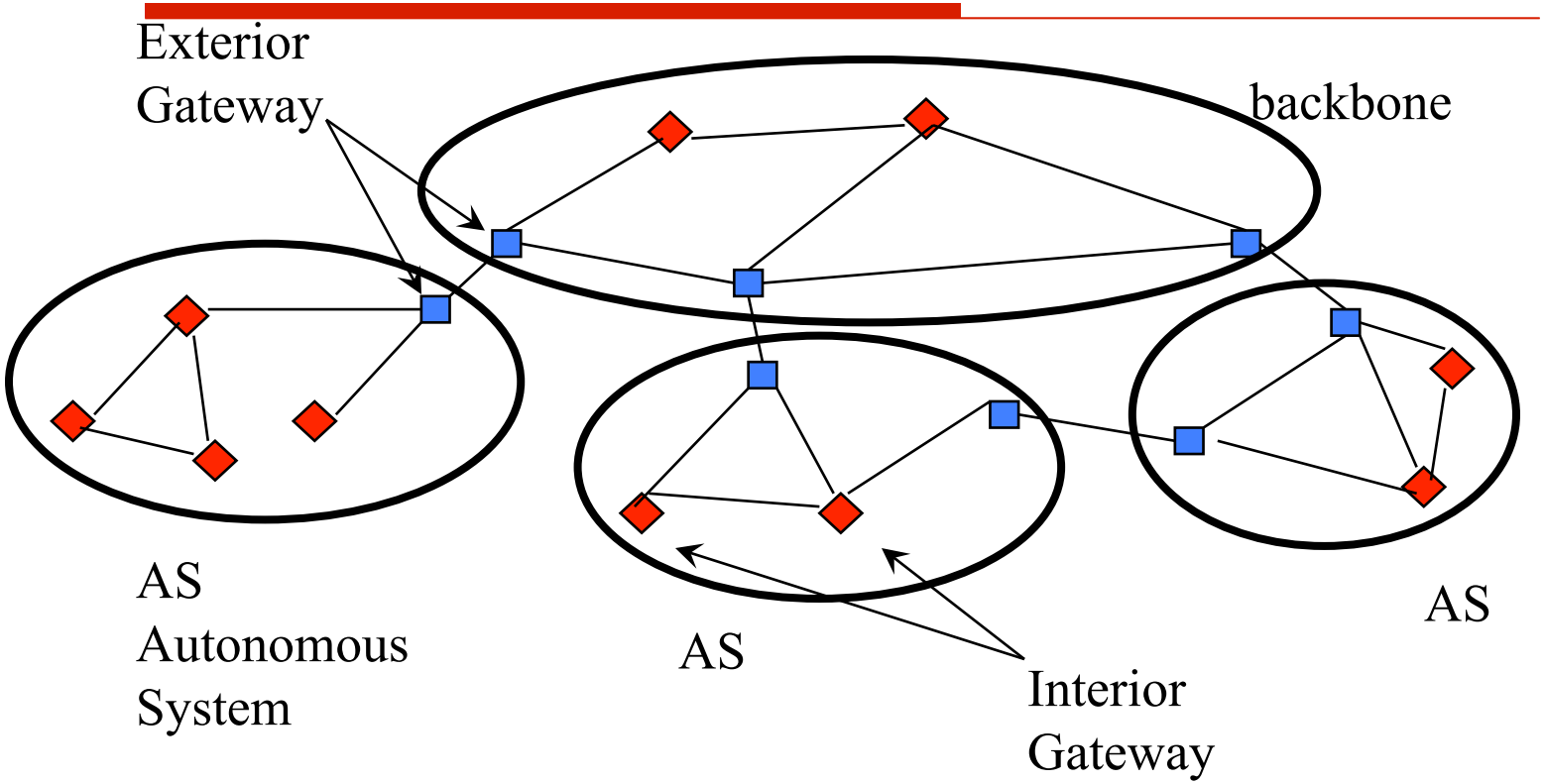
Internet Routing

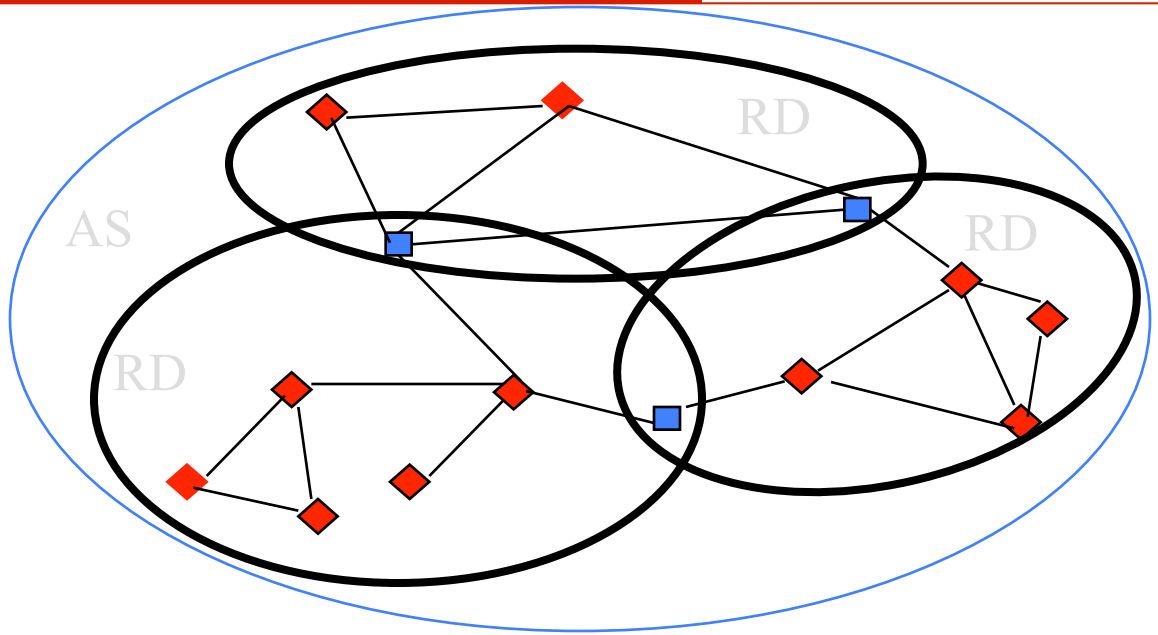
RIP, OSPF, BGP

Routing in Internet



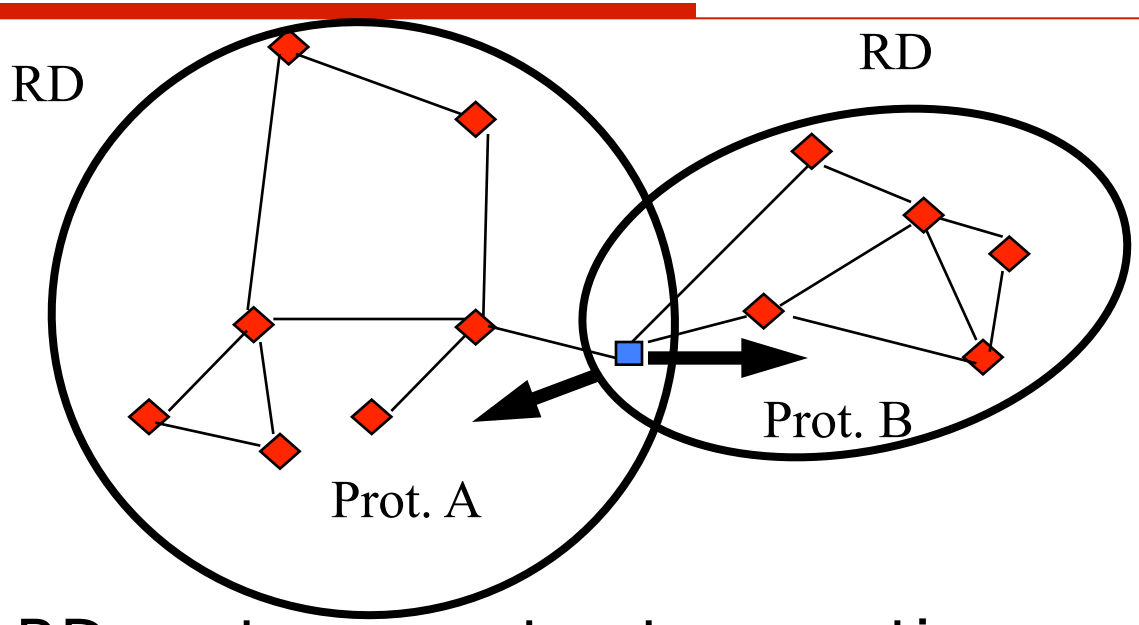
- *Autonomous System*: portion of Network managed by a single organization
- EGP - *Exterior Gateway Protocol*
- IGP - *Interior Gateway Protocol*

Routing Domains



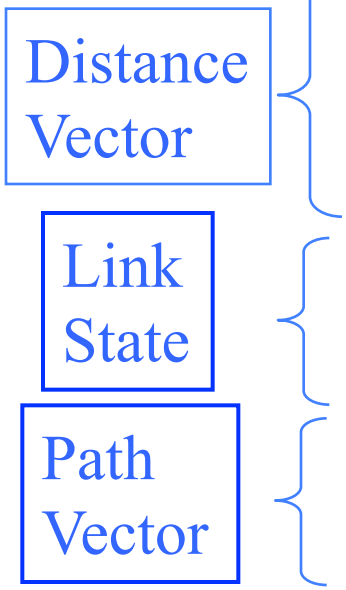
- Routing Domain (RD): portion of an AS running a single routing protocol
- some *routers* belonging to multiple RDs implement multiple routing protocols

Routing Distribution



- Multiple RD routers must act as routing protocols gateways
- Translation from Prot. A to Prot. B depends on the implementation of A and B
- Prot A and B may be one IGP and one EGP (distribution criteria are defined)

The most common routing protocols

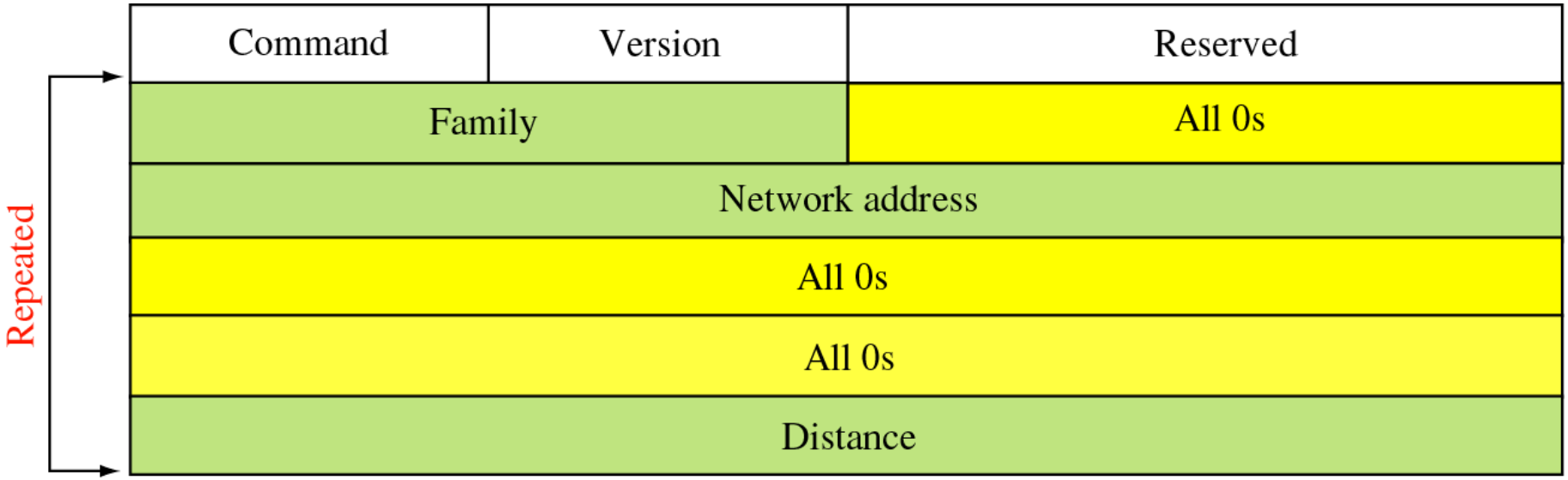


- IGP
  - RIP (Routing Information Protocol), version 1 and 2
  - IGRP (Interior Gateway Routing Protocol) CISCO proprietary
  - IS-IS (Intermediate System Intermediate System)
  - OSPF (Open Shortest Path First)
- EGP
  - BGP (Border Gateway Protocol)

RIP Version 1

- Designed at *Berkeley* (1982) and standardized in RFC 1058
- IGP
- *Distance Vector*, uses *Bellman-Ford* to compute shortest paths
- Metrics: number of hops
- Limited to 16 *hops*
- RIP messages are encapsulated into UDP segments (port: 520) and sent with IP destination address: 255.255.255.255

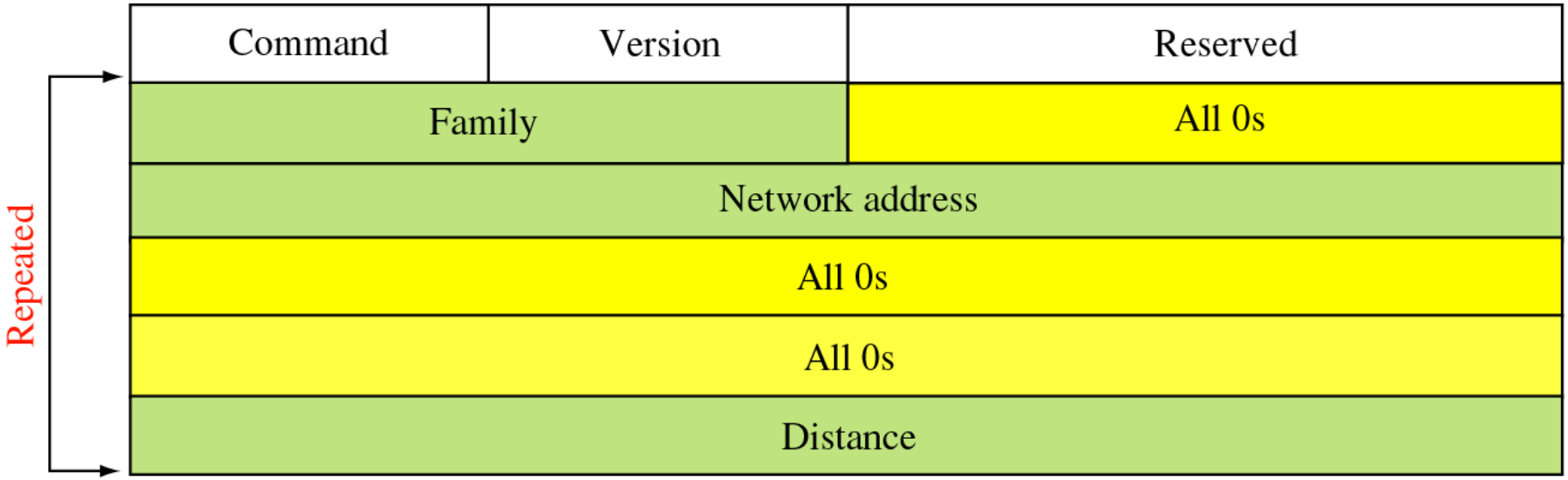
RIP v1: message format



Source: TCP/IP Protocol Suite, B. Forouzan

- RIPv1 messages can be:
  - Requests (of sending a DV)
  - Responses (stimulated/non stimulated)

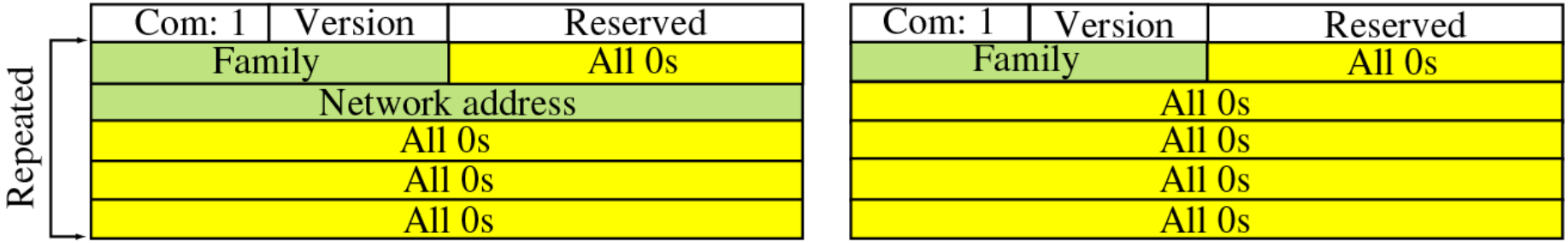
RIP v1: message format



Source: TCP/IP Protocol Suite, B. Forouzan

- Command: 1=request, 2=response
- Version: RIP version
- Family: address family used (2=IP)
- Net. Address: add. Of the destination network
- Distance: cost (from 1 to 15, 16=inf)

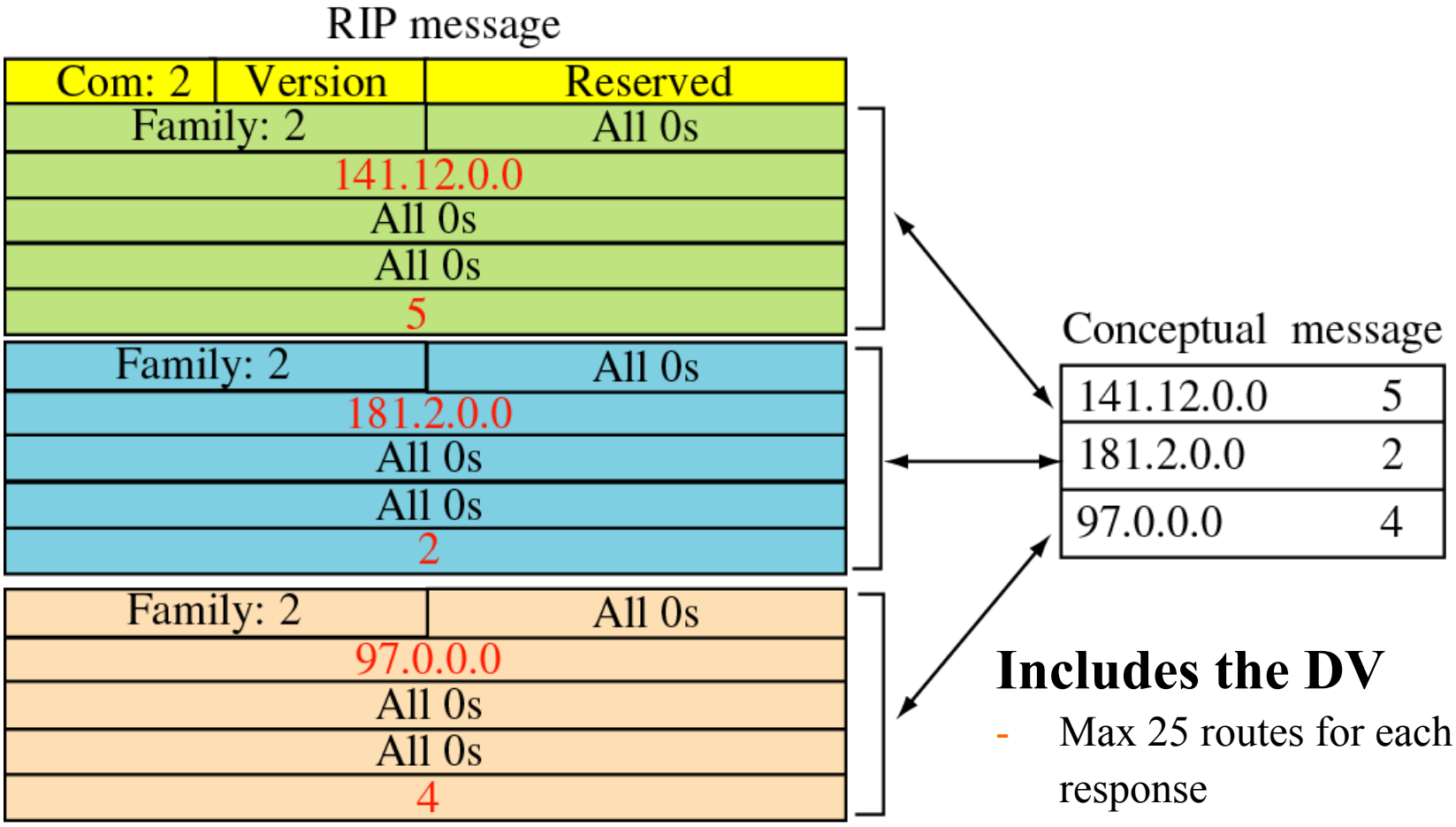
Request Messages



Source: TCP/IP Protocol Suite, B. Forouzan

- Requests may come from
  - "Just-Switched-on" router
  - A router having some destination out of date
- Requests may deal with
  - All the destinations
  - A specific destination

Response Messages



Source: TCP/IP Protocol Suite, B. Forouzan

Includes the DV

- Max 25 routes for each response
- >25 routes transferred with more UDP messages

RIP v1: message timing

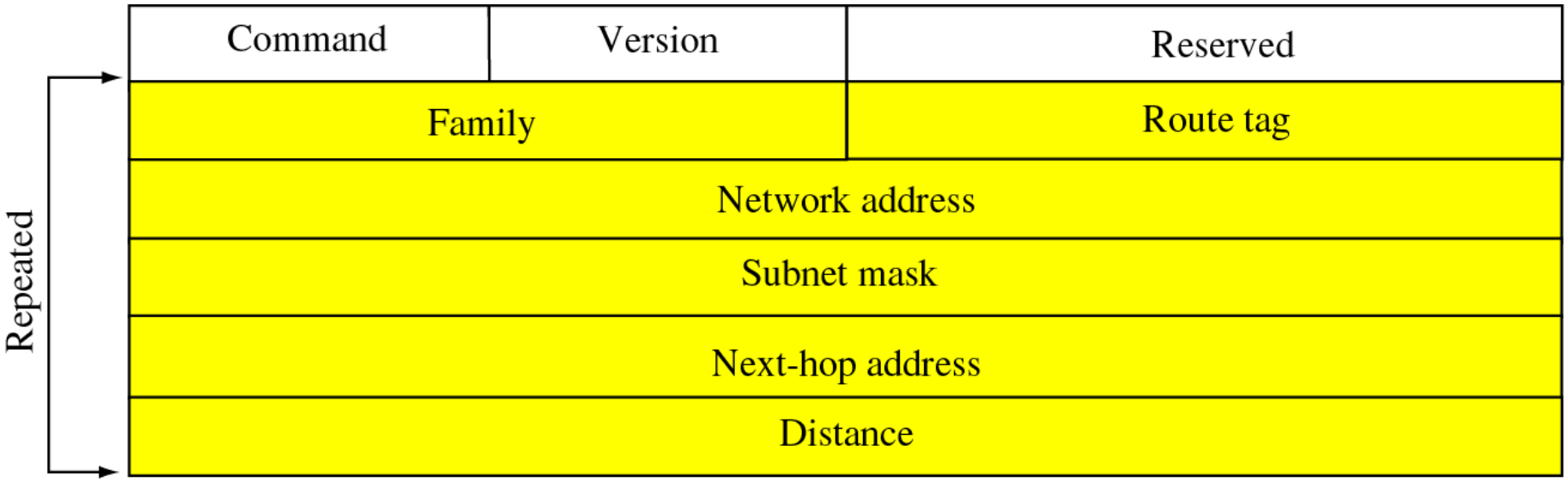
- *routing update timer* (default 30 s)
  - Period of time between the tx of 2 contiguous DV
- *route invalid or duration timer* (default 180 s)
  - If no DV is received from an interface in this interval, the routes are declared invalid. It is still announced, but with distance= 16
- *route flush timer or garbage collection timer* (typically 60-120)
  - Time interval after which an invalid route is erased (if other DVs arrive from other interfaces they are accepted)
  - It is used to announce to neighbors about destination invalidity before cancelling it
- Triggered update: if a metric changes on a route, a DV is immediately sent with only the changed entries

RIP v1: limitations

- *Hop count is a very simplistic metric*
  - *One would like to use more complex metrics like:*
    - *Queue length*
    - *Delays*
    - *Packet error rate*
    - *…*
- *It works only for small-medium networks (up to 15 nodes as network diameter)*
- Convergence time is slow

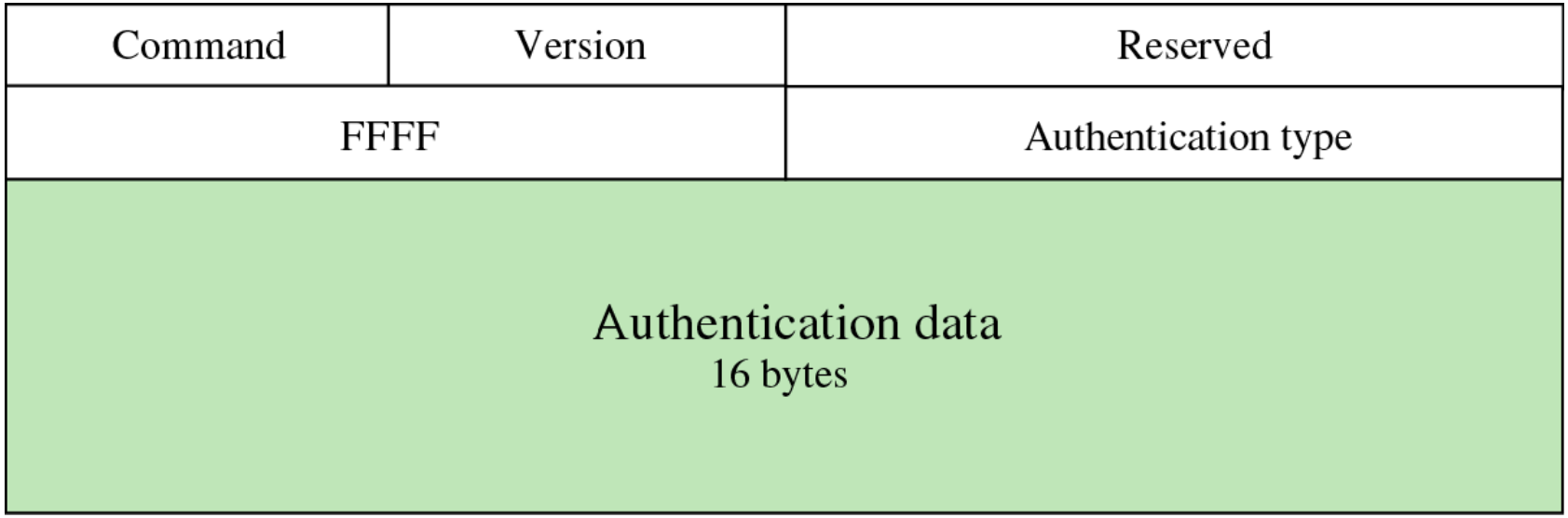
RIP Version 2

- Standardized in RFC 1723
- Added Functionalities
  - Info on connectivity
    - Explicit indication of next hop address)
    - router tag: allows to mark routes based on their origin (discovered by RIP, by other IGP protocols, by EGP …)
  - Authentication
  - Classless routing (subnet mask)
  - Multicasting: uses address 224.0.0.9 as dest. Add.



Source: TCP/IP Protocol Suite, B. Forouzan

RIPv2: Authentication



Source: TCP/IP Protocol Suite, B. Forouzan

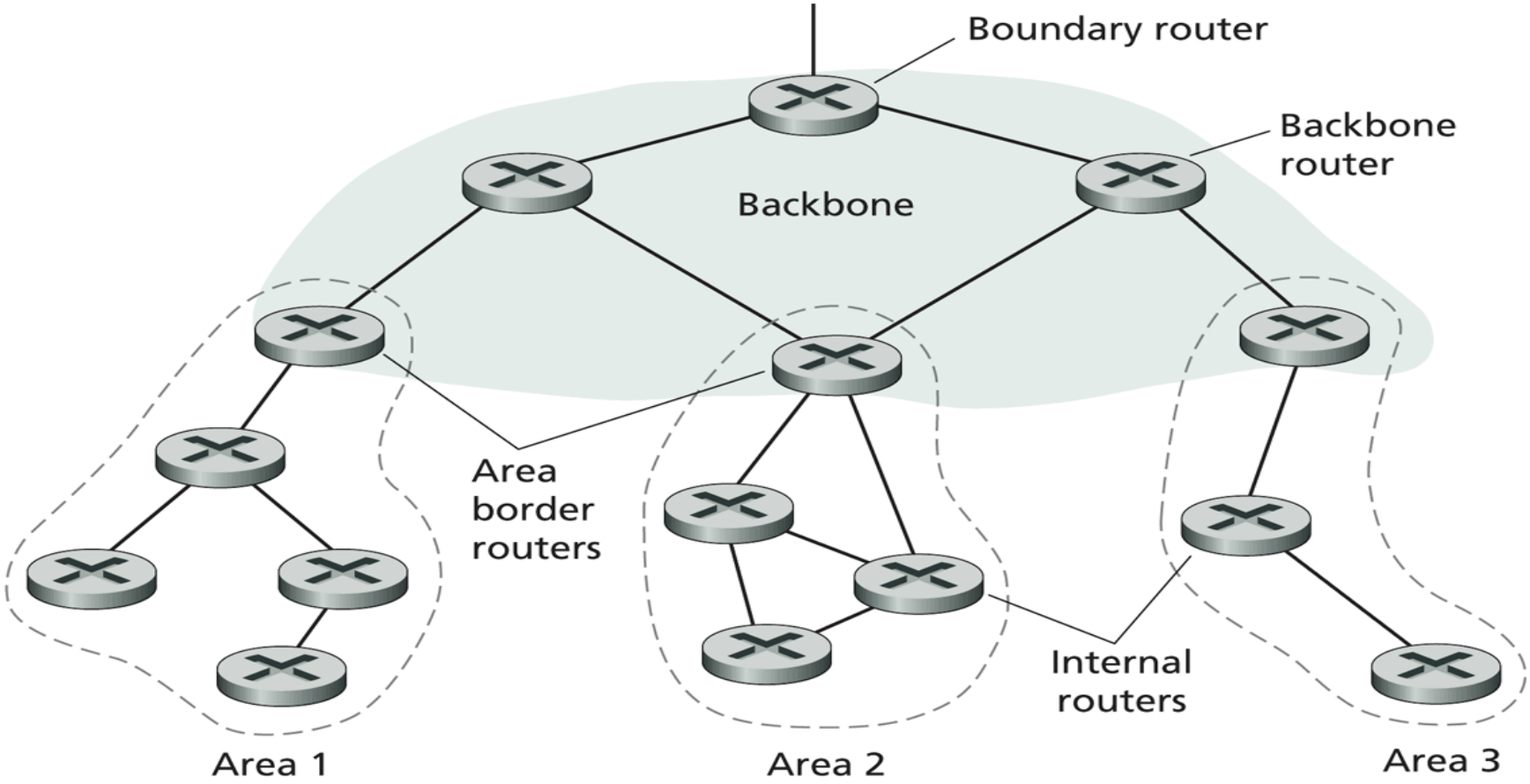
OSPF

- RFC 1247, 1583, 2328
- Link state
  - Execution of Dijkstra algorithm at each node
- OSPF supports Hierarchical routing
  - Routing areas and backbone area
- Generic metrics
  - Cost of traversing an interface can be set by the network administrator
- *Hello protocol* used to monitor the state of neighbors
- LSA (*link state advertisement*)

OSPF (2)

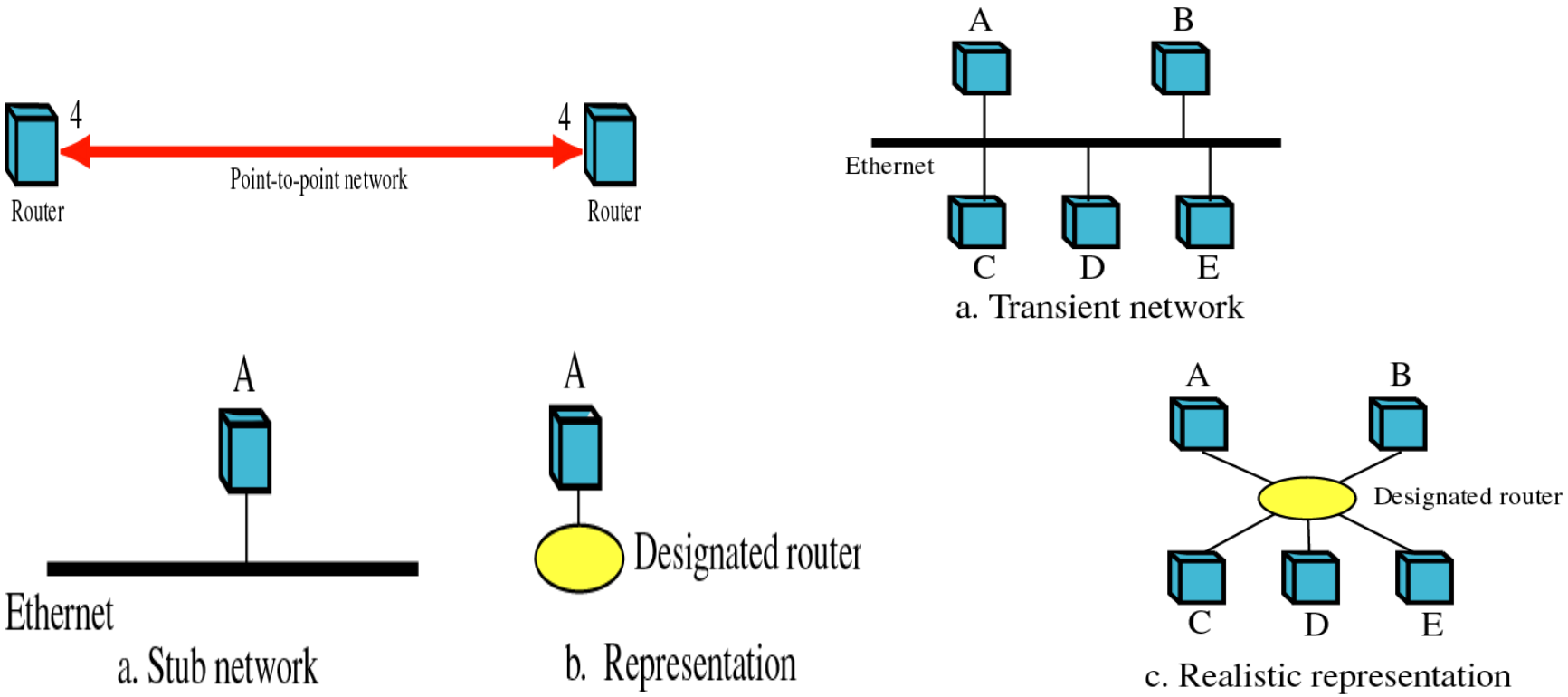
- Trasported directly over IP (Protocol = 89)
  - It must implement transport functions
    - ACK messages
  - Sever types of messages
  - It supports authentication
  - It supports multiple routes towards the destination
    - Routes with the same “length/cost” are used to perform load balancing

OSPF: routers classification



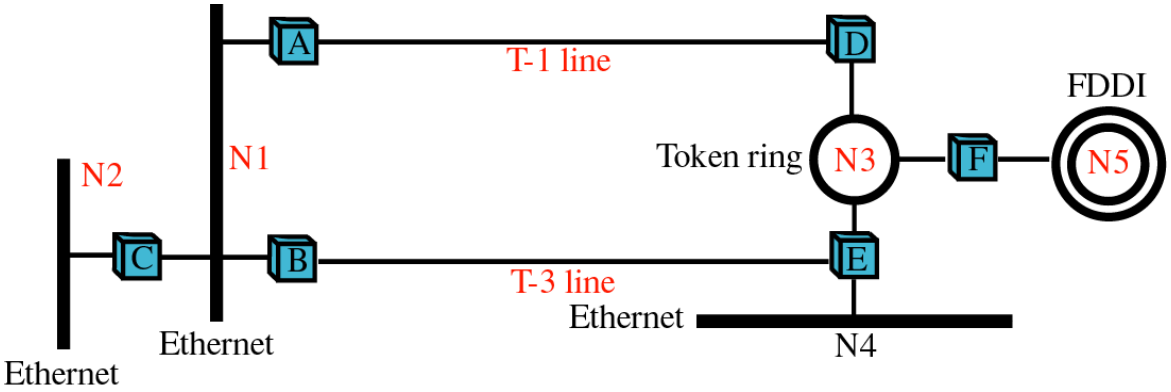
Source: Computer Networking, J. Kurose

Types of links

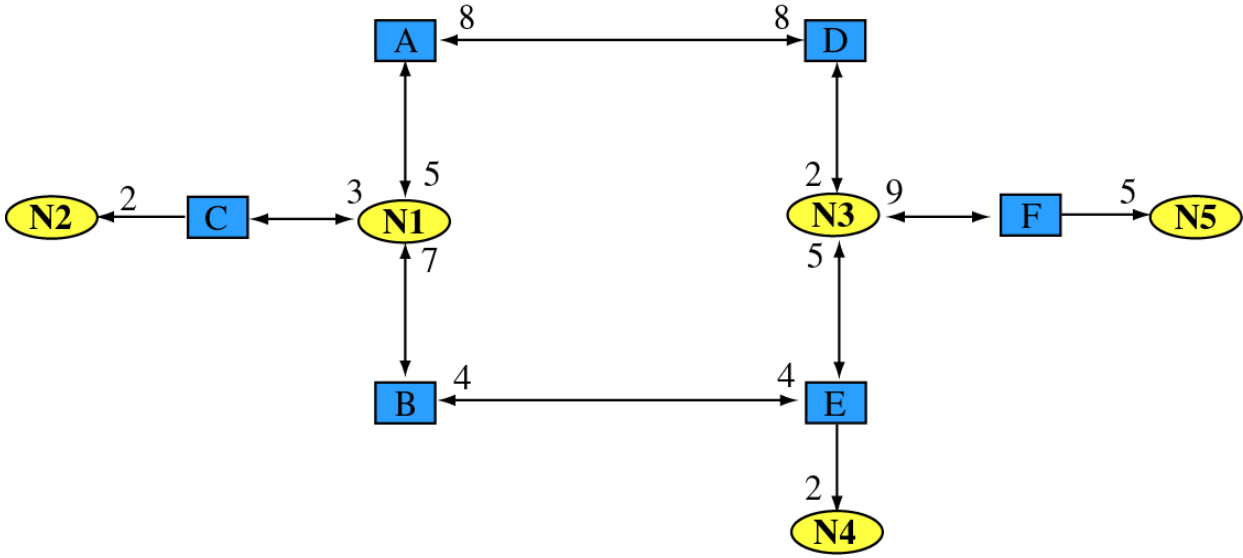


Source: TCP/IP Protocol Suite, B. Forouzan

Topology Representation



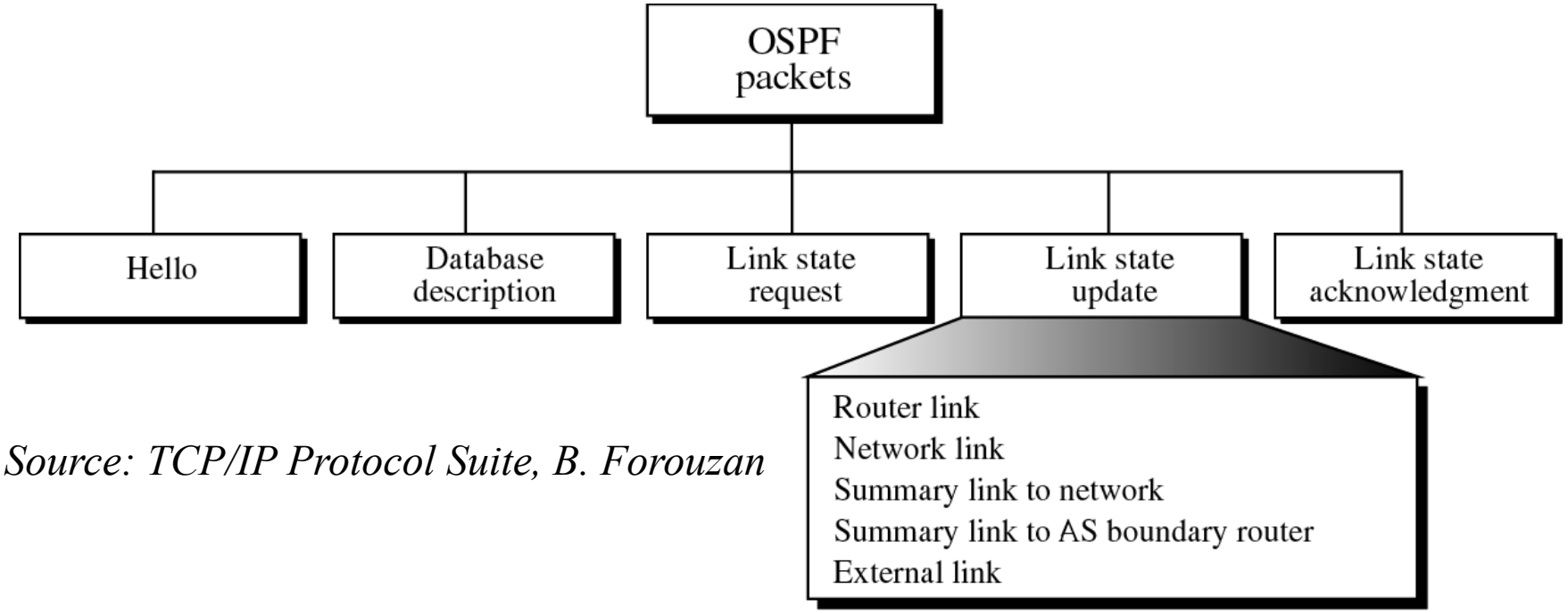
Real Network



Network as represented by OSPF

Source: TCP/IP Protocol Suite, B. Forouzan

OSPF: The Packets



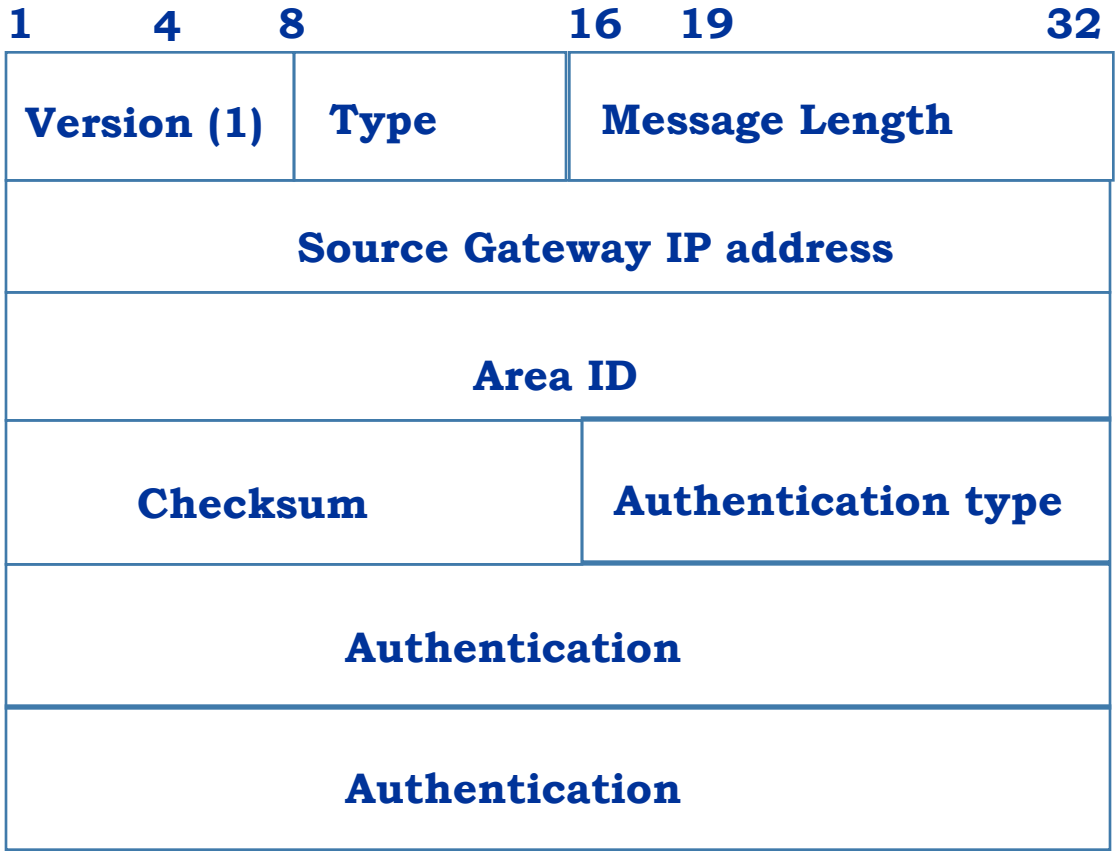
Source: TCP/IP Protocol Suite, B. Forouzan

- Routing Packets are acknowledged

OSPF: the packets

- *Hello*: manages the link state of neighbors
- *DB description*: exhanges the whle network DB (for ex. during initialization phase)
- *LS request*: asks information about a specific route
- *LS update*: Link state messages, both for internal topology and for external destinations
- *LS ACK*: ACKs for LS messages

OSFP: Common Header



OSFP: Open Shortest Path First

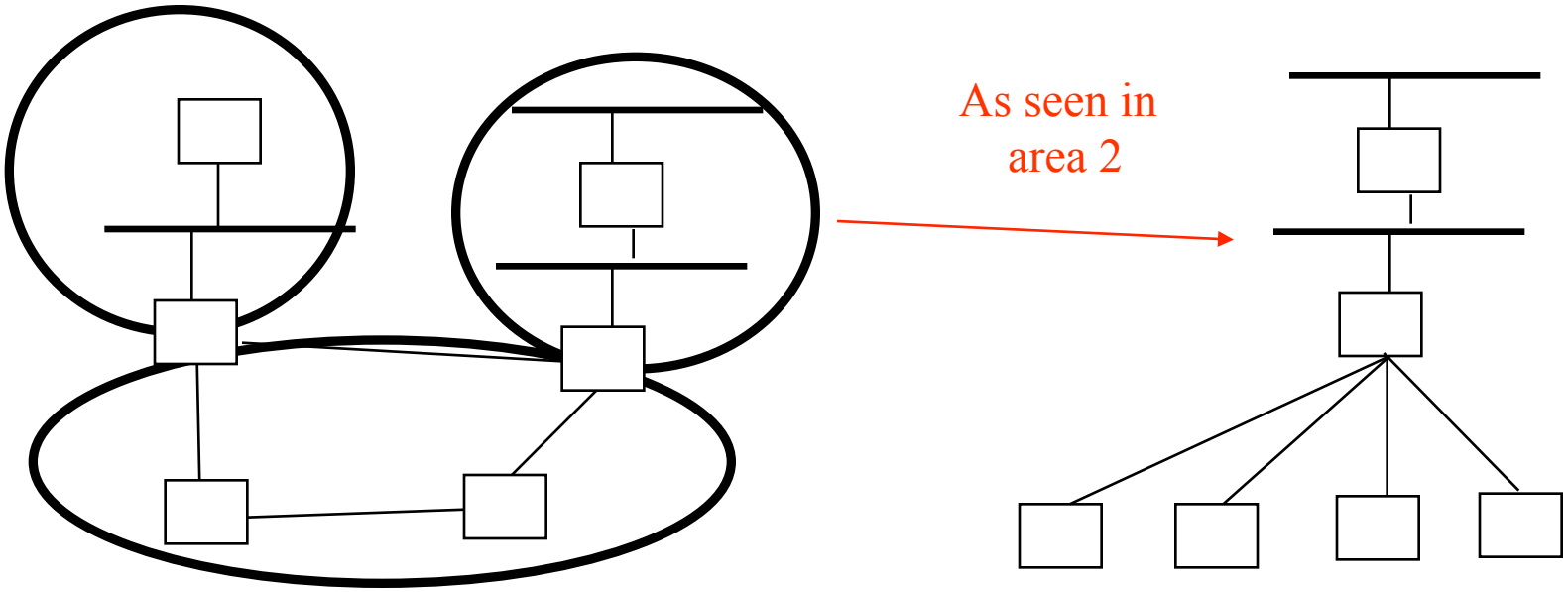
- *Type field:* type of OSPF packets
  - HELLO: neighboring nodes detection
  - DATABASE DESCRIPTION: link state broadcasting
  - LINK STATUS REQUEST
  - LINK STATUS UPDATE
  - LINK STATUS ACKNOWLEDGE: ack for the LSU packets
- *Source gateway IP address* IP address of the sender
- Area ID indicates the area

OSPF: Types of LSA

- Type 1: *router links advertisement*
  - Within the same area (classical LSP)
- Type 2: *network links advertisement*
  - Generated by a LAN pseudo-Node (DR)
- Type 3: *network summary link advertisement*
  - Generated by *area border routers* to summarize the info regarding an area
- Type 4: *boundary routers summary link advertisement*
  - Generated by the *area border routers*, indicates the presence of a AS *boundary router* in the area and the associated cost
- Type 5: *AS external link advertisement*
  - Generated by AS *boundary routers* and propagated to all the routers of all the areas with info on external destinations and the associated costs

OSPF

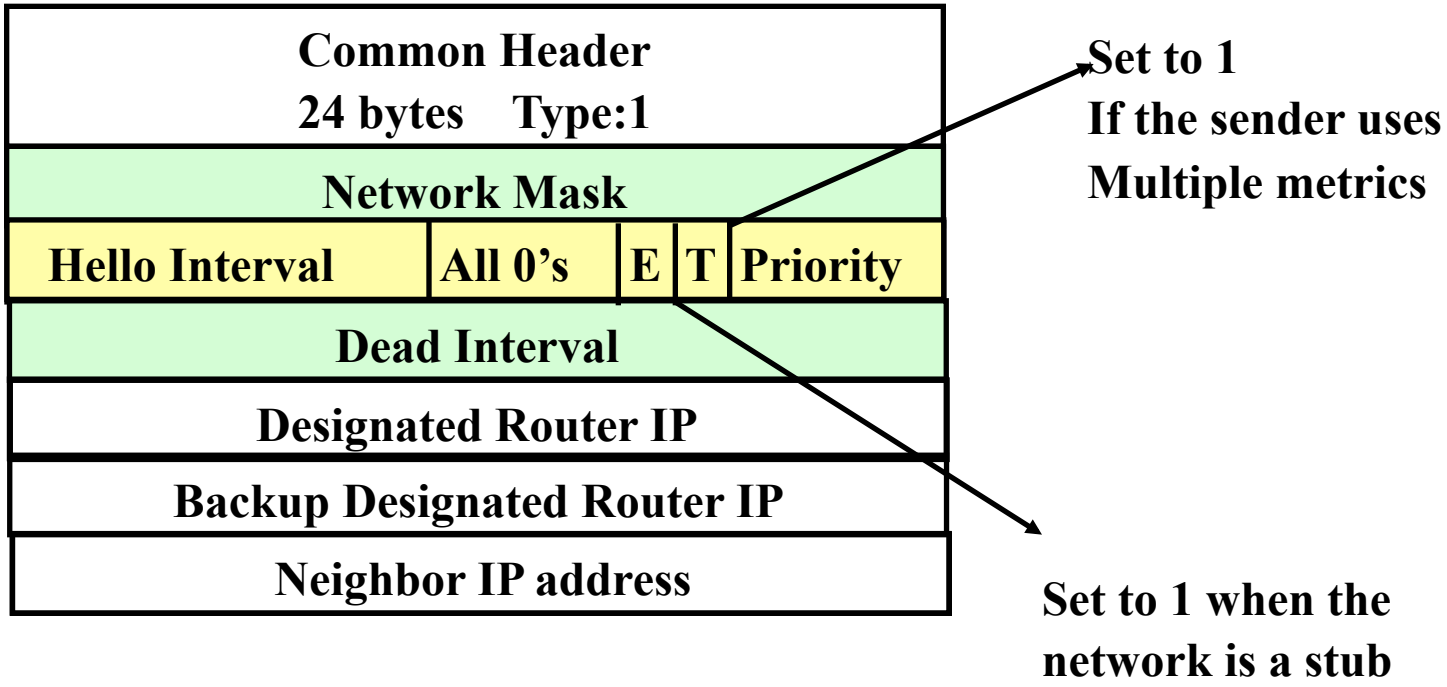
- The *area border router* propagates in every area routing info regarding all the other areas they are connected to
  - distance vector contamination



OSFP: Open Shortest Path First

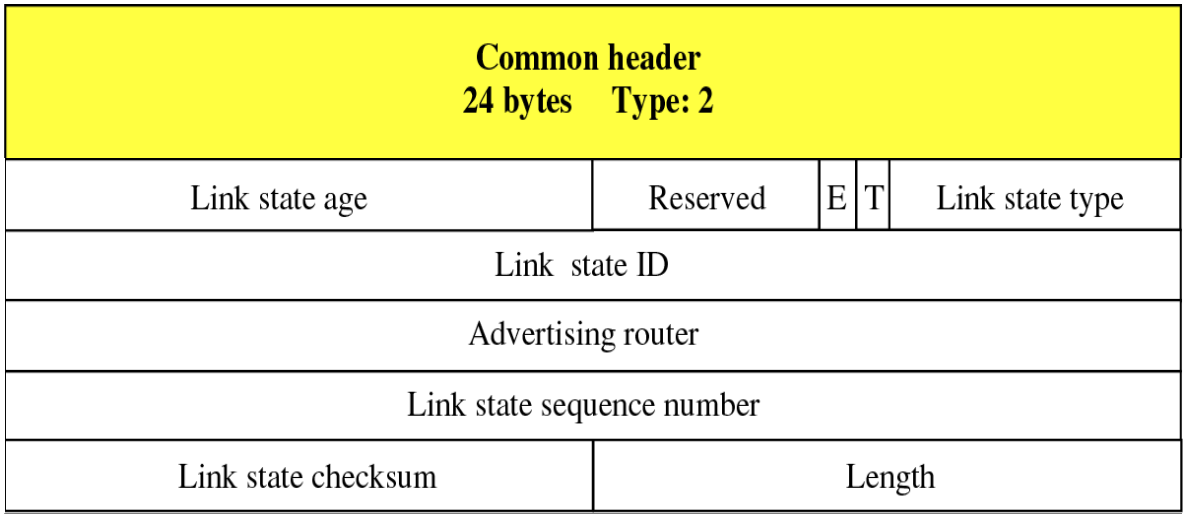
- OSPF sends periodically HELLO messages to test if neighbors are reachable
- *database description* messages are used to initialize the topology data base
- Data on link metrics are broadcast through the *link status update* messages

Hello Packets



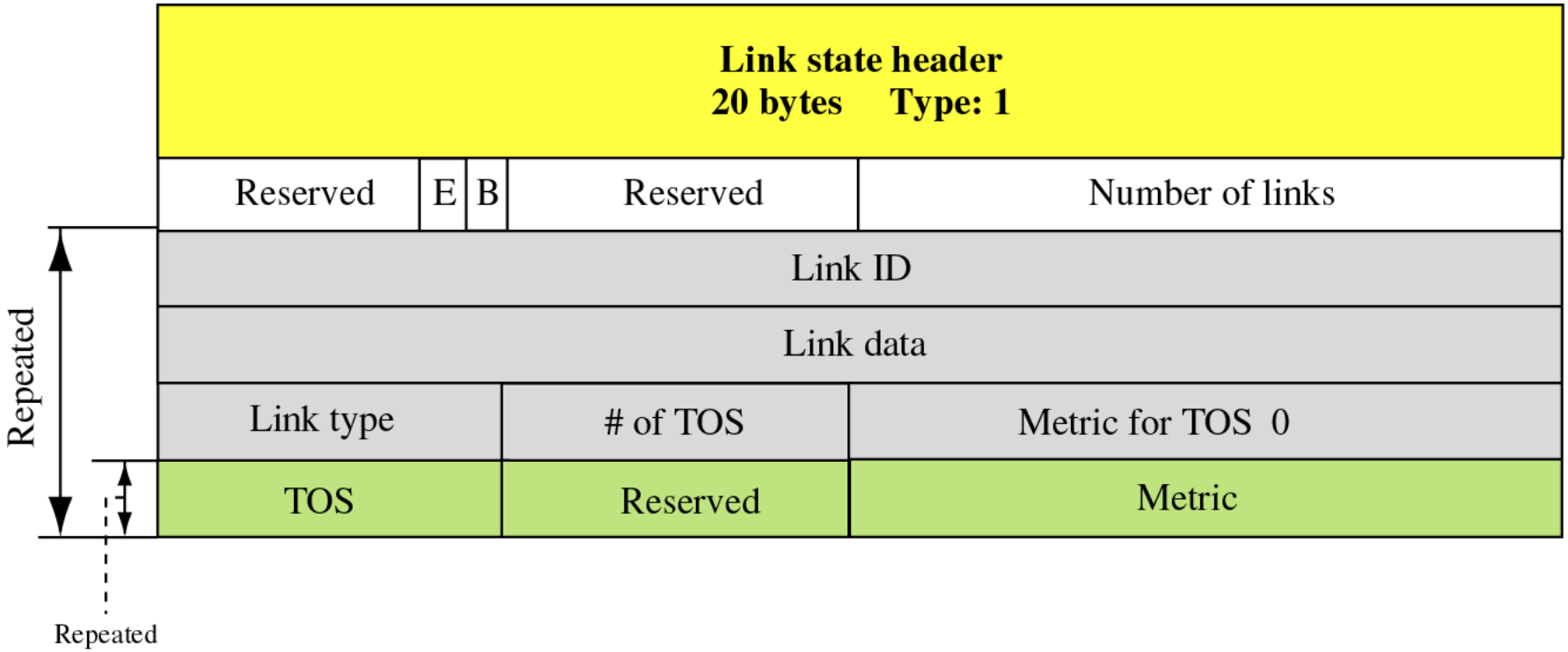
- Used for
  - Neighbors discovery
  - Select a *designated router*

LSU Packets



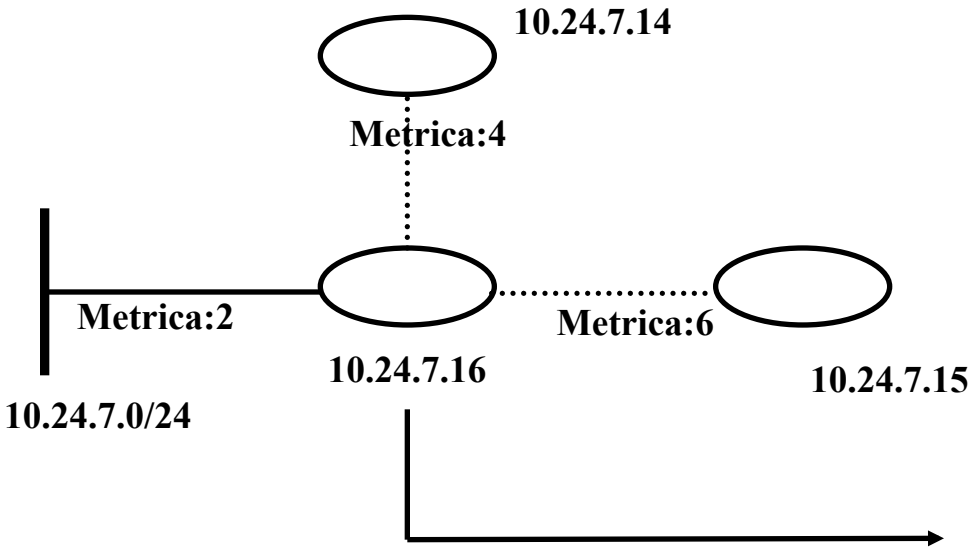
- LSU packets have a common header + *Link State common* header + payload

Router Link LSA



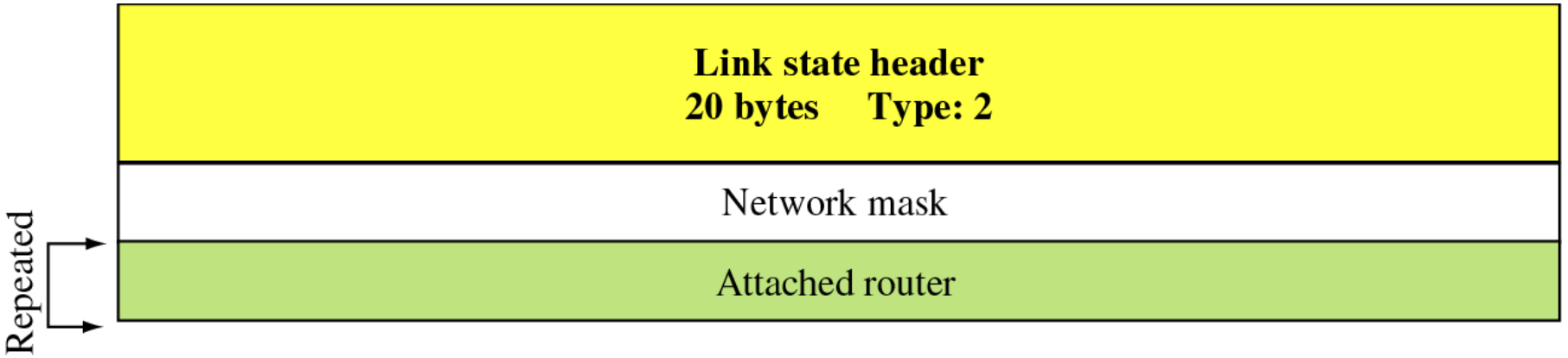
- *Link ID (link address)*
- *Link data/Link Type: depends on the link type (point to point, stub, network)*

Router Link LSA: Example



OSPF Header Type: 4		
LSA Header Type:1		
10.24.7.14		
1		
	1	
		4
10.24.7.15		
2		
	1	
		6
10.24.7.0		
255.255.255.0		
	3	
		2

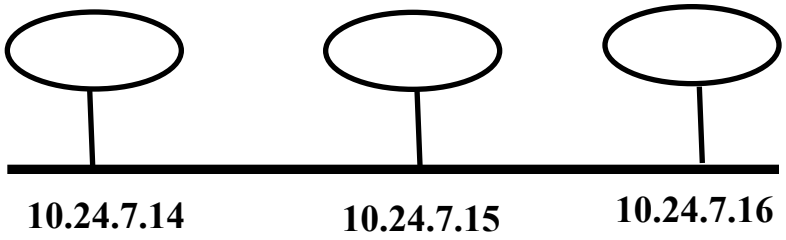
Network Link LSA



- *Network Mask*
- *Attached Router*: all the routers connected to the network

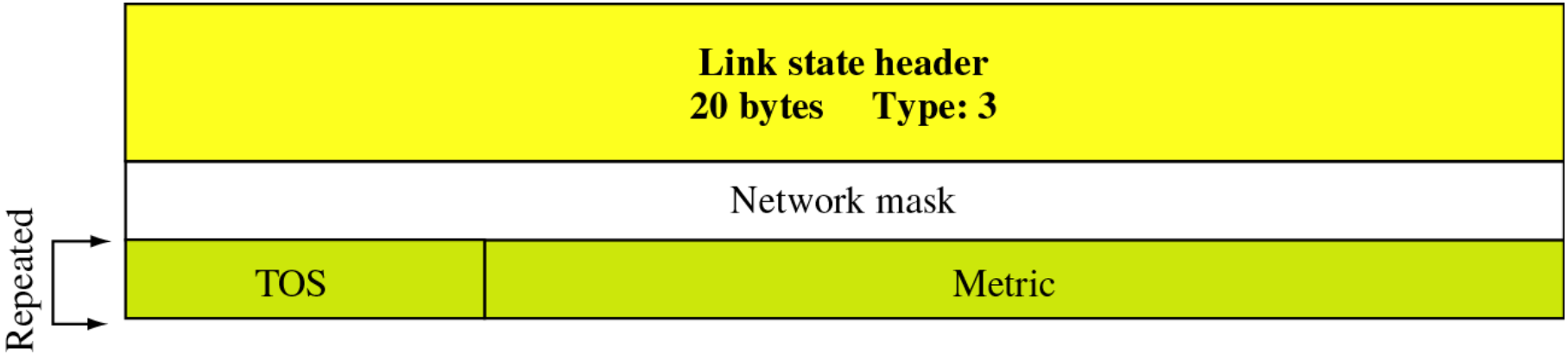
Network Link LSA: example

OSPF Header Type:4
LSA Header Type:2
255.255.255.0
10.24.7.14
10.24.7.15
10.24.7.16



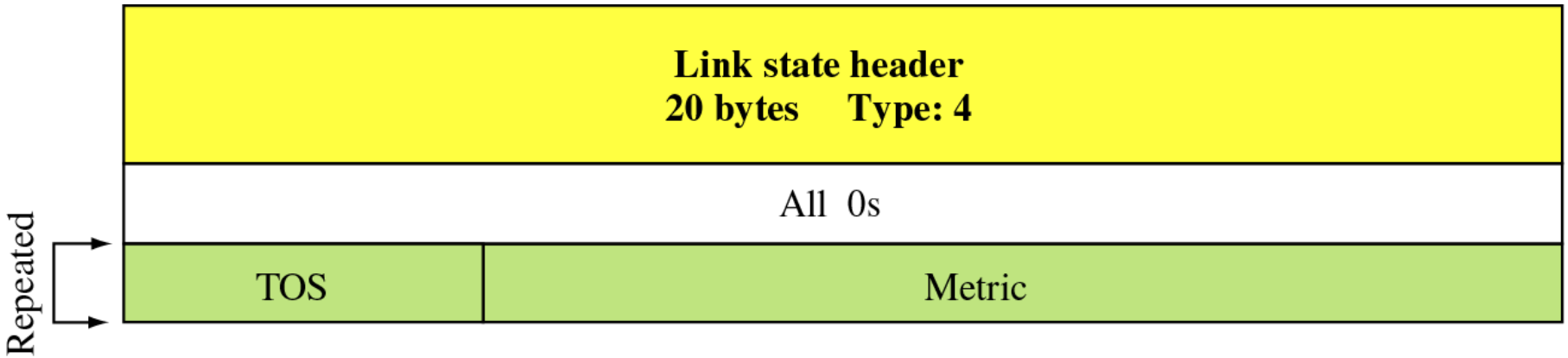
- Only the *Designated Router* (one of the three routers) signals the presence of all the other routers
- Network address is not advertised (can be obtained form the header info)

Summary Link to Network LSA



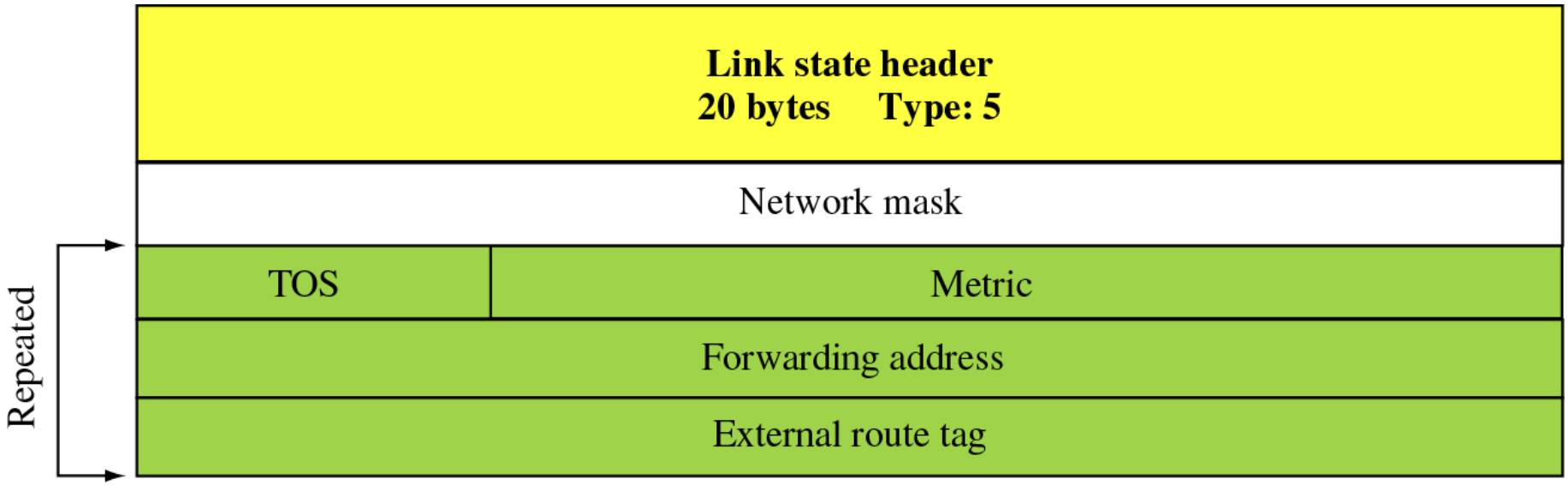
- Used to advertise networks outside an area of a AS
- 1 message for 1 network (multiple messages needed to address more networks)

Summary Link to AS Boundary Router LSA



- Defines the network a border router is connected to

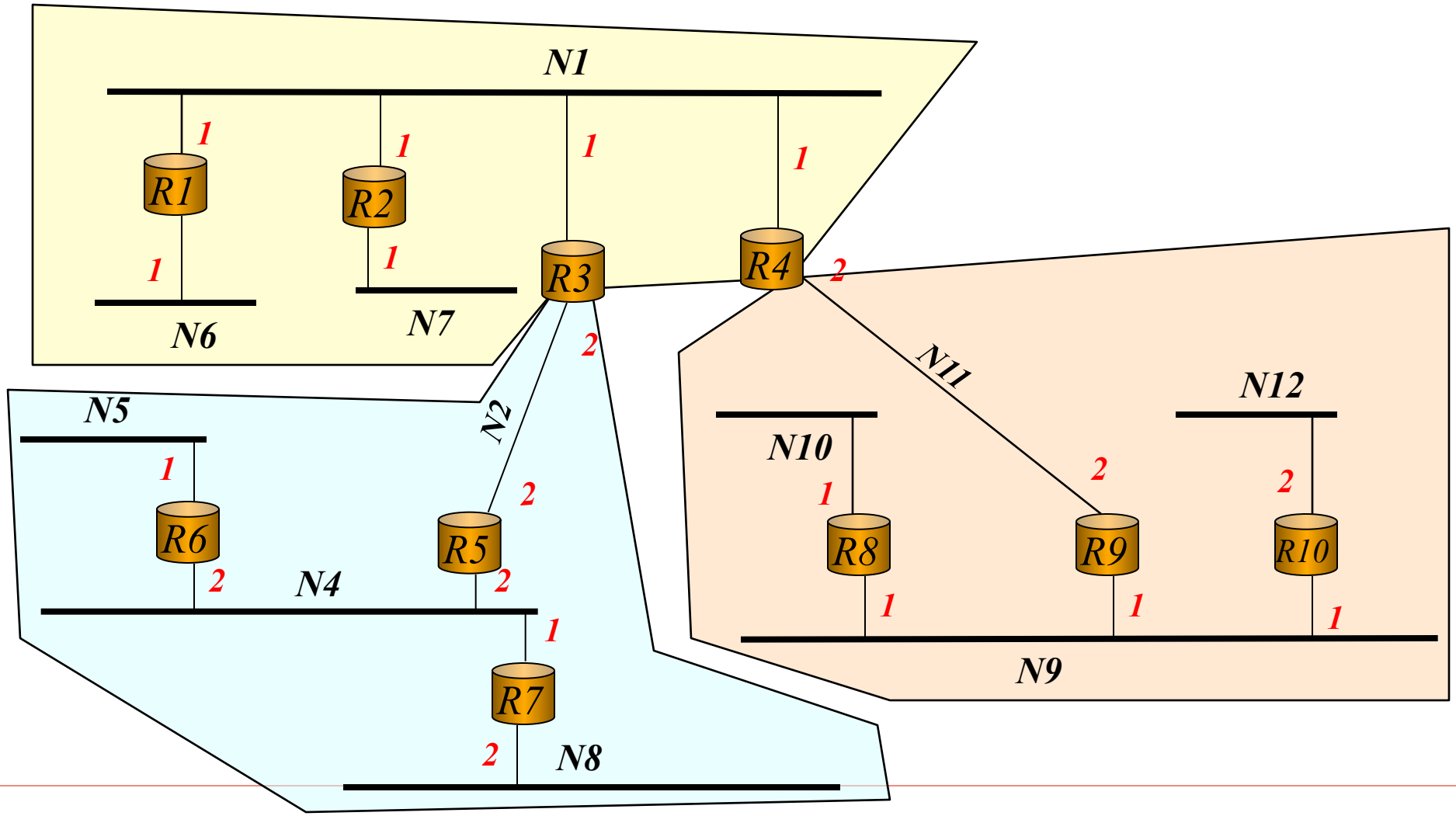
External Link LSA



- Defines external networks
- *Forwarding Address*: to route packets meant for external destinations

Template Activity

- Given the network below with routers, networks and costs associated to the interfaces

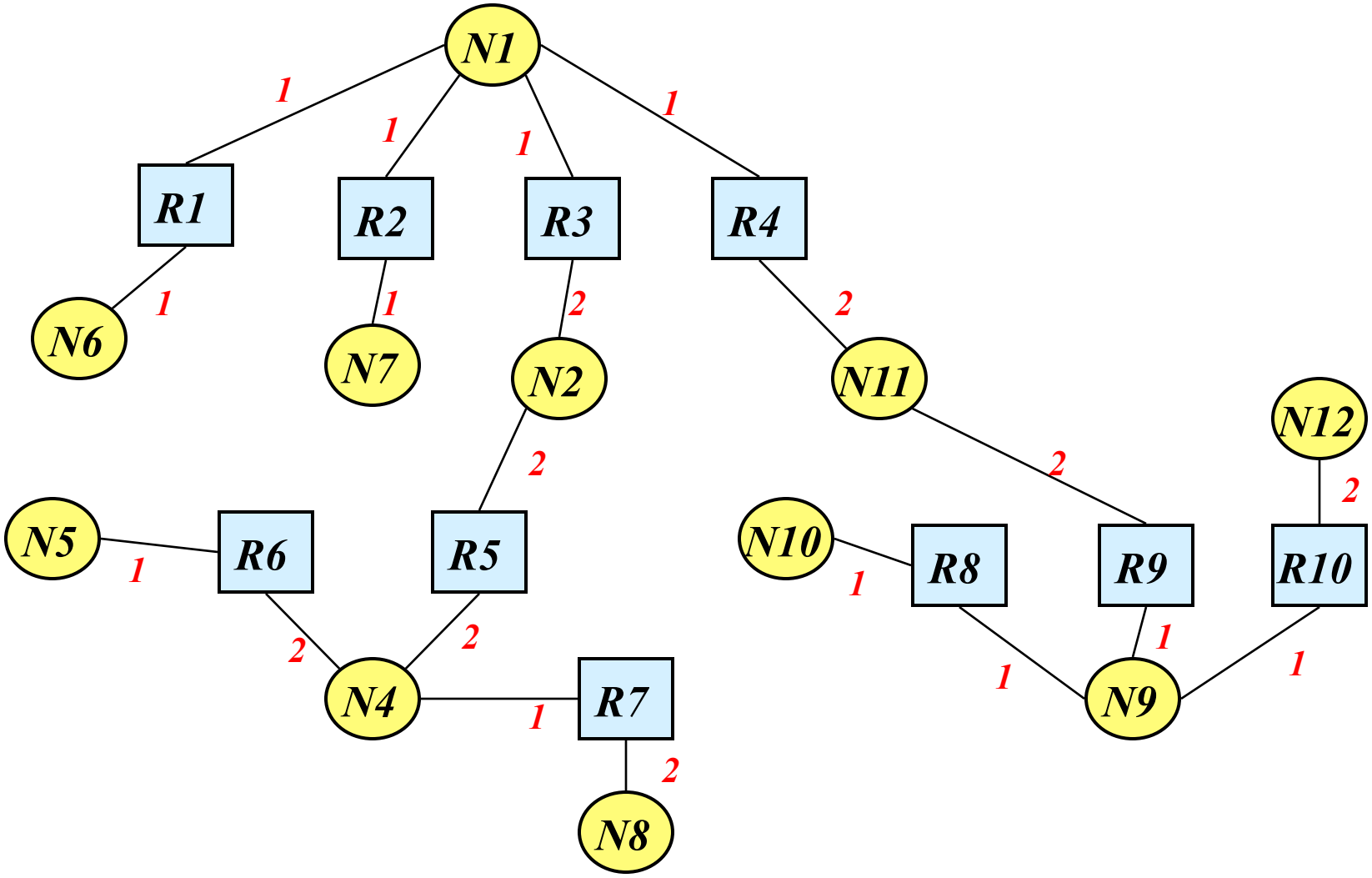


Template Activity

- Assuming the AS runs OSPF
  a) Sketch the graph of the network as represented by OSPF assuming one single area
  b) Assuming the AS divided in areas as in the figure (area 0, area 1 and area 2) sketch the graphs of the AS as seen by routers R1, R7 and R10

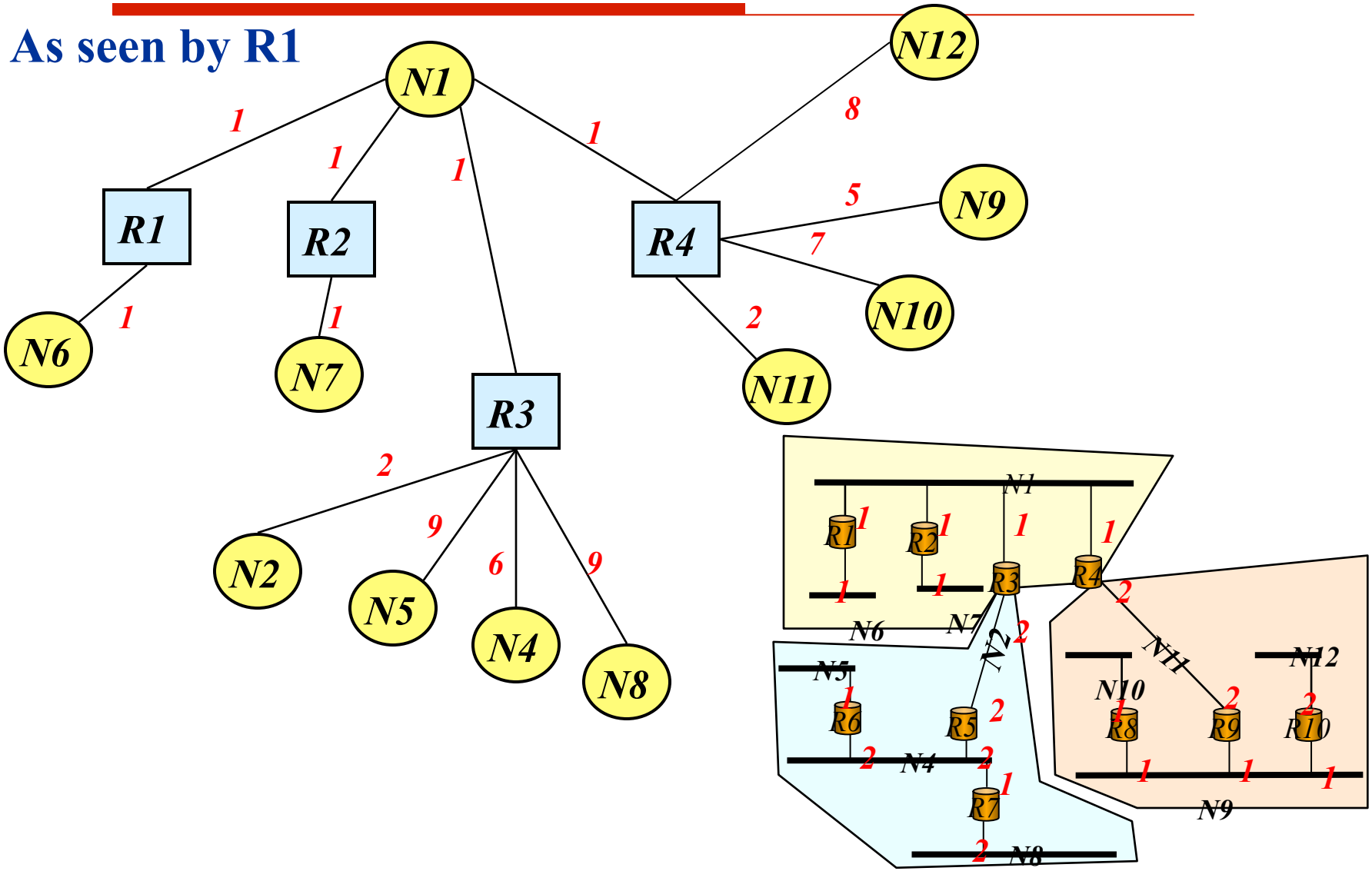
Solution

a)



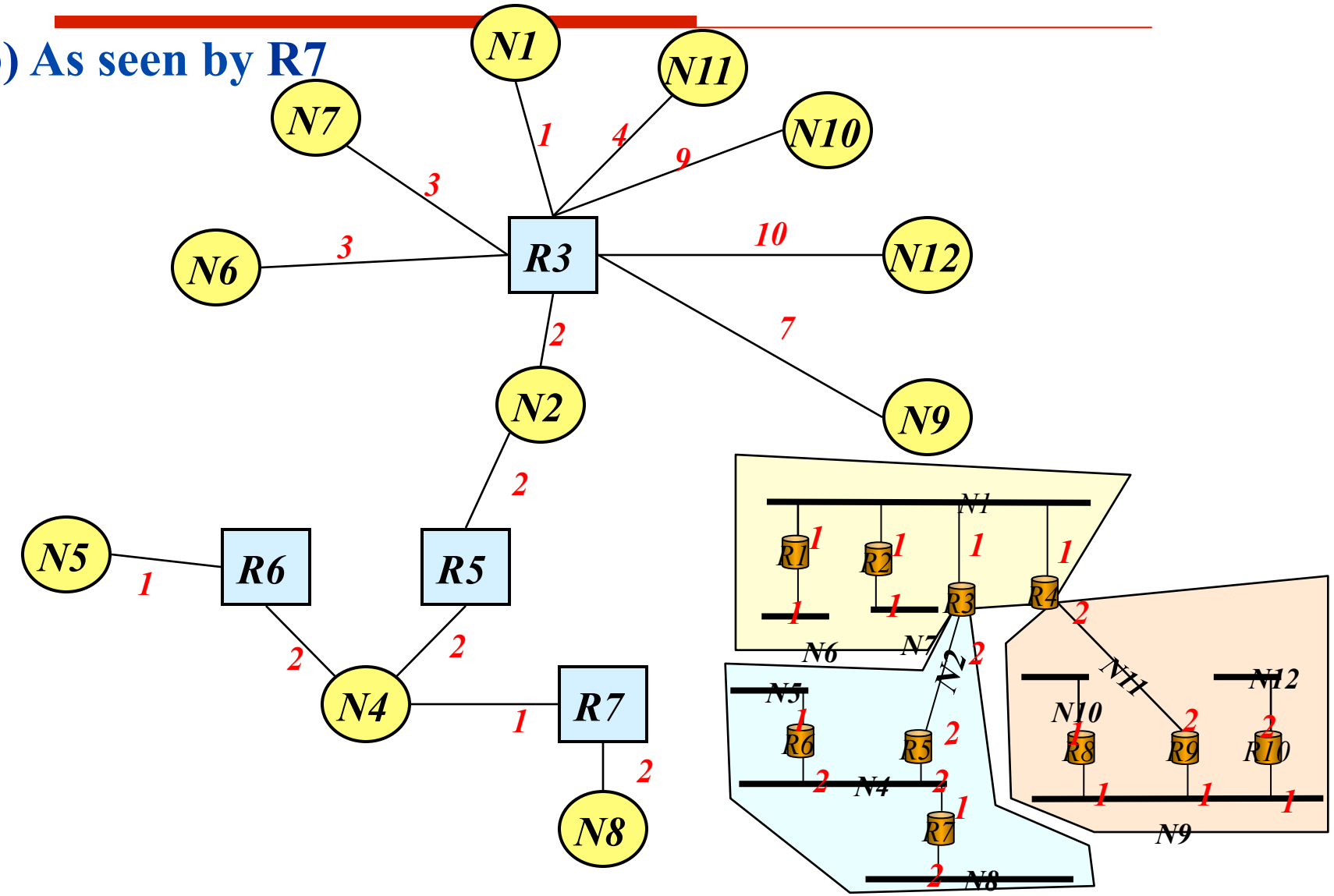
Solution

b) As seen by R1



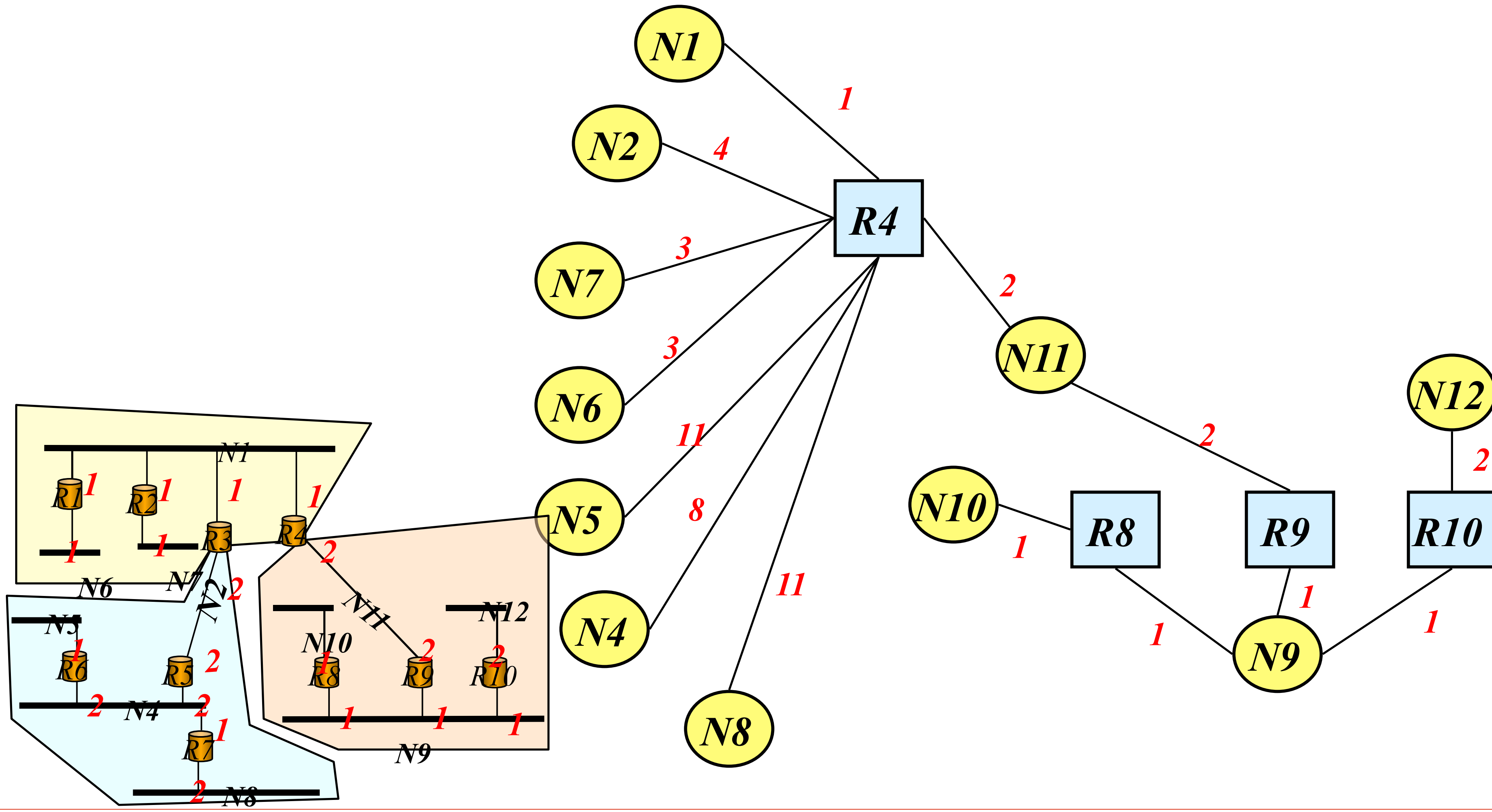
Solution

b) As seen by R7



Solution

b) As seen by R10



BGP

- Most used EGP (standard *de facto*)
- It is the "glue" of the Internet
  - BGP allow to AS to announce their presence in Internet, and to reached
- Inter AS routing problem is different from intra AS one
  - Route decisions criteria are not based on metrics
  - Backbone managers choose the routes according to their own chosen policy
  - Routing choice may need to exploit full knowledge of the path to destination
- Thus:
  - DV does not fit since it has no knowledge of all the path
  - LS does not fit since it will need to build up a database of the entire internet

BGP: Path vector

- BGP is similar to *distance vector*, but:
  - the PVs do not report a "distance to destination", but the entire path to destination

Network	Next Router	Path
N01	R01	AS2,AS5,AS7,AS12
N02	R07	AS4,AS13,AS6,AS9
N03	R09	AS11,AS12,AS8,AS6
…	…	…

BGP: Path vector

- Messages exchanged between 2 routers in a *path vector* do not contain only a path, but a sequence of *attributes*
- Attributes may be mandatory (that must be understood by any BGP implementation) and optional
- Mandatory attributes:
  - ORIGIN: IGP protocol origin of the info (e.g. OSPF, RIP, IGRP)
  - AS_PATH: sequence of traversed AS
  - NEXT_HOP: next router

BGP: messages exchange

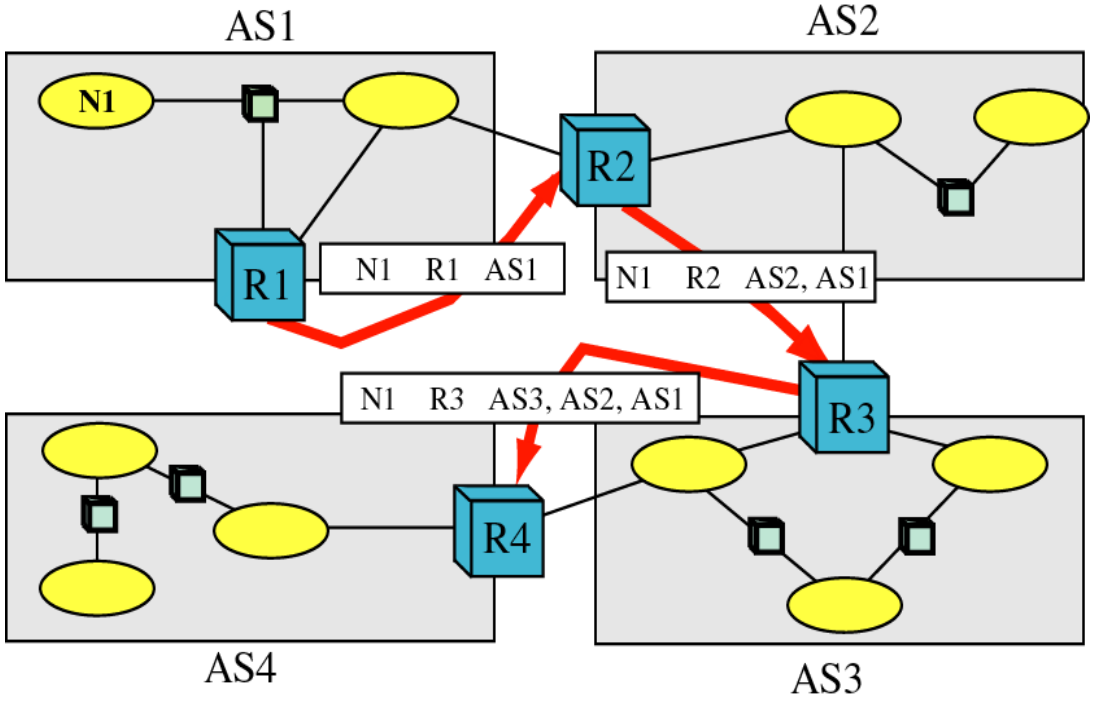
- Each BGP router sends its *path vector* to neighboring nodes (*peers*)
- BGP messages use TCP
- TCP connections are opened by sending routers
- BGP uses port number 179

BGP: Types of messages

- Types of messages:
  - OPEN: opens the TCP connection and manages the mutual authentication of the two routers
  - UPDATE: announces a new route (or erases an old one)
  - KEEPALIVE: mintains active a connection in the absence of UPDATE (used also as an ACK for OPEN messages)
  - NOTIFICATION: notifies errors in previous messages (used also to close a connection)

BGP: Path Vector

- BGP allows the distribution of paths to specific destinations
- …but leaves the routing choice to the network administration (*policy based routing*)

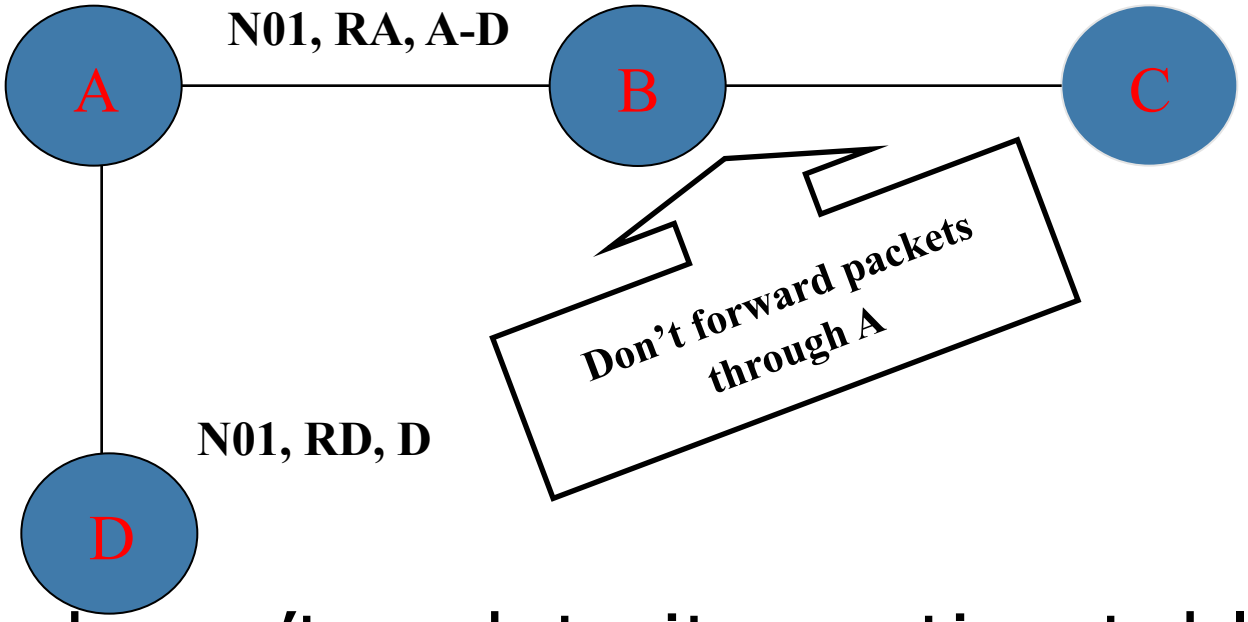


Policy based routing

- A BGP router receiving a *path vector* from a peer may whether or not to:
  - Add to the routing table the destination specified in the PV
  - Forward the PV to its neighbors
- Based on the local routing policy
- To each AS it is assigned an Autonomous System Number (ASN), with global meaning, from IANA (like for IP addresses)

Policy based routing: example 1

Net	Next Router	Path
N01	RD	D

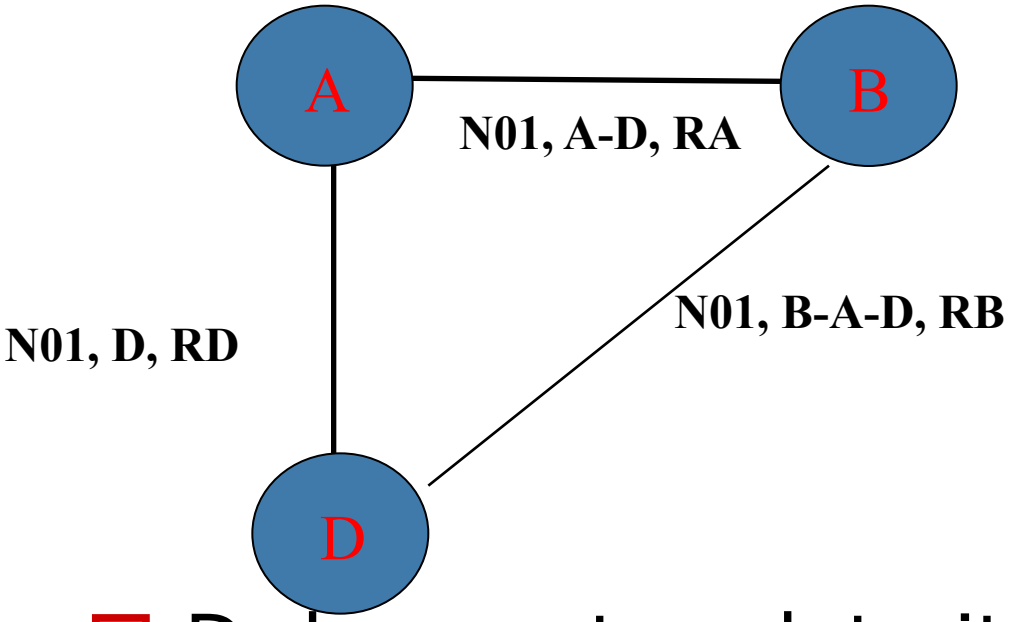


- B doesn't update its routing table and doesn't forward the PV since this goes against the local routing policy

Policy based routing: example 2

Net	Next Router	Path
N01	RD	D

Net	Next Router	Path
N01	RA	A-D

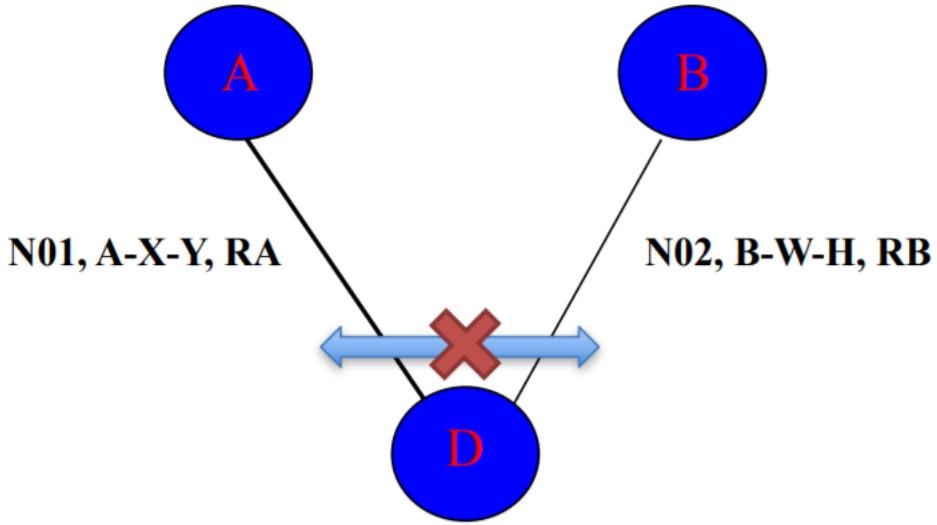


- D does not update its routing table and does not forward the PV since its own AS is specified in the path

Policy based routing: example 3

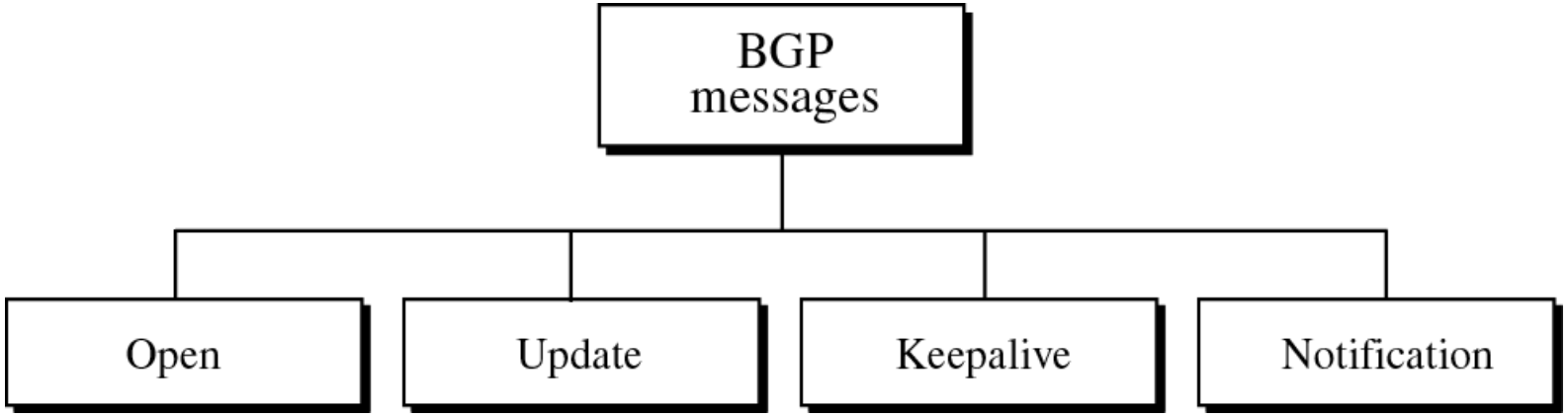
Rete	Router successivo	Percorso
N01	RA	A-X-Y

Rete	Router successivo	Percorso
N02	RB	B-W-H

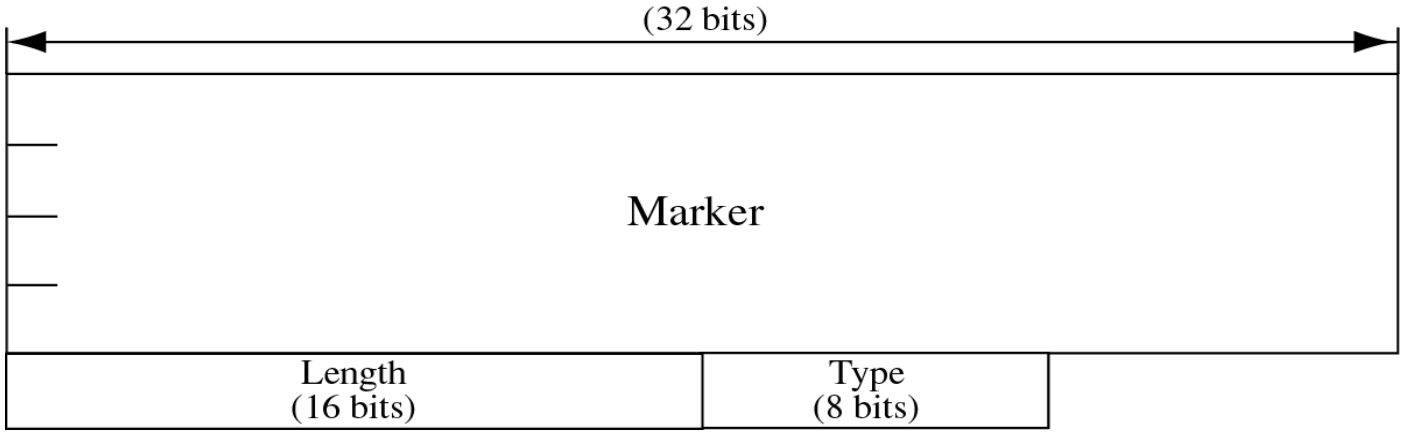


- D does not want to forward the traffic from A to B and viceversa, hence it does not forward the PV received from A to B and viceversa.

BGP Messages

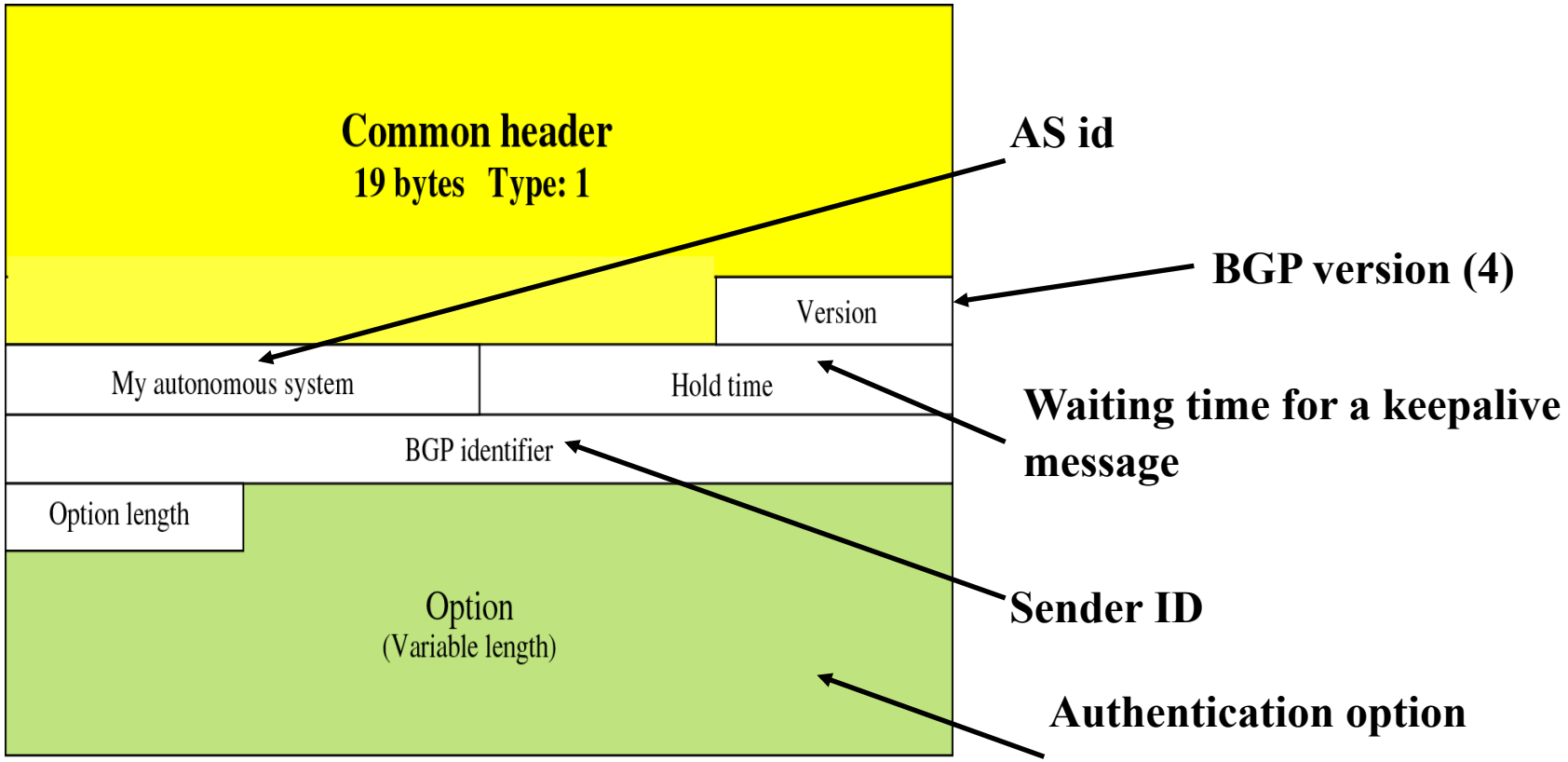


- Common header

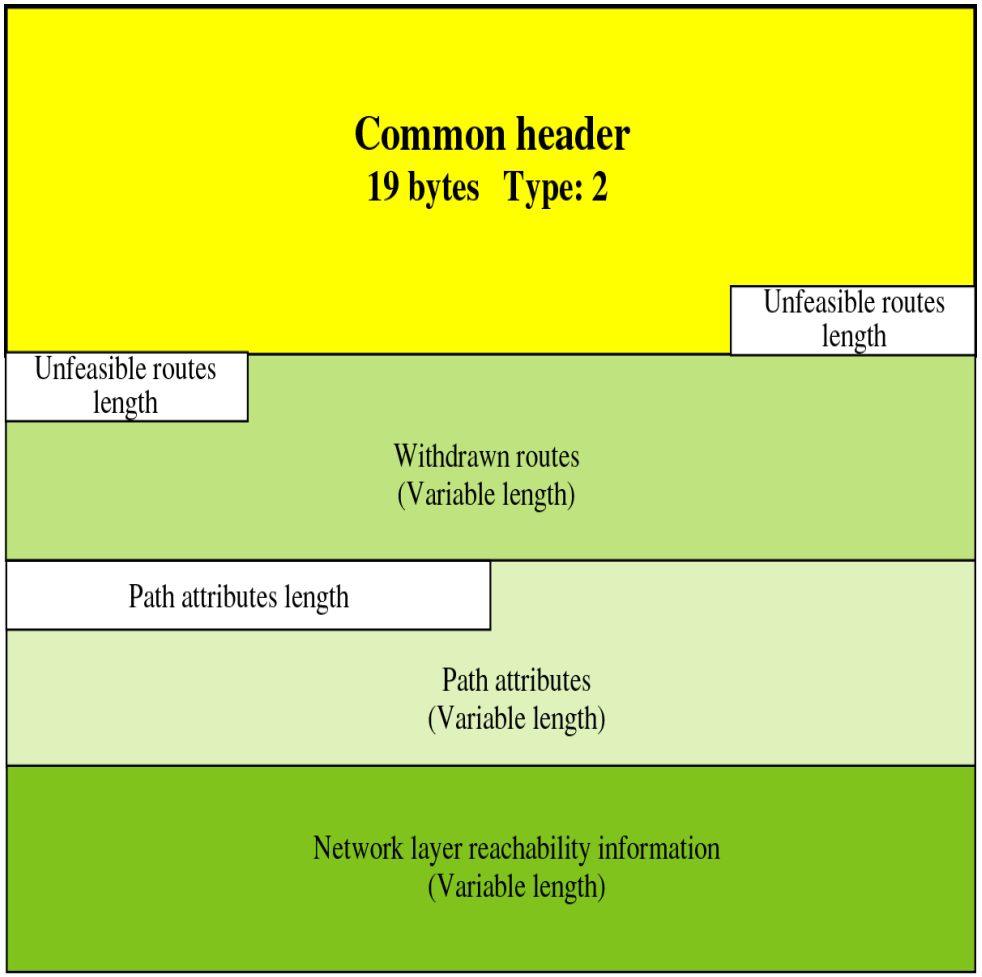


Open Messages

- *Peering set up messages*
- Routers answer with keepalive messages (common header only)

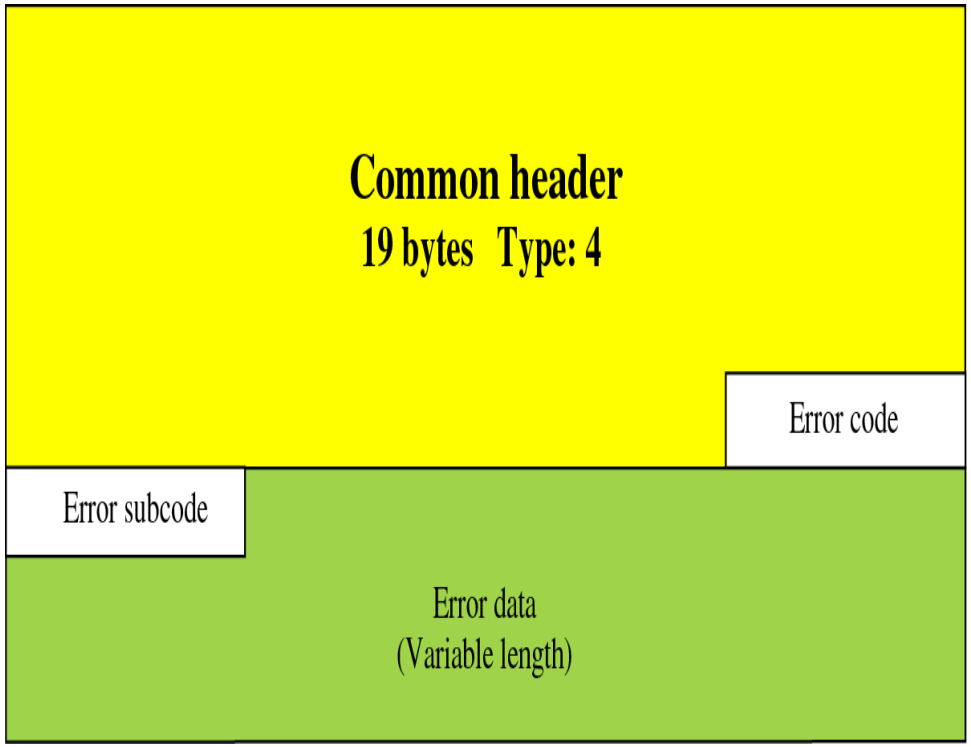


Update Messages



- Contain the *path vector*
- Used to advertise path or to cancel previously advertised paths

Notification Messages



- To notify an error or to close a connection