



# OSPFv2 deep dive

Михайловский Василий – [vamikhai@cisco.com](mailto:vamikhai@cisco.com)

Cisco TAC Poland

Ноябрь 17, 2015

# Содержание

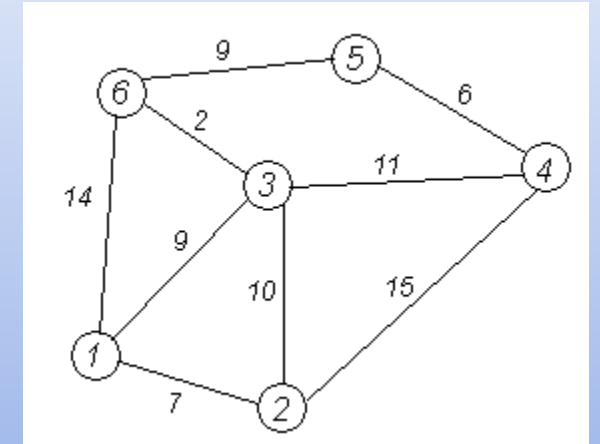
- TBD

# OSPFv2 transport

- Using IP-multicast 224.0.0.5 and 224.0.0.6;
- Protocol number 89;
- Unicast for immediate hellos, retransmits and initial exchange;
- IP MTU from egress interface (announced in DBD);
- DSCP –eq CS6 (dec 48), TOS –eq 192;

# SPF & LSDB concepts

- OSPF is Link-state protocol;
- Based on Dijkstra's algorithm;
- Building shortest path to remote node;
- Graph is build based on LSDB (link-state DB);
- Every link has a cost (cost to use the link);
- Path with lowest cost wins;
- LSDB consists of LSA (types 1,2,3,4,5,6,7);
- LSDB must be consistent across network.



# LSA concept

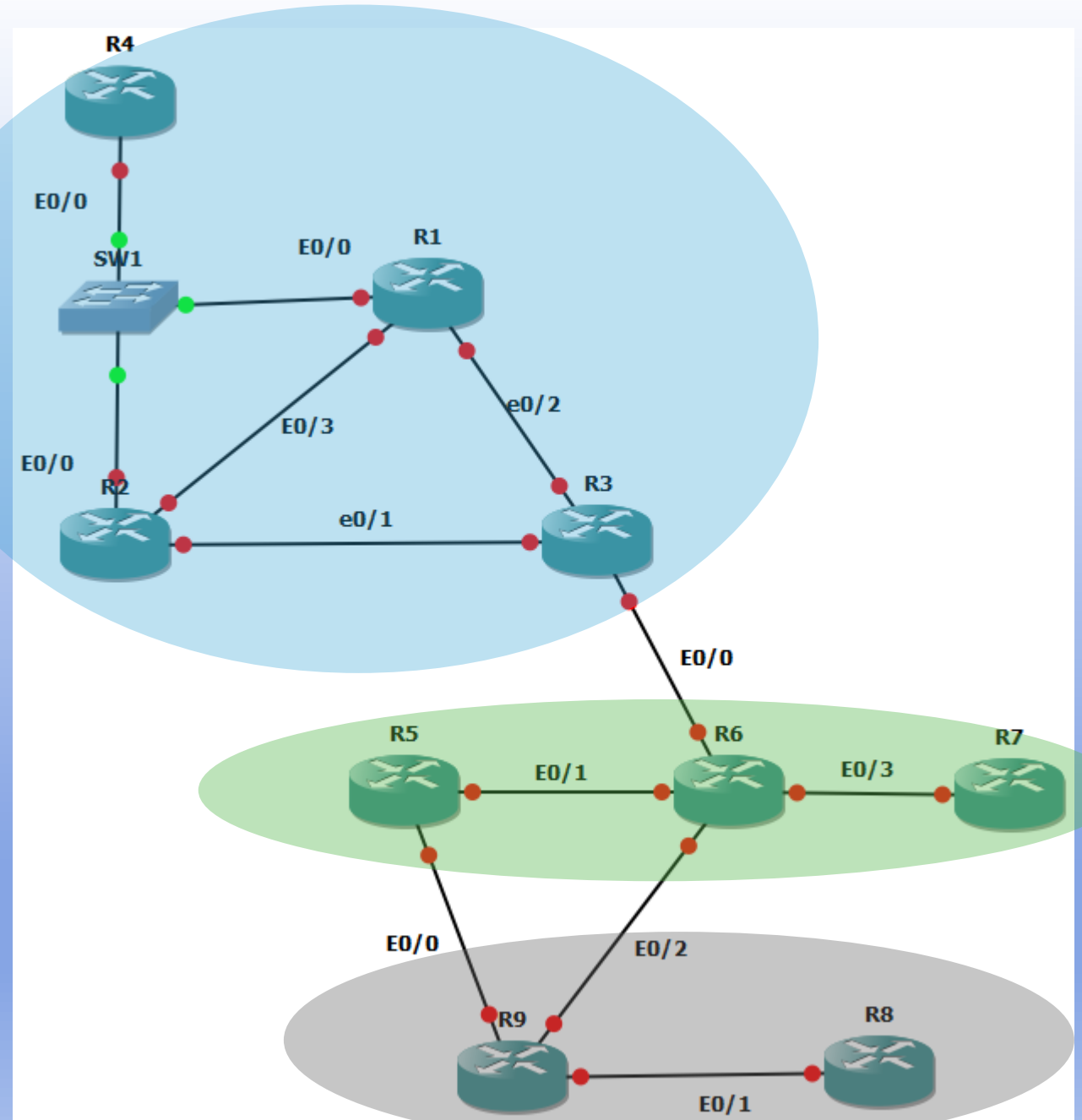
- LSA is identified by ID;
- LSA must have: ID, age, seq number, checksum;
- MAX AGE (3600 sec) – when LSA is purged from LSDB;
- SEQ number – defines “freshness” on LSA (the higher – the fresher);
- Checksum – to verify consistency of LSA payload;
- LSAs 1,2,3,4,5,6,7 are used in LSDB (OSPFv2) to build graph.

# OSPF area concept

- Area – the way to segment your network;
- Area 0 – Backbone area (the rest must be connected);
- Consistency is mandatory within an area;
- Area types: Backbone, normal, stub, totally-stub, nssa, totally nssa;
- ABR is to interconnect areas;
- ASBR is to connect external network to OSPF domain;

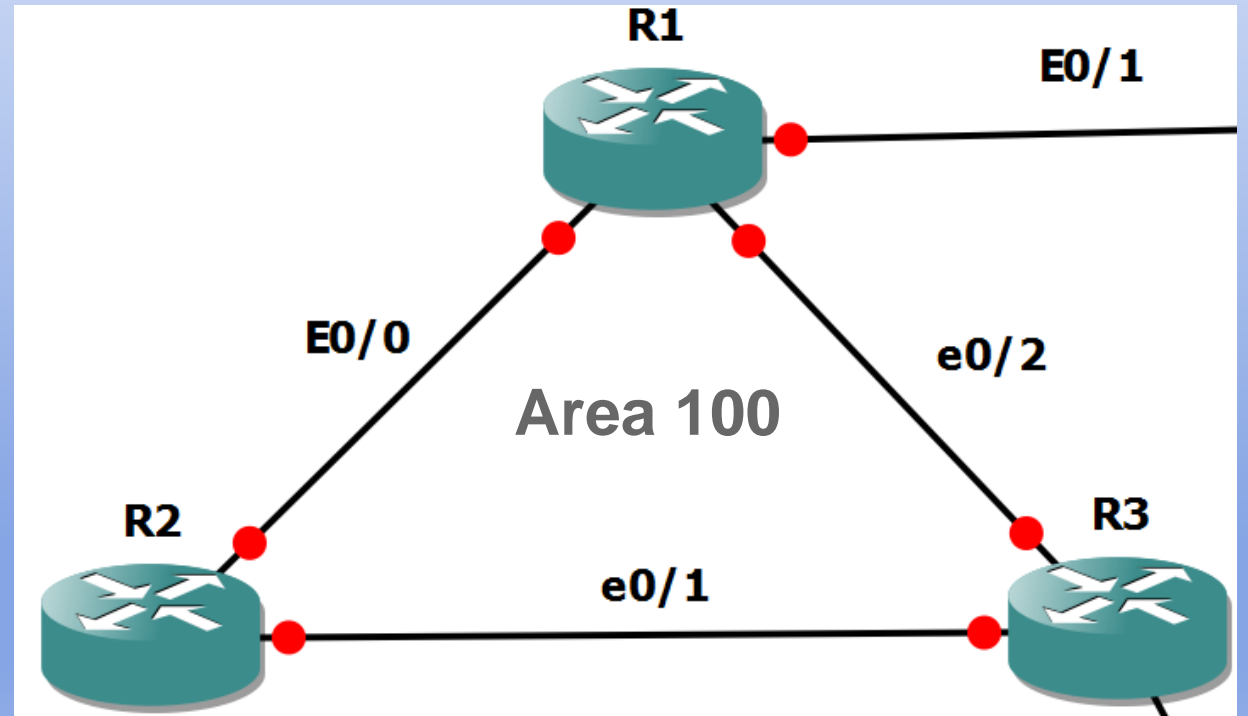
# Network topology

- Lo0 –eq x.x.x.x/32
- Router-id –eq Lo0
- Link –eq 10.x.y.x/24
- R1,R2,R3,R4 – area 100
- R5,R6,R7 – area 0 (ABRs);
- R8,R9 – NSSA area 101.



# LSA type1 – Router LSA

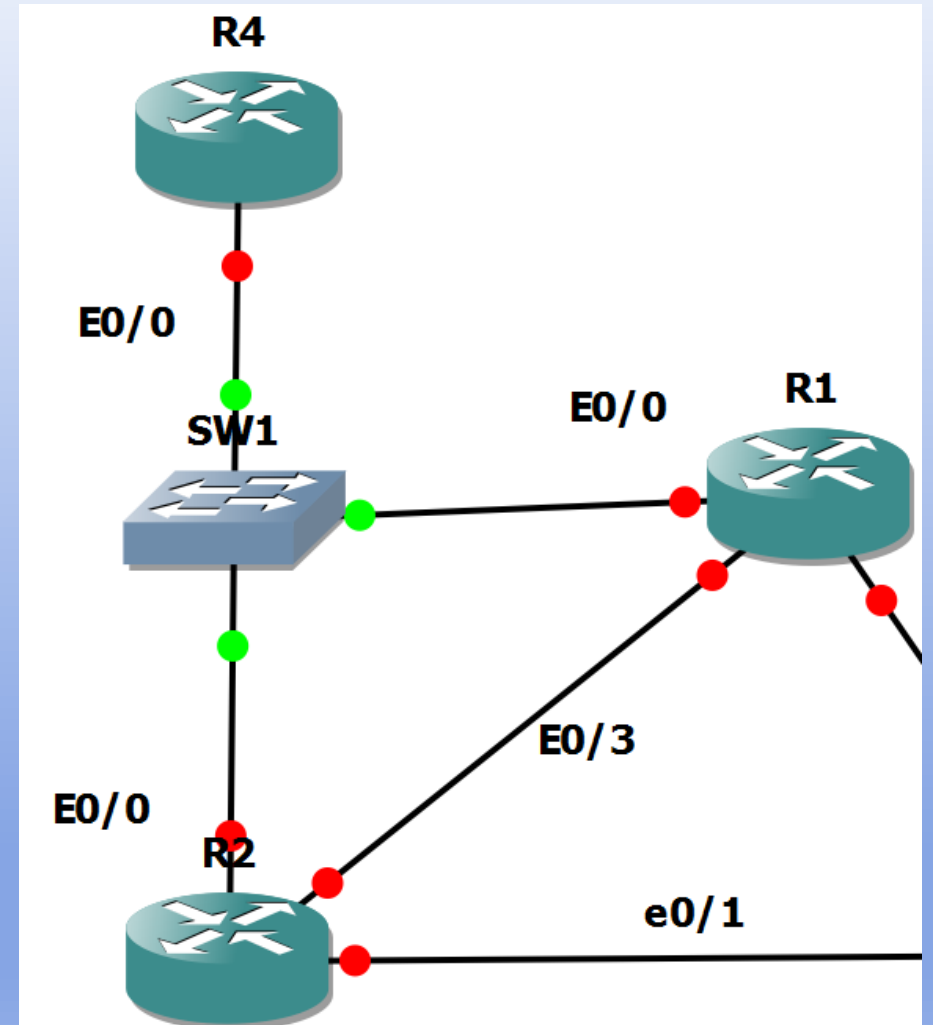
- Describes all the interfaces/adjacencies an owner has;
- Identified by router-id (unique);
- Scope: single area;
- Topology needs 3 LSAs only;





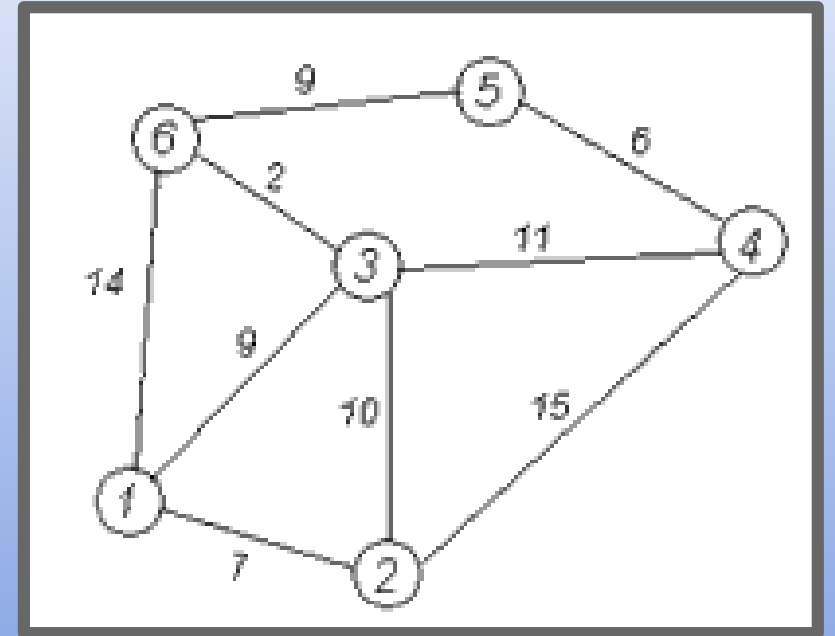
# LSA type2 – Network LSA

- Needed due to SPF restrictions;
- Defines pseudo-node in the network;
- Scope: single area;
- Special roles defined – DR / BDR;
- Slows down convergence time;
- NOT recommended in your CORE!



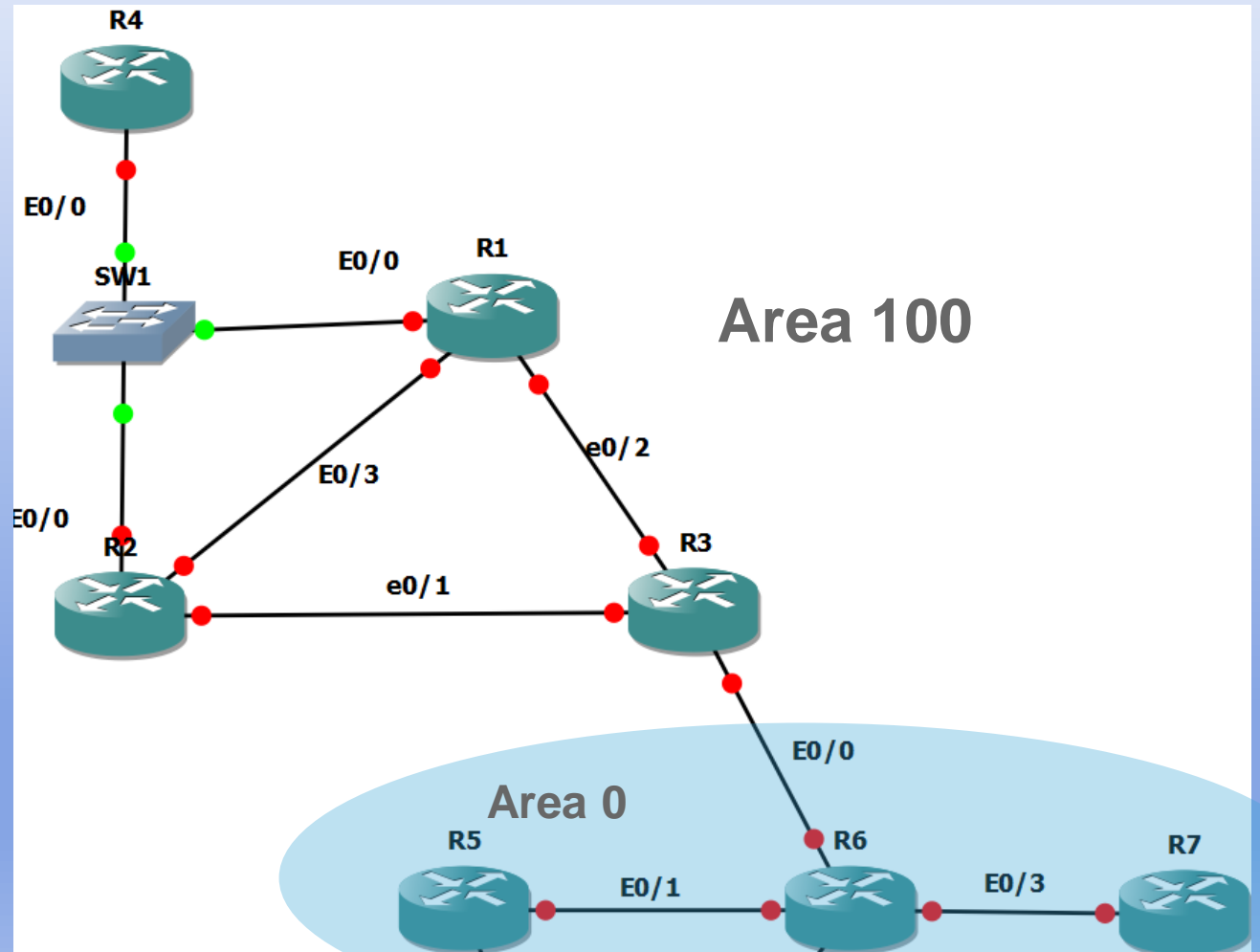
# OSPFv2 wire/segment roles

- Designated router – non-preemptive;
  - Building network LSA;
  - Updating network LSA;
  - Managing segment updates
- Backup designated router - non-preemptive;
  - Standby for DR;
  - Takes over once DR is lost.
- “DROTHER”;



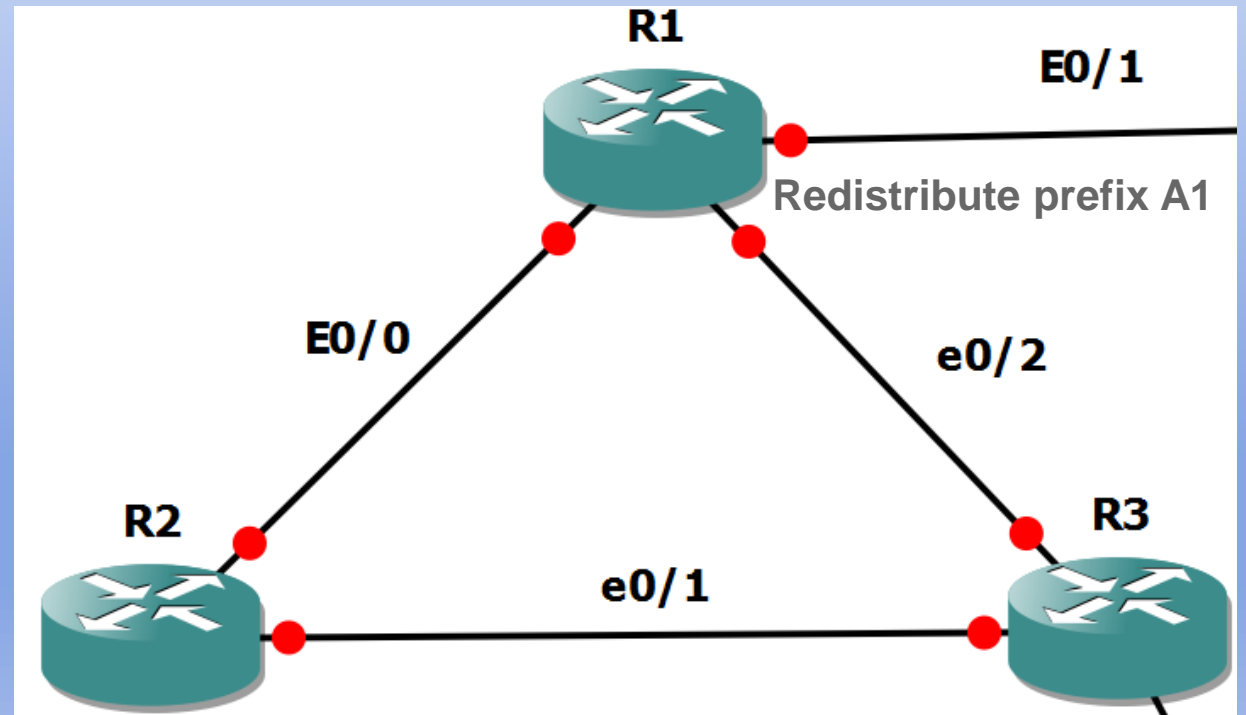
# LSA type3 – Summary LSA

- Created by ABR;
- R5/R7 use LSA3 to reach anything in area 100;
- Scope: local area, may be replicated by area 0 ABR;
- RIB – IA.



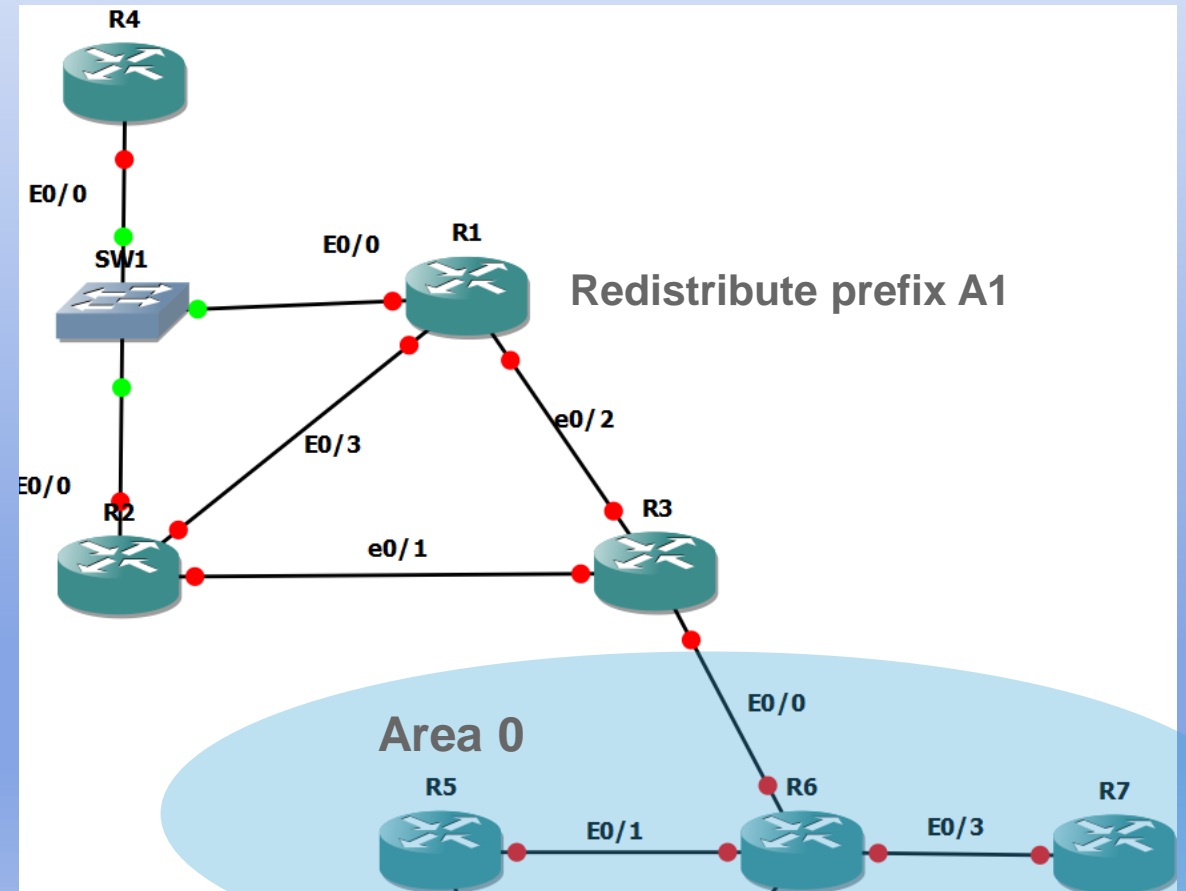
# LSA type5 – External LSA

- Created by ASBR;
- IMPORTANT: concept of Forwarding address;
- Scope: normal areas;
- RIB – E1/E2.



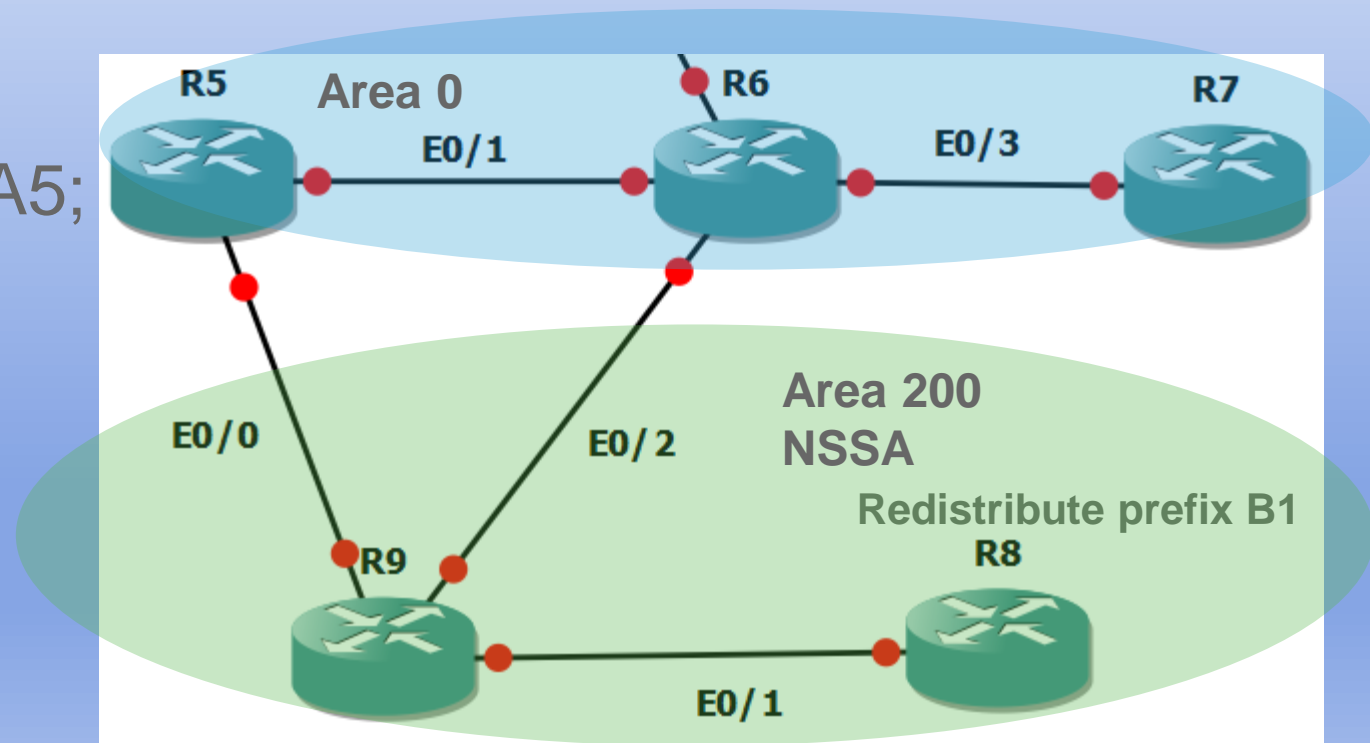
# LSA type4 – Summary router LSA

- Created by ASBR;
- IMPORTANT:  
concept of Forwarding address;
- Scope: normal areas;
- R5/R7 are using LSA5 to find ASBR, LSA4 to find the way to it;



# LSA type7 – NSSA external LSA

- Created by ASBR in NSSA area;
- IMPORTANT: concept of Forwarding address;
- Scope: nssa areas only;
- ABR does translation into LSA5;
- RIB - N1/N2;



# Injecting 0.0.0.0/0

- LSA5 - “default-information originate”;
- LSA3 – [totally] stub area;
- LSA3 – nssa (area 101 nssa no-summary);
- LSA7 – nssa (area 101 nssa default-information-originate);

# Routing decision / same prefix

- Intra-area;
- Inter-area;
- E1
- E2
- N1
- N2

NX-OS behavior differs from IOS (RFC 1583 vs 2328)!

<http://www.cisco.com/c/en/us/support/docs/ip/open-shortest-path-first-ospf/117824-config-ospf-00.html>



# Building adjacency – debug ip ospf adj

- Hello fields to match:
  - Area number;
  - Stub flag;
  - Hello/dead interval (unless sub-second);
  - Authentication type;
  - Authentication data;
  - within the same subnet;
  - MTU (or MTU-ignore – checked in DBD);
- Fields to be unique:
  - Router-id;

# OSPF network types

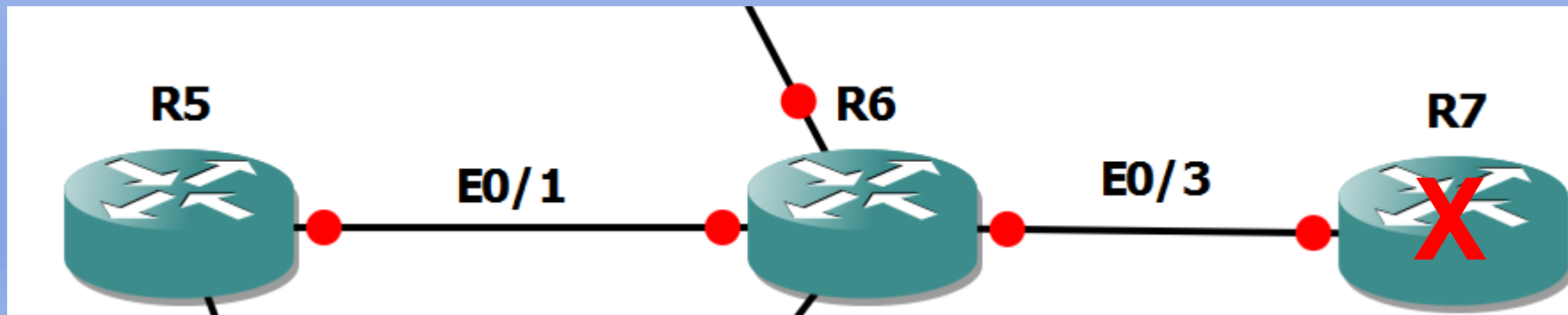
- Broadcast;
- Non-broadcast;
- Point-to-point;
- Loopback;
- Point-to-multipoint;
- Point-to-multipoint non-broadcast;

Network type mix may lead to an issue.

# Issue 1 – stuck in EXSTART/EXCHANGE

# More details on LSDB – stale entry

- LSDB is a database of raw information for SPF algorithm;
- It may have stale entries, that could confuse you!



# Issue 2 – inconsistent network type

```
R7#sh ip ospf ne
```

Neighbor ID	Pri	State	Dead Time	Address	Interface
6.6.6.6	1	FULL/BDR	00:00:30	10.6.7.6	Ethernet0/3

```
R7#sh ip route 6.6.6.6
```

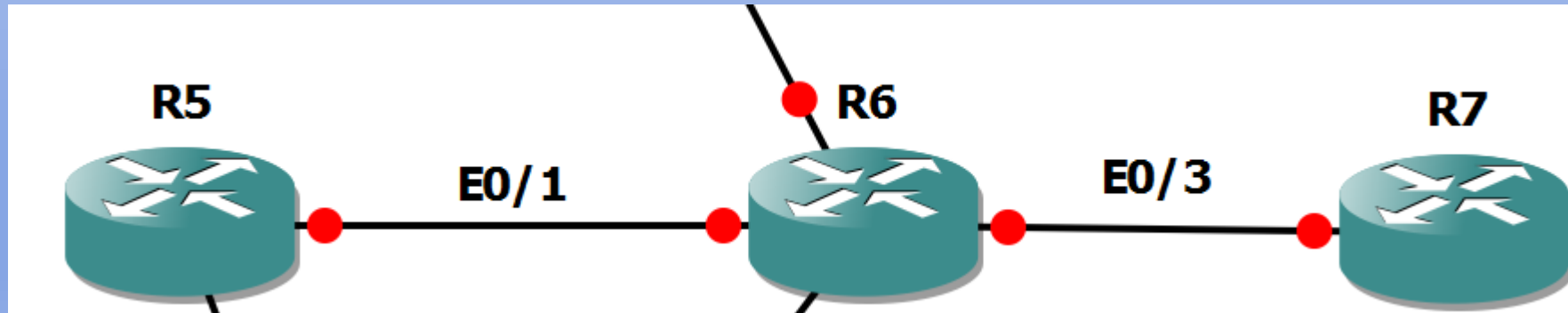
```
% Network not in table
```

```
R6#sh ip ospf ne
```

Neighbor ID	Pri	State	Dead Time	Address	Interface
7.7.7.7	0	FULL/ -	00:00:33	10.6.7.7	Ethernet0/3
5.5.5.5	0	FULL/ -	00:00:35	10.5.6.5	Ethernet0/1

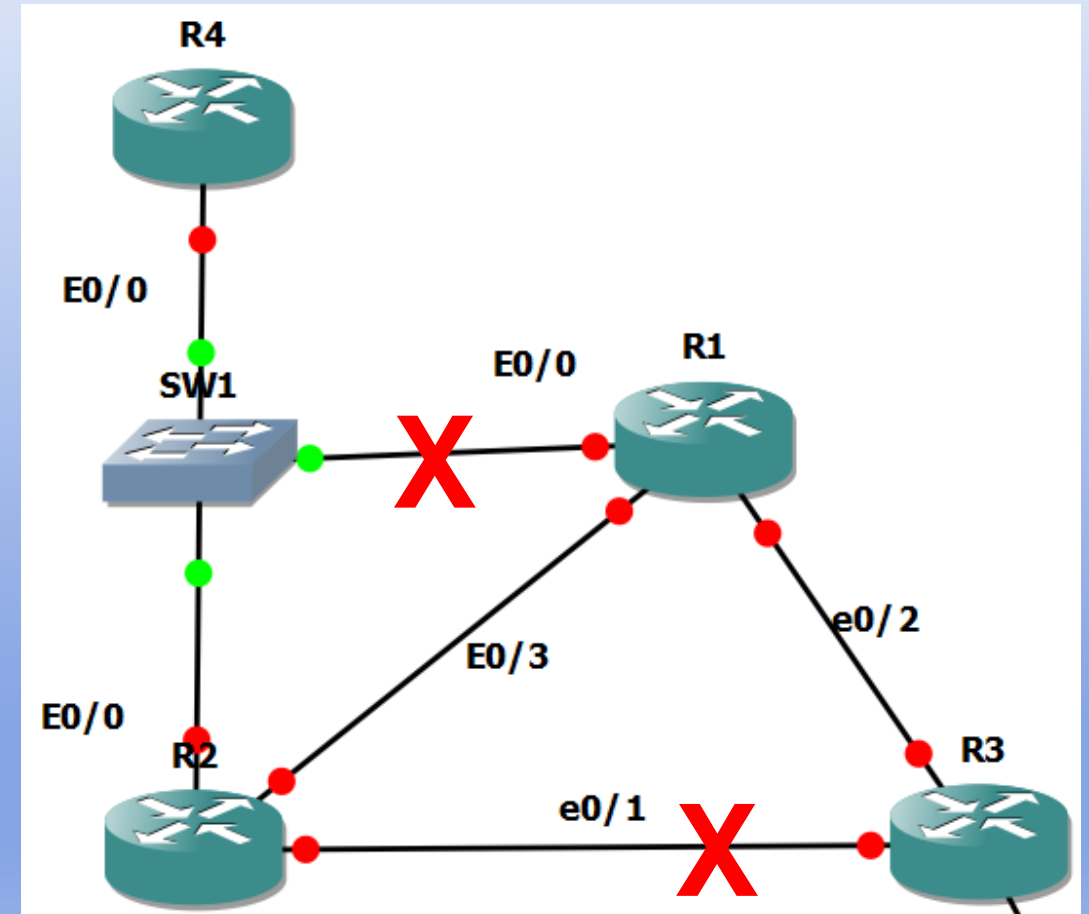
```
R6#sh ip route 7.7.7.7
```

```
% Network not in table
```



# Issue 3 – flood-war/duplicated router-id

- R3/R4 have same router-id 3.3.3.3
- LSDB symptoms:
  - seq number is high;
  - age time is low (under 15 sec).
- %OSPF-4-DUP\_RTRID\_AREA: Detected router with duplicate router ID 3.3.3.3 in area 100
- %OSPF-4-FLOOD\_WAR: Process 1 flushes LSA ID y.y.y.y type-5 adv-rtr x.x.x.x in area 0



# Issue 4 – FA missing in RIB (as internal OSPF)

```
R2#sh ip ospf data ext 1.0.0.0
```

```
LS Type: AS External Link
```

```
Link State ID: 1.0.0.0
```

```
Advertising Router: 1.1.1.1
```

```
LS Seq Number: 80000025
```

```
Checksum: 0xF0C7
```

```
Length: 36
```

```
Network Mask: /8
```

```
    Metric Type: 2 (Larger than  
any link state path)
```

```
Metric: 200
```

```
    Forward Address: 10.0.1.254
```

```
ip access-list standard IN_OSPF  
deny 10.0.1.0  
permit any
```

```
router ospf 1  
  distribute-list IN_OSPF in
```

```
R2#sh ip route 10.0.1.254  
% Subnet not in table
```

```
R2#sh ip route 1.0.0.1  
% Subnet not in table
```

# Caveat 5 – Running two OSPF processes

- Behavior depends on release, so distance manipulation is recommended;
- NX-OS compares only metric -> redistribution issues;
- IOS – depends on a version, but compares prefix type:
  - still does not compare “fwd cost” for E2/N2;
  - oldest wins (older releases);
  - lowest process ID wins (newer releases).

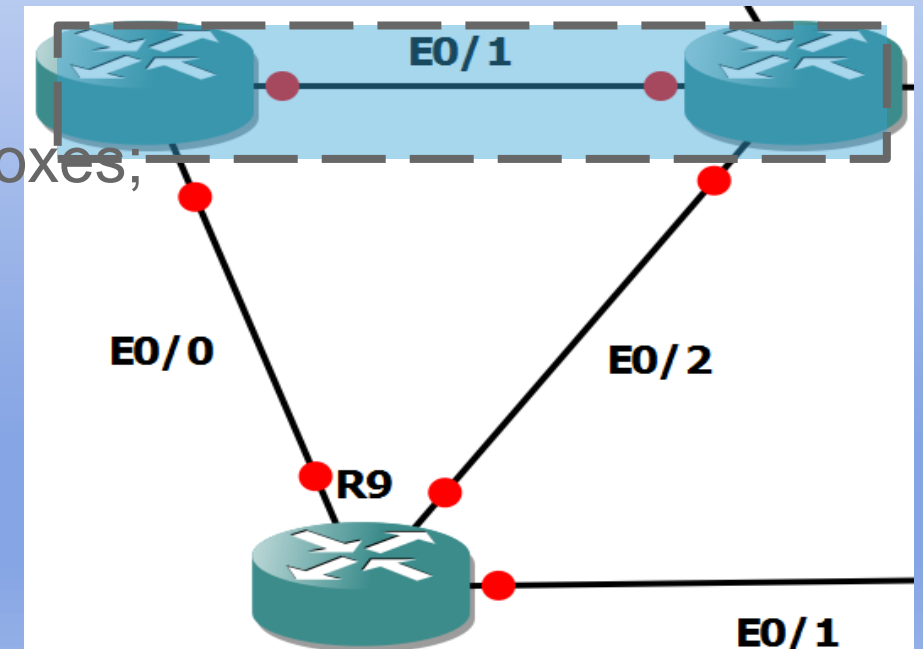


# OSPF - graceful restart capability /NSF

- KEY: data-plane is not broken during control-plane convergence;
- CEF kept intact, while process may restart information exchange;
- Typically useful for RP switchover;
- Modes: router / helper operations;

# Issue 6 – NSF in VSS or stack design

- If we do graceful restart in stack due to master going down;
- Peer may use dead link in CEF;
- NSF → freezes CEF → forwarding broken;
- Solution:
  - Avoid using L3 redundant links to different boxes;
  - LFA FRR (if supported by a platform).



# LFA FRR - Loop-Free Alternate Fast Reroute

- Allows to precompute backup path and build CEF adjacency;
- Fast reroute in case of link failure (without CEF change);
- Analog of EIGRP feasible successor;
- Possible in OSPF due to “link-state” nature of LSDB (possible to run SPF on behalf no other neighbors to find “loop-free” backup path);
- LFA election process is adjustable (by default - avoids SRLG first);

[http://www.cisco.com/c/en/us/td/docs/ios-xml/ios/iproute\\_ospf/configuration/xe-3s/iro-lfa-frr-xe.html](http://www.cisco.com/c/en/us/td/docs/ios-xml/ios/iproute_ospf/configuration/xe-3s/iro-lfa-frr-xe.html)

# Remote LFA FRR

- Allows to precompute backup path and build CEF adjacency;
- Fast reroute in case of link failure (without CEF change);
- Works in case, when LFA could not find one-hop “feasible successor”  
→ uses multi-hop successor;
- Runs on top of MPLS and requires LDP;

[http://www.cisco.com/c/en/us/td/docs/ios-xml/ios/iproute\\_ospf/configuration/xe-3s/iro-xe-3s-book/iro-ipfrr-lfa.html](http://www.cisco.com/c/en/us/td/docs/ios-xml/ios/iproute_ospf/configuration/xe-3s/iro-xe-3s-book/iro-ipfrr-lfa.html)



**CISCO**

*TOMORROW starts here.*