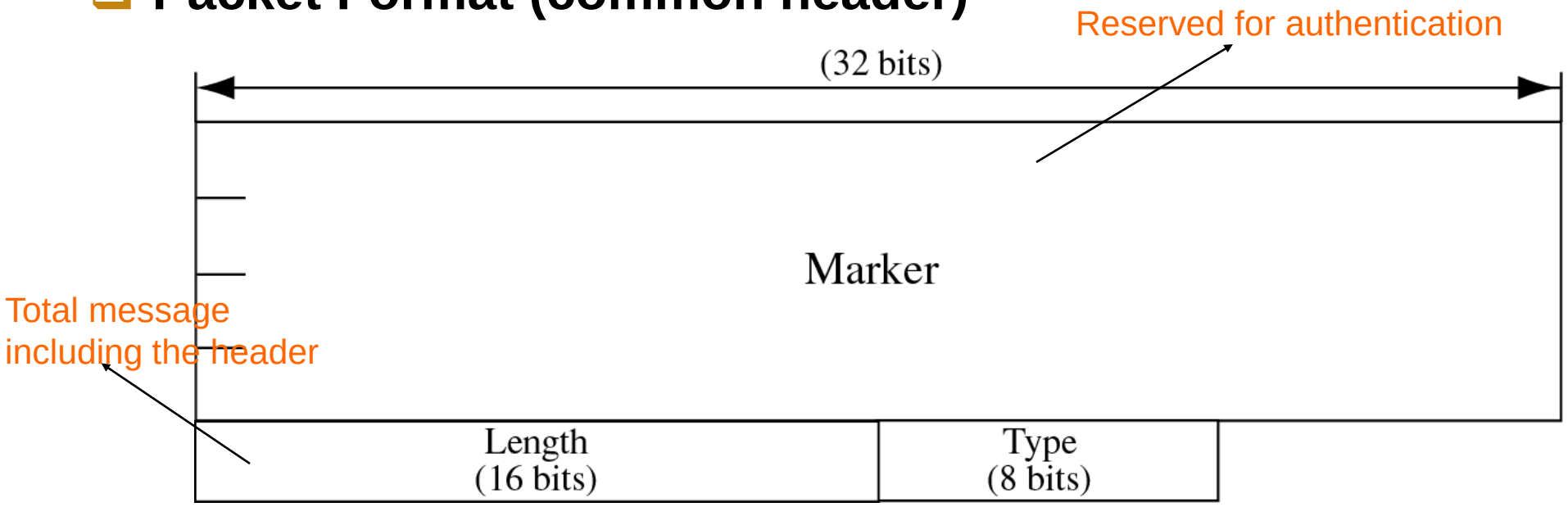


BGP (cont'd)

- **Types of Packets**



- **Packet Format (common header)**



BGP (cont'd)

- Open message
  - To create a neighborhood relationship, a router running BGP opens a TCP connection with a neighbor and sends an open message
- Update message
  - used by a router to withdraw destinations that have been advertised previously, announce a route to a new destination, or both
- Keepalive message
  - exchange keepalive messages regularly (before their hold time expires) to tell each other that routers are alive
- Notification message
  - sent by a router whenever an error condition is detected or a router wants to close the connection

BGP (cont'd)

- **Encapsulation**
  - **BGP messages are encapsulated in TCP segments using the well-known port 179**

Summary(1)

- A metric is the cost assigned for passage of a packet through a network.
- A router consults its routing table to determine the best path for a packet.
- An autonomous system (AS) is a group of networks and routers under the authority of a single administration.
- RIP and OSPF are popular intradomain routing protocols used to update routing tables in an AS.
- RIP is based on distance vector routing, in which each router shares, at regular intervals, its knowledge about the entire AS with its neighbors.
- A RIP routing table entry consists of a destination network address, the hop count to that destination, and the IP address of the next router.
- RIP uses three timers: the periodic timer controls the advertising of the update message, the expiration timer governs the validity of a route, and the garbage collection timer advertises the failure of a route.
- Two shortcomings associated with the RIP protocol are slow convergence and instability.
- Procedures to remedy RIP instability include triggered update, split horizons, and poison reverse.

Summary(2)

- The RIP version 2 packet format contains fields carrying AS information and authentication information.
- OSPF divides an AS into areas, defined as collections of networks, hosts, and routers.
- OSPF is based on link state routing, in which each router sends the state of its neighborhood to every other router in the area. A packet is sent only if there is a change in the neighborhood.
- OSPF defines four types of links(networks): point-to-point, transient, stub, and virtual.
- Five types of link state advertisements (LSAs) disperse information in OSPF: router link, network link, summary link to network, summary link to AS boundary router, and external link.
- A router compiles all the information from the LSAs it receives into a link state database. This database is common to all routers in an area.
- OSPF routing tables are calculated using Dijkstra's algorithm.
- There are five types of OSPF packets: hello, database description, link state request, link state update, and link state acknowledgment.

Summary(3)

- An LSA is a multifield entry in a link state update packet.
- BGP is an inter autonomous system routing protocol used to update routing tables.
- BGP is based on a routing protocol called path vector routing. In this protocol, the ASs through which a packet must pass are explicitly listed.
- Path vector routing does not have the instability nor looping problems of distance vector routing.
- There are four types of BGP messages: open, update, keepalive, and notification.