Subnetting in OSPF

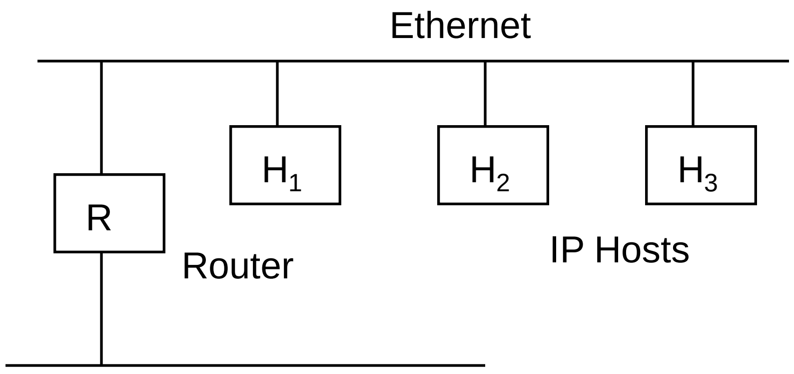
- Separating Host and Routers
  - link states records

  R-H1

  R-H2

  R-H3

  ⋮



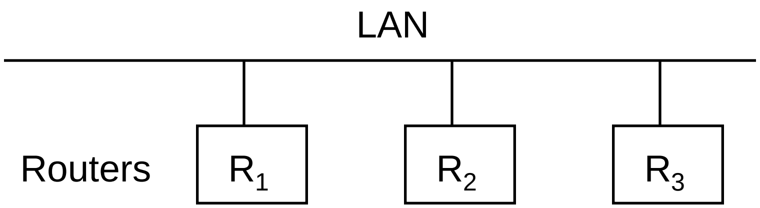
- Simplification:

  => subnet model
  - one link between the router and the subnet

    "link to a stub network"
  - identified by its subnet number

OSPF in Broadcast Networks

- Broadcast Networks
  - Full Connectivity
  - Broadcast Capability

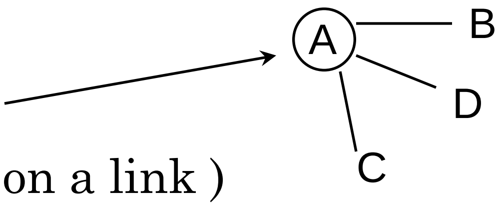


- # Adjacencies : $\frac{N(N-1)}{2}$

- Each router would advertise:

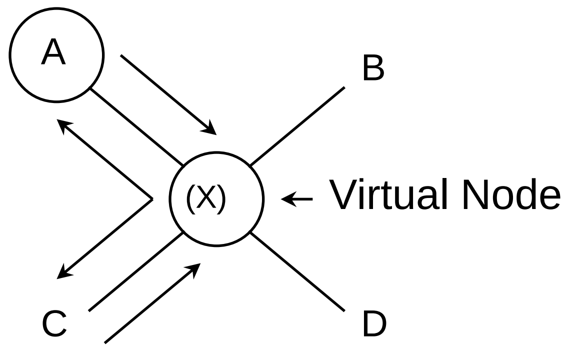
$$\left[\begin{array}{l}\text{N-1 links to other routers}\\ \text{one stub- network link}\end{array}\right\} \Rightarrow N^2$$

- Solution
  - One router is designated through election
  - Hello Protocol (used for discovering peers on a link)



Broadcast Networks (cont.)

- Reducing the number of link state records
- Database will include two links per router:
  - 1) link from router to Virtual node
    - advertised by router
  - 2) link from Virtual node to router
    - advertised by designated router [“network links”]
    - network links have a null metric



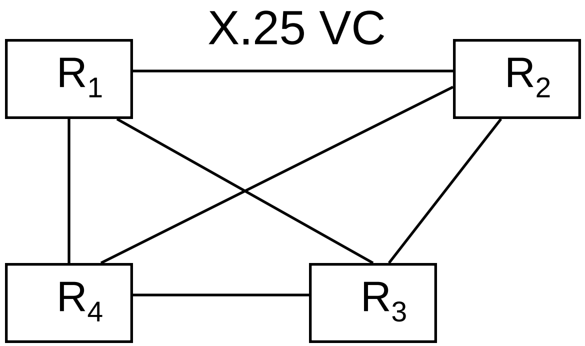
Flooding in Broadcast Networks

- Simplifying flooding
  - A router sends a link state advertisement to the designated router only using **224.0.0.6** “all-designated-routers” multicast address
  - If advertisement is new, designated router floods the link state on all its interfaces (including the network on which it received it) using **224.0.0.5** “all OSPF routers” multicast address

Backup Designated Router

- Designated routers forms adjacencies with all routers in the broadcast network
- Backup Designated Router also does that.
- Reliability of designated router => one back-up designated router
  - must be synchronized
  - listens to 224.0.0.6 (but remains silent in terms of flooding, etc..)
  - failures of Designated Router is discovered by means of Hello Protocol
  - smoother transition since the Backup Router already has formed all the adjacencies

OSPF Over Non-Broadcast Networks



- Non-broadcast Networks
- IP over X.25 Networks
  - popular in the 1980's in Europe
- IP over ATM
  - static configurations
  - avoid N(N-1)/2 overhead; designated router
  - on-demand circuits
  - permanent circuits are for links between Routers and the "Designated Router"

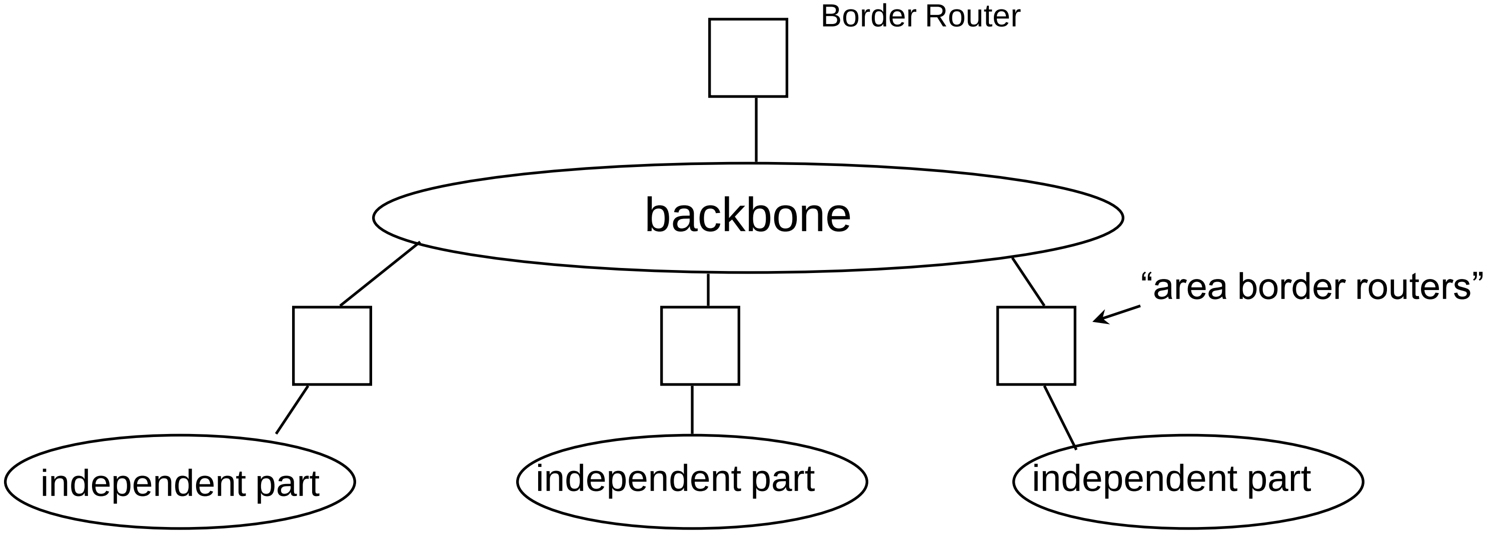
OSPF
Multiple Areas

- Hierarchical routing in order to decrease routing overhead (size of link state database; duration of route computation; volume of messages exchanged)
- Split the network into set of independent parts connected by a backbone
- Each area operates like an independent network
  - Database includes only the state of the area’s links
  - Flooding stops at the boundaries
  - routers compute routes within the area
- Cost of routing proportional to size of area

OSPF
Multiple Areas (cont.)

- Reduce the link state database
  - duration of computation
  - volume of messages

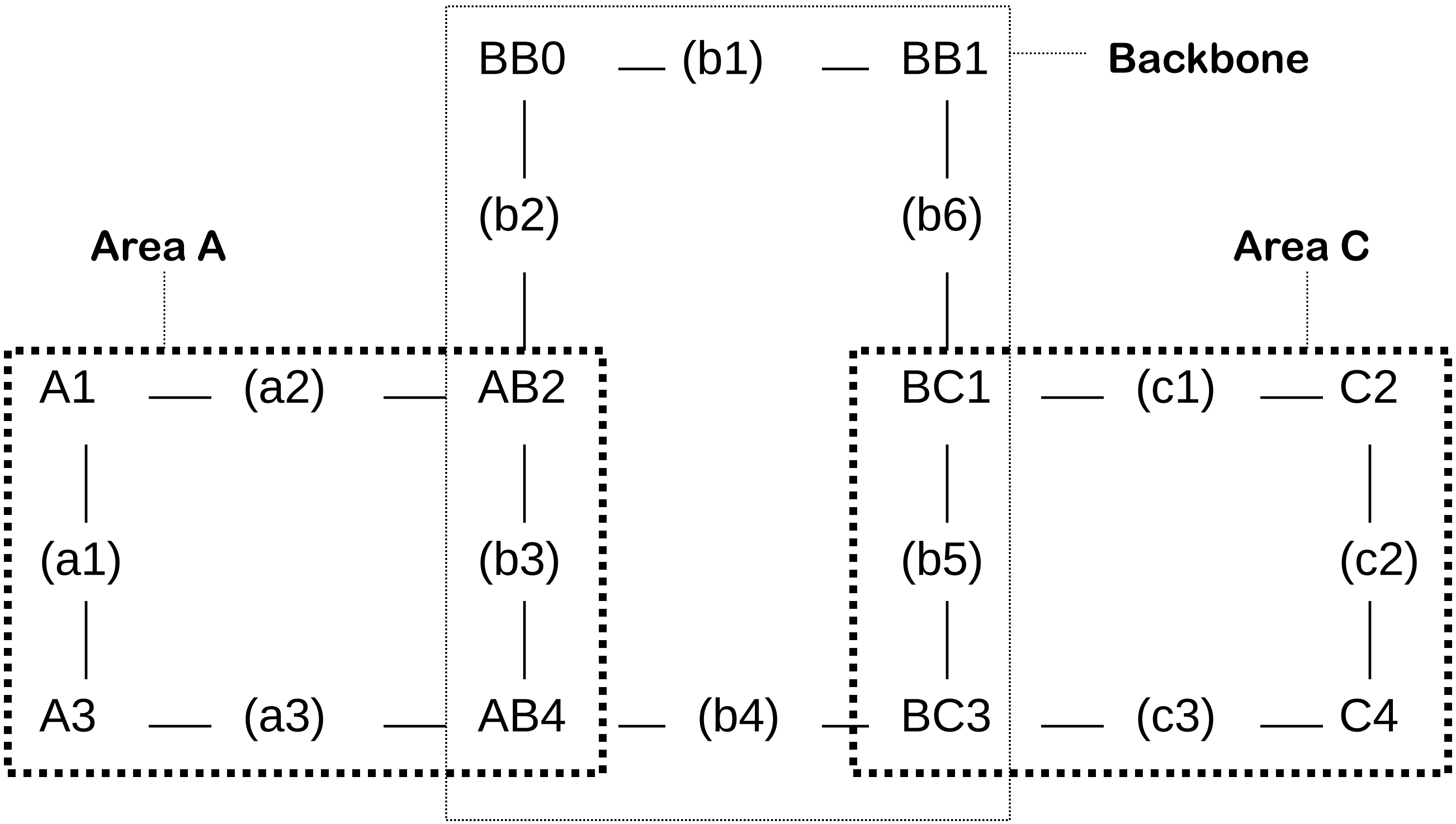
  => hierarchical Routing !



OSPF
Multiple Areas (cont.)

- How to glue the network together?
- Some routers belong to several areas
  - typically to one lower-level area and to the backbone area
  - called area-border routers
  - at least one area border router in each area
- Area border routers
  - maintain several link state database (one for each area to which they belong)
  - emit special link state records to signal reachability of networks in each area

OSPF
Multiple Areas Example



OSPF

Database Contents for the Example

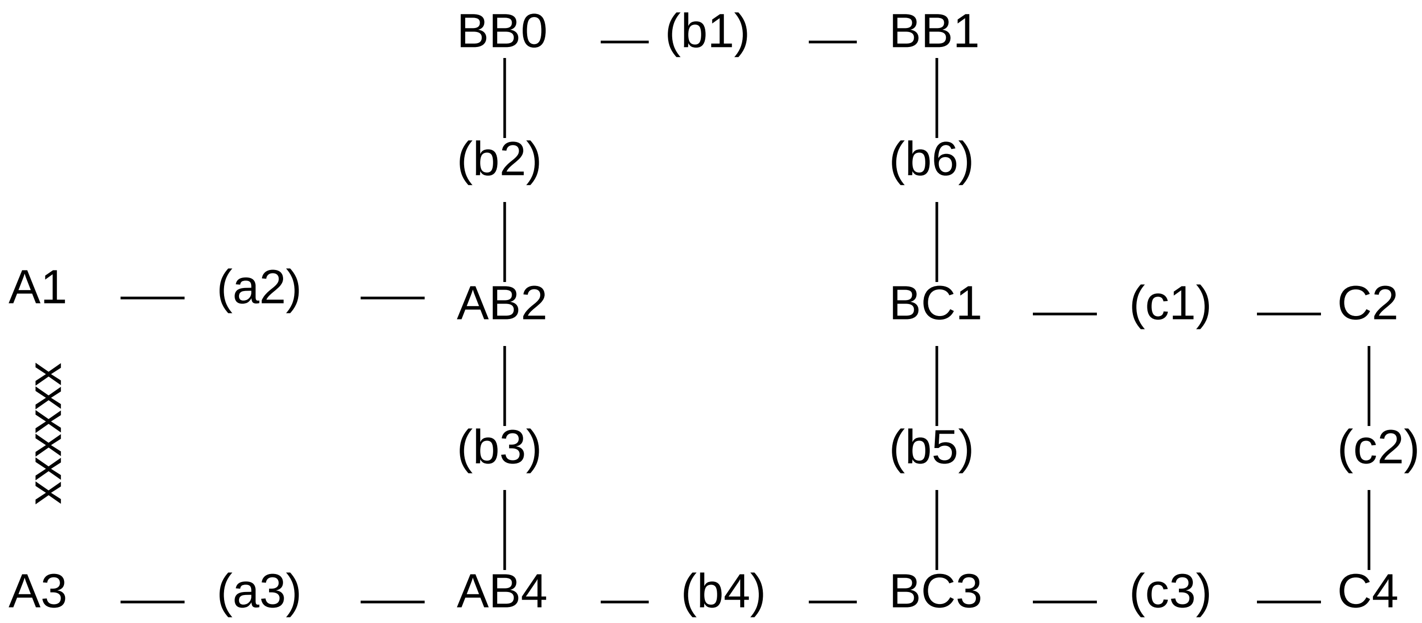
- Database for area A contains :
  - link state records for links a1 a2 a3 sent by A1, AB2, A3, AB3
  - Summary records emitted by AB2 and AB4 for all IP networks and subnetworks that are part of the backbone and of area C
  - External records emitted by BBO and BB1 and relayed by AB2 and AB4

OSPF
Multiple Areas Example

- Assume b1 is an Ethernet LAN with IP subnet number nb1
  - AB2 will build a summary link for nb1 with metric = metric (b1) + metric (b2)
  - AB4 will also build a summary link for nb1
  - metric = metric (b1) + metric (b2) + metric (b3)
- Assume C3 is also an Ethernet nC3
  - BC3 will advertise summary with metric = metric (C3)
  - BC1 will advertise summary with metric = metric (C1) + metric (C2) + metric (C3)
  - AB4 will advertise summary with metric = metric (b4) + metric (C3)
  - AB2 will advertise summary with metric = metric (b3) + metric (b4) + metric (C3)

OSPF
Example: failure inside an area

- Example of link failure:

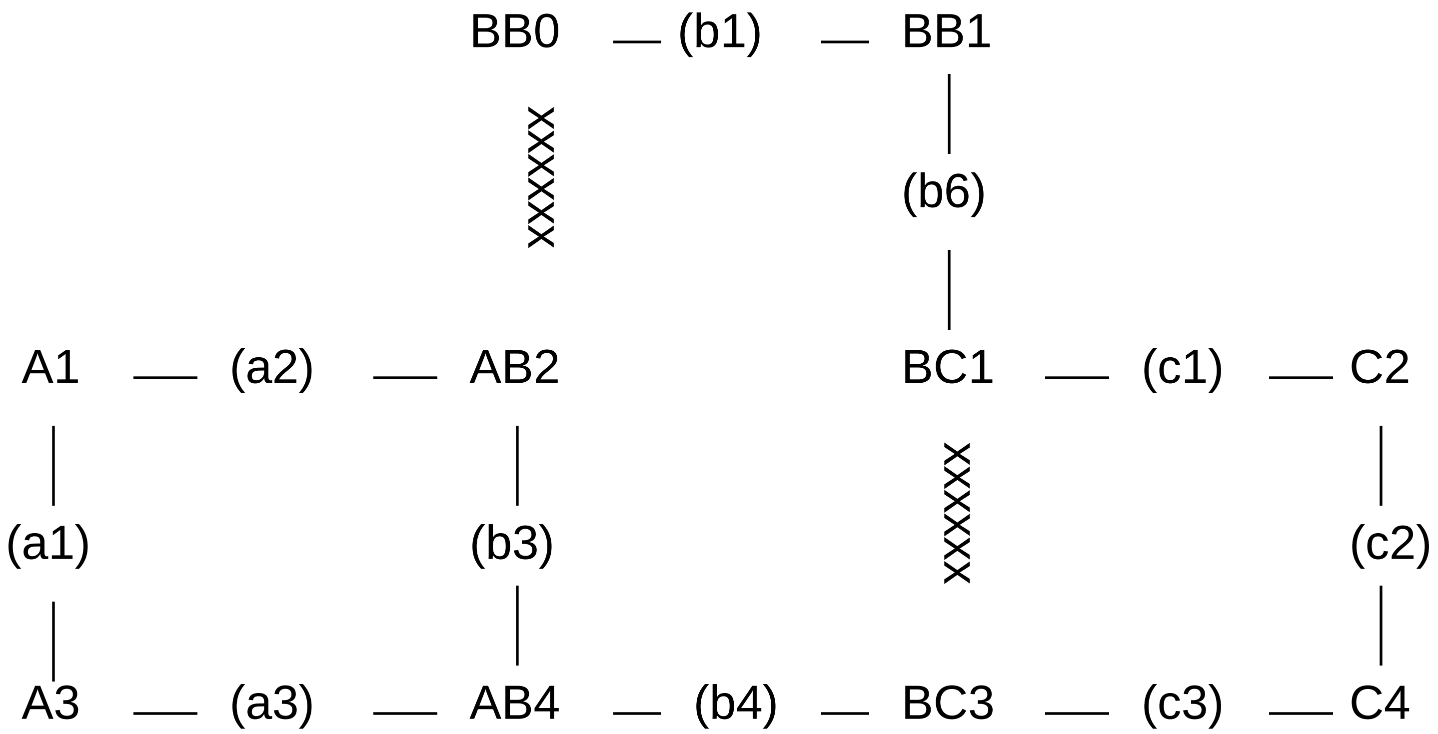


- The area-border routers will issue summary link state records for networks and subnetworks that they can reach

OSPF

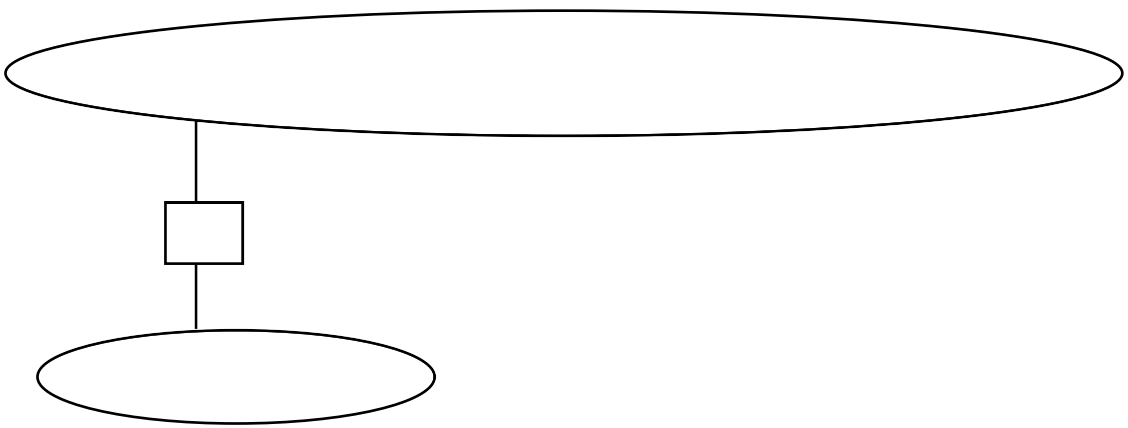
Example: Failure in the Backbone

- Link failure in the backbone area



OSPF
Stub Areas

- Areas where there is only one exit point, or the exit point is not a function of the external destination.
- Stub area does not need to know the topology of the rest of the AS - all external traffic goes to the exit point.
- Obviously, no AS boundary routers can be internal to stub areas.



OSPF Protocol Packets

- OSPF runs directly over IP, using protocol number 89.
- OSPF does not explicitly support fragmentation, but protocol messages can generally be split; this should be used, rather than IP fragmentation.
- OSPF packets are sent with an IP TOS of 0 (zero).
- OSPF packets are sent with IP Precedence set to Internetwork Control
- All OSPF packets use the same header
- OSPF Multicast Addresses (send with TTL=1):
  - 224.0.0.5: All OSPF Routers
  - 224.0.0.6: OSPF Designated and Backup Routers

The Common OSPF Header

<table>
<tr><td>VERSION #</td><td>TYPE</td><td colspan="2">PACKET LENGTH</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">AUTYPE</td></tr>
<tr><td colspan="4">AUTHENTICATION</td></tr>
<tr><td colspan="4">AUTHENTICATION</td></tr>
</table>

OSPF Header Fields

- Version #: Set to 2 to indicate the current version of OSPF
- Type: The OSPF packet type, as follows:
  - **1 Hello**
  - **2 Database Description**
  - **3 Link State Request**
  - **4 Link State Update**
  - **5 Link State Acknowledgement**
- Packet Length: The number of bytes in the packet, including the header
- Router ID: The IP address selected for identifying the router
- Area ID: The identification of the area. The value 0 is reserved for the backbone area. It is common practice to chose an IP network number for identifying an area.
- Checksum: Computed on the whole OSPF packet, excluding the 8-octet authentication field, using the classic IP algorithm

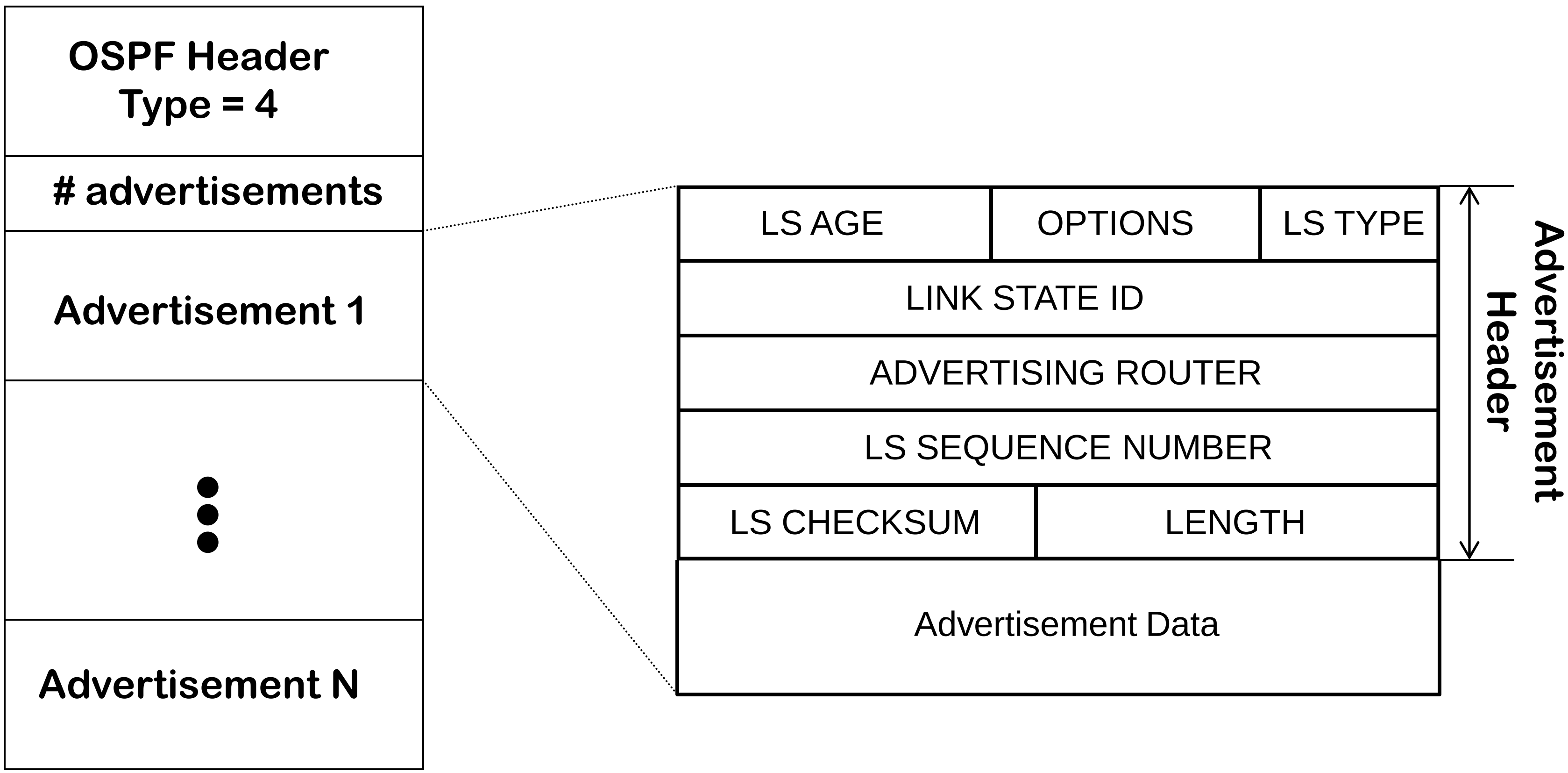
OSPF Header Fields (cont.)

- AUTYPE: Identifies the authentication algorithm. Only three values are defined in the standard itself:
  - 0: No authentication:
    - Exchanges not authenticated.
    - Authentication field ignored; can be set to anything
  - 1: simple authentication
    - “Clear password” type of authentication; all packets must contain the right value, pre-configured for that area.
    - Used to prevent unconfigured routers from joining in.
  - 2: cryptographic authentication
    - Secret key is used to generate a digest of the packet.
    - Digest is added at the end of the packet; size not included in the header
    - 64-bit field is restructured to contain digest size, key ID, and sequence number (to protect against replay attacks)

OSPF
Link State Database Records

- Made of link state records.
- Several Types of links state records:
  - 1. Router link
  - 2. Network link
  - 3. Summary link (IP network)
  - 4. Summary link (to a border router)
  - 5. External link

OSPF
Link State Advertisements

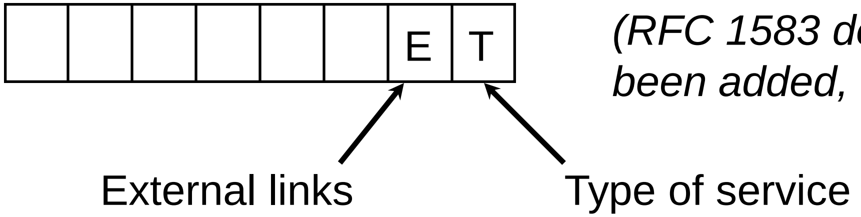


OSPF
Advertisement Header Fields

- LS type :
  - Type of link state record (1,2,3,4,5)
- Link State ID:
  - Chosen by the advertising router
  - Generally an IP address. Interpretation varies with type of link.
  - Note: < adv. router ID; link ID; LS type > uniquely identifies a record.
- Sequence Number:
  - Identifies one particular advertisement.
- Checksum:
  - Protects header as well as content
- Length:
  - Total length of the record (including the 20-byte-header)

OSPF
Advertisement Header Fields (cont.)

- Advertising Router:
  - For each router, one of its IP addresses is selected as the "OSPF identifier".
- Age:
  - 16-bit unsigned integer indicating the time in seconds since the link state record was first advertised.
- Options:



(RFC 1583 definitions; new bits have been added, and the T bit removed)

  - E: used in Hello Protocol (see later)
  - T: Set when router supports nonzero TOS. (0 indicates type 0 metric)
    - (removed in the latest version of the standard)

OSPF
Link State Record

- Router Links State Record
  - Summarizes all the links that start from the advertising router
  - Follows the Link State Advertisement Header, with LS Type =1

--0---VEB	---0----	NUMBER OF LINKS
LINK ID		
LINK DATA		
TYPE	# TOS	TOS 0 METRIC
TOS=X	0	TOS X METRIC
TOS=Y	0	TOS Y METRIC
- - -	- - -	- - -
TOS=Z	0	TOS Z METRIC

Repeated multiple times, one for each link

OSPF
Link State Record Fields

EB: E (external) AS boundary router
B (border) Area Border router

Link ID, link data and link type depend on type.

Type = 1 link is a point-to-point link to another router
link ID = OSPF identifier for router
link data = router’s interface IP address.

Type = 2 link connects router to a transit network
link ID = IP address of the designated router’s interface
link data = router’s interface IP address

Type = 3 link connects to a stub network
link ID = IP network or subnet number
link data = network or subnet mask

OSPF
Network Links Advertisement

- Network Links:
  - Follows the Link State Advertisement Header, with LS Type = 2
  - Advertised by designated routers for transit networks (multi-access)
  - Lists all routers attached to that network

NETWORK MASK
ATTACHED ROUTER
ATTACHED ROUTER
- - -
ATTACHED ROUTER

← *OSPF identifier of attached router (router that has built up an adjacency with designated router)*

OSPF
Summary Links

- Summary Links (generated by area border routers):
  - Type 3 (for IP networks)
  - Type 4 (for AS boundary routers)
  - Advertised by area-border routers (issues a separate advertisement for each destination)
- Link state ID
  - Type 3: IP network or subnet number
  - Type 4: IP address of border router

NETWORK MASK		
TOS=0	0	TOS 0 METRIC
TOS=X	0	TOS X METRIC
⋮		

OSPF
External Links

- External Links
  - LS Type 5
  - Advertised by border routers (one destination advertisement per second)
- Link state ID = IP network or subnet of the destination

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 Z METRIC
EXTERNAL ROUTE TAG (Z)		

External metric may differ from "internal" metric, and thus may not be directly comparable to internal metric.
(Hence, E tag added & set)
If E is set, metric should be considered larger than any internal value.
If E = 0, metric may be added to cost of internal path.
External route tag: used by border routers to exchange information transparent to OSPF.

OSPF
The Hello Packet

- Follows standard OSPF header, type = 1
- The Hello protocol is used for two purposes:
  – 1. To check that links are operational.
  – 2. To elect the designated router and the backup on broadcast and non-broadcast networks.

OSPF PACKET HEADER, TYPE = 1 (HELLO)		
NETWORK MASK		
HELLO INTERVAL	OPTIONS	PRIORITY
DEAD INTERVAL		
DESIGNATED ROUTER		
BACKUP DESIGNATED ROUTER		
NEIGHBOR		
....		
NEIGHBOR		

All the fields in this format are 32 bits, except for the hello interval (16 bits), the option field (8 bits), and the priority (8 bits).

OSPF
Hello Packet Fields

- Network mask:
  - The subnet mask associated with the interface. In the absence of subnetting, it will be set to FF000000 for a class A network, FFFF0000 for class B, FFFFFF00 for class C.
- Hello interval:
  - The hello packets are sent by the router every “hello-interval” seconds.
- Options field:
  - Same definition as the Options in the Advertisement Header

						E	T

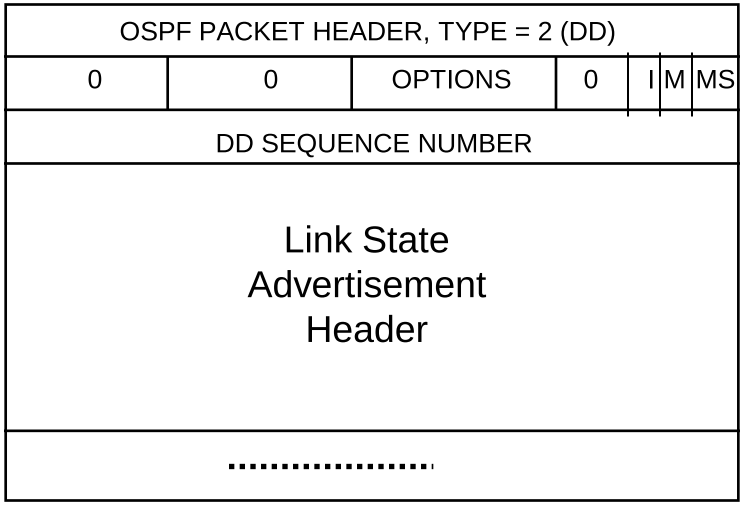
T bit is set when the router is capable of TOS routing.
E bit is set when the router is capable of receiving and sending external routes; it is null when the interface belongs to a stub area.

OSPF
The Hello Protocol

- The link between two routers is declared operational if
  - 1. packets can flow in both directions;
  - 2. both routers agree on the state of the E bit.
- Bi-directional connectivity is easily checked by looking at the "neighbor" list of remote routers.
- Priority:
  - On network interfaces, one must first select the designated router and the backup.
  - The selection procedure uses the "priority" field carried in the hello packets. Each router is configured with a priority, varying between 0 and 255.

OSPF
The Database Description Packet

- The Exchange Protocol
  - When two routers have established two-way connectivity on a point-to-point link, they must “synchronize” their databases
  - The initial synchronization is performed through the “exchange” protocol; the “flooding” protocol will then be used to maintain the two databases synchronized.



OSPF
Database Description Fields

- Option octet:
  - Same as in the Hello packets
- I, M, MS:
  - I: initialize; M: more; MS: master-slave
- Sequence number:
  - The sequence number of the database description packet.
- The content of the packet is set of link state record descriptions headers, without the actual data
  - It just uniquely identifies the records

OSPF
Link State Request Packet

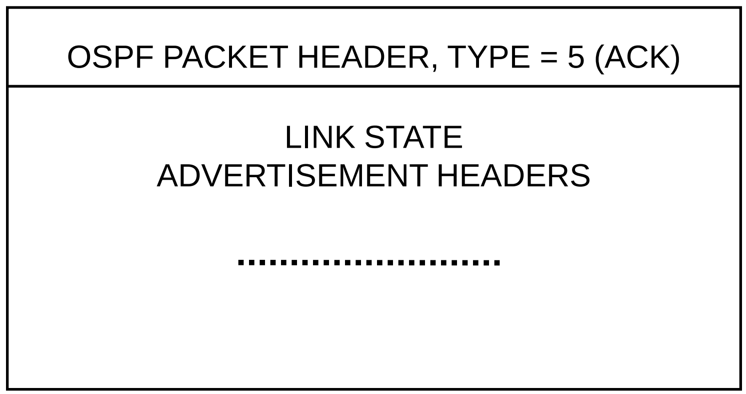
- Link state request packet

OSPF PACKET HEADER, TYPE = 3 (RQ)
LINK STATE TYPE
LINK STATE ID
ADVERTISING ROUTER
....

- These packets contain a set of link state record identifiers. Each record is described by three 32-bit words:
  - the record type,
  - the record identifier
  - the advertising router identifier.
- Upon reception of such request, a router will send a set of link state updates, using exactly the same procedures as for flooding new record values.

OSPF
Link State Acknowledgement

- The advertisements are normally acknowledged in link state acknowledgment packets.



- Each acknowledgment packet may contain a number of advertisement headers, exactly similar to those transmitted in database description packets during the exchange protocol. Since one acknowledgment packet may acknowledge several advertisements it is normal practice to delay its transmission in order to group many link state acknowledgments in a single packet.