Institute of Telecommunications
Warsaw University of Technology
2018

internet technologies and standards

- Piotr Gajowniczek
- Andrzej Bąk

IP routing

principles & protocols

Routing table example (Cisco router)

```
Gateway of last resort is 192.168.1.1 to network 0.0.0.0

C      192.168.1.0/24 is directly connected, Ethernet0
       150.50.0.0/24 is subnetted, 1 subnets
C      150.50.200.0 is directly connected, Loopback1
R      10.0.0.0 [120/1] via 192.168.123.1, 00:00:00, Serial0
                [120/1] via 192.168.111.2, 00:00:00, Serial1
O      192.168.111.0/24 [110/58] via 192.168.131.1, 00:00:00, Serial3
R      192.168.111.0/24 [120/1] via 192.168.123.1, 00:00:00, Serial0
R      192.168.111.0/24 [120/5] via 192.168.5.2, 00:00:00, Serial1
S      192.168.0.0/16 [1/0] via 10.1.1.1
```

- *major components:*
  - route tag– indicates routing protocol that was used to advertise a given prefix
  - dest. prefix and mask– network that has been advertised to the router
  - administrative distance – preference of one routing protocol over another
  - metric – a value used by a routing protocol to calculate the best route to a destination (hop count or link cost)
  - next hop – IP address of the neighbor that advertised the destination prefix to the router.
  - age – how long this entry has been in the routing table.
  - interface – outgoing router interface

Routing in the Internet

- *The Internet is organized as a set of independent Autonomous Systems (AS)*
  - the AS is a part of the network (collection of routers, networks) under single technical administration
  - the AS appears to the outside world as having a coherent routing plan and presents unique view on what destinations are reachable through it
- *The AS can use many different routing protocols*
  - the routing protocols used inside the AS are called the Interior Routing Protocols (IGP)
- *A separate protocol is used to transfer information between ASs*
  - the routing protocol used between the ASs is called the Exterior Routing Protocol (EGP)

Types of dynamic routing protocols

- *Distance vector protocols*
  - Routers periodically exchange routes (with metrics) with neighbours
    - Routers do not have the knowledge about network topology, they only know the cost of links to neighbors and neighbours routes
  - Next routers use Bellman-Ford algorithm to actualize routing tables
    - After receiving update from the neighbour, router adds cost of the link to that neighbour to every received route
    - Finally router selects the paths with smallest cost (the leased cost paths) and adds them to routing table
- *Link state protocols*
  - Routers periodically exchange network topology information (about links, metrics, prefixes)
    - Routers maintain topology database
  - After receiving update from the neighbour, the router adds obtained links (metrics and prefixes) to the topology database
  - Finally the router uses Dijkstra algorithms to calculate least-cost paths and actualize its routing table

IP routing

OSPF protocol

OSPF - Open Shortest Path First

- *OSPF is an IGP protocol based on the Dijkstra algorithm*
- *OSPF uses IP as its transport protocol*
  - ❑ OSPF is protocol number 89
- *OSPF main characteristics*
  - ❑ Hierarchical routing
    - two-level hierarchy: local areas, one backbone area
  - ❑ Support for CIDR
    - subnet and mask propagation
  - ❑ No network topological limits
  - ❑ Single dimensionless metrics - link cost (weight)
    - links costs are usually set in proportion to the bandwidth
  - ❑ Load balancing over equal cost paths (up to 6 paths)
  - ❑ Security: all OSPF messages authenticated (to prevent malicious intrusion)
  - ❑ Integrated uni- and multicast support:
    - Multicast OSPF (MOSPF) uses same topology database as OSPF

OSPF terminology

- *Interface – the connection between router and the network (represented as link in the topology database, routers and networks are represented as vertex)*
- *LSA (Link State Advertisement) – data describing the status of the router's links or the status of the network*
- *Link State database (LSDB) – database that stores the topology information*
- *Cost – weight assigned to the interface (link)*
- *Area – a collection of routers that have the same area identification*
- *Neighbours – two routers on the same network*
- *Adjacency - relationship between neighbour routers established for the purpose of exchanging routing information*
- *Designated router (DR) - the router that generates LSA for given broadcast or NBMA network*

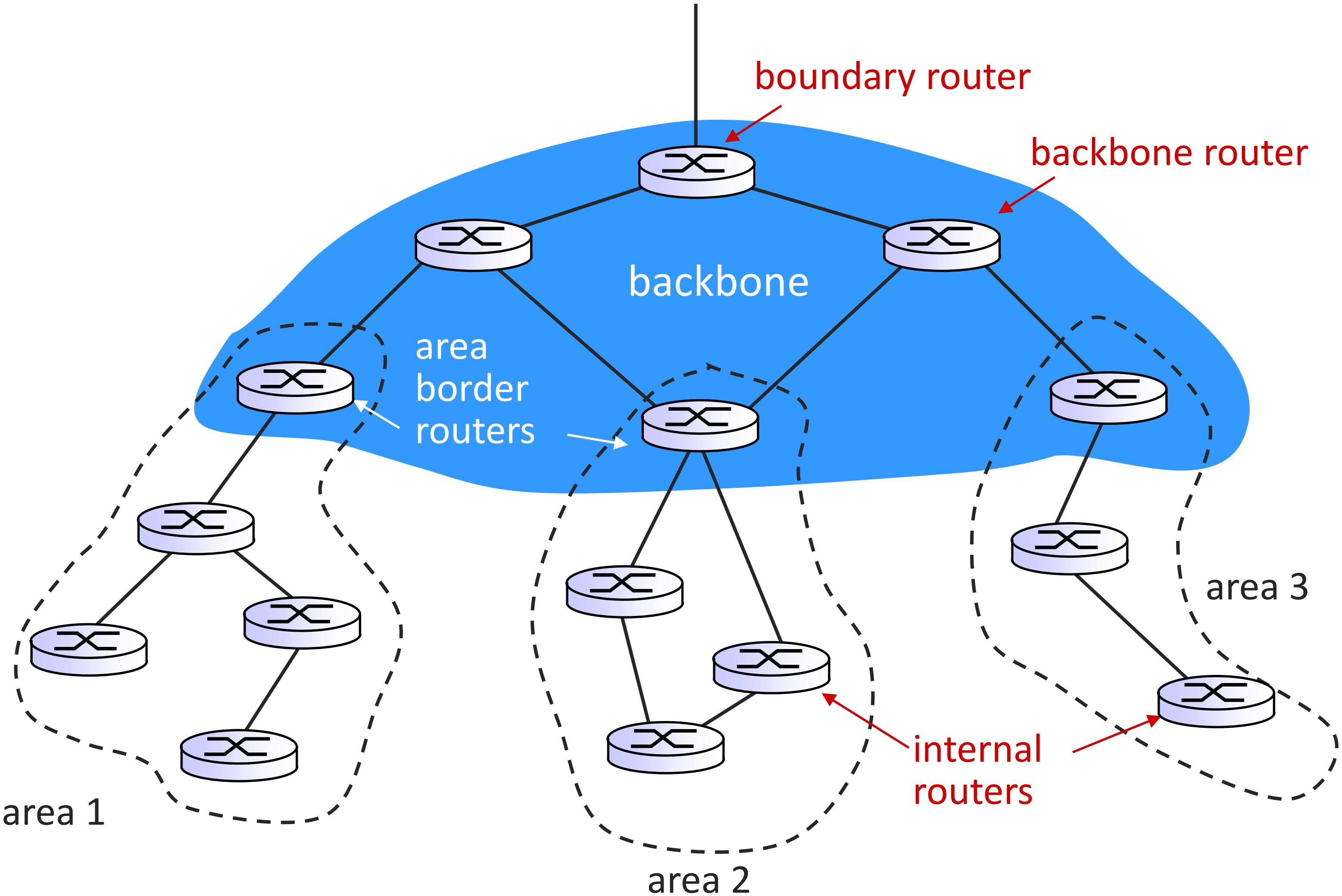
OSPF single area limitations

- In large networks the following issues may arise
  - *Frequent routing table recalculation*
    - in large network link state changes are inevitable
    - the larger the network the more frequently a link in the network will go out of service causing recalculation
    - high CPU utilisation
  - *Large topology database*
    - more LSA data to store - larger LSDB
  - *Large routing tables*
    - each router has at least one entry in the routing table per each destination
    - large memory requirements
  - *Significant routing traffic in the network*
    - link bandwidth is consumed by routing traffic
    - less bandwidth available to data packets

Hierarchical OSPF

- In OSPF large network can be split into multiple areas
- OSPF areas allow to implement hierarchical routing and the following benefits can be achieved
  - *individual LSA updates are not propagated out of the area so the change in one area does not incur routing table recalculation outside of this area*
    - reduced routing traffic in the network – more bandwidth available to data packets
    - decreased router CPU utilization
    - smaller LSDB
  - *smaller routing tables*
    - the detailed routing information is kept within an area
    - the route summarisation can be done for the routes that are sent out of the area, with proper addressing even only one route can be propagated out of the area
  - *better convergence time*

Hierarchical OSPF



Link State Advertisement (LSA) types

- *LSA describes the routers local topology information*
- *Exchanged by OSPF routers in order to build network topology database (LSDB)*
- *Types of the LSA*
  - ❑ Router LSA (1) – describes the router's links to other routers or networks, builds area topology
  - ❑ Network LSA (2) – lists the routers interconnected by the L2 network, builds area topology
  - ❑ Summary LSA (3) – describes prefixes external to the given area, hides the details about external areas topology, provides for route summarization
  - ❑ ASBR Summary LSA (4) – provides information about the presence of ASBR router
  - ❑ External LSA (5) – describes prefixes external to the OSPF domain

Router LSA (type 1)

- *Generated by each router in the OSPF domain*
- *Flooded throughout a single area (the area of the orginator)*
- *Identified by*
  - ❑ Link State ID = Router ID
- *Describes the router's links*
  - ❑ Each link is identified by Link ID
  - ❑ Types of links
    - Point-to-point link : Link ID = Neighbor Router ID
    - Link to transit network: Link ID = network address of the DR router
    - Link to stub network: LinkID = network address of the interface
    - Virtual link: Link ID = Neighbor Router ID
- *In LSDB*
  - ❑ Router LSA represents vertex
  - ❑ Each link is interpreted as pointer to another Router LSA or Network LSA

Router LSA example

R4# **show ip ospf database router**

OSPF Router with ID (4.4.4.4) (Process ID 1)

Router Link States (Area 1)

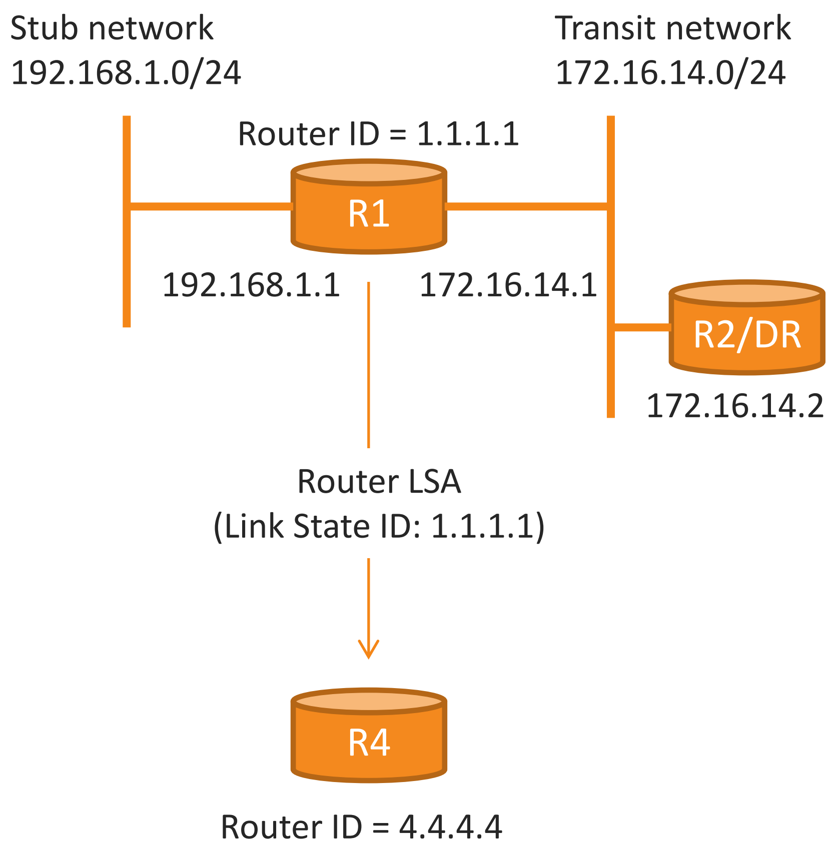
LS age: 321
Options: (No TOS-capability, DC)
LS Type: Router Links
Link State ID: 1.1.1.1
Advertising Router: 1.1.1.1
LS Seq Number: 8000000B
Checksum: 0x966C
Length: 48
Area Border Router
Number of Links: 2

Ip network address

Link connected to: a Stub Network
(Link ID) Network/subnet number: 192.168.1.0
(Link Data) Network Mask: 255.255.255.0
Number of TOS metrics: 0
TOS 0 Metrics: 1

IP address of DR router interface

Link connected to: a Transit Network
(Link ID) Designated Router address: 172.16.14.2
(Link Data) Router Interface address: 172.16.14.1
Number of TOS metrics: 0
TOS 0 Metrics: 10



Router LSA example

R4# **show ip ospf database router**

OSPF Router with ID (4.4.4.4) (Process ID 1)

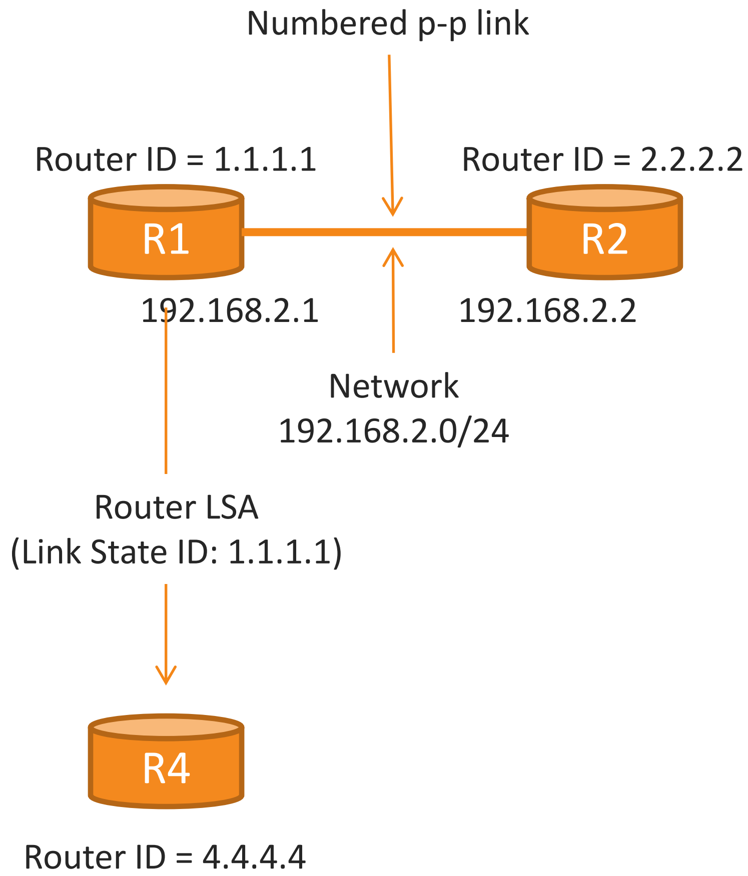
Router Link States (Area 1)

LS age: 321
Options: (No TOS-capability, DC)
LS Type: Router Links
Link State ID: 1.1.1.1
Advertising Router: 1.1.1.1
LS Seq Number: 8000000B
Checksum: 0x966C
Length: 48
Area Border Router
Number of Links: 2

Link connected to: another Router (point-to-point)
(Link ID) Neighboring Router ID: 2.2.2.2
(Link Data) Router Interface address: 192.168.2.1
Number of TOS metrics: 0
TOS 0 Metrics: 64

Link connected to: a Stub Network
(Link ID) Network/subnet number: 192.168.2.0
(Link Data) Network Mask: 255.255.255.0
Number of TOS metrics: 0
TOS 0 Metrics: 64

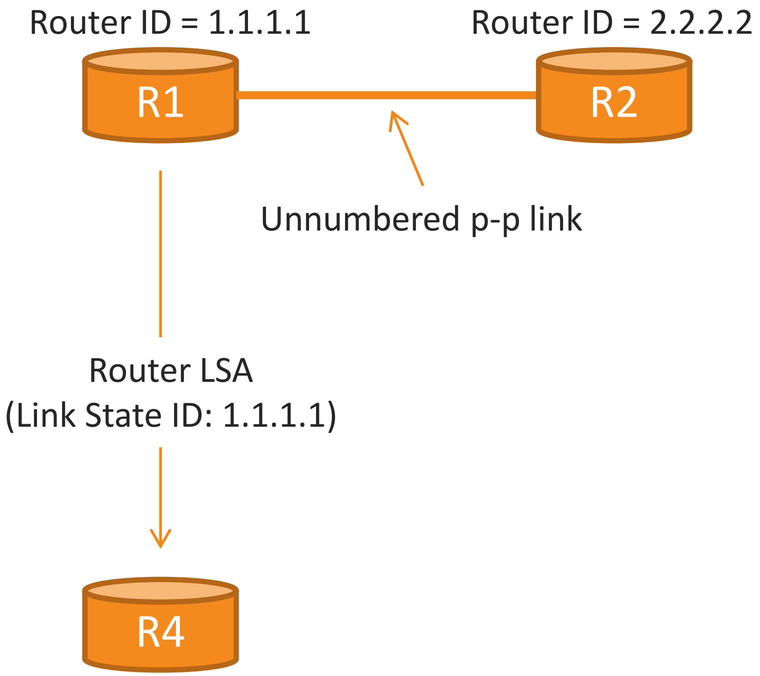
IP network address on P-P link



Router LSA example

```
R4# show ip ospf database router

        OSPF Router with ID (4.4.4.4) (Process ID 1)

                Router Link States (Area 1)

LS age: 321
Options: (No TOS-capability, DC)
LS Type: Router Links
Link State ID: 1.1.1.1
Advertising Router: 1.1.1.1
LS Seq Number: 8000000B
Checksum: 0x966C
Length: 36
Area Border Router
Number of Links: 1

  Link connected to: another Router (point-to-point)
   (Link ID) Neighboring Router ID: 2.2.2.2
   (Link Data) Router Interface address: 0.0.0.10
    Number of TOS metrics: 0
     TOS 0 Metrics: 64
```

MIB-II ifIndex of P-P link

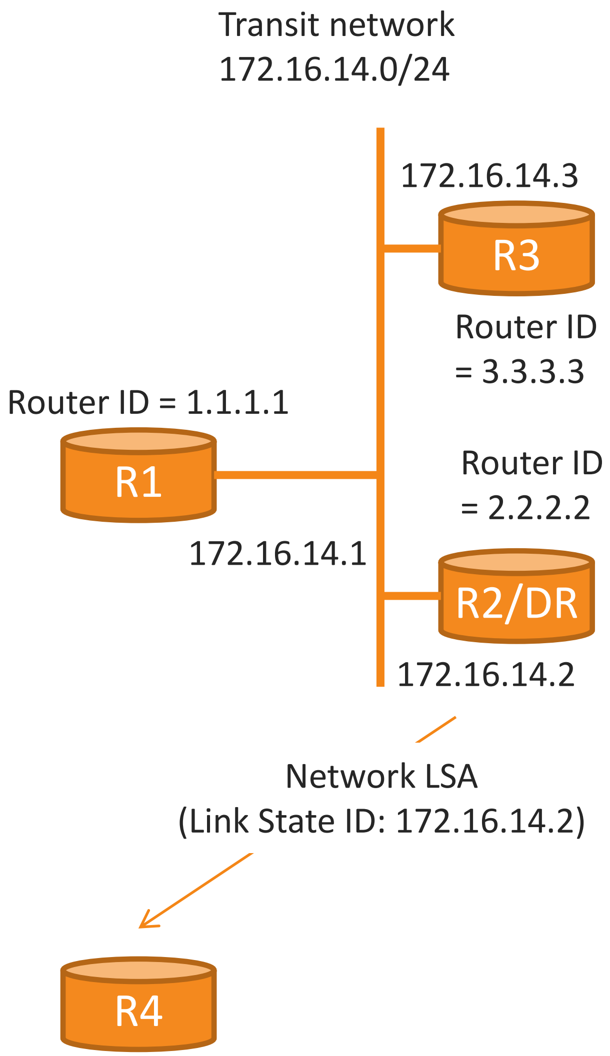


Network LSA (type 2)

- *Originated for broadcast and NBMA networks by the DR router*
- *Flooded throughout a single area (the area of the orginator)*
- *Identified by*
  - Link State ID = network address of the DR router interface
- *Contains Router IDs of all routers connected to the transit network (including the DR router ID)*
- *In LSDB*
  - Network LSA represents vertex
  - Each Router ID is interpreted as pointers to Router LSA(s)

Network LSA example

```
R4# show ip ospf database network

      OSPF Router with ID (4.4.4.4) (Process ID 1)

                Net Link States (Area 0)

Routing Bit Set on this LSA in topology Base with MTID 0
LS age: 170
Options: (No TOS-capability, DC)
LS Type: Network
Link State ID: 172.16.14.2 (address of Designated Router)
Advertising Router: 2.2.2.2
LS Seq Number: 80000007
Checksum: 0x8FB6
Length: 36
Network Mask: /24
      Attached Router: 2.2.2.2
      Attached Router: 3.3.3.3
      Attached Router: 1.1.1.1
```

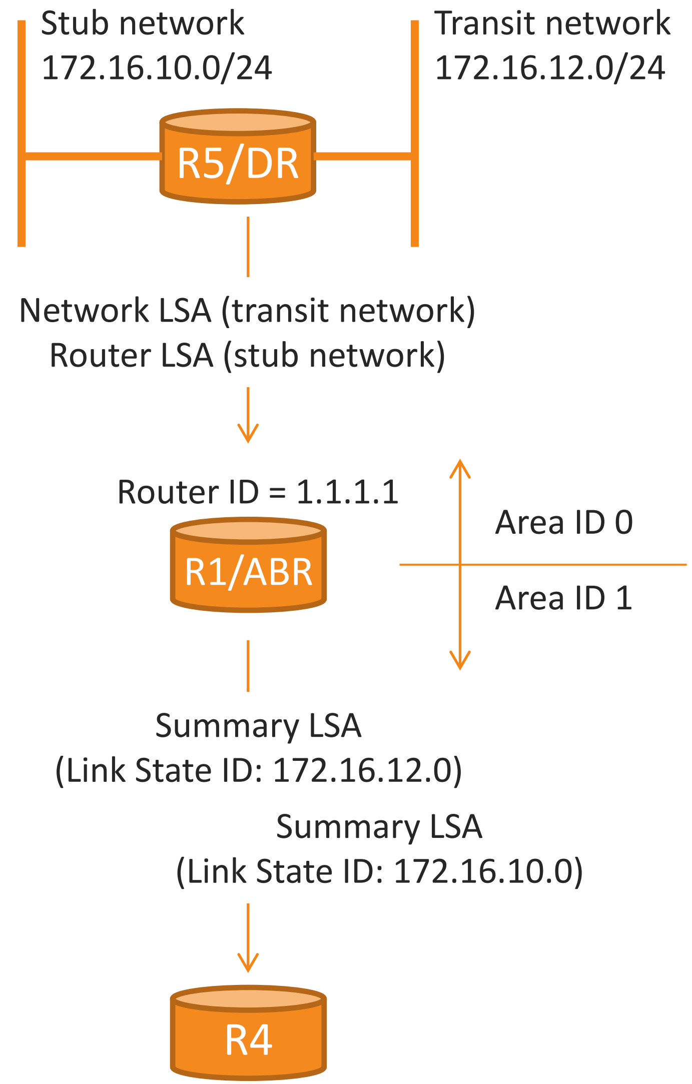


Summary LSA (type 3/4)

- *originated by area border routers (ABR)*
- *flooded throughout single area*
- *Type 3 - describes a route to a destination outside the area (but inside the AS)*
  - ❑ ABR router learns the destination addresses form Router LSA and Network LSA of given area
  - ❑ Link State ID = IP network address (IP network number)
- *Type 4 – informs about the presence of ASBR outside the area*
  - ❑ ABR learns the presence of ASBR from its Router LSA
  - ❑ Link State ID = ASBR Router ID

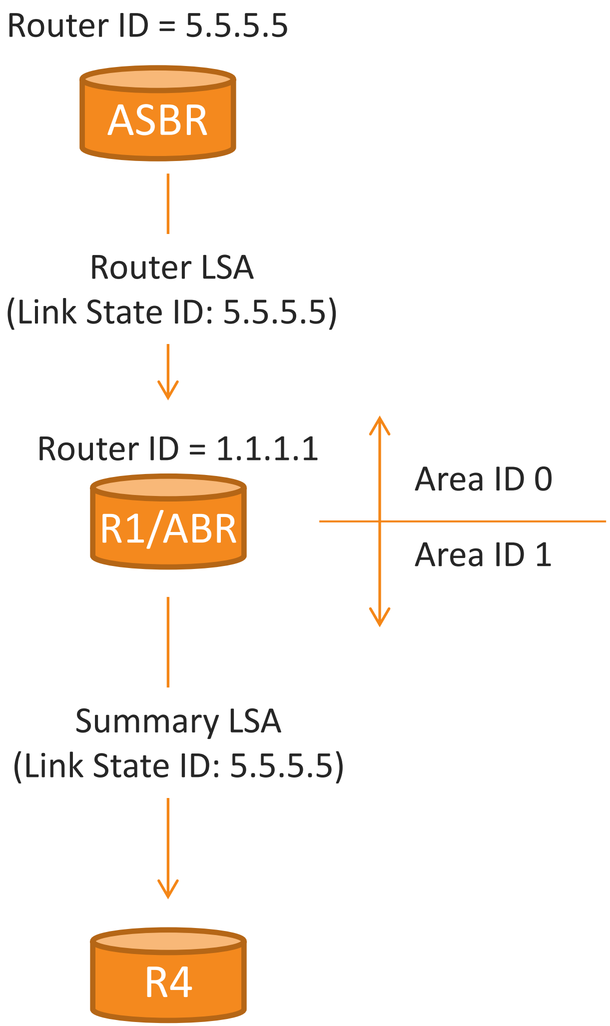
Summary LSA type 3 example

```
R4# show ip ospf database summary

OSPF Router with ID (4.4.4.4) (Process ID 1)

Summary Net Link States (Area 1)

LS age: 608
Options: (No TOS-capability, DC, Upward)
LS Type: Summary Links(Network)
Link State ID: 172.16.12.0 (summary Network Number)
Advertising Router: 1.1.1.1
LS Seq Number: 80000007
Checksum: 0xC567
Length: 28
Network Mask: /24
Metric: 64

LS age: 1023
Options: (No TOS-capability, DC, Upward)
LS Type: Summary Links(Network)
Link State ID: 172.16.10.0 (summary Network Number)
Advertising Router: 1.1.1.1
LS Seq Number: 80000007
Checksum: 0xC567
Length: 28
Network Mask: /24
Metric: 64
```



Summary LSA type 4 example

```
R4# show ip ospf database asbr- summary

    OSPF Router with ID (4.4.4.4) (Process ID 1)

        Summary ASB Link States (Area 1)

LS age: 608
Options: (No TOS-capability, DC, Upward)
LS Type: Summary Links(AS Boundary Router)
Link State ID: 5.5.5.5 (AS Boundary Router address)
Advertising Router: 1.1.1.1
LS Seq Number: 80000007
Checksum: 0xC567
Length: 28
Network Mask: /0
Metric: 10
```

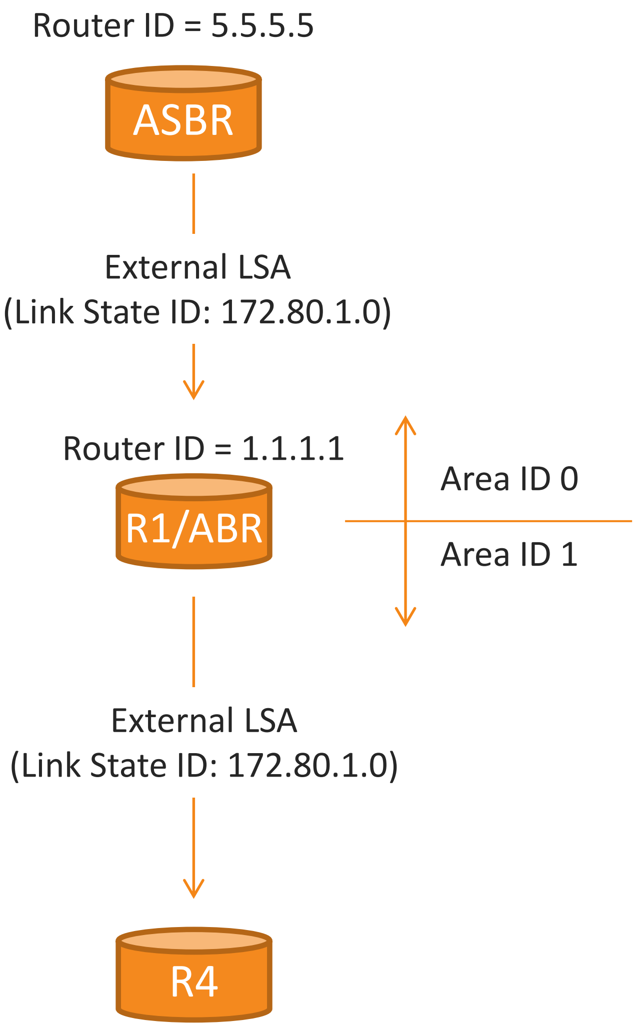


External LSA (5)

- *originated by AS boundary routers*
- *flooded throughout the AS (except stub area)*
- *describes a routes to a destination in another Autonomous System or default route*
- *The location of advertising router (ASBR) is obtained from Summary LSA type 4*
- *External route metric*
  - ❑ type 1 - sums the external metric with the cost of internal path to ASBR
  - ❑ type 2 - only external metric is considered

External LSA (5) - example

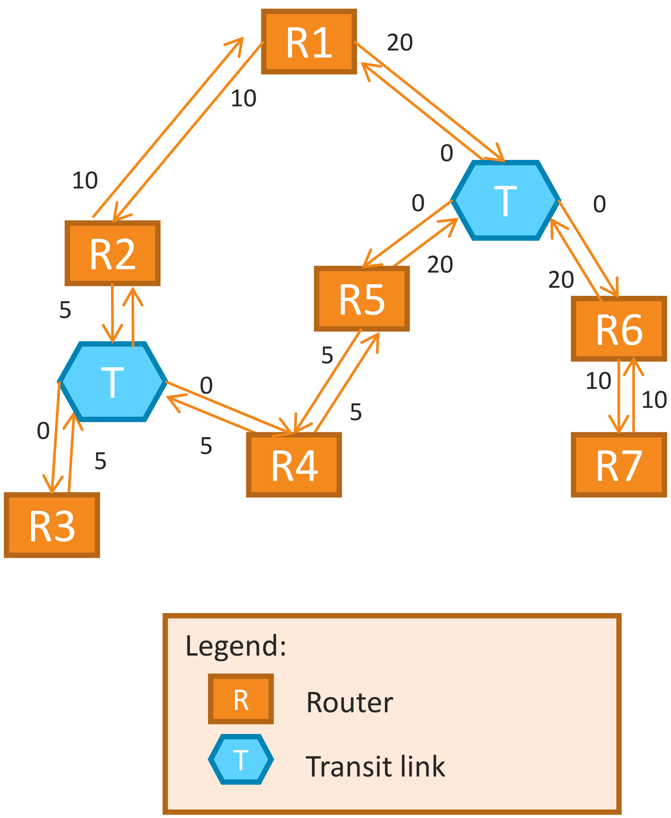
```
R4# show ip ospf database external

    OSPF Router with ID (4.4.4.4) (Process ID 1)

              Type-5 AS External Link States

LS age: 1434
Options: (No TOS-capability, DC, Upward)
LS Type: AS External Link
Link State ID: 172.80.1.0 (External Network Number)
Advertising Router: 5.5.5.5
LS Seq Number: 80000002
Checksum: 0x980
Length: 36
Network Mask: /24
    Metric Type: 2 (Larger than any link state path)
    Metric: 20
    Forward Address: 0.0.0.0
    External Route Tag: 0
```



LSDB – building network topology graph

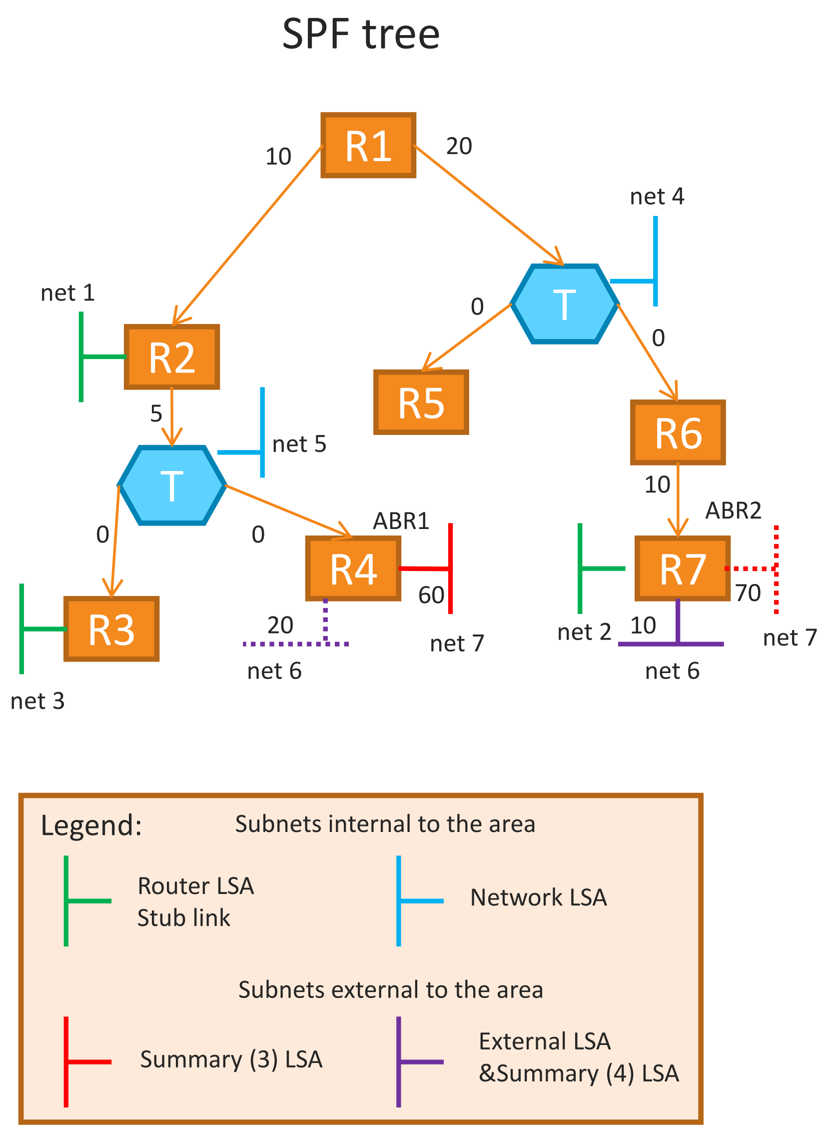
- *Vertexes*
  - ❑ Router LSA - represents routers
  - ❑ Network LSA - represents transit links (broadcast or NBMA networks)
- *Directed Edges*
  - ❑ Router-to-router
    - Obtained from point to point links in Router LSA
    - In LSDB represented as pointer between Router LSA(s)
  - ❑ Router to transit link
    - Obtained from transit link in Router LSA
    - In LSDB represented as pointer between Router LSA and Network LSA
  - ❑ Transit link to router
    - Obtained from Network LSA
    - In LSDB represented as pointer between Network LSA and Router LSA

Network topology



LSDB – building routing table

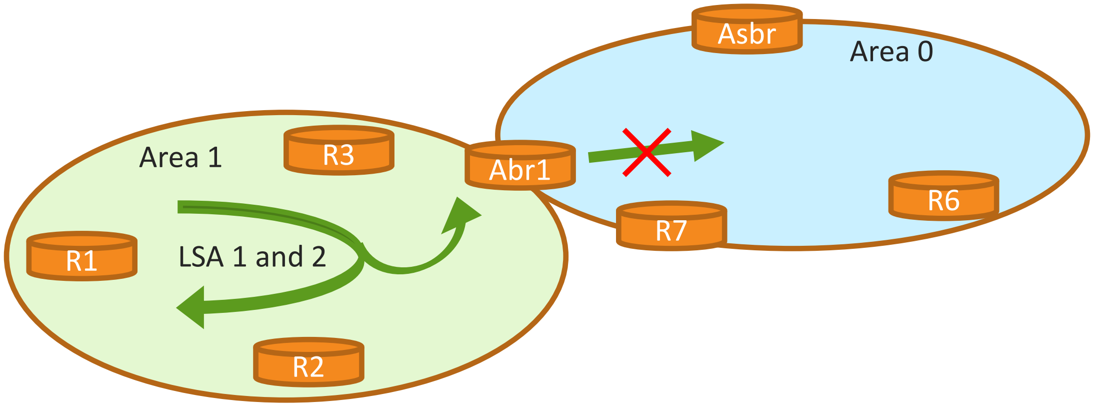
- *Router runs Dijkstra algorithm to find SPF*
- *Router adds information about IP subnets to the SPF tree*
  - ❑ Subnets internal to the area
    - Obtained from Router LSA (stub link) and Network LSA
  - ❑ Subnets external to the area
    - Obtained from Summary LSA 3 and External LSA
  - ❑ ASBR routers external to the area
    - Obtained from the Summary LSA 4 and External LAS 5



LSA flooding

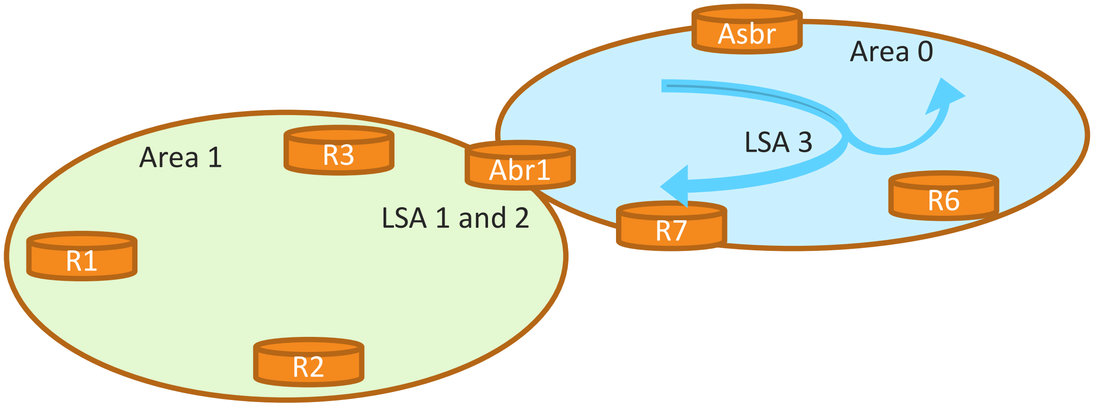
- *OSPF routers send LSA messages periodically every 30 seconds or whenever there is a change in network topology (links state changes)*
  - LSA are generated with sequence numbers
  - LSA are broadcasted over the network (according to their scope)
  - Routers track the LSA sequence numbers to determine which LSA is newer

LSA flooding(LSA 1&2)



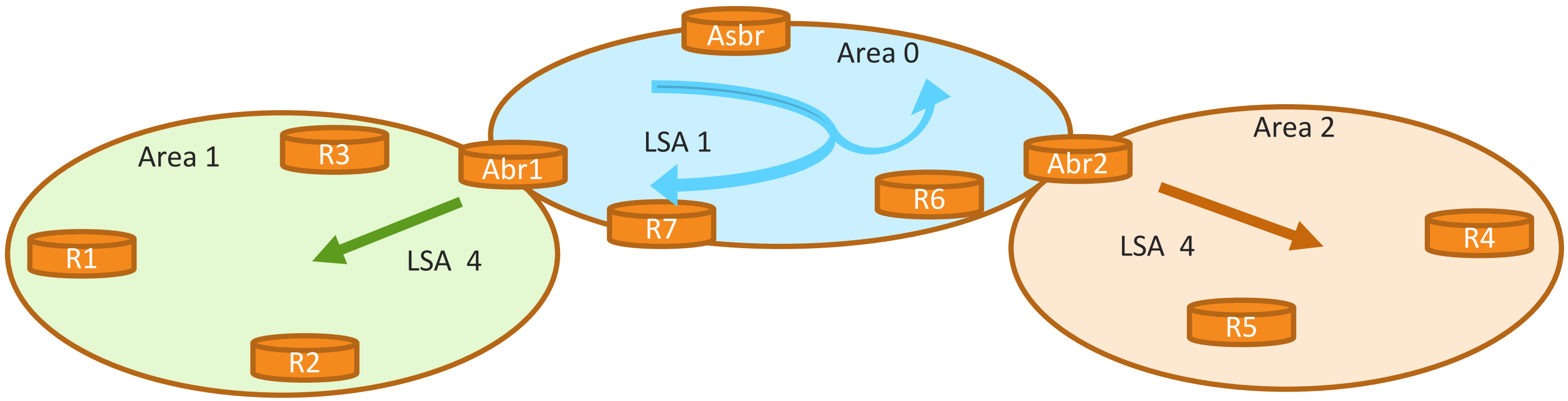
- *Router LSA (1) and Network LSA (2) are flooded only in the area of the originating router*
  - ❑ ABR routers filter them out
- *Routers build the area LSDB using the information received in Router and Network LSA*
  - ❑ Routers don't know the topology of the other areas
  - ❑ Inter area traffic is routed via ABR routers

LSA flooding (LSA 3)



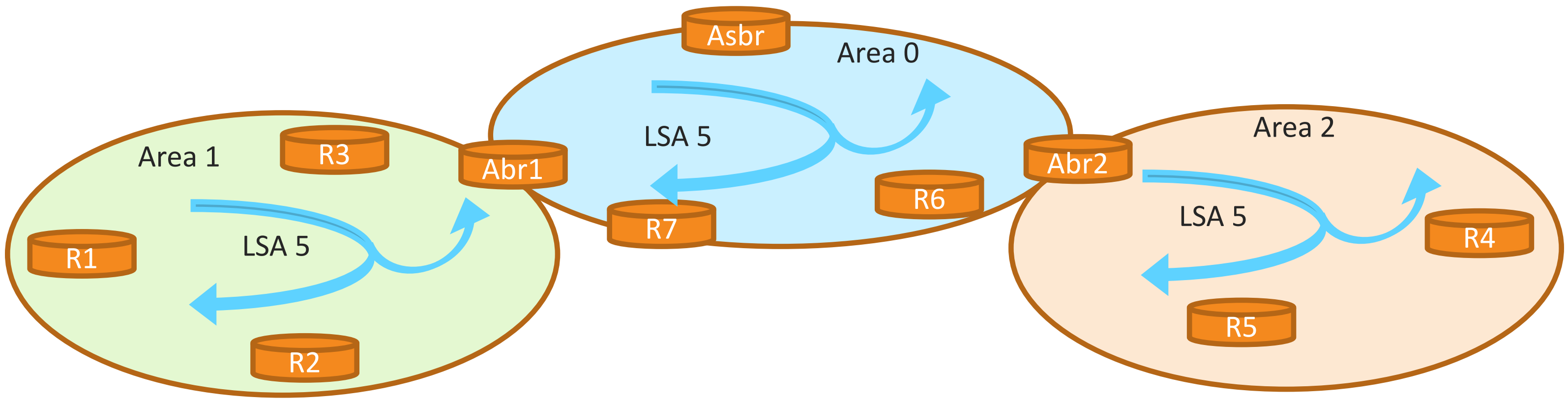
- *ABR routers learn the prefixes available in the give area form LSA type 1 and 2 received from given area*
- *ABR routers send Summary LSA (3) to another areas to distribute prefixes between areaa*
  - ❑ ABR can optionally summarise the routing information
- *In the area LSDB all external prefixes are represented by ABR routers*

LSA flooding (LSA 4)



- *ABR learns the presence of ASBR routers from Router LSA (1) received from given area*
- *ABR routers send Summary LSA (4) to other areas to inform them about the presence of ASBR*

LSA flooding (LSA 5)



- *AS External LSA (5) originated by the ASBR are sent over entire AS*
  - ❑ The Exteranl LSA carries information about the external routes to the OSPF domain
  - ❑ The location of ASBR is known from the Summary LSA type 4 or Router LSA (in area where the ASBR is located)

OSPF packets

- Hello packet – used to establish and maintain the adjacency relationship
  - *Neighbour discovery*
  - *DR/BDR routers election*
  - *Establishing master/slave relationship*
- Database Description – used to exchange sequence numbers of the LSA messages contained in the router LSDB
- Link State Request – used to request a certain LSA from neighbor LSDB
- Link State Update – used to send LSA messages
- Link State Acknowledgement – used to acknowledge the reception of Links State Update packet

Hello protocol

- *Hello protocol is used for establishing neighbour relationship (adjacency) between routers*
  - ❑ Only adjacent routers can exchange routing information
  - ❑ Adjacency on point-to-point links is established automatically
  - ❑ Adjacency on bradcast networks is established via DR and BDR router election
- *Hello protocol is used for keep alive procedure*
  - ❑ Hello packets are periodically send out from each router interface with the multicast IP address 224.0.0.5 (allOSPFRouters)
  - ❑ typically the hello packets are sent every 10 seconds
- *Hello protocol is used to discover changes in the network topology*
  - ❑ link failures/deletion
  - ❑ link addition

Hello protocol

- *DR/BDR routers election*
  - via the exchange of hello packets router learns about its neigbours on the network
  - the router with the highest priority becomes DR router
  - the router with the second priority in order becomes the BDR router
  - the BDR router becomes active when the DR router goes out of service
- *Each router on the network establishes adjacency with the DR and BDR routers*
- *Only the DR router sends LSA messages for given network*
  - DR router represents the broadcast network to other routers
  - this decreases the routing traffic

Hello protocol

Down state

Hello packet — I am router ID 172.16.5.1 and I see no one →

Init state

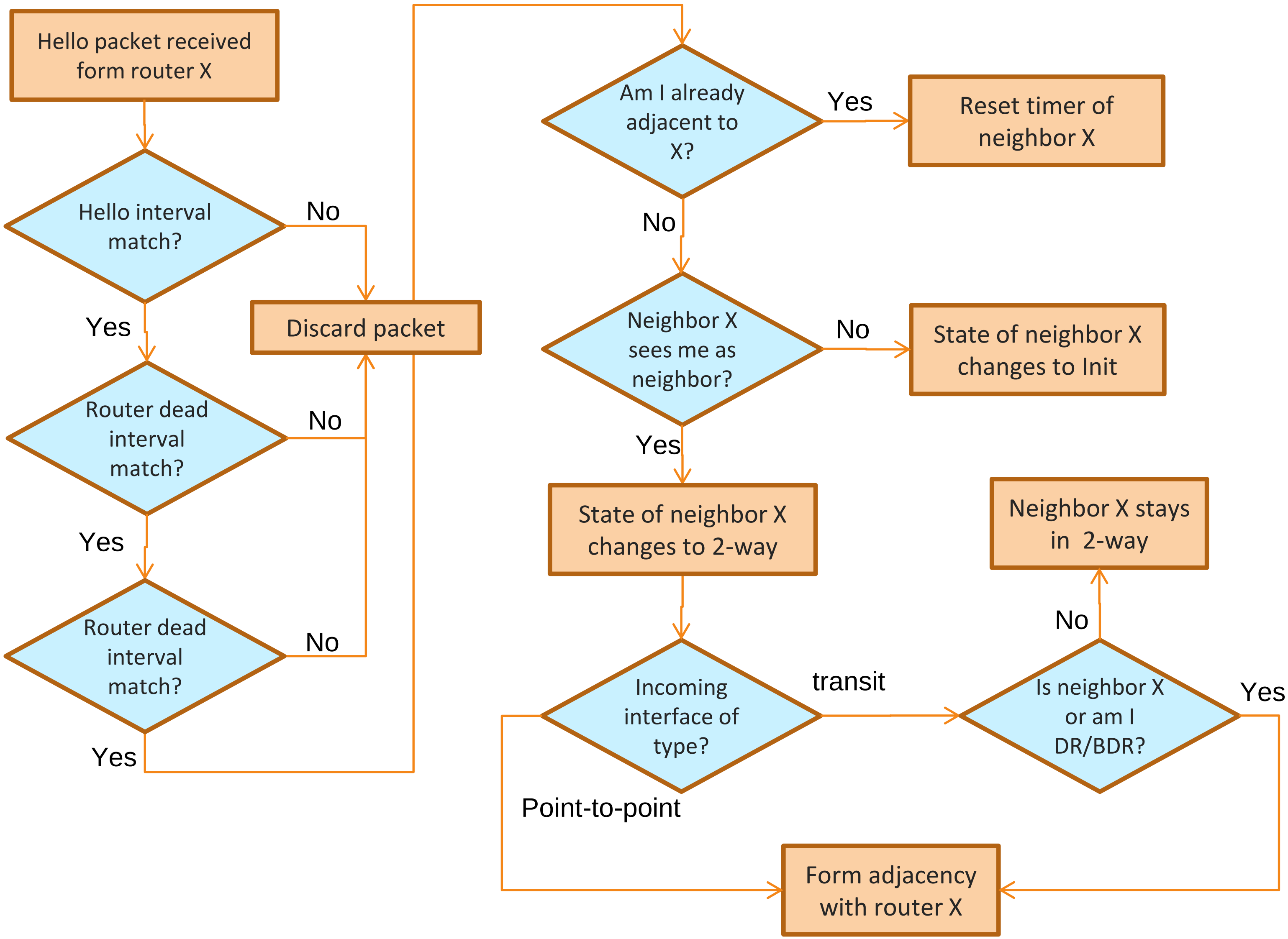
Router B
Neighbour list
172.16.5.1/24 int E1

← I am router ID 172.16.5.2 and I see 172.16.5.1 — Hello packet

Router A
Neighbour list
172.16.5.2/24 int E0

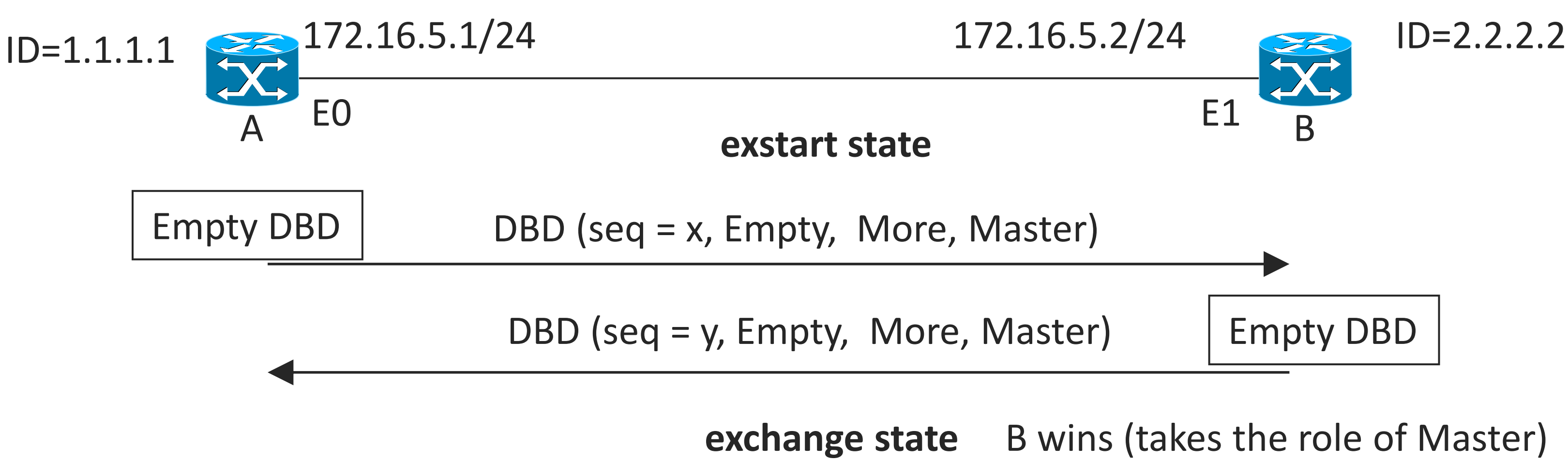
Two way state

Hello Protocol



Exchange protocol

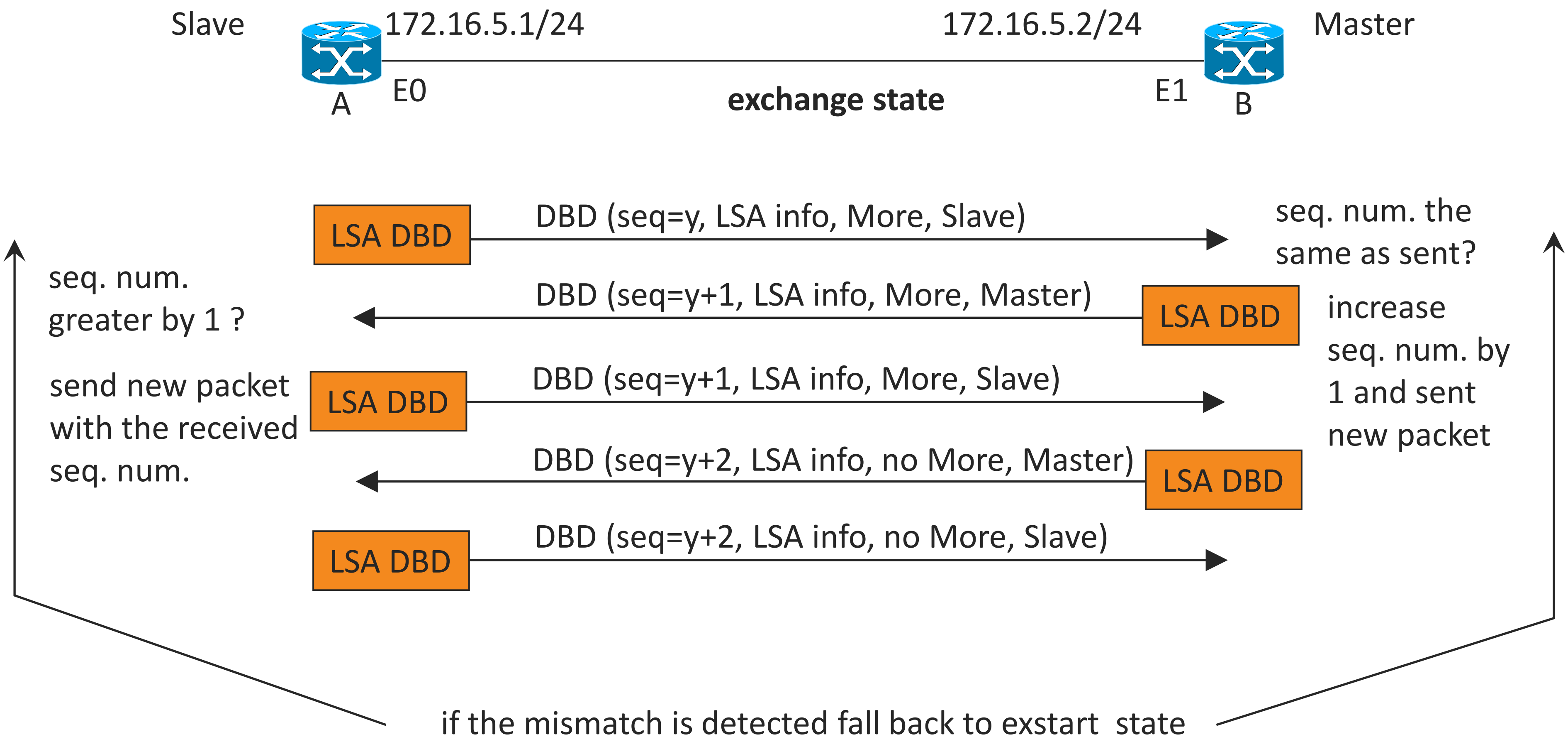
- *Exchange protocol is used to synchronise the neighbour routers LSDB databases*
- *Adjacent routers start to exchange routing information*
  - ❑ the master-slave relationship is established via DBD packets exchange
    - It is used to synchronise the sequence number used to provide reliable packet exchange
  - ❑ the router with the higher Router ID becomes the master



Exchange protocol

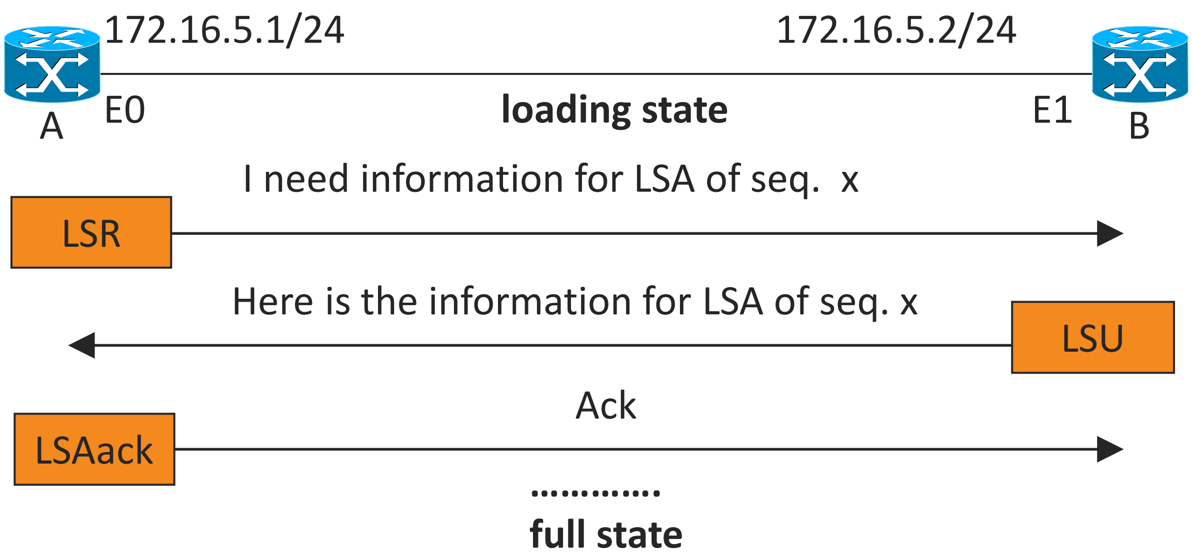
- Adjacent routers exchange DBD packets to learn the LSA sequence numbers in the neighbour LSDB
  - *the slave always sends DBD packets with the sequence number received from last master DBD packet*
    - master always expects packets with the same sequence number as sent to the slave
    - In this way slave acknowledge packet received from master
  - *the sequence number is increased by master every time the new packet is send*
    - the slave always expect packet with sequence number higher by one then the previous one
    - In this way master acknowledge packet received from slave
  - *if mismatch is detected on the master or slave the whole procedure is restarted*

Exchange protocol



Exchange protocol

- Each router compares the received LSA sequence numbers with those it has, if some LSAs are older it sends the LSR (Link State Request) packet for these LSAs
- The other router responses with the LSU (Link State Update) packet that contains full LSAs information
- The router acks the LSU packet
- After the exchange protocol reaches full state (no differences in LSDB) each router pre-computes its routing table entries



Routing table maintenance

- LSA are flooded to all routers in the OSPF area each time there is a topology change on one of the links directly connected to the router or there is change in the broadcast network configuration
- If there are no topology changes, the router will flood its LSAs every 30 minutes.
  - *every LSA has a maximum age of 60 minutes.*
  - *an OSPF router will age all LSAs in its link state database and will purge any LSAs for which it has not received a refresh in the last 60 minutes.*
- LSAs are send to
  - *multicast address 224.0.0.5 on point-to-point links*
  - *multicast address 224.0.0.6 on bradcast networks when sending to DR/DBR routers*
  - *multicast address 224.0.0.5 on bradcast networks when sending by DR to non DR routers*
- After router receives LSU it re-computes its routing table

OSPF metric

- *OSPF metric for an interface is automatically calculated based on the OSPF reference bandwidth which, by default, is 100 Mbps.*
  - ❑ the metric is calculated by dividing the reference bandwidth by the actual bandwidth of the link
  - ❑ example: 10 Mbps link -> metric = 10
  - ❑ with default reference bandwidth the cost for links of capacity higher then 100 Mbps will be always 1
- *Alternatively, the OSPF metric of an interface can be configured manually*
- *The default metric of loopback interface is zero*