Open Shortest Path First

Fulvio Risso

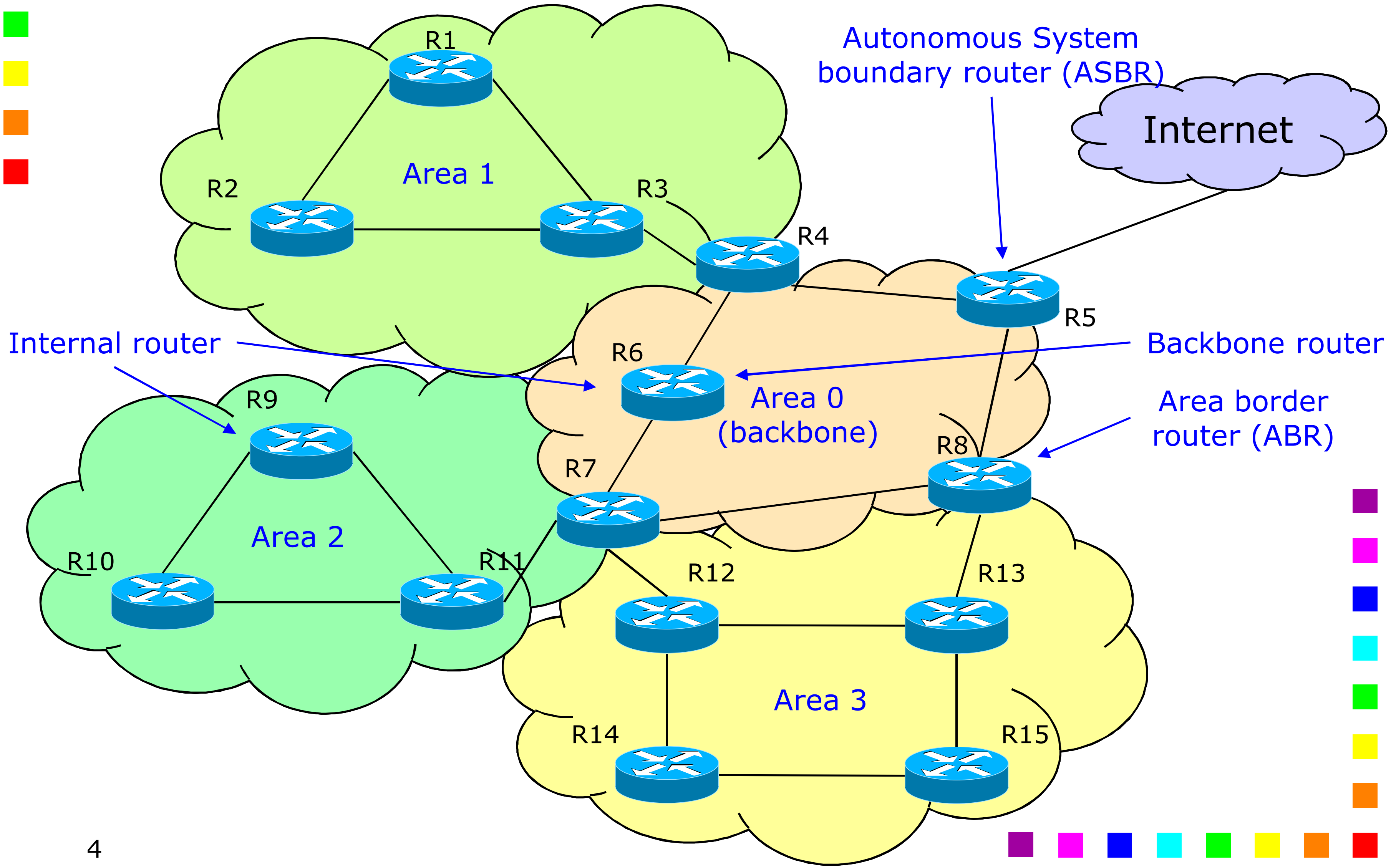
Politecnico di Torino

This set of slides is based on a previous version created by Mario Baldi and Giorgio Valent

General concepts

- Link State protocol defined by the IETF
  - OSPF v.1: RFC 1247 (1991)
  - OSPF v. 2: RFC 2328 (1998)
- Newest versions exist for IPv6 (OSPF v.3: RFC 5340, 2008)
- OSPF has the hierarchical concept:
  - Can handle large networks
  - An AS is subdivided in areas
  - Each area contains a group of contiguous networks
  - Backbone: special area, not necessarily contiguous, which is connected to all other areas

Example of a possible OSPF network



Terminology (1)

- OSPF defines its own terminology, which is not always aligned to what other protocols do
  - Sometimes names are misleading, e.g. “autonomous system boundary router” is not necessarily the router at the border of an AS
- Backbone level-2 network
- Backbone router router in a level-2 network
- Area border router router that has interfaces in two or more areas (one of them must be the backbone)
  - This router executes many copies of the link state algorithm (one per area)
- Internal router router in a level-1 network
- Note: level-1 and level-2 refer to the hierarchical placement of those network, not to the OSI layer

Terminology (2)

- Autonomous System (AS) OSPF routing domain
- AS boundary router router that has at least one interface connected to another routing domain (not necessarily into another AS)
- Link State Advertisement (LSA) data structure that contains the "core" link state information
  - LSA are not packets
  - A router can put several LSAs within the same packet

Metrics

- OSPF supports more than one metrics simultaneously on a single link
- The best path may be, depending on the packets
  - The shortest path
  - The one with the best capacity
  - The one with the lowest delay
  - …
- OSPF allows to define metrics depending on the TOS field of the IP packet
  - In theory, 64 possible types of service
  - In practice, this feature is almost unused

Costs

- Differently from IGRP, OSPF does not define an unambiguous way to calculate the cost of a link
- Assigned by the manufacturer of the network device
  - Between 1 and 65535
  - Each manufacture has his own default values

    Cost (Cisco) = 10^8/bandwidth

  - Better to customize it on most important links (on both ends)

```
interface <name>
          ip ospf cost <value>
```

Interface	OSPF default cost (Cisco)
Fast Ethernet	1
T3 (45 Mbps)	2
T1 (1.5 Mbps)	64

Equal Cost Multi-Path Routing

- Possible when more than one path have the same cost
  - Also known as Load Sharing
  - Note that paths that have the same cost may not be, in fact, equivalent from the point of view of the network provider
- OSPF does not mandate the implementation of ECMP Routing
  - OSPF does not preclude the possibility to use several paths in parallel if they have the same cost but it does not give further details
- Almost all implementations support ECMP Routing
  - Each implementation may have different criteria for distributing the traffic across the equivalent links

OSPF and aggregation

- The routers of the backbone are usually configured (manually!) in order to aggregate network addresses
  - Network summaries are propagated in the other areas
  - For example
    - An area contains addresses 5.12.*.* and 5.7.*.*
    - The level 2 router propagates the connectivity toward 5.*.*.*
- Backbone routers propagate summaries of the networks received from the other areas
  - Routers inside an “edge” area may not choose an optimal exiting point from the area
- The aggregation must be specified manually by the operator in order not to have troubles with network summarization

Router ID

- Unique “name” of the router
- OSPF does not specify how it should be determined
  - It mandates that the algorithm produces unique identifiers
- Cisco
  - Bigger addresses present on the loopback interface
    - Loopback interfaces do not depend on the state of the physical interfaces and are thus more stable
  - If there is no loopback interface, the bigger IP address configured on the OSPF interfaces is chosen
  - The RouterID is computed at the beginning of the OSPF process and is not modified even if the IP addresses on the router are modified
    - In such a case, the router may have a new RouterID at the following reboot

Authentication

- OSPF can store information for authenticating the other end party
- To each link is associated one (and only one) password
  - Only the OSPF packets containing the password associated to the link they come from are accepted
  - It is difficult to change the password of a link
    - You have to change on both ends of the link
  - Easy to intercept a message to have the password of a link and to use it in both directions
  - Almost unused, due to its weakness

Propagating LSA on the network

- Routers have to propagate LSA on the network
  - Needed e.g., to propagate topological changes to the rest of the network
- In practice, LSAs are generated:
  - When a change is detected in the network (e.g., cost, link up/down, etc.)
  - When a timer expires (default: 30 min)
    - This makes the network much more stable, even if LSAs are acknowledged
    - It prevents an LSA of a dead router to keep staying in the network forever
- Please note that LSAs in the OSPF database expire if not renewed within a given time (default: 60 min)

Selective flooding protocol

- Selective Flooding protocol is used to propagate LSA
- Flooding requires includes two types of packets
  - Link State Update (Header OSPF: Type = 4)
  - Link State Acknowledgement (Header OSPF: Type = 5)
    - Needed to implement a reliable transmission of the LSAs and guaranteeing that the database is coherent across all the routers
- In practice, two fields are needed in the OSPF packet:
  - Sequence Number (to recognize the newest LSA)
  - Age (to delete old LSA that have not been renewed)

OSPF Sequence Numbers (1)

- OSPF v.1 uses a lollipop space
  - Studies demonstrated that the overflow in OSPF was not a real problem
  - Sending a new LSA every 1 sec. with a SeqNum over 32 bits translates into 136 years of continuous running before getting an overflow
- OSPF v. 2 uses a linear space
  - A special rule applies when the SeqNum reaches the end of the numbering space

OSPF Sequence Numbers (2)

- Value -N (0x80000000) is reserved (and unused)
- Value -N + 1 (0x80000001) represents the Initial Sequence Number
- Afterwards, the LSA sequence number is incremented each time the router originates a new instance of the LSA
- If a router R1 receives a LSA with an age < of the age of the LSA currently in its database, the "newest" LSA is flooded to all the network (hence it will reach also the sender R2)
  - In case R2 is the originating router for that LSA, it recognizes that an "older" LSA exists in the network, hence it updates its SeqNum to the one of the received LSA + 1 and it re-floods it

OSPF Sequence Numbers (3)

- If a router has to send a LSA with value N - 1 (0x7fffffff; also referred to as MaxSequenceNumber), the current instance of the LSA must first be flushed from the routing domain, which is done by prematurely aging the LSA
  - The LSA is sent with Age equal to the maximum value MaxAge
  - Any receiving router has to flood that LSA as usual, then it keeps it in memory until all the ACKs from its neighbors have been received
  - At that point, that LSA is deleted from the memory of the router
- As soon as this flood has been acknowledged by all adjacent neighbors, a new instance can be originated with sequence number of InitialSequenceNumber

LSA on broadcast networks (1)

- OSPF has to implement a special algorithm for propagating LSA on broadcast networks
  - High complexity for Dijkstra, a lot of traffic exchanged over the link, many (useless) synchronizations when a new router joins the network
- On each broadcast network, two routers are selected
  - Designated Router (DR)
  - Backup Designated Router (BDR)
- Election process based on the two routers that have the best RouterID
  - In practice, the two routers that boot first become DR and BDR
    - The election process is initiated as soon as two routers are found on that network, and DR/BDR do not decade unless an explicit configuration command is given

LSA on broadcast networks (2)

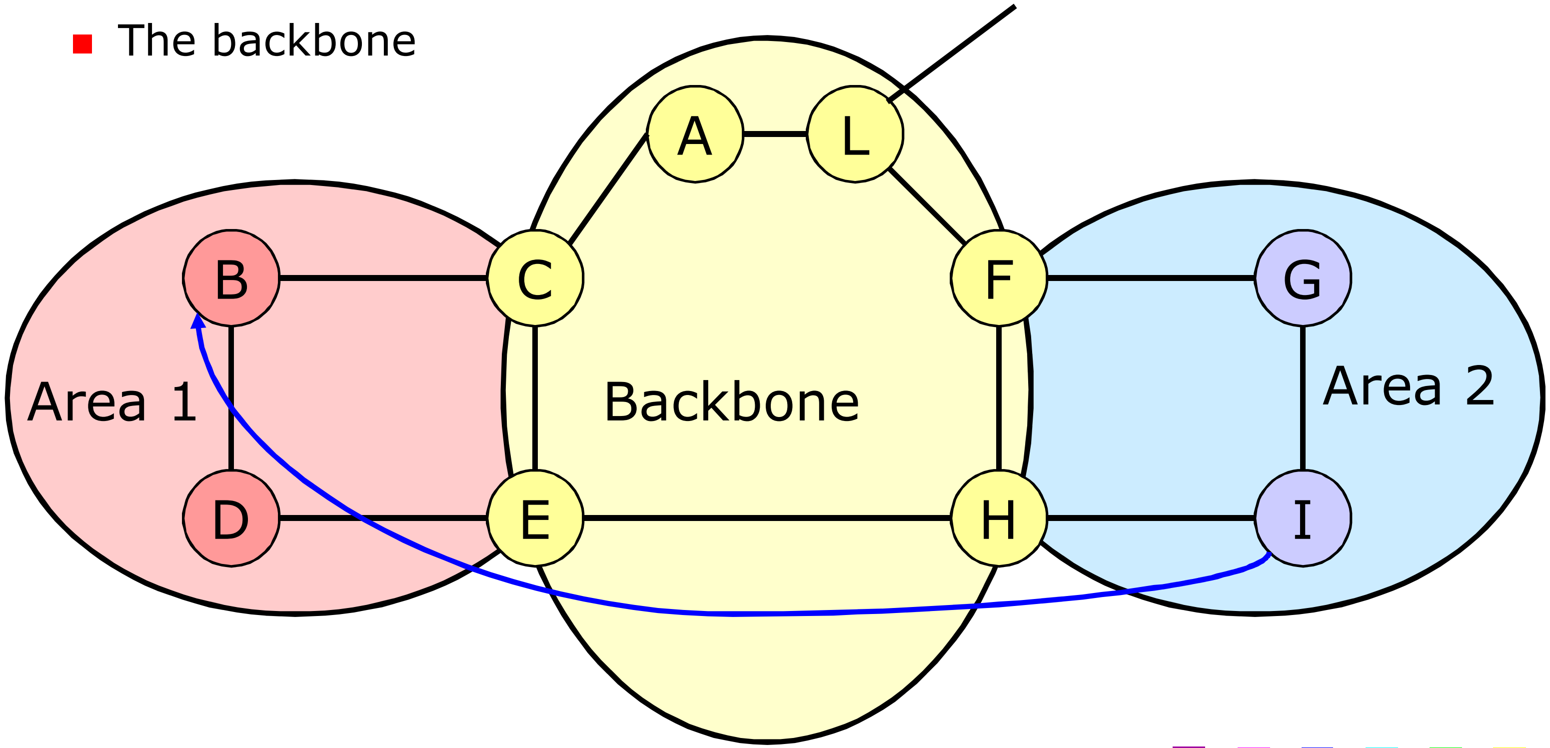
- When a router R has to transmit a LSA on the LAN, it sends the packet to the multicast data-link address AllDrouters
  - Only the DR and the BDR join that group, hence they receive traffic at the AllDrouters multicast address
  - The ACK is generated by the DR and sent back to the original sender at the AllSPFrouters multicast address
- Now, the DR forwards the LSA to the multicast data-link address AllSPFrouters
- The routers send the acknowledgment (that confirms reception) to the address AllDrouters
  - This is needed in order to allow the BDR to receive the ACK
- If the DR does not receive acknowledgment from a subset of routers, it sends copies of the LSA to each of these routers (in unicast)

Propagation of the LSA on LAN

- The DR keeps a lot of information about the state of the other routers
- The replacement of the DR would need:
  - A huge exchange of messages needed by the new DR to gather this information
  - A “large” amount of time before being actually operative
- To avoid so, a Backup Designated Router (BDR) is elected
  - Sends messages directly to the DR
  - Records the same information gathered by the DR
- DR and backup DR are never replaced by another router, unless they stop working
- Beware: the function of DR/BDR is determined per-LAN
  - A router can be DR in a LAN and nothing in another
  - This is a property of the router interface, based on the RouterID and the Priority

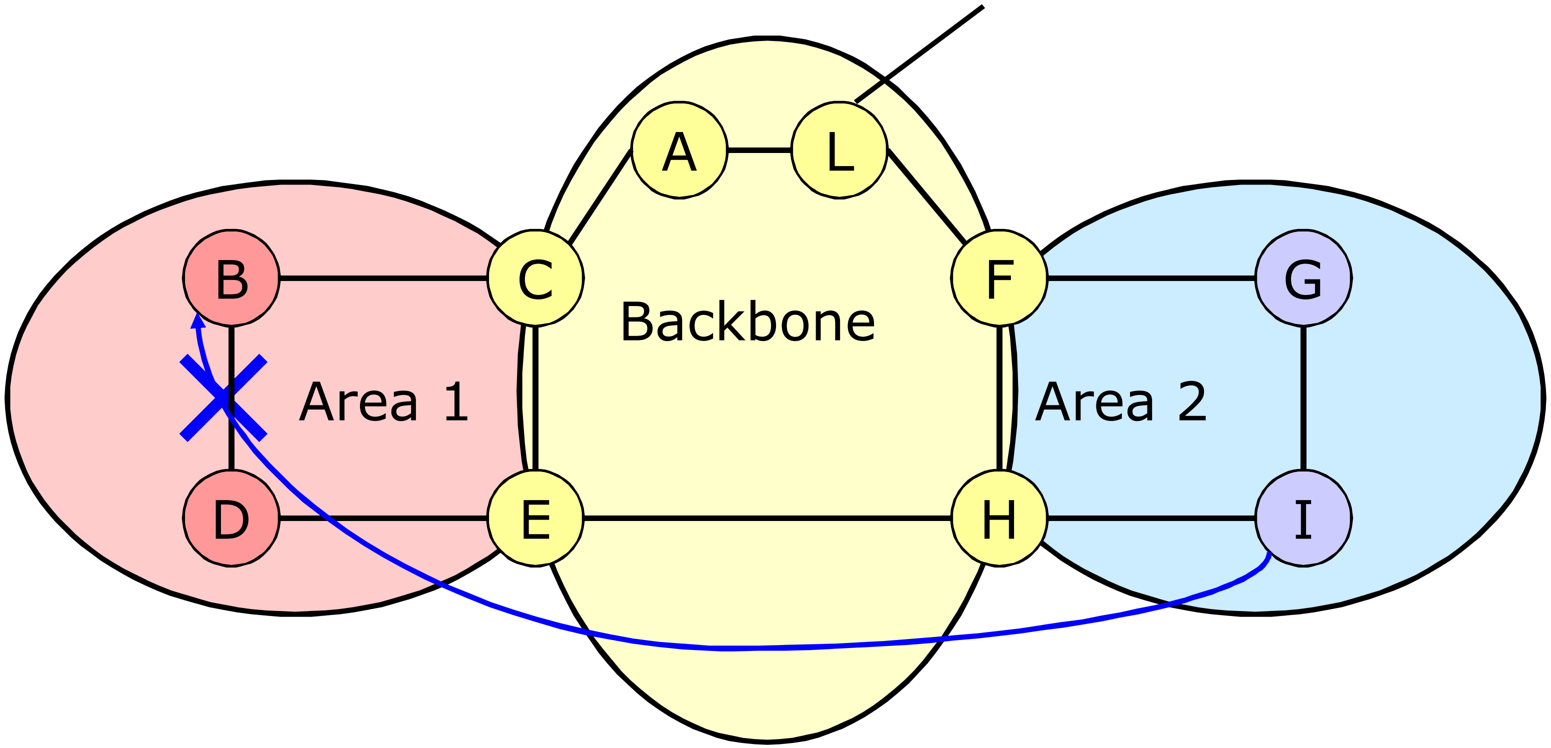
Partitioned areas (1)

- Classical problem of the hierarchical routing
- In OSPF, this is handled differently depending on the type of area
  - A normal area
  - The backbone



Partitioned areas: Area (2)

- Once the routes are summarized, a packet from I to B may be injected in the Area 1 from C or from E
  - The best path is I-H-E-D-B
  - If it is injected from E, the packet cannot reach its destination

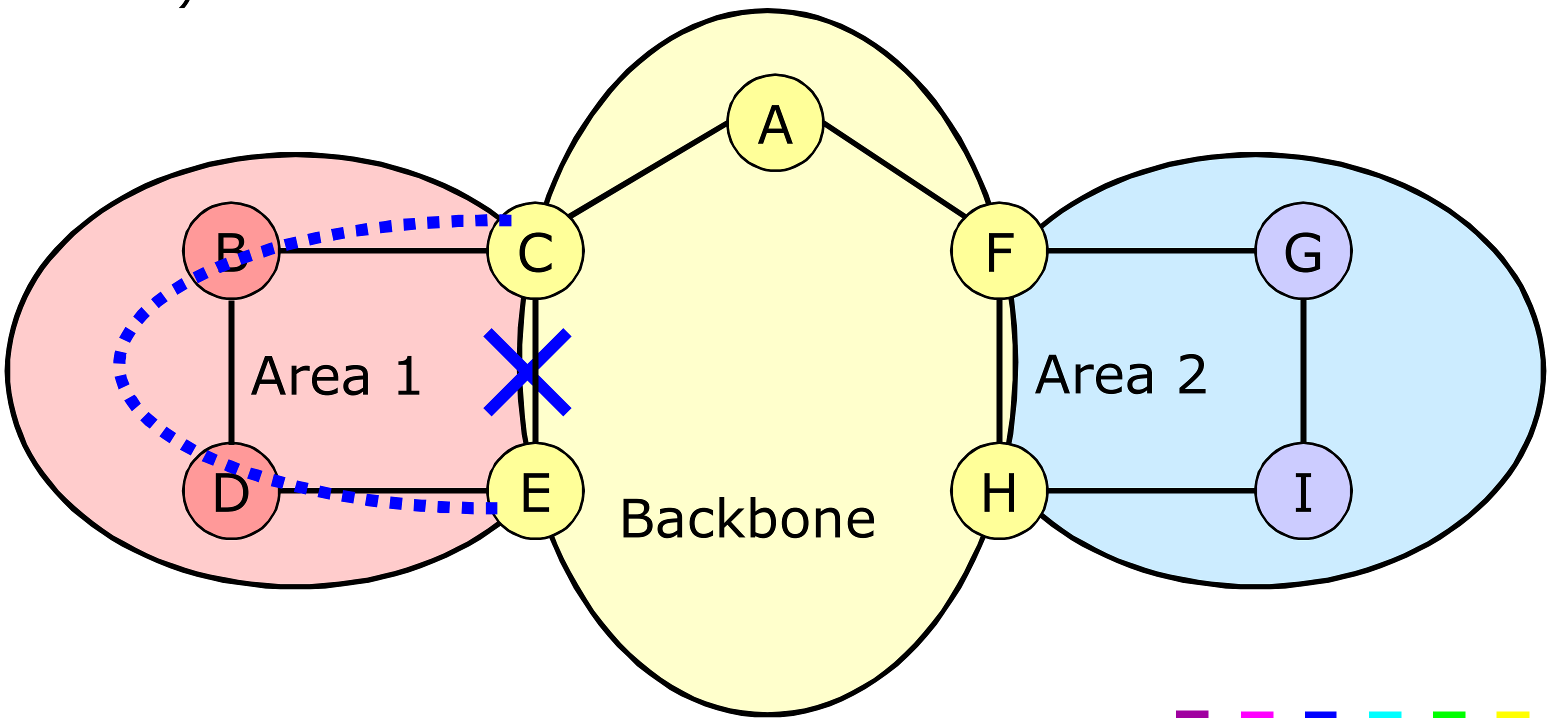


Partitioned areas: Area (3)

- Solution: the Area Border Router does not summarize the information about all the networks present in the edge area; instead, it announces only those network that are reachable from the node itself
  - In the backbone there is still only one entry for each network, in order to be able to select the best path when 2 border routers announce the same network (as it is the case for the network B when there is no failure)

Partitioned areas: Backbone (1)

- A failure in the backbone may isolate the router E even ifan alternative route through area 1 may exist
- The solution is to create a Virtual Link between E and another node of the backbone (for example C, by going through D and B)



Partitioned areas: Backbone (2)

- The Virtual Link looks like a tunnel whose endpoints are automatically configured by the OSPF
  - Configuration is automatic, but…
  - …activation must be done by hand
- Virtual Links are a sort of automatic tunnels
  - In fact, a GRE tunnel looks similar, but the entire configuration has to be done by hand
  - OSPF will advertise one more link crossing the backbone (the Virtual Link, in fact), although its cost is usually rather high

- Advanced use of Virtual Links: bring into the backbone a router that is not directly connected to other backbone routers

Partitioned areas: Virtual Link

- OSPF routing messages are encapsulated in IP unicast packets crossing the link
  - The destination and the source are the endpoints of the virtual link (C and E)
- The data packets sent from C to E have to go through area 1
  - The level 1 routers have to know that the packets are heading to E
    - The information must come from the level 2 information propagated inside the area
- Note: the configuration of the VL requires only the RouterID of the two backbone routers involved, and not the IP addresses of their interfaces
  - OSPF will derive the correct IP address automatically

Timers

- Key values announced in the Hello packet
  - HelloInterval: 10s
  - RouterDeadInterval: 40 sec
    - Ignored if the router detects the death of one of its neighbors through a signal coming from the physical level
  - The adjacency of a router with another peer is established only if the values of the timers are the same for both
- Those parameters have an impact on the convergence speed of a network
- Non specified (only suggested) by the standard
- Other timers
  - LSA Refresh: every 30 min
  - MaxAge: 1 hour

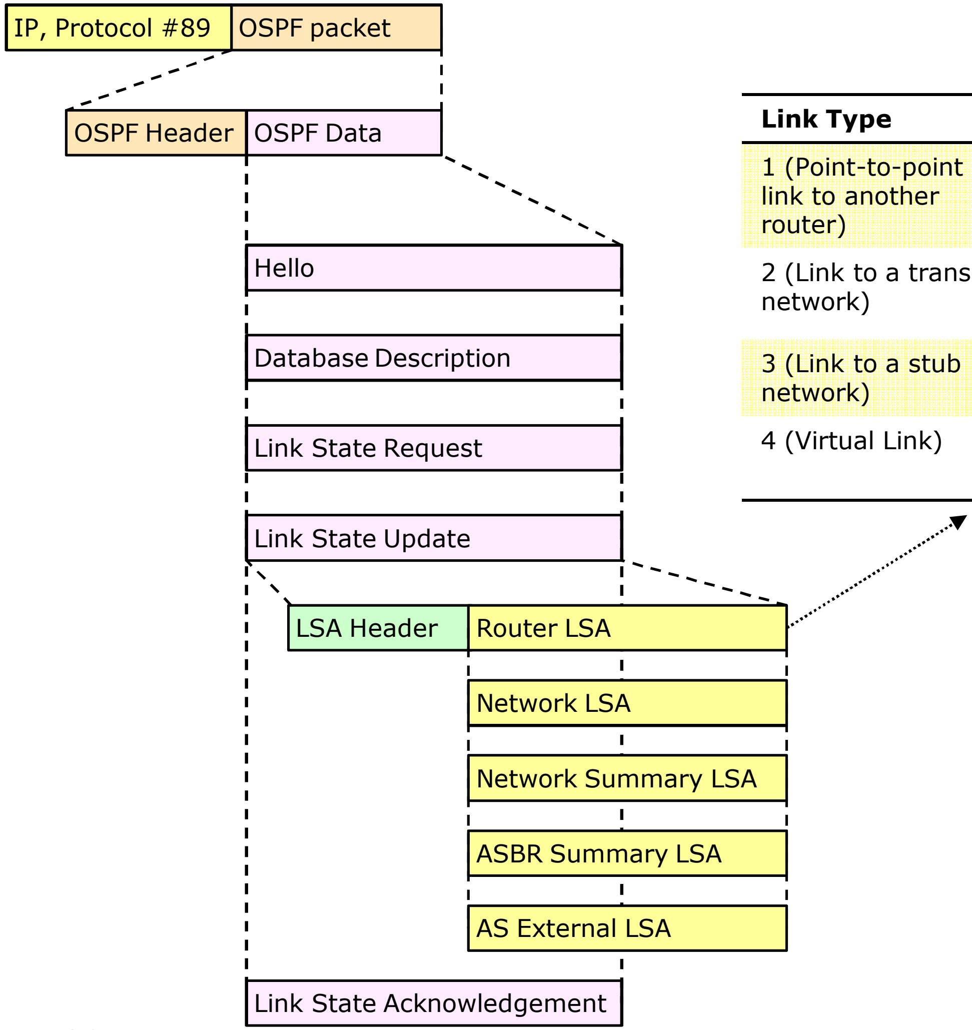
Content of the OSPF database

- The database of a router belonging to an area contains:
  - Link States belonging to all the routers in this area
  - Summary Records, generated by the ABR and related to all the networks belonging to the other areas (including the backbone); those routes can be summarized
    - Previous Links States are never summarized
  - External Records, generated by the ASBR related to all external routes
- This splitting contributes to the scalability of OSPF
- Database entries are removed if replaced by a new entry, if invalidated by special message (an entry with MaxAge) or in case of timeout

Link State Records: types

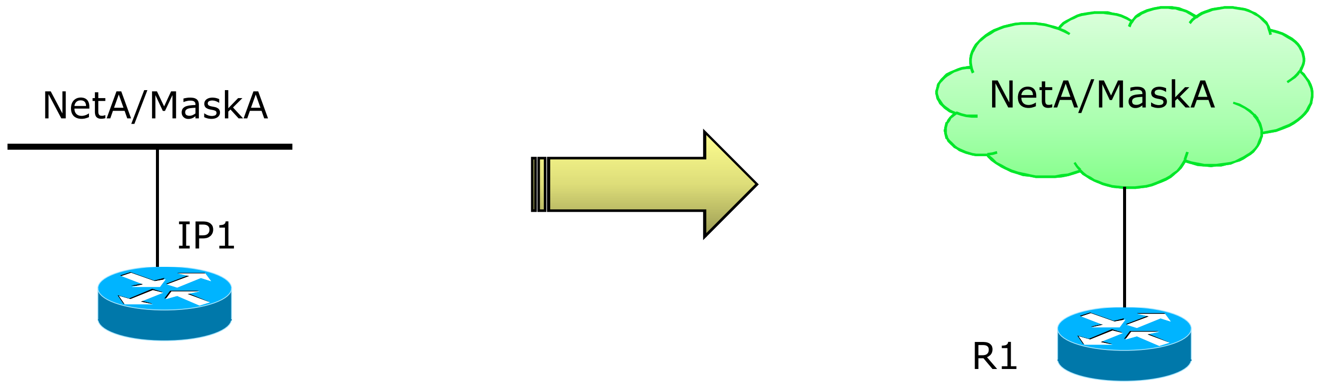
- In the LS-like algorithms, every adjacency is modeled as a point-to-point link
  - Adjacency = adjacent router, adjacent network IP
- In the case of a LAN, the number of adjacencies may explode
  - A particular adjacency is defined with the Designated Router
- OSPF defines two types of link
  - Router Link
    - Point-to-point connection between two routers (e.g., serial link)
    - Point-to-point connection between a router and an adjacent IP network
  - Network Link
    - Point-to-point connection between a router an a transit network that hosts two or more OSPF routers

General structure of OSPF packets



Link Type	Link ID	Link Data
1 (Point-to-point link to another router)	Neighboring router's RouterID	IP address of the originating router's interface to the network
2 (Link to a transit network)	IP address of the DR's interface	IP address of the originating router's interface to the network
3 (Link to a stub network)	IP network address	IP network mask
4 (Virtual Link)	Neighboring router's RouterID	The MIB-II ifindex value for the originating router's interface

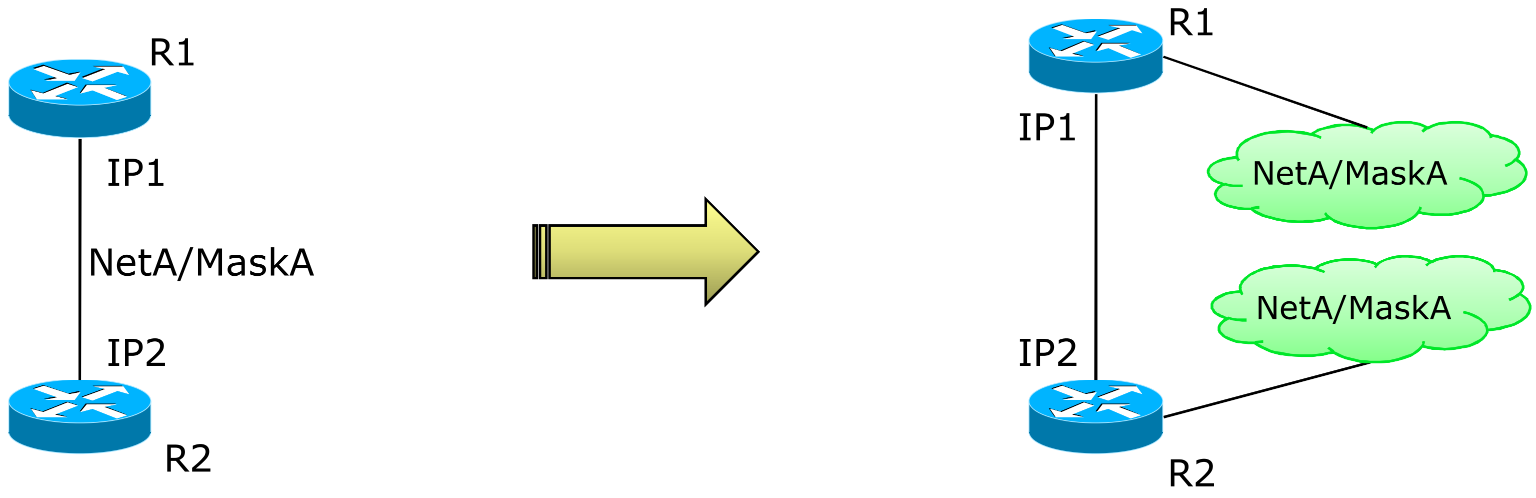
Link State in OSPF (1)



Link State Database:

Router LSA (Link State ID: R1, Adv Router: R1, #links: 1)
Router Link: Link ID (Network): NetA Link Data (Netmask): MaskA - Link to a stub network

Link State in OSPF (2)

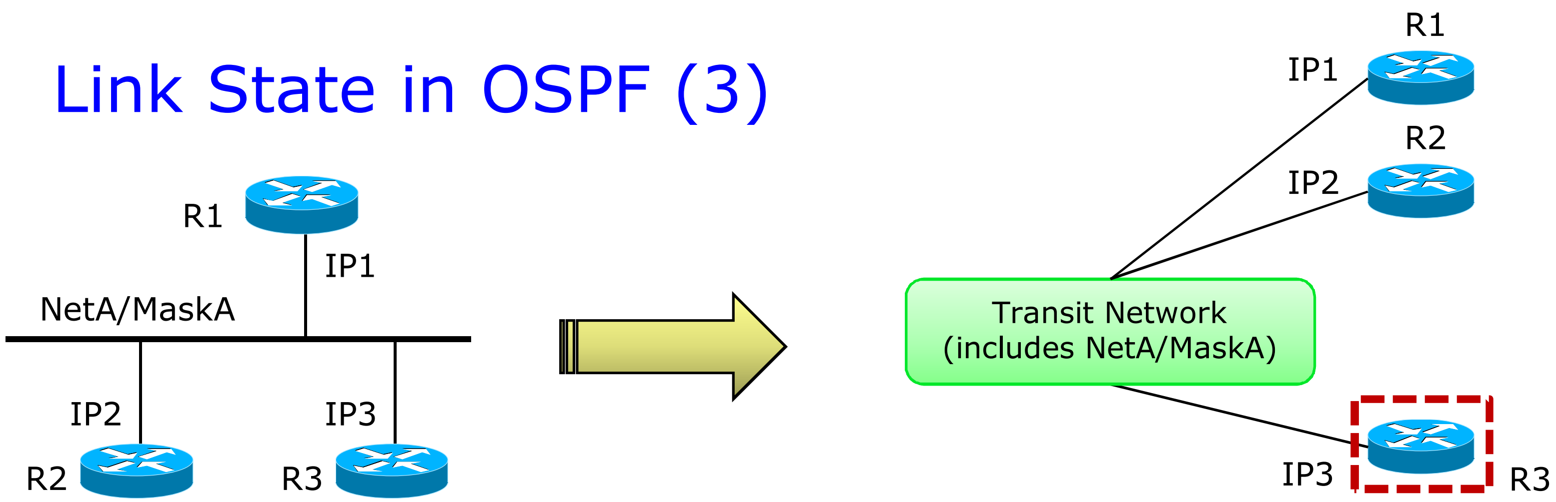


Link State Database:

Router LSA (Link State ID: R1, Adv Router: R1, #links: 2)		
Router Link: Link ID (Network) : NetA	Link Data (Netmask): MaskA	- Link to a stub network
Router Link: Link ID (Neighbor RouterID): R2	Link Data (Router If. Addr.): IP1	- Point-to-point link to another router

Router LSA (Link State ID: R2, Adv Router: R2, #links: 2)		
Router Link: Link ID (Network): NetA	Link Data: MaskA (Netmask)	- Link to a stub network
Router Link: Link ID (Neighbor RouterID): R1	Link Data: IP2 (Router If. Addr.)	- Point-to-point link to another router

Link State in OSPF (3)



Link State Database:

Router LSA (Link State ID: R1, Adv Router: R1, #links: 1)
Router Link: Link ID (DR IP Address): IP3 Link Data (Router If. Addr.): IP1 - Link to a transit network

Router LSA (Link State ID: R2, Adv Router: R2, #links: 1)
Router Link: Link ID (DR IP Address): IP3 Link Data (Router If. Addr.): IP2 - Link to a transit network

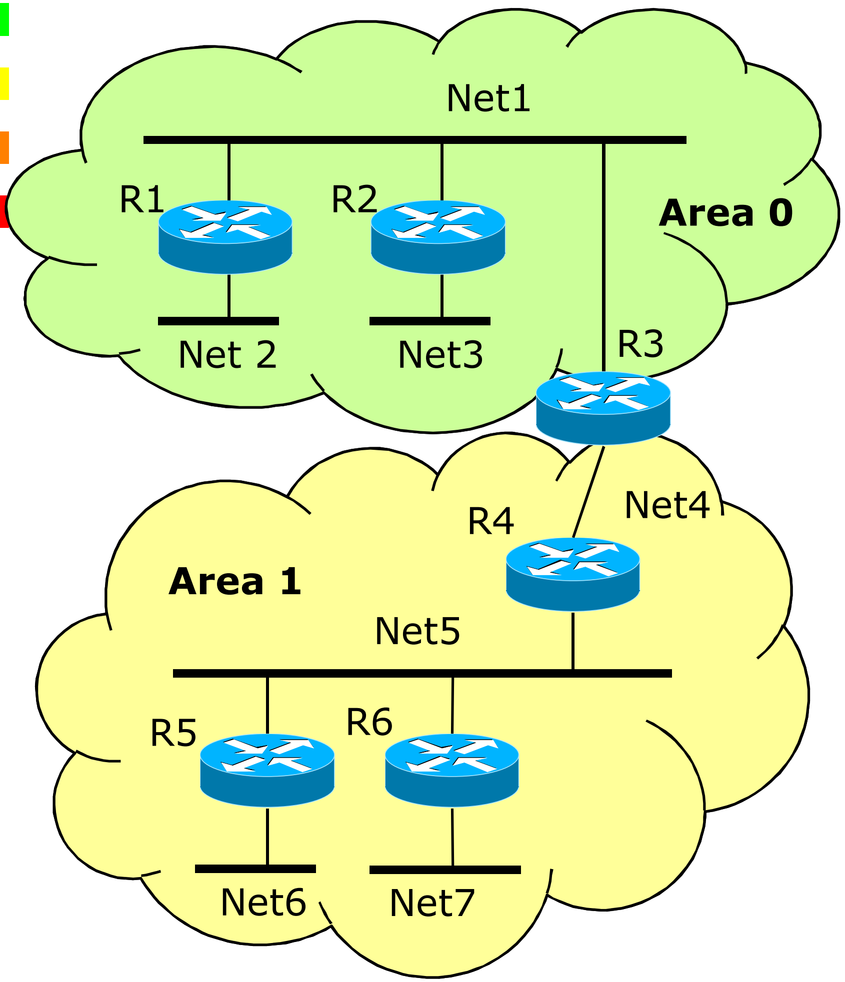
Router LSA (Link State ID: R3, Adv Router: R3, #links: 1)
Router Link: Link ID (DR IP Address): IP3 Link Data (Router If. Addr.): IP3 - Link to a transit network

Network LSA (Link State ID: IP3, Adv Router: R3)
Netmask: MaskA Attached Routers: R1, R2, R3

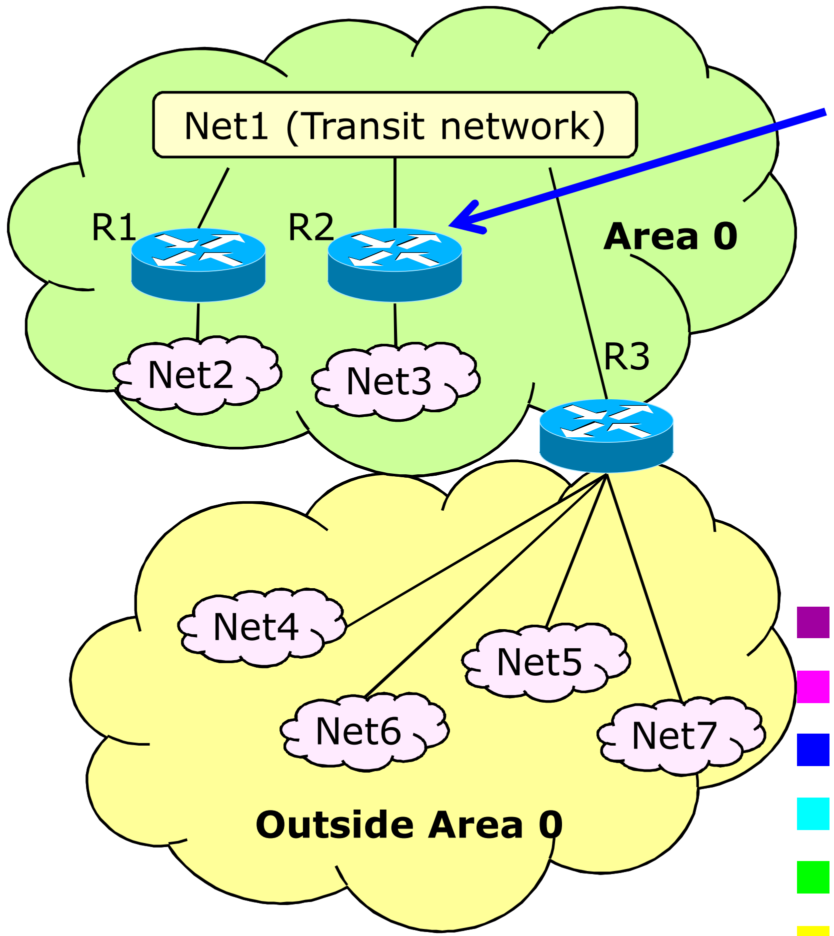
OSPF: LS topology and hierarchy

- Each LS router knows perfectly the topology of the area it belongs to
- With respect to other areas:
  - The precise topology is unknown
  - The router *can* know the list of destinations reachable outside its area
    - Some destinations may be summarized (e.g., default route for stub areas)
- A router that belongs to many areas has many OSPF databases, one per each area
  - Of course, this will originate a *single* routing table
- Some examples of the view from a router are the next slides

View of the network from a single router (1)

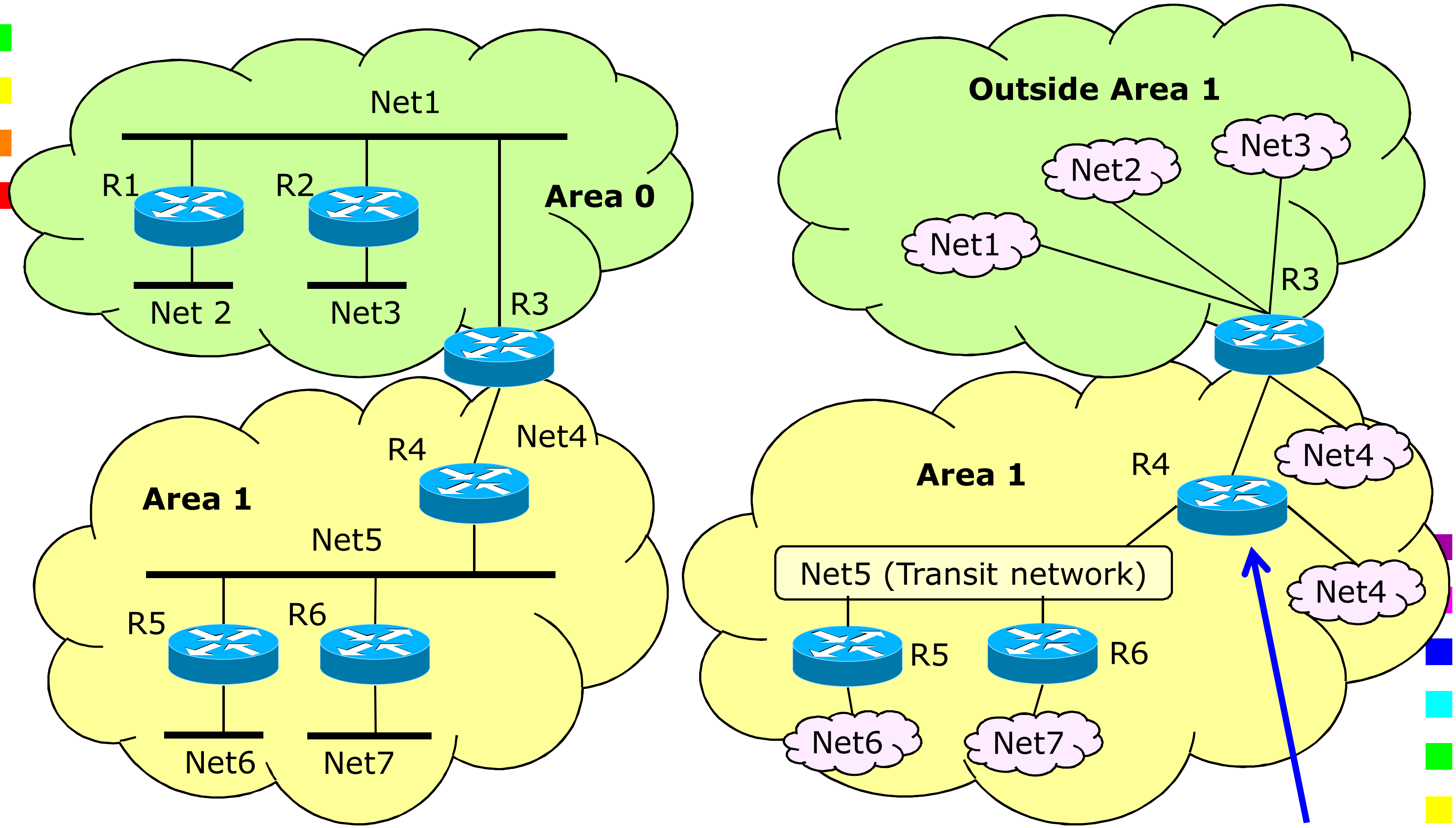


Complete network topology



View from R2

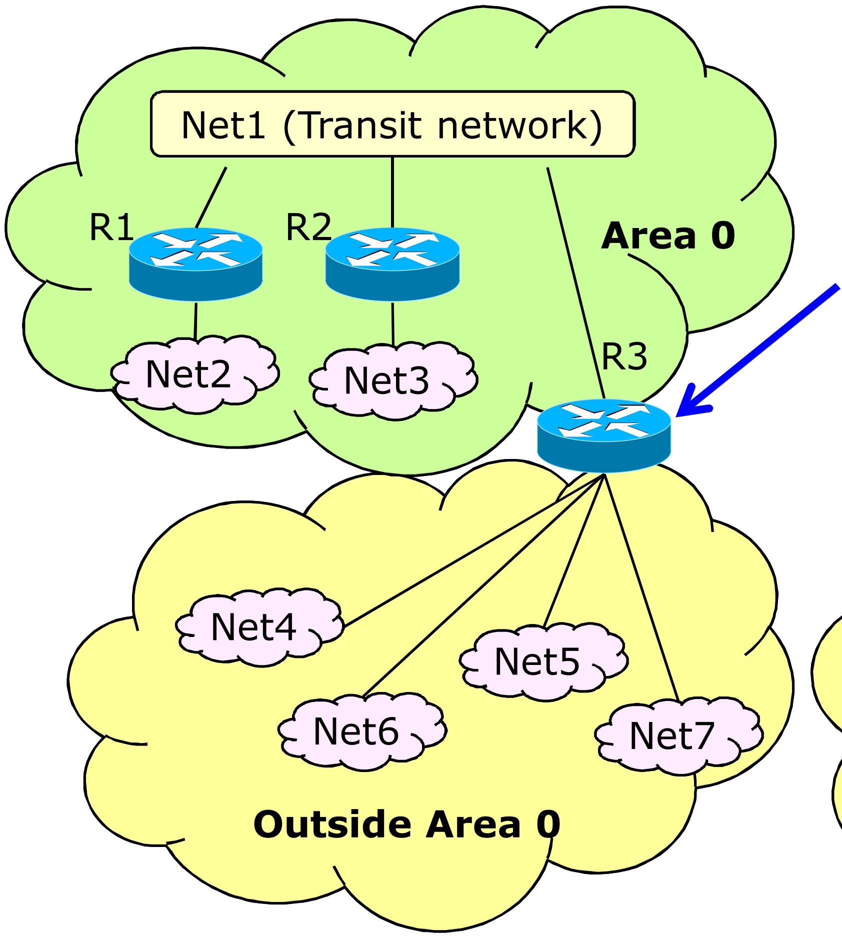
View of the network from a single router (2)



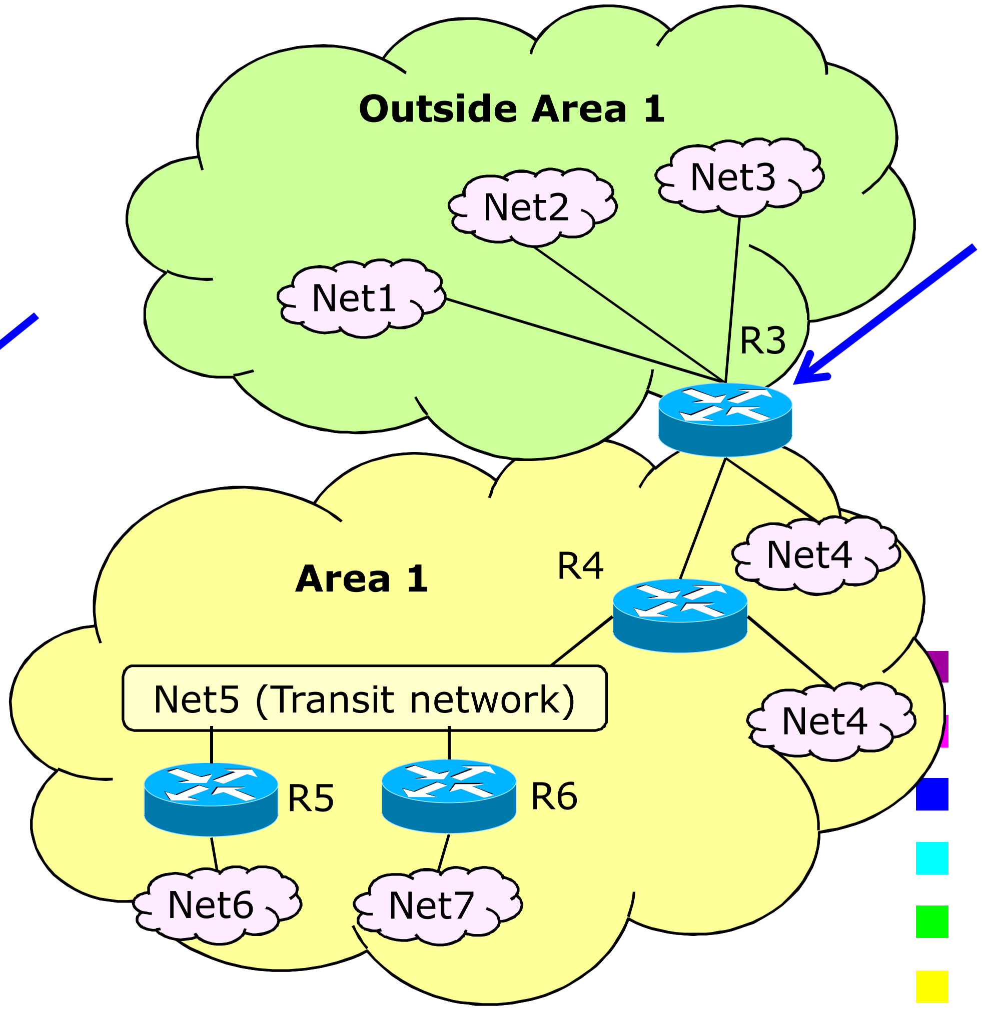
Complete network topology

View from R4

View of the network from a single router (3)



View from R3 (in area 0)



View from R3 (in area 1)

OSPF packet format

- It is encapsulated directly in IP (protocol Type = 89)
- All fields are fixed length
- There is no room for keeping future extensions
  - Executable code (and parsing) is more compact
  - Packets processing is faster
  - More rigid to evolution
- It is in fact split in 3 sub-protocols
  - Hello
  - Exchange
  - Flooding
- All packets begin with the same header

Packet format: header (1)

<table>
<tr><th>0</th><th>8</th><th>16</th><th>31</th></tr>
<tr><td>Version</td><td>Type</td><td colspan="2">Packet Lenght</td></tr>
<tr><td colspan="4">Router ID</td></tr>
<tr><td colspan="4">Area ID</td></tr>
<tr><td colspan="2">Checksum</td><td colspan="2">Authentication Type</td></tr>
<tr><td colspan="4">Authentication</td></tr>
</table>

Packets format: header (2)

- Version: nowadays 2
- Type: may have the following values:
  - Hello, Database Description, Link State Request, Link State Update, Link State Acknowledgement
  - Represents the type of the transported OSPF packet
- RouterID: Router ID of the device that *propagates* the message
  - E.g., in case of flooded packets, the RouterID represents the router that propagates the message, not the one that actually *generated* the LSA

Packet format: header (3)

- AreaID: number that identifies uniquely the area inside the OSPF domain
  - An IP address is often chosen
  - The value 0 identifies the backbone
- Authentication Type: two possible options exist:
  - No Authentication
  - Simple Password

Hello packets (1)

- Used to:
  - Test whether the link is fully active, thus to build the table of adjacencies
    - In order for a link to be active, bidirectional connectivity must be verified
    - By-product: OSPF does not work on unidirectional (e.g., some satellite) links
  - Elect the Designated Router and the Backup DR on the LAN
- Hello packets are only transmitted to neighboring nodes and never propagated
- Header OSPF: Type = 1

Hello packets (2)

<table>
<tr><th>0</th><th>8</th><th>16</th><th>31</th></tr>
<tr><td colspan="4">Network Mask</td></tr>
<tr><td colspan="2">Hello Interval</td><td>Options</td><td>Priority</td></tr>
<tr><td colspan="4">Dead Interval</td></tr>
<tr><td colspan="4">Designated Router</td></tr>
<tr><td colspan="4">Backup Designated Router</td></tr>
<tr><td colspan="4">Neighbor1</td></tr>
<tr><td colspan="4">. . . . . . .</td></tr>
<tr><td colspan="4">NeighborN</td></tr>
</table>

Hello packets (3)

- Network Mask: the netmask associated with the interface that generated the Hello
- Hello Interval: time between two consecutive Hello packets
- Options: Only the two last bits are defined
  - E: if the router is able to send and receive external routes; it is equal to 0 of the interface is part of a stub area
    - If two routers have different values for this bit, the adjacency cannot be established
  - T: if the router is able to handle packets with different TOS
- Priority: used for the election of the DR
  - A router without the Priority bit cannot become DR
    - The best router with the Priority bit is elected DR
  - Set (manually) by the network manager

Hello packets (4)

- DeadInterval: maximum temporal validity of the Hello packet
  - If no Hello packets are received from a neighbor passed this time, the router is considered dead
- DR, BDR: address of the Designated Router – Backup DR
  - Set to 0 if the election process has not been completed (or no need to elect any DR/BDR)
- Neighbor: list of RouterID that indicate the routers that are directly reachable from that network interface
  - I.e., the routers whose Hello packets have been received on that interface over the last DeadInterval seconds
- The link between two routers is declared operant if
  - Packets can flow in both directions
  - Both routers have the same value of the bit E

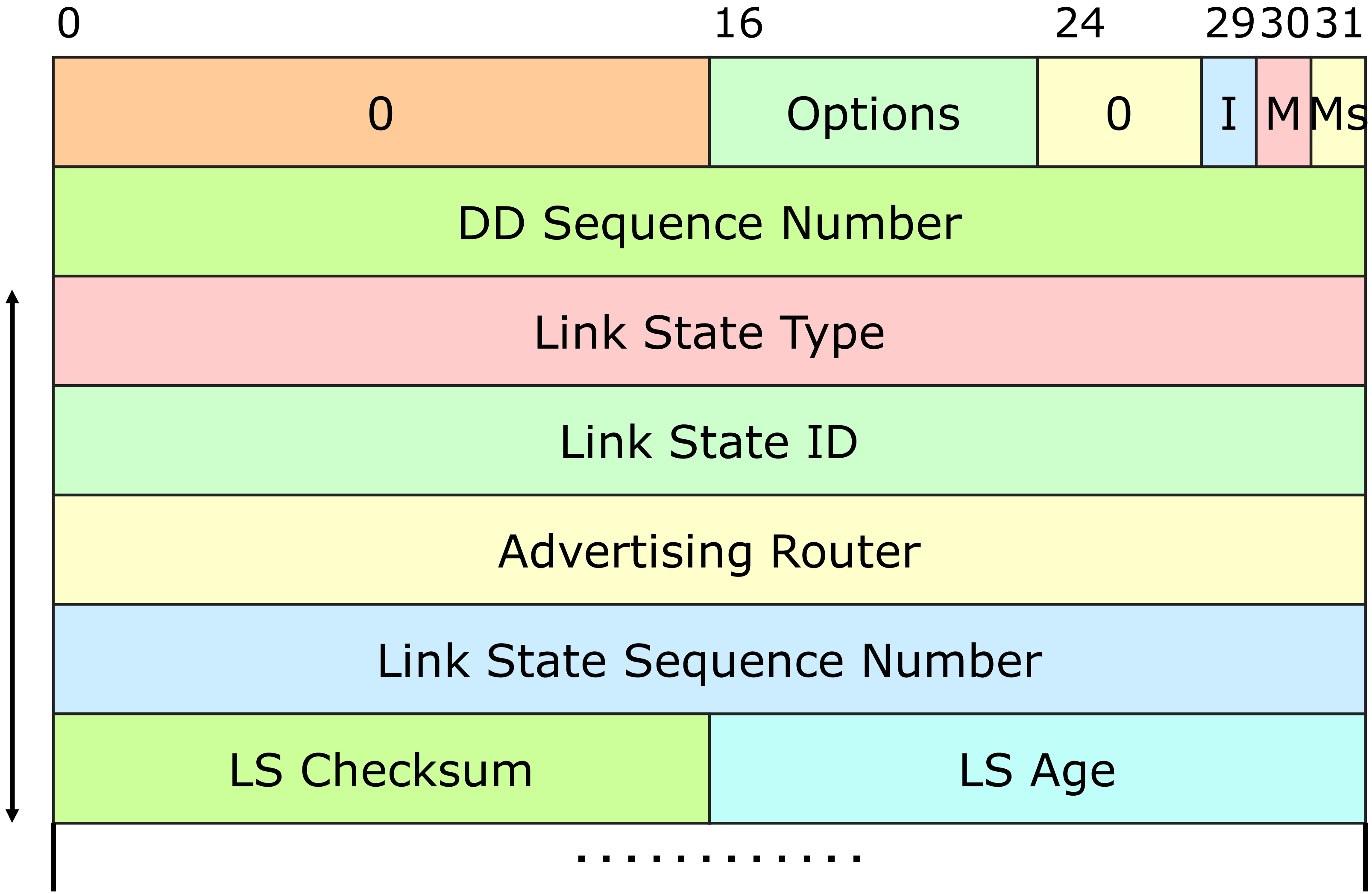
Exchange protocol (1)

- Requires four types of packets
  - Database Description (Header OSPF: Type = 2)
  - Link State Request (Header OSPF: Type = 3)
  - Link State Update (Header OSPF: Type = 4)
  - Link State Acknowledgement (Header OSPF: Type = 5)
- Used to synchronize the database of two routers when they become adjacent
  - For example at boot time or when a new link becomes active
- It is an asymmetrical protocol
  - The first step consists in the definition of which router becomes the Master and which one acts as Slave

Exchange protocol (2)

- Phases of the Exchange protocol
  - Definition of the Master/Slave roles
  - Exchange of a summary of all the LSA available on the routers on both sides of the connection
  - Request to the other router of old/missing LSA
  - Transmission of the requested LSA
    - The complete LSA is transferred only at this step (if needed)
- The normal sending of the LSA (outside the exchange protocol) looks definitely similar to how this protocol operates
  - In this case, only two types of packets are used
    - Link State Update
    - Link State Acknowledgement

Database Description Packets (1)



Database Description Packets (2)

- Options: as in the Hello packet
  - E: if the router is able to send and receive external routes; it is equal to 0 if the interface is part of a stub area
  - T: if the router is able to handle packets with different TOS
- I: Initialize
- M: More
- MS: Master - Slave (1= Master)
- DD SN: sequence number in the Database Description packet
- The remaining fields (that can be repeated) are the description of the header of a LSA and thus have the same meaning already presented for the LSA

Exchange Protocol: initialization (1)

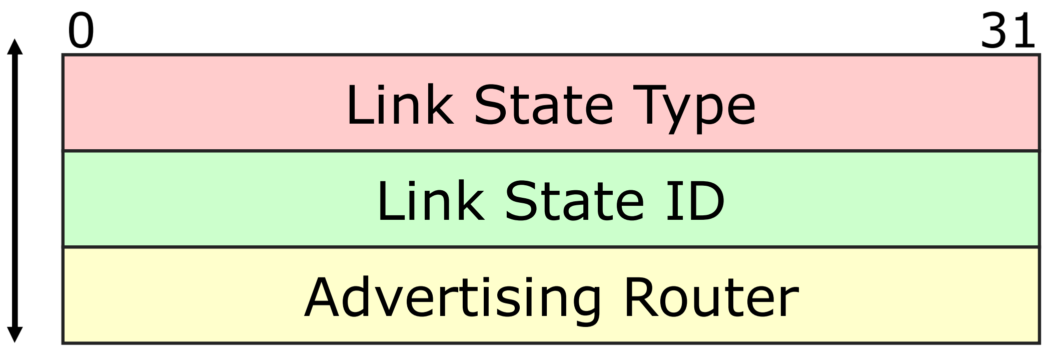
- The router that wants to initialize the procedure
  - Emits an empty DD packet with I, M, and MS set.
- The other router answers
  - Emits a DD packet of "acknowledgment" with I, M
  - If its RouterID is better than the one received, it sets the MS bit as well, and it becomes the Master
  - If not, it accepts the other as Master and it keeps the MS bit to 0
- The first router begins sending the summaries of its LSA
  - DD packets with M, MS, till last packets that has only MS
- The slave answers
  - Emits a DD packet of "acknowledgment" with M, reporting its "own" description of the database
- If the master does not receive the Ack within a given timeout, it resends the original DD packet

Exchange Protocol: initialization (2)

- If vice-versa the slave has not finished to transmit its descriptions
  - In correspondence to the packet of the Sender with M=0, it emits a DD with M=1
  - The master continues to send empty packets with M = 0, and to accept the Acks that come from the slave
  - The procedure of synchronization ends when also the slave sends a packet with M = 0
- During the exchange
  - Both the master and the slave checks that they have the LSA sent by the counterpart and that this LSA is not older than the one received
  - If this is not verified the LSA is inserted in the list of the LSA that needs to be fully exchanged

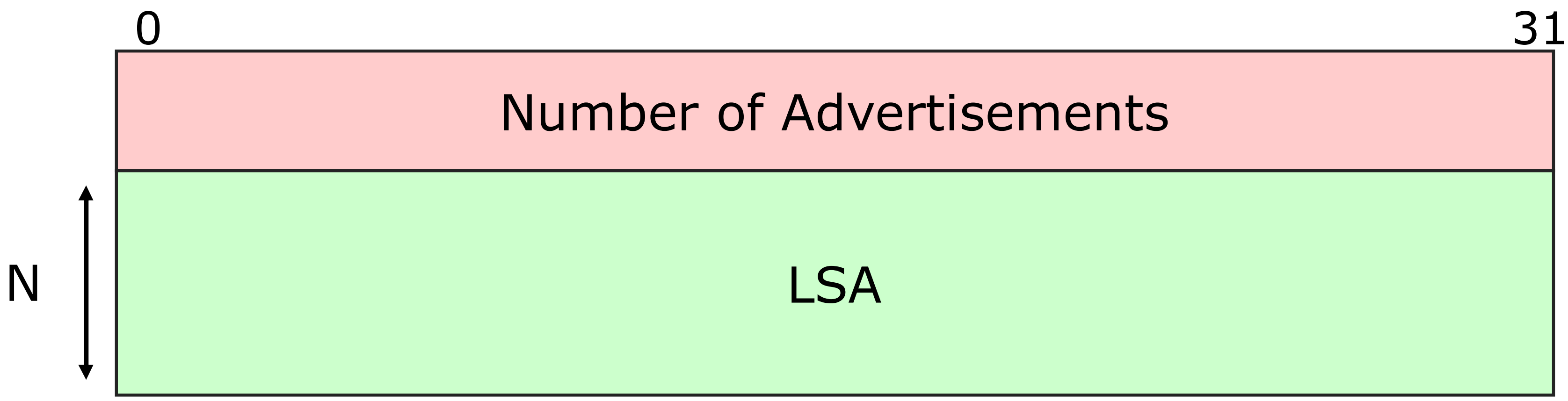
Link State Request packets

- They are sent at the end of the DD if there are some LSA that need to be synchronized
  - They request the other router to send the complete LSA corresponding to the field Link_State_Type, Link_State_ID and Advertising_Router indicated
- More requests can be grouped together
  - The three field may be repeated more than once in an OSPF packet
- Requested LSA are sent using the flooding protocol
  - In this way it is possible to update the whole network with the missing information



Link State Update

- Number of Advertisement: the number of LSA that are transported in the current packet
  - A single OSPF can contain multiple LSA, even of different types



LSA Header (1)

- All LSA have the same header
- Advertising Router: RouterID of the router that *generated* the LSA
  - Not the one that *propagated* the LSA

0	8	16	31
LS Age		Options	LS Type
Link_State_ID			
Advertising Router			
LS Sequence Number			
LS Checksum		Length	

LSA Header (2)

- Age: age of the current LSA (in seconds)
- Options
  - E: External Link
  - T: set when the router supports the Routing TOS
- Type: type of transported LSA (Router LSA, Network LSA, etc.)
- LinkID: identifier whose meaning depends on the type of LSA
  - Chosen by the Advertising Router, but the exact meaning may change according to the Type
  - The combination of RouterID, LinkStateID and LSA Type identifies univocally a LSA

OSPF LSA types

- LSA 1 – O, **Router LSA**
  - Contains all Link IDs – network, generated by every router and is local to the area
- LSA 2 – O, **Network LSA**
  - Contains all routers attached to the segment, generated by DR and is local to the area
- LSA 3 – O IA, **Network Summary LSA**
  - Describes a network from another area, generated by ABR and is propagated between areas
- LSA 4 – O IA, **ASBR Summary LSA**
  - Generated by the ABR and is propagated between areas
- LSA 5 – O E1, O E2, **External LSA**
  - Generated by ASBR and is propagated between areas
- LSA 7 – O N1, O N2, **NSSA External LSA**
  - Generated by ASBR into NSSA area and is propagated into area 0 as E1 or E2

Link State Advertisement

- It represents the real data structure that transports the Link State (as presented in the theory)
- Each LSA begins with a common header
- The remaining part is specific for each LSA type

Link State Acknowledgement

- Message used to confirm the successful reception of an LSA (through the LS Update)
- Sent in unicast (unless we are on a broadcast network)
- Used both in the normal flooding process and in the Exchange phase

0	16 ... 31
Link State Type	
Link State ID	
Advertising Router	
Link State Sequence Number	
LS Checksum	LS Age

LSA Type 1: Router Link (1)

- It keeps the information about all the links connected to the router that is currently creating the LSA
  - The information includes all the adjacent routers and all connected LANs
  - The LSA includes N repeated sections, one for each “link” of the router
- Propagated only inside the current area
  - No matter if the area is a backbone or an edge area

LSA Type 1: Router Link (2)

<table>
<tr><th>0</th><th>6</th><th>7</th><th>8</th><th>16 … 31</th></tr>
<tr><td>0</td><td>E</td><td>B</td><td>0</td><td>Number of links</td></tr>
<tr><td colspan="5">Link ID</td></tr>
<tr><td colspan="5">Link Data</td></tr>
<tr><td colspan="3">Link Type</td><td># TOS</td><td>TOS 0 metric</td></tr>
<tr><td colspan="3">TOS = x</td><td>0</td><td>TOS x metric</td></tr>
<tr><td colspan="5">. . . . .</td></tr>
<tr><td colspan="3">TOS = z</td><td>0</td><td>TOS z metric</td></tr>
</table>

LSA Type 1: Router Link (3)

- E: set if the router is an Area Border Router (External)
- B: set if the router is an AS Boundary Router (Border)
- Link Type, LinkStateID, LinkData: presented in Slide 30
- A specific metric for all the TOS that are used in that network is stored
  - This will originate N Shortest Path Trees, with the obvious consequences in terms of memory occupation and CPU consumption
  - Only the TOS whose metric is different from the value associated to TOS 0 (which represents the default 0) is explicitly stored in the packet
  - TOS=0 always exist
  - Other TOS are assumed to be equal to the cost of TOS=0

LSA Type 2: Network Link (1)

- Type = 2
- Generated by the Designated Router on behalf of the transit network
  - Lists all routers present on the LAN

0 … 31
Network Mask
Attached Router 1
. . . .
Attached Router N

LSA Type 2: Network Link (3)

- Network Mask: netmask of the transit network
- Attached Router: IP address of all routers that it is interfaced with in the LAN (or rather, all the routers that have created an adjacency with the given router)
  - It is not mandatory to indicate the number of adjacent routers because the length of the message is given by the field LENGTH of the header
- LinkStateID (in the LSA header): IP address of the interface of the Designated Router connected to the transit network

LSA Type 3-4: Summary Links (1)

- Generated by the ABRs and aim at providing summaries for IP networks and ASBRs
- Both deal with summarized information
- Each LSA includes only one destination (differently from LSA Type 1 - Router Links)
  - Each ABR generates a lot of them, inserted in the same packet

LSA Type 3-4: Summary Links (2)

- Summary Links for IP Networks (Type = 3)
  - Propagates the distance between the ABR and the IP networks present in the AS
  - Generated only by ABRs
  - An ABR generates different LSA for each area it belongs to
    - Obviously, summaries are different according to the area we are looking at
- Summary Links for Border Routers (Type = 4)
  - Propagate the distance between the ABR and the ASBR
  - Used to determine the best ABR for reaching external destinations (i.e. networks outside the AS)
  - Generated only by ABRs
  - An ABR generates exactly the same Type-4 LSA on all the areas it belongs to

LSA Type 3-4: Summary Links (3)

0	8	16	31
Network Mask			
TOS = 0	0	TOS 0 metric	
TOS = x	0	TOS x metric	
. . . .			
TOS = z	0	TOS y metric	

LSA Type 3-4: Summary Links (4)

- Network Mask
  - Type 3: netmask associated to the summarized IP network
  - Type 4: set to 0xFFFFFFFF
- TOS = x, TOS x metric: same meaning as the one of the Router Links
  - The field # TOS is not used because the length can be found using the LSA header
- LinkStateID (in the LSA header)
  - Type 3: the network address of the summarized network
  - Type 4: IP address of the Area Border Router

LSA Type 5: External Links (1)

- Type =5
- Keeps the cost to reach the external destinations, calculated from the ASBR
  - Generated by the AS Border Router
  - Propagated to all routers of the OSPF domain
- It includes one destination for each LSA (as for the Summary Links)

LSA Type 5: External Links (2)

0	8	16 … 31
Network Mask		
E, TOS = 0	0	TOS 0 metric
External Route Tag (0)		
E, TOS = x	0	TOS x metric
External Route Tag (x)		
. . . .		
E, TOS = z	0	TOS y metric
External Route Tag (z)		

LSA Type 5: External Links (3)

- Network Mask: netmask of the destination network
- E: bit 0 of the field TOS
  - Indicates if the external routes are learned by protocols such as BGP, which do not necessarily supplies a metric comparable to the one of OSPF
  - If it set, it indicates that the metric is not compatible with OSPF and thus the distance must be considered “bigger than any other internal route”
  - In the case in which the metric is compatible, it is possible to add it to all the costs inside the OSPF domain, hence getting the real cost to the destination

LSA Type 5: External Links (4)

- External Route Tag: used by the Border Router to exchange information related to this route
  - It is neither used nor analyzed by OSPF
- TOS = x, TOS x metric: same meaning as the one of the Router Links
  - The field # TOS is not used because the length can be found using the LSA header
- LinkStateID (in the LSA header): IP address of the destination network

OSPF: different types of areas

- Normal area
  - ABRs forward all LSAs from Area 0, including external LSAs
- Stub Area
  - Eliminates external routers/LSA present in area (Type 5)
  - External routes are replaced by a single Default Route

```
!
! Normal area
router ospf 100
!
!
```

```
!
! Stub area
router ospf 100
  area 10 stub
!
```

OSPF: different types of areas (Cisco extensions)

- Totally Stubby Area
  - Keeps only internal LSA (Type 1 and 2) and a single default route
  - Used for stable-scalable internetworks
  - Minimizes the number of LSAs and the need for any external area SFP calculation
- Not-so-stubby area (NSSA)
  - Can import AS external routes and send them to other areas, but still cannot receive AS-external routes from other areas
  - Allows the injection of external routes in a limited fashion into the stub area
- Cisco extension, although implemented by many vendors

```
!
! Totally Stubby area
router ospf 100
  area 20 stub no summary
!
```

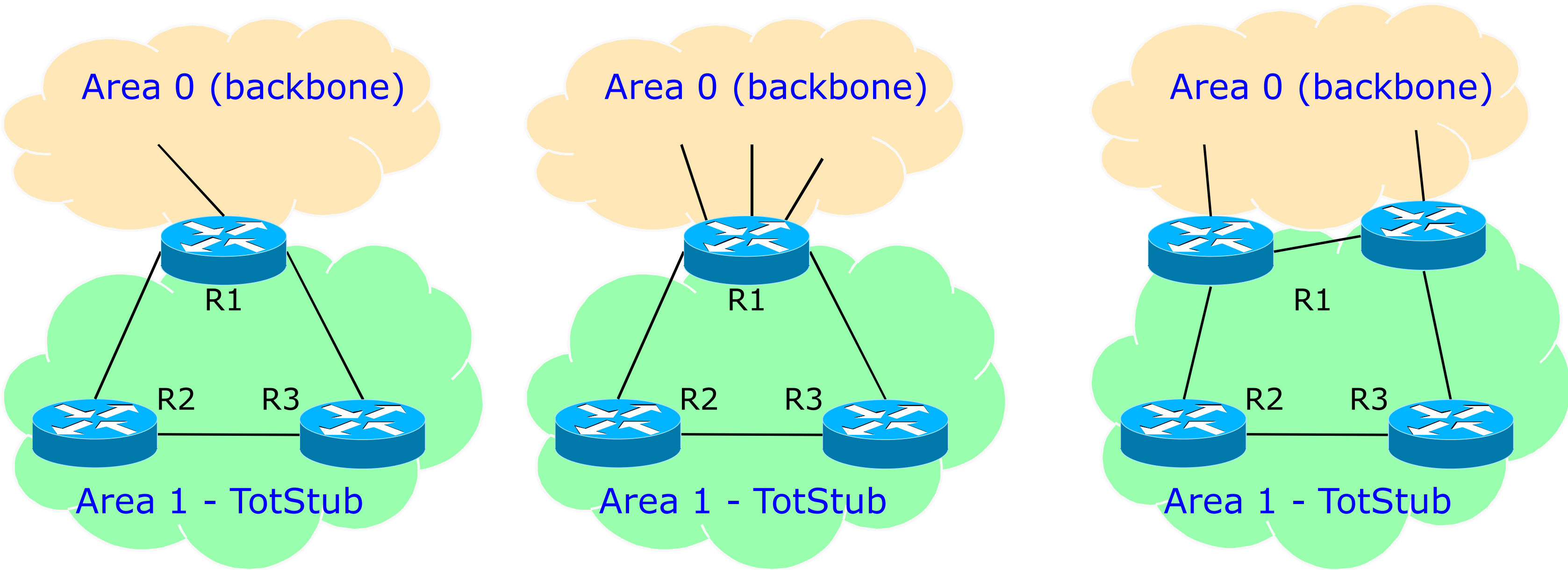
```
!
! Totally Stubby area
router ospf 100
  area 20 nssa
!
```

Stub areas

- It is often used when an area has only one ABR
  - In this case, we do not need to propagate external routes because there is only one path that connects to the rest of the network
    - In fact, external routes in non-stub areas are useful only if more than one Egress Router exists
- Stub areas are activated upon an explicit configuration coming from the network manager

Totally stubby areas: example

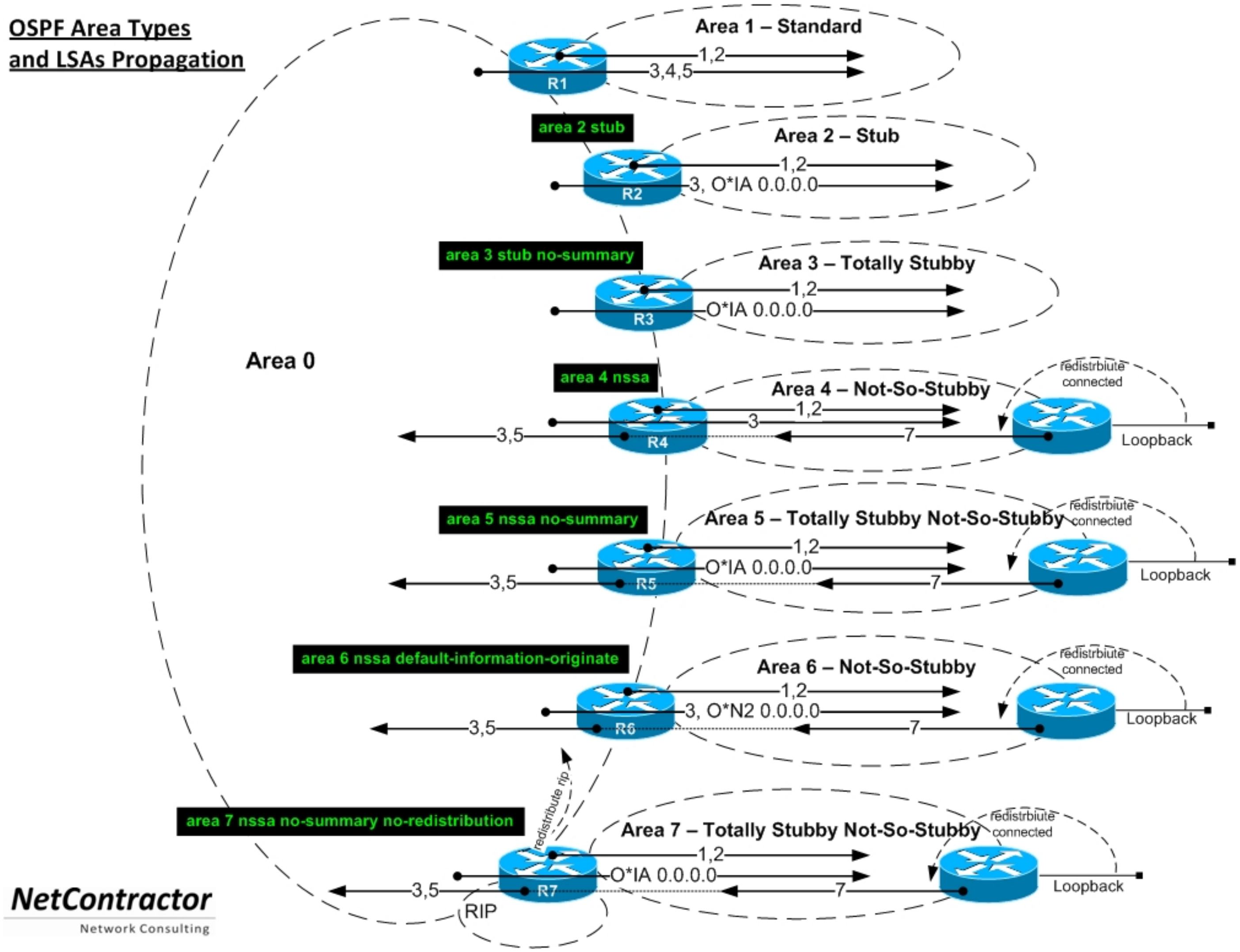
- OSPF does not prevent to have a *totally stubby area* with more than one Area Border Router
  - Internal routers reach the external destinations based on the best ABR



Stub/totally stubby areas: limitations

- A stub/totally stubby area can not be used to reach other domains
  - It is thus not possible to place an AS Boundary Router into a stub area
    - Let’s assume we have a Boundary Router into a Stub area
    - If the packets that enter the area are directed to another area inside the AS, the Default Route is used
      - Only internal routes are known inside the stub area
    - The packets following the Default Route exit through the Boundary Router, that is to say through where they come from
- It is forbidden to have Virtual Links through a stub/totally stubby area
  - More on Virtual Links later

Summary of LSA propagation



Conclusions

- Very powerful (and very complex) protocol
- Widely used in modern networks