

# OSPF

Mattias Jansson, KTHNOC/SUNET  
`fimblo@nada.kth.se`

February 13, 2004

# Contents

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Link State Protocols</b>               | <b>3</b>  |
| 1.1      | How does a LS protocol work? . . . . .    | 4         |
| 1.2      | Dijkstra's algorithm . . . . .            | 5         |
| 1.3      | LS: a look at Pros & Cons . . . . .       | 6         |
| <b>2</b> | <b>OSPF Basics</b>                        | <b>7</b>  |
| <b>3</b> | <b>OSPF in more detail</b>                | <b>13</b> |
| 3.1      | OSPF Packet Types . . . . .               | 13        |
| 3.2      | OSPF header . . . . .                     | 14        |
| 3.3      | Hello: Finding neighbours . . . . .       | 17        |
| 3.4      | Database Exchange . . . . .               | 23        |
| 3.5      | Link State Advertisements (LSA) . . . . . | 25        |

|          |                               |           |
|----------|-------------------------------|-----------|
| 3.6      | Reliable Flooding . . . . .   | 34        |
| 3.7      | Various Peer States . . . . . | 37        |
| 3.8      | External routes . . . . .     | 41        |
| 3.9      | OSPF Areas . . . . .          | 43        |
| <b>4</b> | <b>Bibliography</b>           | <b>50</b> |

# 1 Link State Protocols

1. How does a link state (LS) protocol work?
2. Dijkstra's algorithm
3. A look at some pros & cons

## 1.1 How does a LS protocol work?

- Each router spreads information concerning its links to its neighbours.
- This information is flooded to every router in the routing domain. Now all routers have knowledge of the entire network topology.
- Using Dijkstra's algorithm, the shortest path to each prefix in the network is calculated.

## 1.2 Dijkstra's algorithm

Dijkstra's algorithm is used to calculate the shortest path between two nodes in a directed graph where each edge has an associated non-negative weight.

Loosely speaking, the weights can be seen as the “cost” to traverse the edge. This cost could represent bandwidth, economic cost etc.

## 1.3 LS: a look at Pros & Cons

### Pros:

- Very fast convergence, when the network changes, new routes are found quickly.
- Can route based on more than just number of hops.
- Can be used in a larger network, thanks to possibilities of partitioning routing domains.
- Less information in updates → less traffic on the net

### Cons:

- More complex- (can be) Harder to administrate, and harder to implement.
- Needs much more memory.
- Demands more processing power.

## **2 OSPF Basics**

There are four basic phases an OSPF router goes through from cold start to a converged state.

Note:

- For a detailed description of how OSPF is defined, see RFC 2328 [1].
- For a more detailed description of OSPF with examples, see Moy [3].

## OSPF Basics- stage one

Hello! Is there there anyone there?

- Neighbour discovery
- Exchange of basic information

## OSPF Basics- stage two

So, what can you tell me of the  
world?

- Summary of Link State Database is exchanged.

## OSPF Basics- stage three

Please tell me more about Router  
XYZ's links

- Link State Requests / Updates

## OSPF Basics- stage four

Ok, now I know where everything is!

- The router can now start Dijkstra's algorithm on its link state database, creating a routing table.
- The router can now route traffic!

## OSPF Basics

What happens if a network goes down?

- New information is sent out by the router who notices this, and it is flooded out in the routing domain.
- Link State Updates / Acknowledgements

## **3 OSPF in more detail**

### **3.1 OSPF Packet Types**

- OSPF Protocol Packet Types
  1. Hello
  2. Database Description
  3. Link State Request
  4. Link State Update
  5. Link State Acknowledgement

## 3.2 OSPF header

|                       |          |
|-----------------------|----------|
| Version               | Pkt Type |
| Len                   |          |
| Source OSPF Router ID |          |
| OSPF Area ID          |          |
| Packet checksum       |          |
| Authentication Type   |          |
| Authentication Data   |          |

## OSPF Header cont'd

| Field Descriptions    |  |
|-----------------------|--|
| Version               | 2  |
| Pkt Type              | Hello<br>Database Description<br>Link State Request<br>Link State Update<br>Link State Acknowledgement |
| Len                   | Length of OSPF packet  |
| Source OSPF Router ID | (see next slide)   |
| OSPF Area ID          | *  |
| Packet Checksum       | *  |
| Authentication Type   | MD5  |
| Authentication Data   | *  |

\* Self-explanatory entries

## Router ID

- Choose the first configured interface as the Router ID.
- After a reboot, choose the highest IP address among the loopback interfaces.

### 3.3 Hello: Finding neighbours

- How does a router know who its neighbours are?  
*By sending each other Hello packets.*
- Sent every 10 seconds
- 3 failed Hello attempts result in a link failure report.

# OSPF Hello Header

|                        |                          |                       |           |   |                    |
|------------------------|--------------------------|-----------------------|-----------|---|--------------------|
| IP Header              | Version                  | header len            | TOS       | V4, 5, Internetwork control<br>68 bytes |                    |
|                        | Datagram len             |                       |           |   |                    |
|                        | Identification           |                       |           |   |                    |
|                        | Fragmentation            |                       |           |   |                    |
|                        | TTL                      | Protocol              |           |   | TTL=1, protocol=89 |
|                        | Header Checksum          |                       |           |   |                    |
|                        | Source IP ADDRESS        |                       |           |   |                    |
| Destination IP address |                          |                       | 224.0.0.5 |   |                    |
| OSPF Header            | OSPF Version             | OSPF Packet Type      |           | v2, Hello<br>48 bytes                   |                    |
|                        | OSPF len                 | Source OSPF Router ID |           |   |                    |
|                        | OSPF Area ID             |                       |           |   |                    |
|                        | Packet Checksum          |                       |           |   |                    |
|                        | Authentication Type      |                       |           | 0 = none                                |                    |
|                        | Authentication Data      |                       |           |   |                    |
| Hello Packet body      | Network Mask             |                       |           |   |                    |
|                        | Hello Interval           |                       |           | 10 sec                                  |                    |
|                        | Options                  | Router Priority       |           | E-bit, priority 1<br>40 sec             |                    |
|                        | Router Dead Interval     |                       |           |   |                    |
|                        | Designated Router        |                       |           |   |                    |
|                        | Backup Designated Router |                       |           |   |                    |
|                        | 1st Neighbor ID          |                       |           |   |                    |

## Hello: Point to point connections

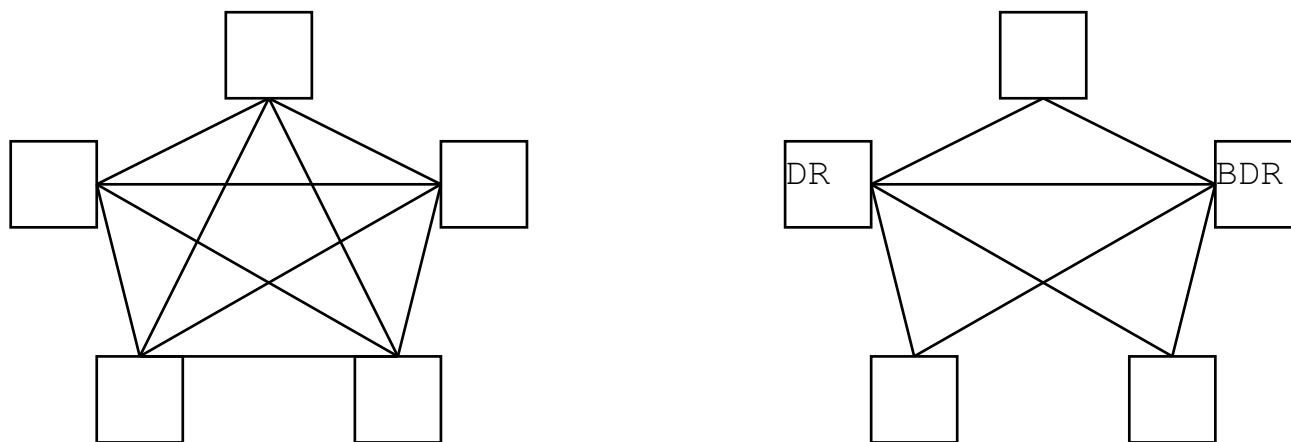
- Send Hello packet to multicast address:  
224.0.0.5 OSPF ALL-SPF-ROUTERS
- Bi-directional Adjacency (ACK!)

## Hello: Multiple access broadcast networks

What problems do we encounter in broadcast networks?

- Number of adjacencies:  $\frac{n(n-1)}{2}$
- Why both Designated and Backup Designated Router?  
(DR & BDR?)

e.g.



Note:

- The problem with broadcast networks is that there are too many adjacencies if the network is fully-meshed. Also called the N-squared problem.
- For a in depth description on the N-Squared problem, see pp132-134 Huitema [5]

## **Hello: Multiple access broadcast networks cont'd**

How is DR chosen?

- Router with highest priority
- Router with highest Router ID

Why do we need a BDR?

- The designated router holds important information on which neighbours has received Hello packets.

## **Hello: Multiple access broadcast networks cont'd**

- Send HELLO messages to multicast address 224.0.0.6  
(OSPF ALL-DR-ROUTERS)  
DR & BDR sends HELLO to 224.0.0.5  
(OSPF ALL-SPF-ROUTERS)
- Each node in the network holds adjacency only to  
DR & BDR.

## 3.4 Database Exchange

When a connection between two routers has just come up, the two routers need to get their databases synchronized.

They send a summary of their databases to each other, so that they can see if they need to ask each other for updates.

# Database Exchange (cont'd)

Sample Database exchange

```

10.1.1.4
-----
OSPF Hello
-----
10.1.1.6
-----
OSPF Hello: I heard 10.1.1.6
-----
DE descr. seq=X
-----
DE descr. seq=X, 5 LSA headers
router-LSA, 10.1.1.1 0x80000004
router-LSA, 10.1.1.2 0x80000007
router-LSA, 10.1.1.3 0x80000003
router-LSA, 10.1.1.4 0x8000003b
router-LSA, 10.1.1.5 0x80000039
router-LSA, 10.1.1.6 0x80000005

DE descr. seq=X+1, 1 LSA headers
router-LSA, 10.1.1.6, 0x80000001
-----
DE descr. seq=X+1
-----
Link state request packet, LSAs
router-LSA, 10.1.1.1
router-LSA, 10.1.1.2
router-LSA, 10.1.1.3
router-LSA, 10.1.1.4
router-LSA, 10.1.1.5
router-LSA, 10.1.1.6

Link state update packet, LSAs
router-LSA, 10.1.1.1, 0x80000004
router-LSA, 10.1.1.2, 0x80000007
router-LSA, 10.1.1.3, 0x80000003
router-LSA, 10.1.1.4, 0x8000003b
router-LSA, 10.1.1.5, 0x80000039
router-LSA, 10.1.1.6, 0x80000005

Link state update packet, LSAs
router-LSA, 10.1.1.6, 0x80000006

```

Note:

- Note that this sample skips the ExStart state.
- For more information on Database Exchange, see p91 in Moy [3]

## 3.5 Link State Advertisements (LSA)

- Every router spreads information about its links to all of its peers. This information is called an LSA.
- All routers forward received LSAs on to its peers.
- LSAs are acknowledged.
- Incremental updates.
- Periodic updates every 30 minutes.

**LSA: header**

|                    |         |
|--------------------|---------|
| LS Age             |         |
| Options            | LS Type |
| Link state ID      |         |
| Advertising router |         |
| LS Sequence number |         |
| LS Checksum        |         |
| Length             |         |

Note:

- For more information on LSA headers, see p74 Moy [3]

## **LSA: Link State Age**

- 0-30 minutes
- When an LSA reaches 30 minutes, it is refreshed by the originating router by flooding a new instance of the LSA throughout the network.
- If the originating router does not refresh the LSA, the LSA will be deleted from the database after one hour (MaxAge 1hr).
- Premature aging is used to flush LSAs from the Link-State Database.

## Notes:

- LSA Age indicates the number of seconds since the LSA was originated. Normally, if LSA Age reaches 30 minutes, the source router floods a new LSA with a new sequence number.
- If the originating router fails to do this, the age of the LSA continues to increase until the value of MaxAge is reached. At that point, the LSA is deleted from the database.
- To ensure that all other routers also delete the LSA at the same time, the deleted LSA (with the MaxAge) is reflooded onto the network. All other routers will then see this LSA and delete it from their database.
- *Premature Aging*: If a router wants to delete an LSA from the network despite the fact that it has not yet reached MaxAge, it can set the LSA Age to MaxAge and reflood it onto the network. This will force the LSA to be deleted from all the other routers on the network.
- See p79 Moy [3] for more information.

## LSA: Link State Type

### 1. Router LSA

Stub networks, and transit networks, and p2p connections between routers, are described with Router LSAs.

### 2. Network LSA

Originated by DR. Contains list of routers connected to the shared medium.

### 3. Network Summary LSA

Originated by BR. Describes a route to a network outside the area.

### 4. ASBR Summary LSA

Originated by BR. Describes the route to ASBRs.

### 5. AS External LSA

Originated by ASBR. Describes a route to a network outside the AS.

### 6. Group Membership LSA (MOSPF)

Describes the location of the area's multicast group members.

### 7. Nssa external information LSA

Describes an injected (external) network in an NSSA. Converted to type 5 at the border to area 0 by the BR.

### 8. External attributes LSA

(Carries information unknown to ospf)

Note:

- For more information on LSA types, see p74, Moy [3]
- MOSPF: Multicast Extensions to OSPF, RFC 1584.

## LSA: Link State ID

Link ID information, depending on the type of LSA, this field can take on different values.

- LSA Type 1: Router-ID
- LSA Type 2: The ip interface address of the DR.
- LSA Type 3: The destination network's ip address.
- LSA Type 4: The Router-ID of the ASBR
- LSA Type 5: The destination net's ip address.

## **LSA: Advertising Router**

Router ID of the advertising router. (The originating router)

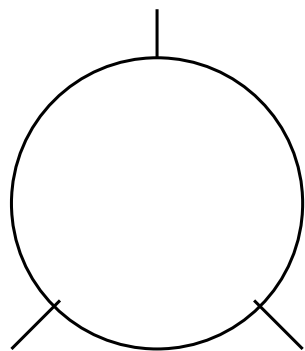
## **LSA: Sequence number**

When a router has two instances of a particular LSA, it detects which instance is more recent by comparing the instances' LS Sequence Numbers.

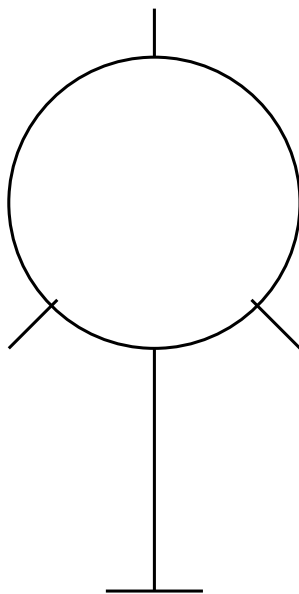
The instance with the *larger* LS Sequence number is the more recent.

## LSA: Sequence number cont'd

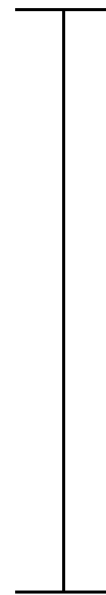
The definition of the word “larger” depends on the implementation of the sequence number space.



Circular sequence



Lollipop shaped  
sequence



Linear sequence

**Circular** Used in the original ARPANET, famous for the ARPANET sequence bug.

**Lollipop** Used in OSPF version 1.

**Linear** Used in OSPF version 2.

## **LSA Checksum**

Excludes the LSA Age field for obvious reasons...

## **LSA Length**

Length of both LSA header & contents in bytes (16 bit field).

*Why can't LSA Length be 65,535 bytes long in practice???*

Answer:

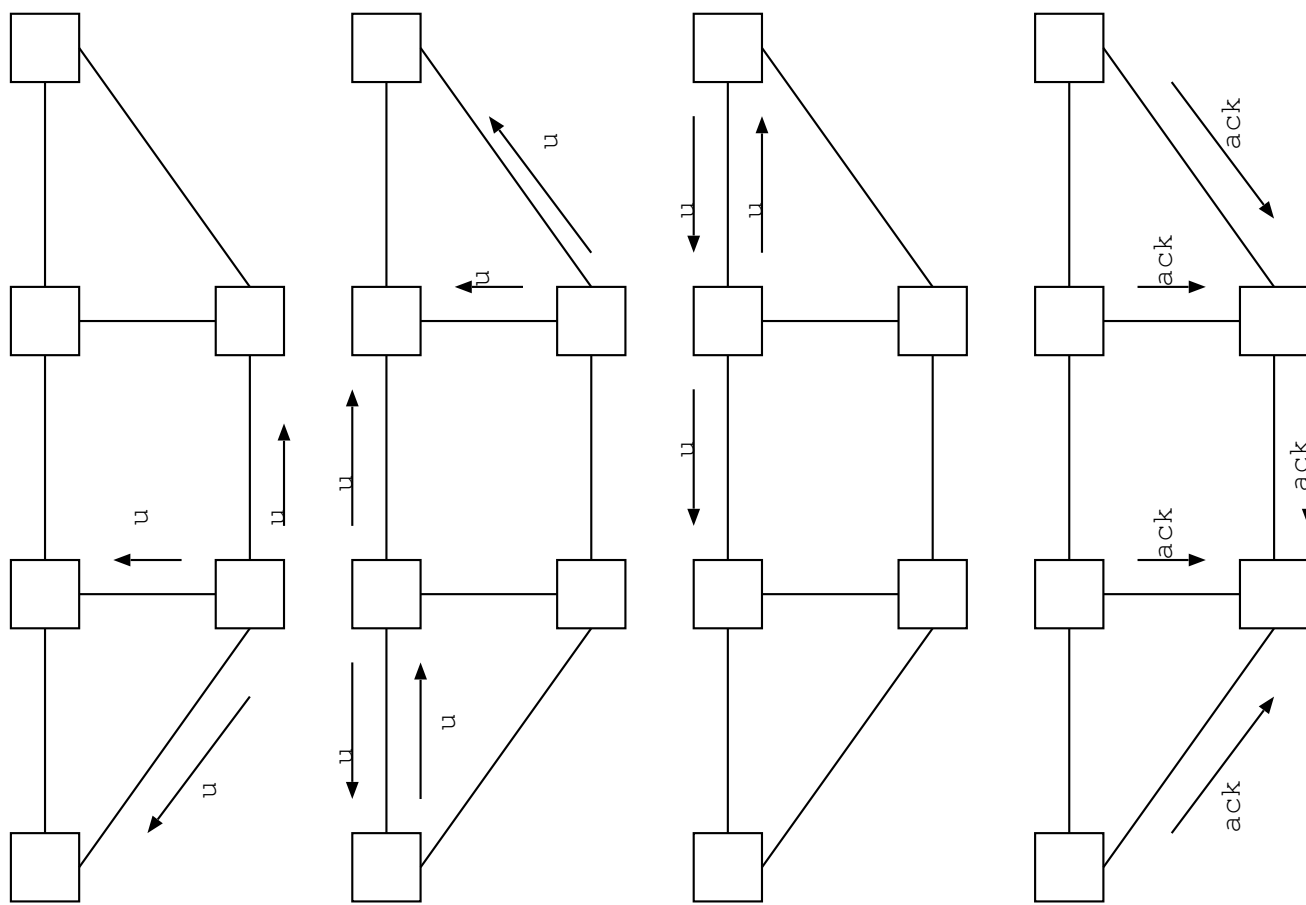
- Because the LSA packet must eventually be transported within an IP packet that is itself restricted to 65,535 bytes in length.
- LSA field can be between 20 bytes to a bit over 65,000 bytes...

## 3.6 Reliable Flooding

A link state protocol can function only if each node has complete knowledge of the network. Therefore we must devise a way to assure ourselves that LS updates are flooded everywhere.

# Reliable Flooding in P2P Networks

Implicit and Explicit acknowledgement

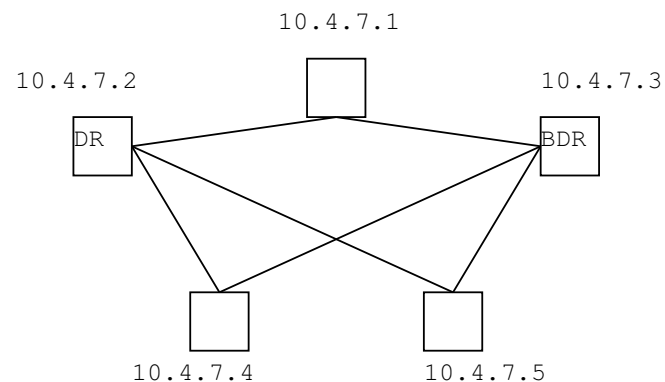


Note:

- For more information, see p92 Moy [3]

## Reliable flooding in Shared Medium Networks

Consider the following: 10.4.7.5 gets an LSA from one of its links and wants to forward it on the multiple access network below.



10.4.7.5 installs the LSA in its database and thereafter floods it to DR (10.4.7.2) and BDR (10.4.7.3) by multicasting an LSA to 224.0.0.6 (ALL DR ROUTERS).

DR (10.4.7.2) then multicasts the update on 224.0.0.5 (ALL SPF ROUTERS).

If BDR does not hear an update from the DR on 224.0.0.5, it assumes that the DR has crashed and it takes over as the new DR.

DR waits for ACKnowledgements from all of its neighbours. If it doesn't get an ACK from one of them, it unicasts the packet specifically to that router.

Note:

- More information on Reliable flooding over broadcast on p107, Moy [3]

## 3.7 Various Peer States

**Down** Initial State; No information has been received from neighbour.

**Attempt** Only NBMA (Nonbroadcast Multiaccess) Networks; No information has been received from neighbour, but an attempt will be made.

**Init** A Hello packet has been received from a neighbour, but bidirectional communication has not been established.

**2-Way** Communication between the two routers is bidirectional. Adjacency is not yet established. The DR & BDR are selected from this or greater states.

## Various Peer States cont'd

**ExStart** First stage in establishing an adjacency with another router. Decision on who in the relationship is master.

**Exchange** The router sends its LS database description to its neighbour. At this stage, all adjacencies are fully capable of discussing all types of OSPF routing protocol packets.

**Loading** Link State Request packets are sent to neighbour asking for more recent LSAs discovered but not received in the Exchange state.

**Full** Full adjacency.

Note:

- For more juicy details on Peer states and the events which make the router switch from one state to another, see section 10, RFC 2328 [1]

## Example - OSPF database

```
s-gw>sh ip ospf database router
      OSPF Router with ID (193.10.252.10) (Process ID 2603)
      Router Link States (Area 0.0.0.0)
LS age: 1372
Options: (No TOS-capability, DC)
LS Type: Router Links
Link State ID: 193.10.68.1
Advertising Router: 193.10.68.1
LS Seq Number: 8000287C
Checksum: 0xD97D
Length: 96
Number of Links: 6
  Link connected to: a Stub Network
    (Link ID) Network/subnet number: 193.10.68.1
    (Link Data) Network Mask: 255.255.255.255
    Number of TOS metrics: 0
    TOS 0 Metrics: 1
  Link connected to: another Router (point-to-point)
    (Link ID) Neighboring Router ID: 193.10.68.2
    (Link Data) Router Interface address: 193.10.68.17
    Number of TOS metrics: 0
    TOS 0 Metrics: 1
```

## Example - OSPF Neighbours

```
s-gw>sh ip ospf neighbor
```

| Neighbor ID   | Pri | State        | Dead Time | Address        | Interface |
|---------------|-----|--------------|-----------|----------------|-----------|
| 193.10.252.5  | 1   | FULL/ -      | 00:00:34  | 193.10.252.129 | POS6/0    |
| 193.10.68.4   | 1   | FULL/ -      | 00:00:06  | 193.10.68.38   | POS3/0    |
| 193.10.68.5   | 1   | FULL/ -      | 00:00:36  | 193.10.68.42   | POS5/0    |
| 193.10.252.2  | 1   | FULL/DROTHER | 00:00:30  | 193.10.252.212 | SRP1/0    |
| 193.10.252.5  | 1   | FULL/BDR     | 00:00:34  | 193.10.252.211 | SRP1/0    |
| 193.10.252.11 | 1   | FULL/DROTHER | 00:00:38  | 193.10.252.209 | SRP1/0    |

Notes:

- POS: Packet over SONET (P2P)
- SRP: Spatial Reuse Protocol (Shared Medium)

## 3.8 External routes

- An external route is a prefix which OSPF has been informed of from another protocol (or a static route)
- There are two types, External Type 1 and External Type 2 (E1 & E2).

**E1** Assumes that in the path from OSPF router to destination, the internal component and external component are of the same order.

**E2** Assumes that the external part of the path is more significant than the internal cost to the AS boundary router. (default)

## Examples of E1 and E2

**RIP-OSPF** If OSPF uses hop count as its metric and RIP routes are imported as E1 metrics, the two systems would work more or less seamlessly.

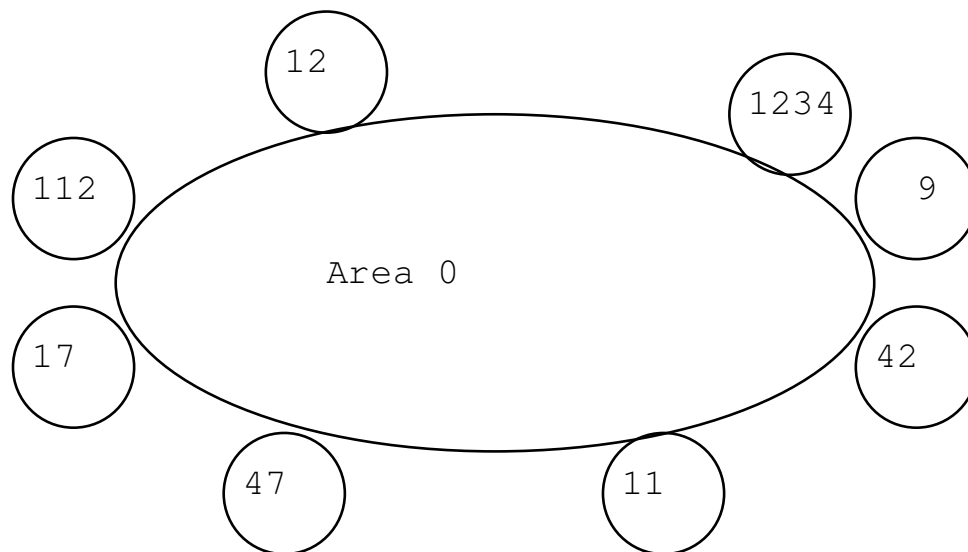
**BGP-OSPF** BGP routes are imported as E2 metrics, with metric set to equal to the BGP route's AS path length.

## Example - routing table

```
s-gw>sh ip rou ospf
    212.1.196.0/24 is variably subnetted, 33 subnets, 2 masks
O E2    212.1.196.248/32 [15/20] via 193.10.252.129, 14:43:49, POS6/0
    195.219.88.0/24 is variably subnetted, 2 subnets, 2 masks
O       195.219.88.8/30 [15/3] via 193.10.252.209, 14:43:49, SRP1/0
    195.67.208.0/30 is subnetted, 1 subnets
O       195.67.208.156 [15/3] via 193.10.252.212, 14:43:49, SRP1/0
    195.54.123.0/30 is subnetted, 1 subnets
O       195.54.123.4 [15/3] via 193.10.252.212, 14:43:49, SRP1/0
```

## 3.9 OSPF Areas

- Divides the OSPF domain into smaller zones
- Makes the link state databases smaller in each zone.



Note:

- See pp122-134 in Moy [3] for more information on Areas.

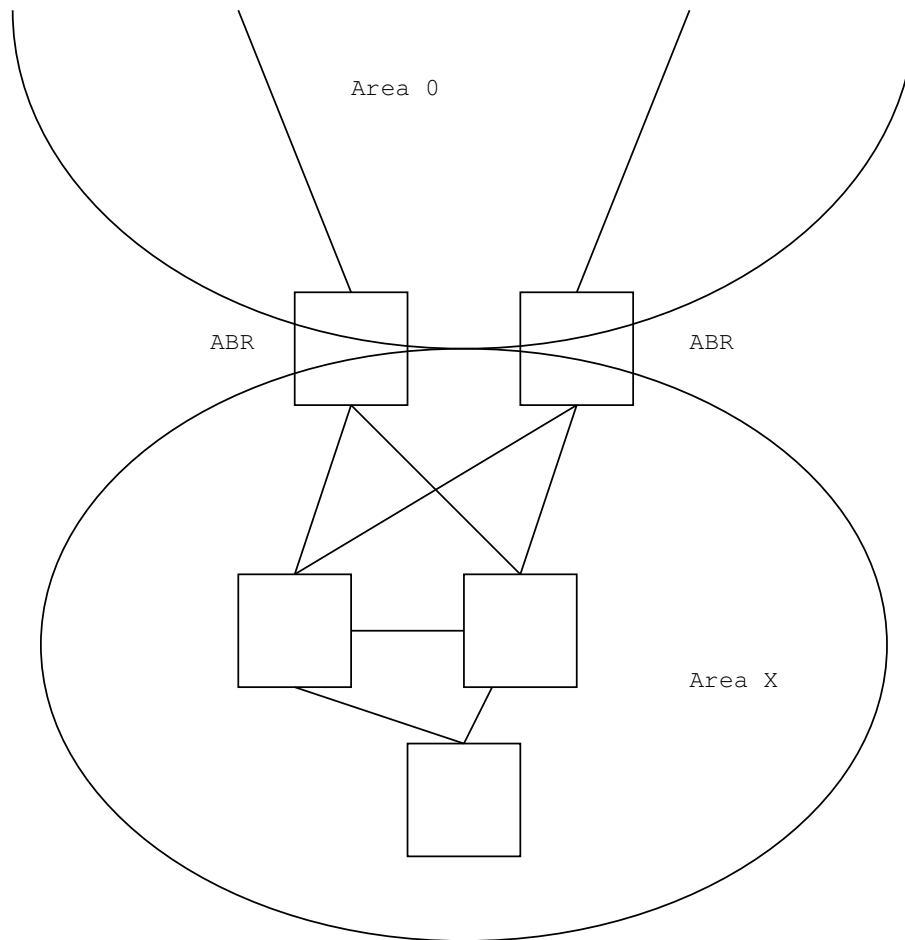
## OSPF Areas cont'd

- When splitting a routing domain into areas, all areas must attach directly to a special backbone area called Area 0.
- A Summary LSA is passed between areas. This special LSA does not contain all link state information of the area, but only a summary. In this way it is similar to RIP.
- Five different kinds of areas:
  1. Area
  2. Stub Area
  3. Totally Stub Area
  4. Not So Stubby Area (NSSA)
  5. NSSA Totally Stub Area

Note:

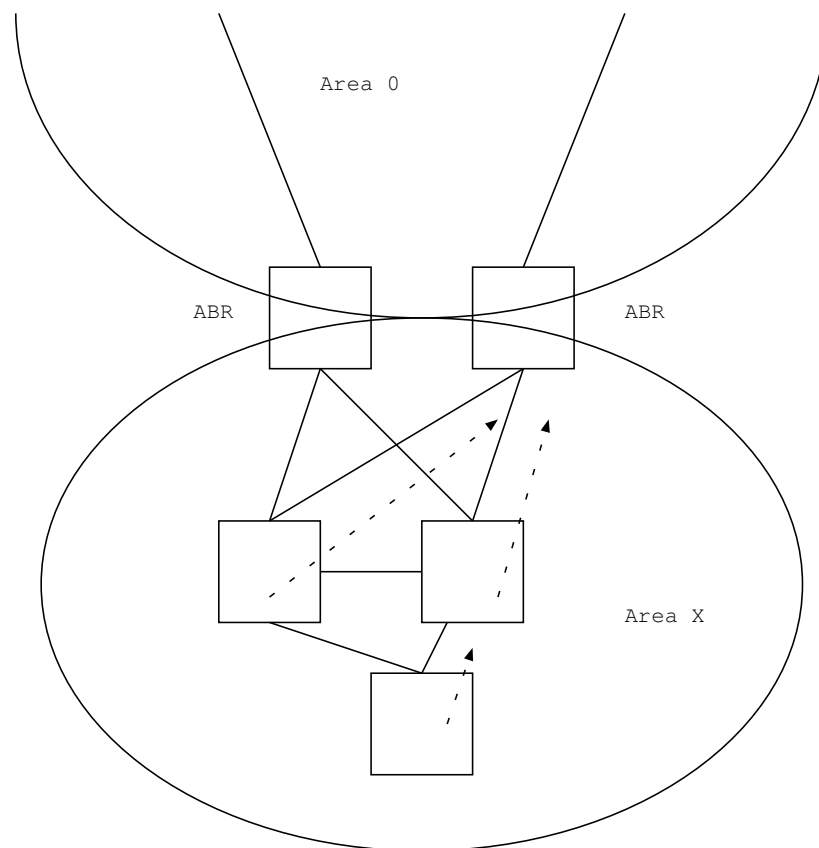
- Though all areas must be directly connected to Area 0, an area does not have to be physically connected. It can be connected through a virtual tunnel.

# Area



An area spreads all information, even external routes.

## Stub Area

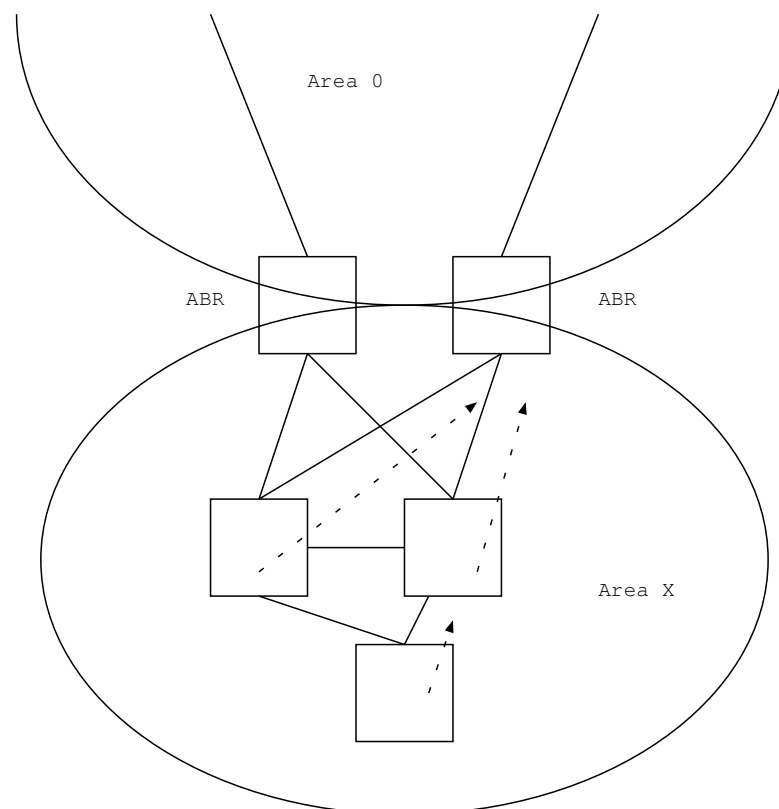


Minimizes size of Link State Database  
No injection of external routes  
All information on other areas are injected  
Default route out of the area (to external routes)

Note: In the above slide, the last point:

- Not only will the internal routers pass packets destined for the external routes to the ABR. All packets destined for non-existing ip addresses will be passed on to the ABR before they are dropped.

## Totally stub area



Minimizes size of Link State Database (even more)

Cisco specific

No injection of external routes

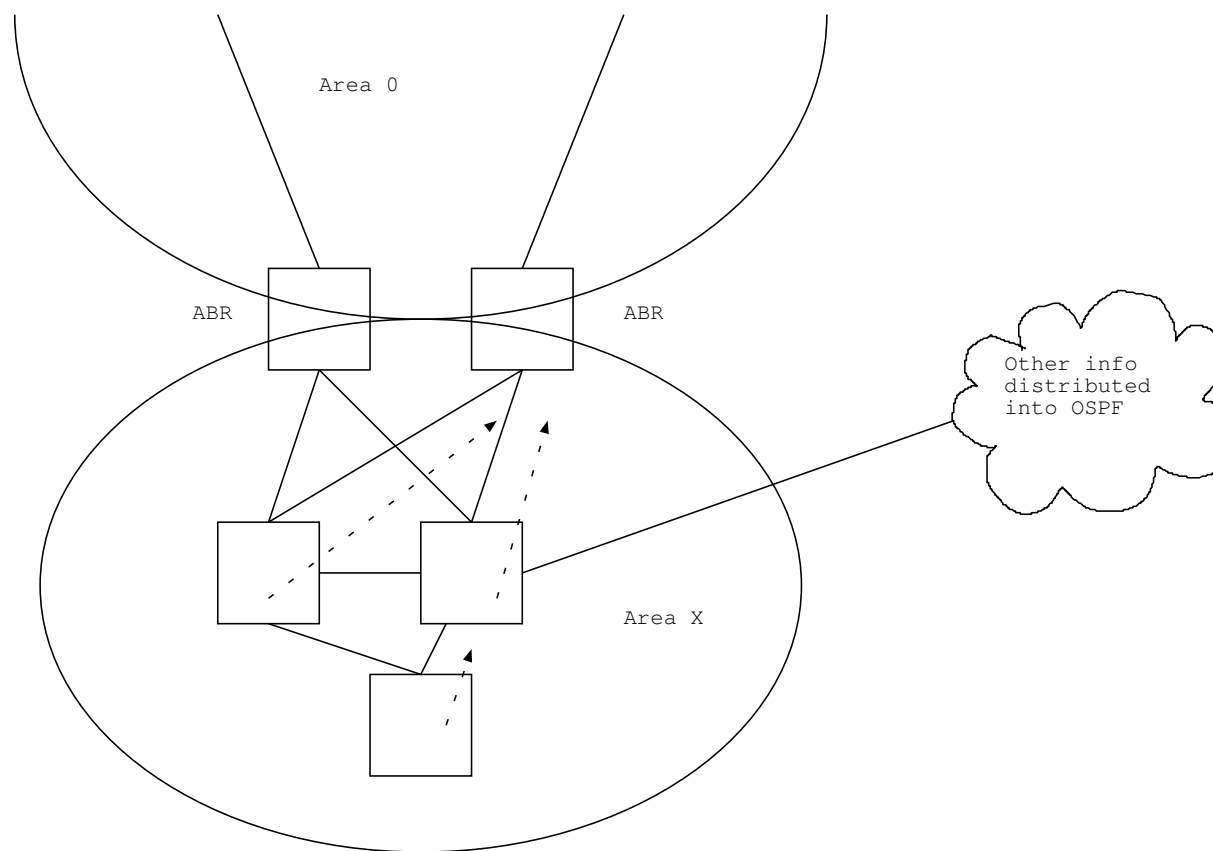
No information on other areas are injected

Default route out of the area (to anything outside the area)

Note:

- See Sackett [4], p490

## Not So Stubby Area (NSSA)



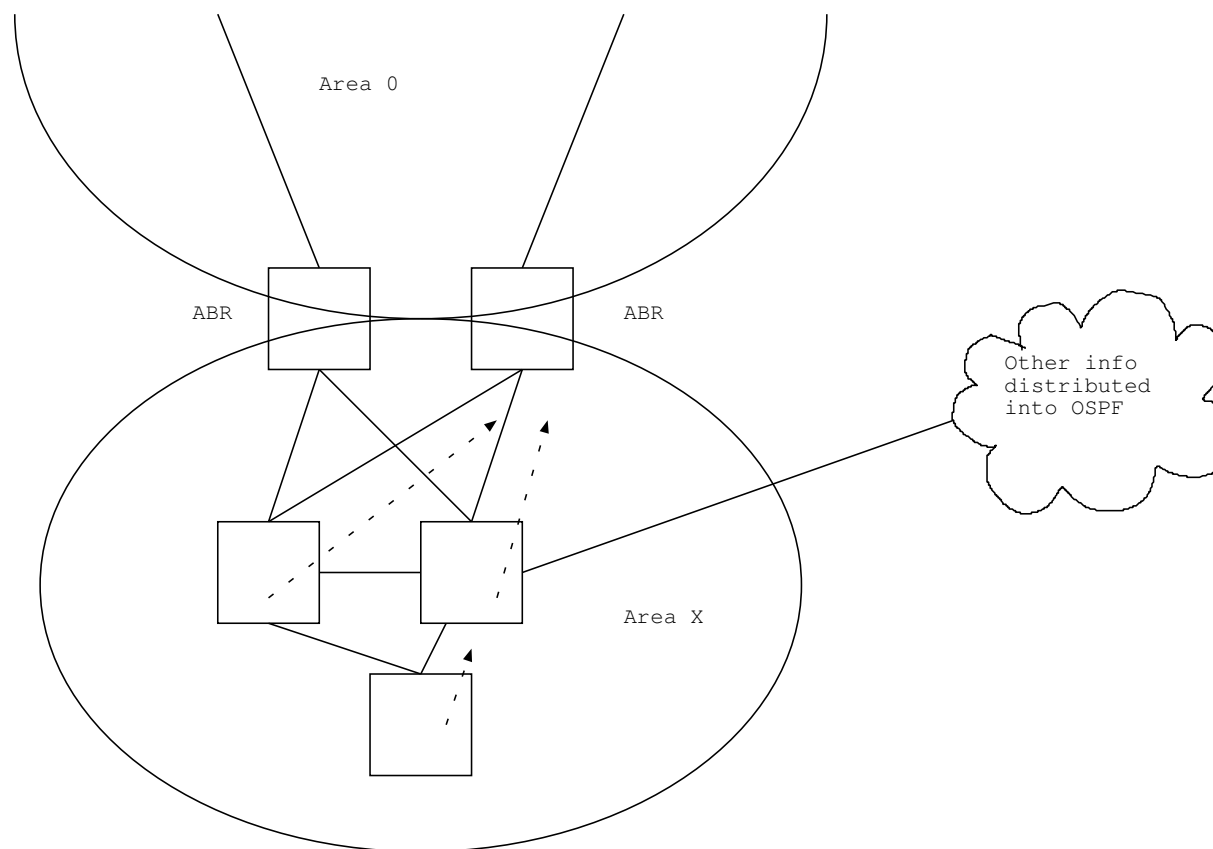
No injection of external routes

Allows external routes out from the area (LSA type 7 that convert to type 5)

All information on other areas are injected

Default route out of the area (to external routes)

## NSSA Totally Stub Area



Cisco Specific

No injection of external routes

Allows external routes out from the area (LSA type 7 that convert to type 5)

No information on other areas are injected

Default route out of the area (to anything outside the area)

## 4 Bibliography

- 1 RFC 2328, J. Moy. *OSPF Version 2*. 1998 The Internet Society, <http://www.ietf.org/rfc/rfc2328.txt>
- 2 Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest. *Introduction to Algorithms*. 1990 The MIT Press, ISBN 0-262-03141-8
- 3 John T. Moy. *OSPF: Anatomy of an Internet Routing Protocol*. 1998 Addison Wesley Longman Inc. 5th printing, May 2000. ISBN 0-201-63472-4
- 4 George Sackett. *The Cisco Router Handbook*. 2000 The McGraw-Hill Companies Inc. ISBN 0-07-058098-7
- 5 Christian Huitema, *Routing in the internet*. 2000, Second Edition, Prentice Hall Inc. ISBN 0-13-022647-5