OSPF protocol

Successfully pass the free certification exam at IW Academy and become an Infinet Certified Engineer.

To the certification exam

Table of contents

Terminology

- **ABR** - a router located at the border of an OSPF area.
- **ASBR** - a router located at the autonomous system border and connected to an external network.
- **DR** - designated router.
- **BDR** - backup designated router.
- **LSA** - link state advertisement.
- **LSDB** - link state database.
- **DBD** - Short description of the LSDB.
- **LSR** - link state advertisement request.
- **LSU** - link state update, reply on LSR.
- **LSAck** - acknowledgment upon receiving an LSU.

The OSPF protocol

OSPF (Open Shortest Path First) - a dynamic routing protocol based on an algorithm that constructs a shortest path tree. The OSPF protocol has the following features:

- OSPF was developed by the IETF community in 1988. Since it is an open protocol, it can be used in heterogeneous networks built using equipment from different manufacturers.
- Today, two versions of the OSPF protocol are relevant: version 2 for IPv4 networks, described in RFC 2328, and version 3 for IPv6 networks, described in RFC 2740. The InfiNet devices support the operation of the IPv4 protocol, therefore, in this article only OSPF version 2 will be described.
- OSPF is a link state dynamic routing protocol.
- OSPF is an internal routing protocol, i.e. used to exchange routing information within an autonomous system (AS).
- The OSPF service messages are encapsulated in IP packets. The upper layer protocol field is set to 89.
  Two multicast addresses are reserved for OSPF: 224.0.0.5 and 224.0.0.6. These addresses are described below (see setting up neighborhood relations and DR and BDR selection algorithm).
- The distance value for the OSPF protocol is 110.

OSPF area

The number of autonomous system routers that use OSPF to exchange routing information can be large. This leads to a high load of the communication channels because of the large number of OSPF service messages. To reduce the amount of the transmitted service information, the OSPF protocol divides the autonomous system into areas.

Each area has a 32-bit identifier, which is usually written in two formats:

- four octet format: used in the device's configuration. For example, areas 0 and 2 (Figure 1a) will be written as 0.0.0.0 and 0.0.0.2 when configuring the devices;
- number format: used in schemes (Figure 1a-b) in order to make it easier to understand and easier to remember.

It is not necessary to use sequential identifiers for the areas. For example, the network can include areas with the identifiers 0, 2 and 7 (Figure 1a).

An interface belongs to an area, not the device itself. Thus, one router can be connected to multiple areas through its interfaces (Figure 1a).

The area with the identifier 0.0.0.0 has a special role - this area is called the backbone area. The backbone area is a requirement for the OSPF operation. Each area must be directly connected to the backbone area, i.e. a scheme in which some area is connected to another one without having a direct connection to the backbone is prohibited (Figure 1b).

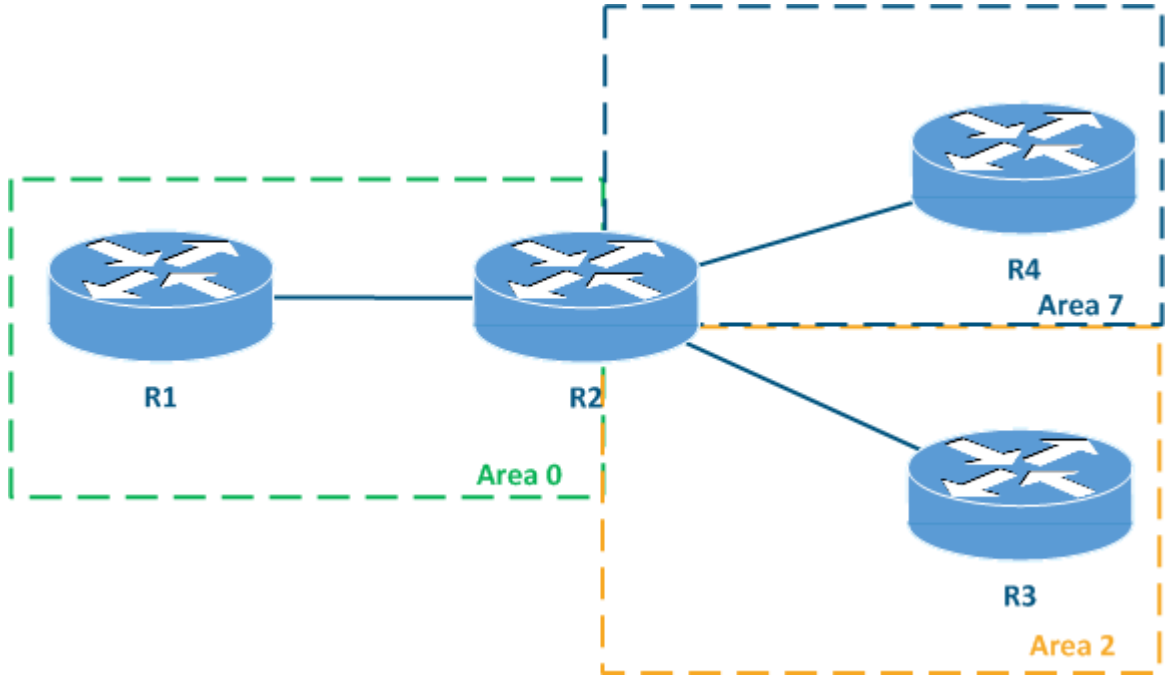

Figure 1a - Allowed network scheme with multiple OSPF areas

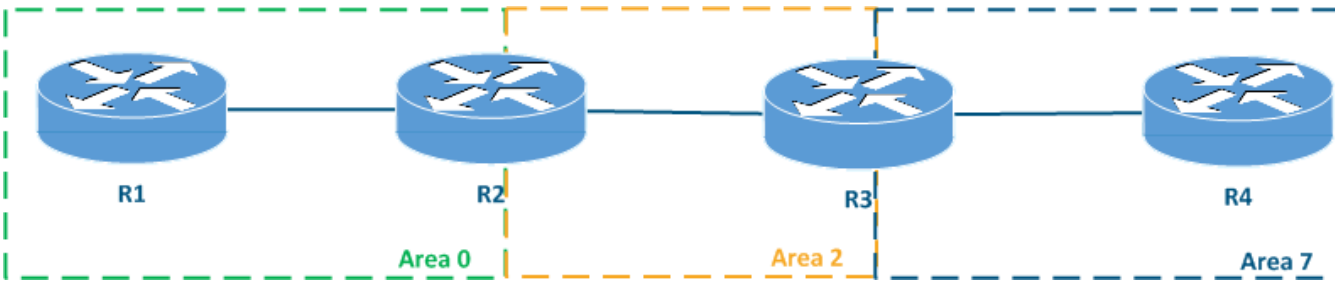

Figure 1b - Prohibited network scheme with multiple OSPF areas

Router types

Depending on the router's place in the network, the following types of devices are distinguished (Figure 2):

- **Internal router** (IR): a router which has all its interfaces associated with the same area. Routers R2 and R4 are internal.
- **Backbone router** (BR): a router with an interface connected to the backbone area. Routers R1, R2 and R3 are backbone routers.

- **Area border router** (ABR): a router having interfaces associated with different OSPF areas. Router R3 is ABR because it is located at the border of areas 0 and 2.
- **‚Autonomous system border router** (ASBR): a router connected to an external network. Router R1 is ASBR because it is connected to a third party LAN.

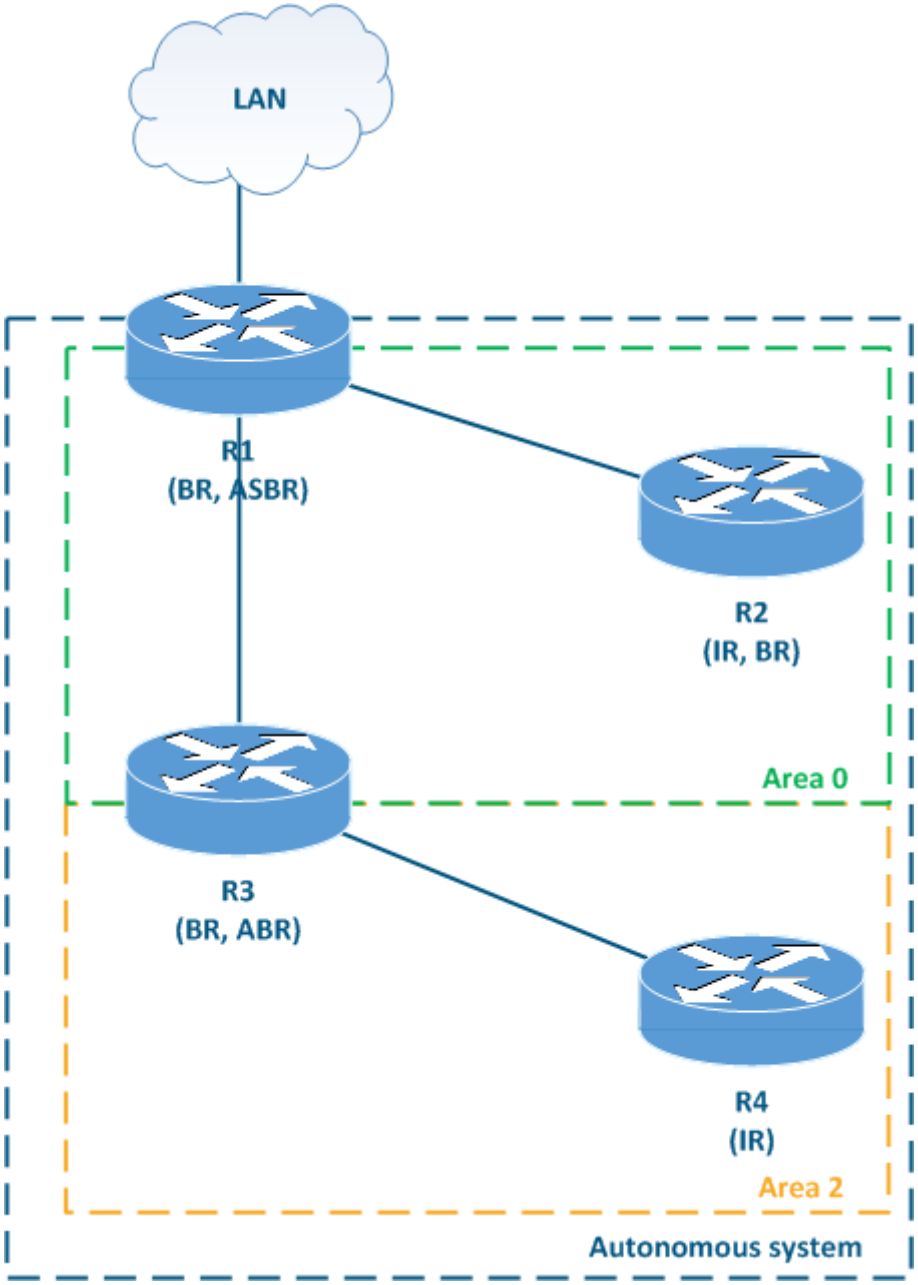

Figure 2 - Network scheme with different router types

OSPF's operation

The OSPF process follows the below steps. Some steps will require a detailed explanation which is provided in the sections below.

- **Step 1:** OSPF protocol launching. The configuration of the devices includes a list of interfaces that will participate in the OSPF's protocol operation, associated with the area identifiers to which these interfaces are connected. Upon this configuration, OSPF is launched.
- **Step 2:** Setting up neighboring relations. The device makes an attempt to find other routers and establish neighboring relations using the list of interfaces defined in step 1.
- **Step 3:** Role distribution. To reduce the service traffic volume in the broadcast network segments, a designated router (DR) is elected, which will be the central point for routing information exchange inside the broadcast segment.
- **Step 4:** Link state database (LSDB) synchronization. OSPF requires that each router has the same set of routing information, which implies the synchronization of the link state databases.

- **Step 5:** Building the shortest paths tree (SPT). Dijkstra's algorithm is applied to the routing information obtained in step 4 in order to build the shortest path tree. The root of the tree is the device on which the algorithm is running and the branches are the known destination networks, obtained from the other routers. Thus, each device has a set of paths to each network, optimized using the metric.
- **Step 6:** Export of the routes to the FIB. The set of routes obtained in step 5 is stored in the RIB, so that the device can perform additional optimizations by comparing the Distance values for the routing information obtained from different sources. The best routes obtained during the comparison are placed in the FIB and used to transfer the user and the service data.
- **Step 7:** Continuous monitoring of the network's state. Dynamic routing protocols perform a constant link state monitoring, because the routing table of all the devices must be kept up to date.

OSPF protocol launching

Two actions are performed when the OSPF service is launched: the selection of the router identifier and the definition of a list of interfaces that will participate in OSPF.

The router has a 32-bit identifier, which is usually written in the IP address format. Usually, the identifier is not related with the device's IP address and can be set manually. If the identifier is not set manually, it will be automatically selected as the highest IP address of the device. In case of manual setting of the ID, it is recommended to set it in the IP address of the loopback0 interface. This will help to identify the devices easier and to speed up the diagnostic of the network problems.

During the automatic router ID selection, the InfiNet device generates a special address from the 224. *. *. * multicast subnet, associated with the router's serial number. This helps to avoid the necessity of redefining the router ID when the IP address or the network interface are removed.

The set of interfaces that will take part into the OSPF protocol is determined based on the following rules:

- the range of IP addresses (or subnet) and their association with a specific area are specified in the configuration of the device (router);
- the network interfaces having IP addresses included in the specified range will take part into the OSPF process and become associated with the specified area. Note: not only the IP address of the interface is checked to see if it included in the specified range, but the whole network associated with the interface (see the example below).

If OSPF has not been started on a network interface, this does not mean that the network associated with this interface will not be advertised to the other routers. Launching OSPF on an interface only impacts the discovery of the neighbors.

Let's take a look at some examples of how to start the OSPF service on router R1 (Figure 3). The table below contains the configuration commands and their correspondence with the router's interfaces; if a match is found, a neighbor discovery process will be performed on the interface.

Command	Correspondence to eth1	Correspondence to eth2	Description
network 0.0.0.0/0 area 0	yes	yes	The 0.0.0.0/0 network includes all IP addresses, so the networks associated with eth1 and eth2 are in this range. Such a configuration has a hidden behavior: if a new IP address appears in the device's configuration, then OSPF will be launched on the interface associated with it. This is because the 0.0.0.0/0 network includes all the networks.
network 10.10.30.0 /24 area 0 network 192.168.6.0 /28 area 1	yes	yes	The command contains the networks associated with the eth1 and eth2 interfaces, so OSPF will use both interfaces.
network 10.10.30.0 /25 area 0 network 192.168.6.0 /28 area 1	no	yes	Although the IP address of the eth1 interface of R1 belongs to the 10.10.30.0/25 network, OSPF will not be launched on this interface. This is because the network associated with interface eth1 contains addresses in the range 10.10.30.0-255, which is not fully included in the 10.10.30.0/25 (10.10.30.0-127) network range. OSPF will be launched only on eth2.

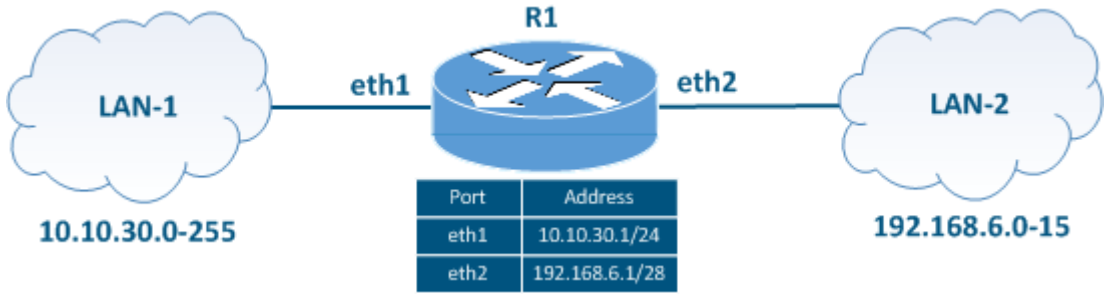

Figure 3 - Router with two network interfaces

Passive interfaces

After the router has determined the list of interfaces where OSPF is running, it starts looking for neighbors connected to these interfaces. Besides that, all the networks assigned to these interfaces will be advertised to the other routers. This behavior can be exploited by an attacker: the router will establish neighboring relations with the attacker's device and will transmit all the routing information about the network.

This type of attack can be prevented by using passive interfaces. Any interface participating in OSPF can be configured as passive. In this case, the search for neighbors via such an interface will not be performed, however, the network assigned to this interface will be advertised to the other routers.

External routes

The list of networks that are assigned to the interfaces is defined when OSPF starts. in addition, OSPF can advertise routes towards other networks, that were added to the device's routing table. The announcement of such routes is called redistribution. These routes are external to OSPF.

The routing sources for redistribution can be other dynamic routing protocols, static entries or directly connected networks not added to OSPF.

Setting up neighboring relations

The routing information exchange is possible only after the establishment of the neighboring relations between the routers. Two routers having a common link will establish a neighboring relation if the following parameters match:

- the network address of the interface towards a potential neighbor;
- MTU value on the interfaces towards a potential neighbor;
- area ID and area type;
- authentication parameters;
- Hello messages interval and Router dead interval (see step 1 of setting up neighboring relations).

The neighboring relations are established in several steps. Let's look at an example (Figure 4a): the network consists of three routers R1, R2 and R3 connected to the switch. Neighboring relations are established between the routers (the R2 router is selected as the designated router (DR), R3 as the backup designated router (BDR)). Router R4 will be added to the network scheme and let's assume that the conditions for establishing neighboring relations are met.

- **Step 1:** The R4 router sends Hello messages to the multicast address 224.0.0.5 (Figure 4b). This address is supported by all the devices running OSPF. Hello messages are sent from all the interfaces defined during the OSPF launching with a specified periodicity. The default Hello message broadcast interval is 10 seconds. Hello messages are an indicator of the connection with the neighbor, therefore, if no Hello messages are received from the neighbor during the Router dead interval, the device is marked as unavailable. By default, the Router dead interval is equal to four Hello message intervals.
- **Step 2:** The R1, R2 and R3 routers receive the Hello message from R4 and add it to the list of neighbors with the Init status (Figure 4b).
- **Step 3:** According to the internal timers, the R1, R2, R3 routers send Hello messages to router R4 (Figure 4c). Since the Hello messages contain a list of neighbors, the messages sent to R4 contain its ID. This means that router R4 can add all therouters to the list of neighbors with the 2-Way status, skipping the Init status. Then R4 will generate Hello messages for the routers, where it will indicate routers R1, R2 and R3 as neighbors, which will allow R1, R2 and R3 to change the status for R4 from Init to 2-Way (Figure 4d).
- **Step 4:** in broadcast segments (Ethernet, MINT, etc.), a primary router (DR) and a backup router (BDR) must be elected. The rest of the routers will be set with the DROther role. This mechanism is intended to reduce the amount of the overhead traffic: each DROther will exchange routing information only with the DR and the BDR. The DR and the BDR election algorithm is describer below. Note that the roles are not assigned to a device, but to an interface, so a router that has multiple interfaces in different broadcast segments may be DR in one segment and DROther in the other.
    - **Step 4a:** let R2 be DR and R3 - BDR as it was before R4 has been added to the network. The routers R1 and R4 have the DROther role, so the status between them will remain 2-Way.
- **Step 5:** The pairs of routers R2-R4 and R3-R4 distribute the roles of master and slave among themselves, the status of their relation becoming ExStart.
- **Step 6:** The Master device first begins the exchange of service messages with a brief DBD route database description. During the exchange of such messages, the relation's status is set to Exchange.
- **Step 7:** The devices receive the short route database description from the neighbor and generate requests for detailed information about the unknown networks. These messages are called LSRs.
    - **Step 7a:** An LSU is the answer to the LSR. LSUs contain detailed information about the requested routes.
    - **Step 7b:** The device receiving an LSU will generate an acknowledgment of the received information. This message is called LSAck.

- **Step 7c:** The routing information base containing all the gathered routing information is called LSDB and the exchange of LSDB service messages changes the relation's status to Loading.

- **Step 8:** After the LSDB is synchronized on all the devices, the relationship between routers R4-R2 and R4-R3 is set to the Full status (Figure 4e). Note that the DR and BDR establish Full relations with all the routers in the segment.

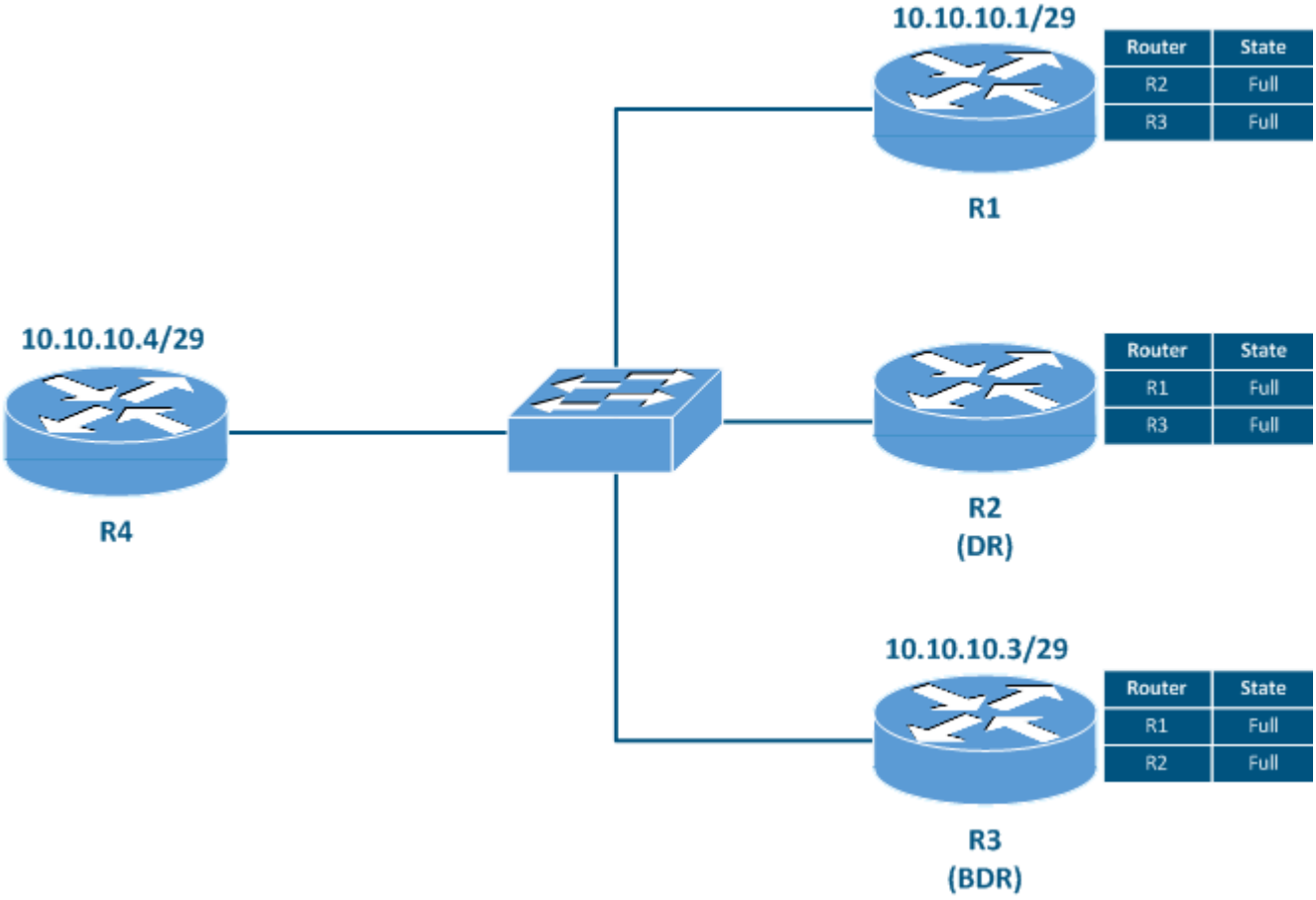

Figure 4a - The R4 router was added to the network scheme

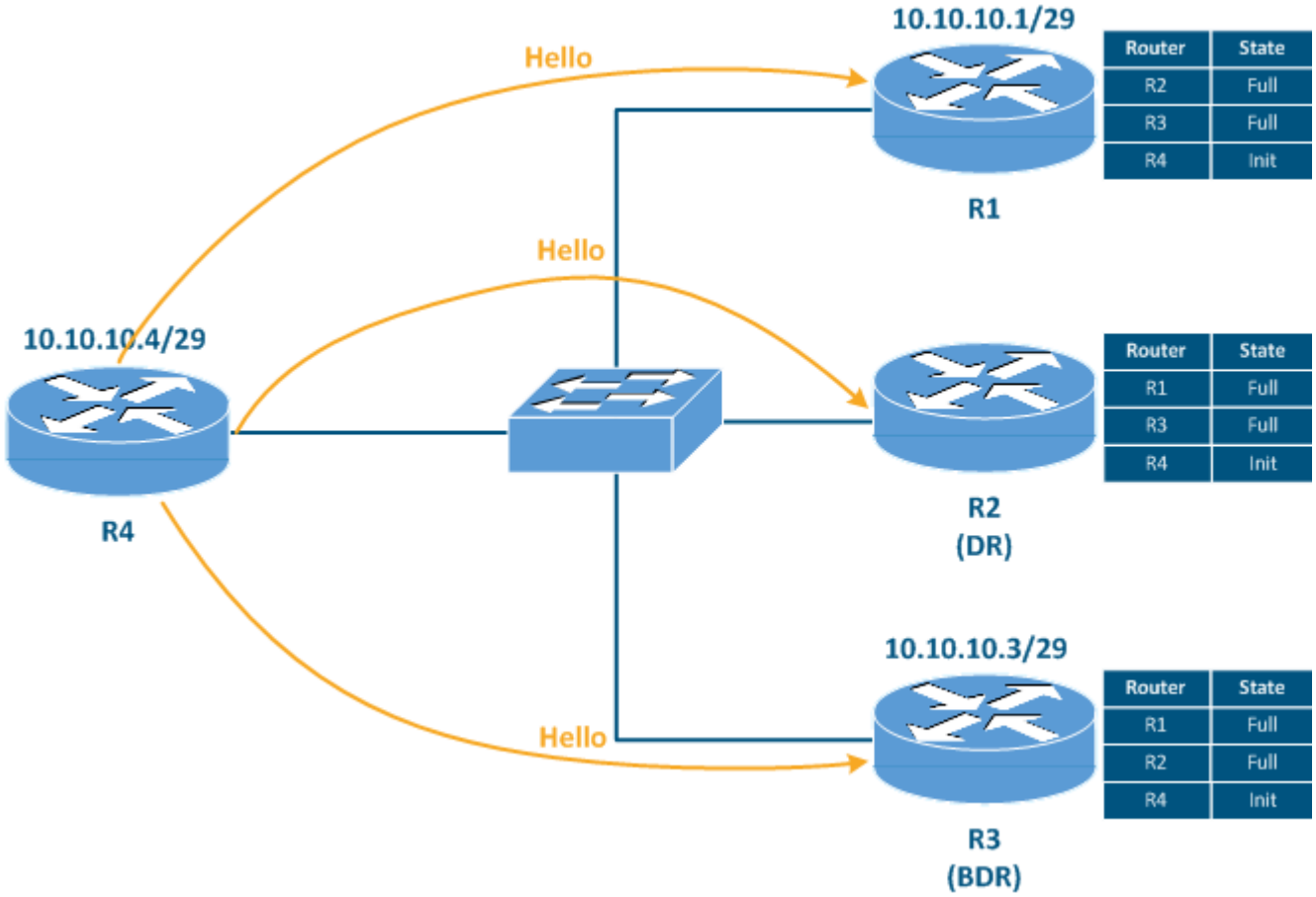

Figure 4b - R4 sends Hello messages

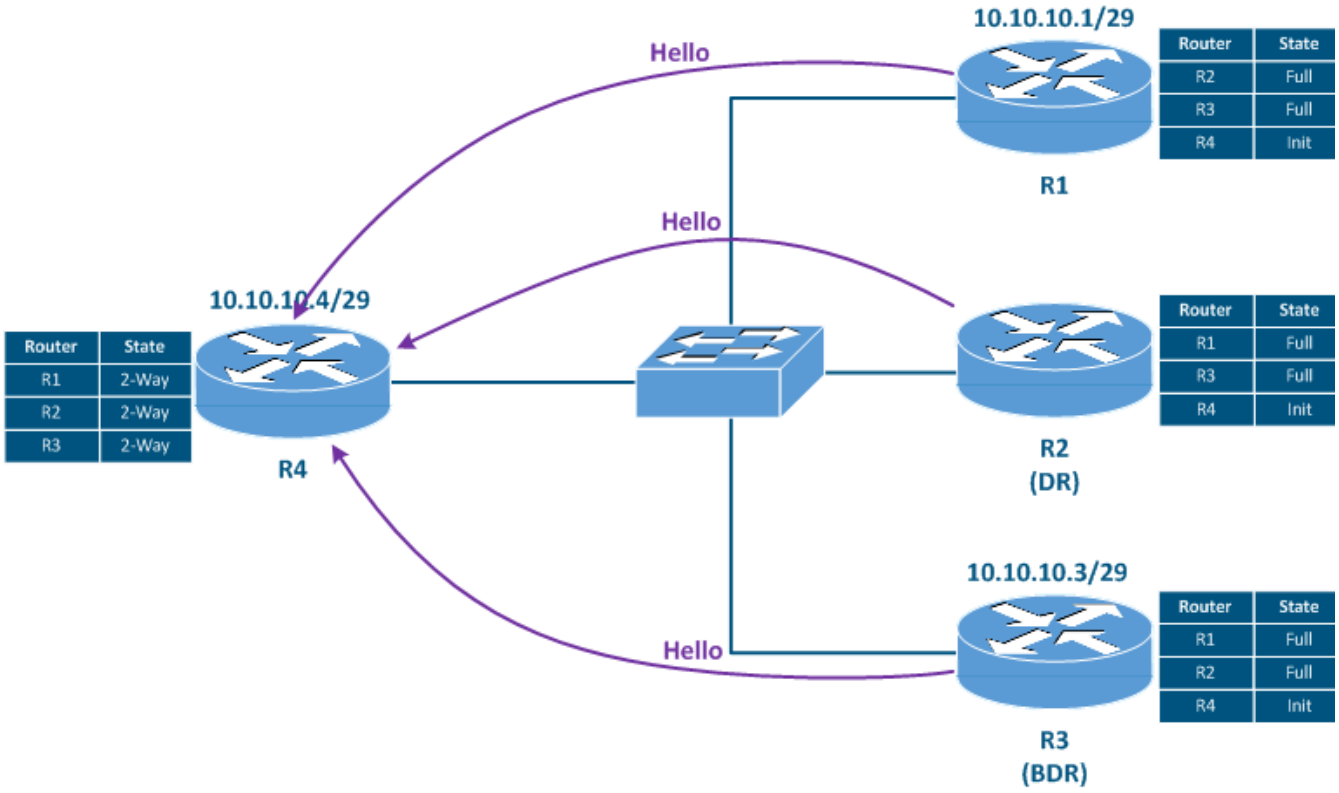

Figure 4c - R1, R2, R3 send Hello messages

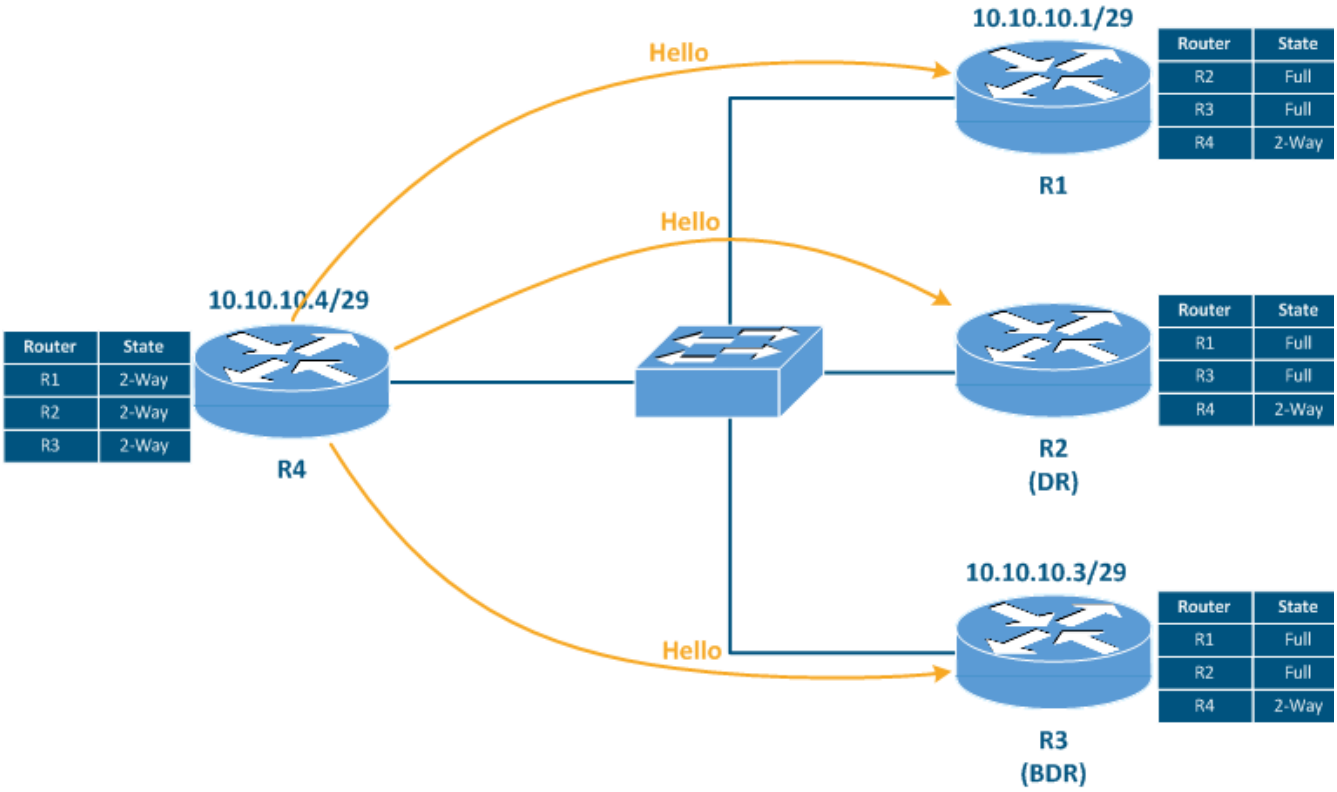

Figure 4d - 2-Way relations were established

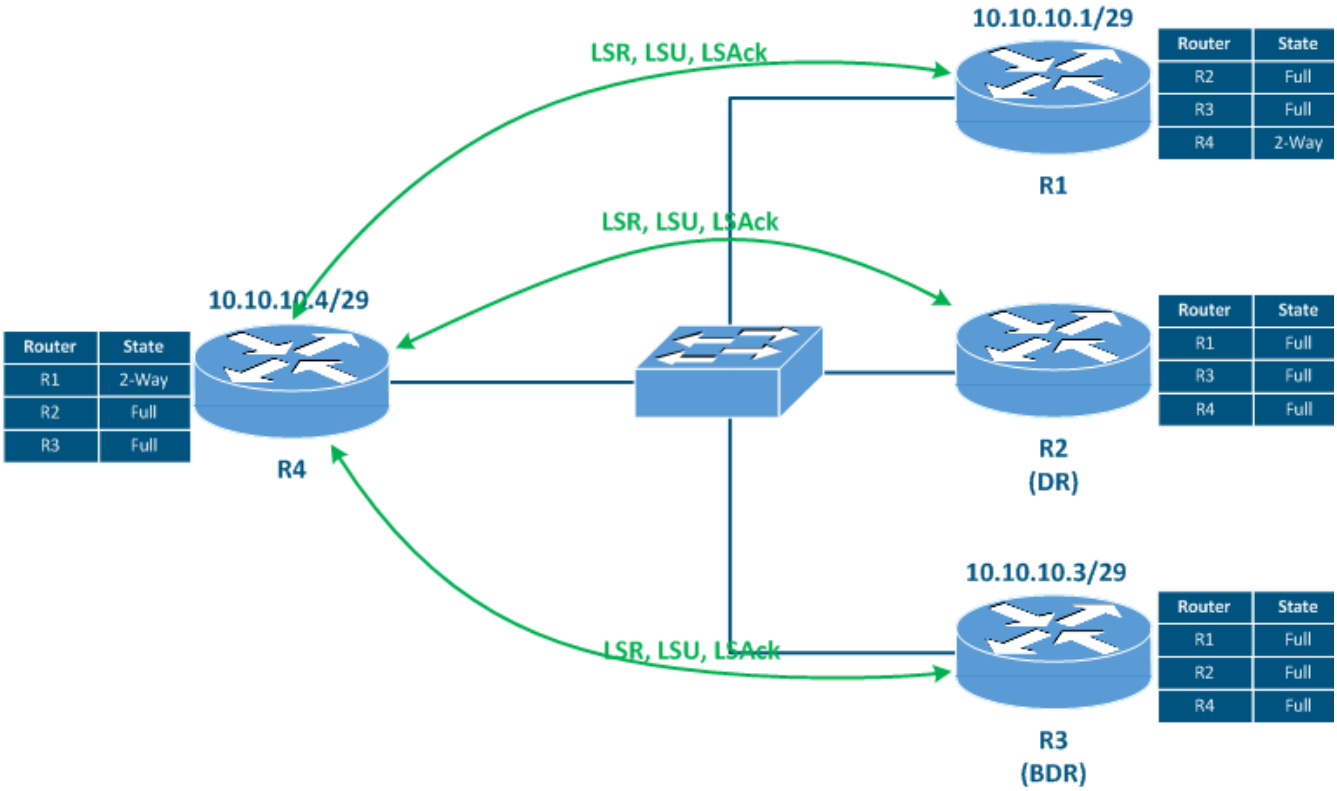

Figure 4e - Full relations were established by R4 with the DR and the BDR

Role distribution

In each broadcast segment where OSPF is running, DR and BDR elections are performed. The elections are carried out according to the following criteria:

- **Interface priority value:** the DR is the router with the highest priority value, the BDR is the router following the DR in priority value, the DROthers are the remaining routers. The priority parameter is configured on the interface that is connected to the broadcast segment. The priority is set manually by the network administrator and it can be in the range from 0 to 255. By default, the priority is 1. If the router interface priority value is set to 0, then that router does not participate in the DR and BDR elections.
- **Router-id value:** The DR is the router with the highest Router-id value, the BDR is the router following the DR in the Router-id value, the DROthers are the remaining routers. The Router-id is unique, so the router ID comparison is used when the priorities are equal, which ensures the distribution of the roles.

The group address 224.0.0.6 is associated with the DR and the BDR devices, which is used for LSDB synchronization. The devices having DR and BDR roles establish a Full relation with each router in the broadcast segment and require higher device performance demands compared to the DROthers. Since the device's hardware performance can become a bottleneck, it should be taken into account during network planning. Interface prioritization should be set in order to ensure a predictable selection of the highest performing devices as DR and BDR.

The main function of the DR is the routing information exchange in the broadcast segment. The main function of the BDR is to monitor the DR's state and, if it fails, to change the role to DR. Since each DROther establishes a Full relation with both the DR and the BDR, the LSDB on the BDR is synchronous with the DR, so the BDR can start performing the DR's functions without database synchronization timing delays. If the BDR becomes DR, then the BDR is selected among the DROthers according to the algorithm described above.

LSDB synchronization

The routing information in OSPF is categorized in different types of LSAs. The LSDB is a set of LSAs. The LSA is not an OSPF service message, therefore, the DBD, LSR, LSU and LSAck messages are used to transmit it according to steps 6-8 used in the neighboring relations establishment algorithm.

OSPF version 2 which is supported by the WANFleX OS defines 7 LSA types as described in the table below. To explain the purpose of different LSA types, the network scheme in Figure 5a will be used: the network consists of 6 routers and three OSPF areas are defined. This scheme describes the LSA types generated by the devices regardless of the neighboring relations establishment stages.

Type	Name	Description	Example
1	Router LSA	This type of LSA is distributed by all the routers within the same area. The LSA contains the following routing information: • a description of all the routes to the networks included in this area; • the costs of the routes; • a list of routers inside the routing area, specifying the established neighboring relations.	This type of LSA is distributed by all the routers in the network (Figure 5b). This LSA type has the following features: • R3 will include only the 10.10.234.0/29 network in the LSA type 1 broadcasted in area 0, and the 192.168.36.0/24 network in the LSA broadcasted in area 36. This behavior is explained by the fact that LSA type 1 is designed to exchange information within a single area; • Router R5 does not generate an LSA type 1 with information about the external network 172.16.0.0/16; • The type 1 LSA generated by R4 will be received by R2 and forwarded to R1 with an increased metric value. Thus, the LSA type 1 is propagated over the entire area with metric increments, the rest of the parameters remaining unchanged.
2	Network LSA	This type of LSA is distributed by the DR within the same area. This LSA contains the following routing information: • the broadcast's segment network address; • the broadcast's segment network mask; • a list of routers with the established neighboring relations.	This type of LSA is generated only by the routers having a DR role - R1, R2, R3 and R4 (Figure 5c). Similar to LSA type 1, LSA type 2 is distributed across the entire area with metric value increments.
3	Summary Network LSA	This type of LSA is distributed by the ABR and contains a summary of the routes in one area, that it is intended to be sent through the interfaces included in a different area. LSA types 1 and 2 allow the router to build a topology of the area and calculate the data transmission paths. Type 3 LSAs are not sources of topology data, they only contain routing information about the neighboring areas. Thus, at the area borders, OSPF behaves as a distance vector protocol. The ABR generates one LSA type 3 for each network. The number of LSA type 3 messaged can be reduced by using route summarization.	This type of LSA is generated by ABR routers - R3 and R4 (Figure 5d). The R3 router generates the following LSA type 3 messages: • route to the 192.168.36.0/24 network of area 36 that will be sent to area 0 through the eth0 interface. • route to the 10.10.234.0/29 network of area 0 that will be sent to area 36 through the eth1 interface. • route to the 10.10.21.0/30 network of area 0 that will be sent to area 36 through the eth1 interface. This network information is taken from LSA types 1 and 2 received from R1; • route to the 192.168.45.0/24 network of area 45 that will be sent to area 36 through the eth1 interface. The route to this

			network is taken from the LSA type 3 received from router R4. When advertising this network, router R3 sets itself as the route source in the LSA. Source substitution is necessary, since the routers in area 36 are unaware of the R4's location. Router R4 generates LSA type 3 messages in the same way as R3.
4	ASBR Summary LSA	This type of LSA is distributed by the ABR in addition to LSA type 5. This type of LSA contains information about the ASBR location.	See the example for LSA type 5.
5	External LSA	This type of LSA is generated by the ASBR for the external routes, including default routes. Such messages are distributed throughout the AS unchanged. Similar to the ABR, the ASBR can summarize the external routes, i.e. replace several routes with one. This reduces the size of the routing table and the amount of service information during the route distribution.	LSA type 5 is generated by R5 as it is the only ASBR in the network scheme (Figure 5e). The type 5 LSA generated by R5 contains information about the 172.16.0.0/16 network which is distributed through the entire autonomous system unchanged. Thus, each router in the scheme receives the information that the 172.16.0.0/16 network is external and is available via the R5 router. The hidden problem is that R1, R2, R3, and R6 do not know R5's location. LSA type 1, where R5's ID is specified, is only propagated within area 45. To solve this problem, border routers R4 and R3 generate an LSA type 4 in addition to the transmitted LSA type 5. In the LSA type 4, the routers advertise that all the traffic directed to R5 can be sent to the ABR.
6	Group Membership LSA	The LSAs of this type are used in Multicast networks and contain a list of groups whose consumers are in the network segment. This type of LSA will not be described in this article.	-
7	Type 7 LSA	This type of LSA is similar to LSA type 5 and is used in NSSA areas. The use of LSA type 7 is necessary for the compatibility between Stub areas and NSSA areas. LSA type 7 is converted to LSA type 5 by the ABR during the export from the NSSA area.	An example of LSA type 7 generation is presented in the NSSA area description.

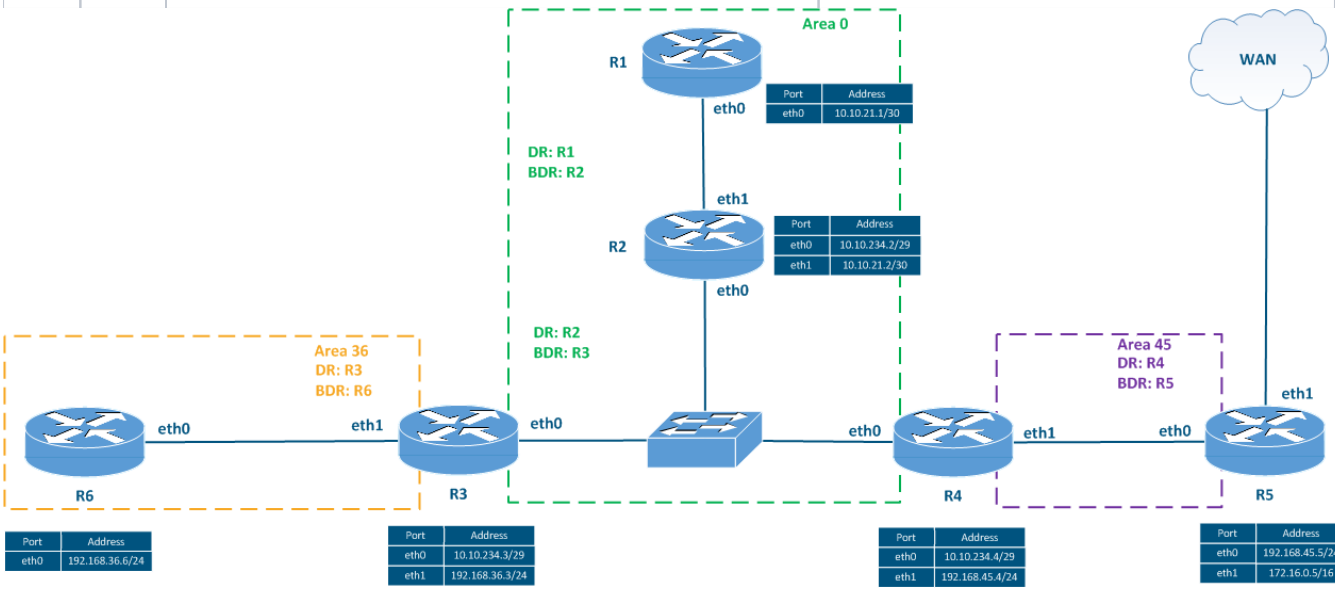

Figure 5a - Network scheme used for analyzing the LSA types

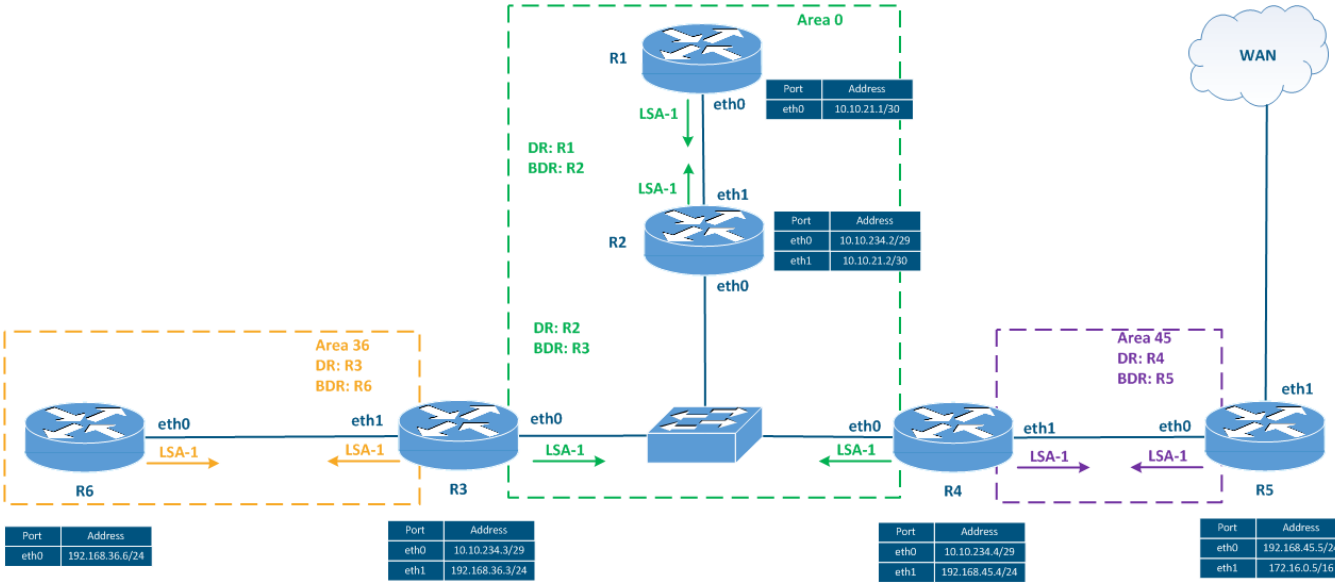

Figure 5b - Distribution of the LSA type 1

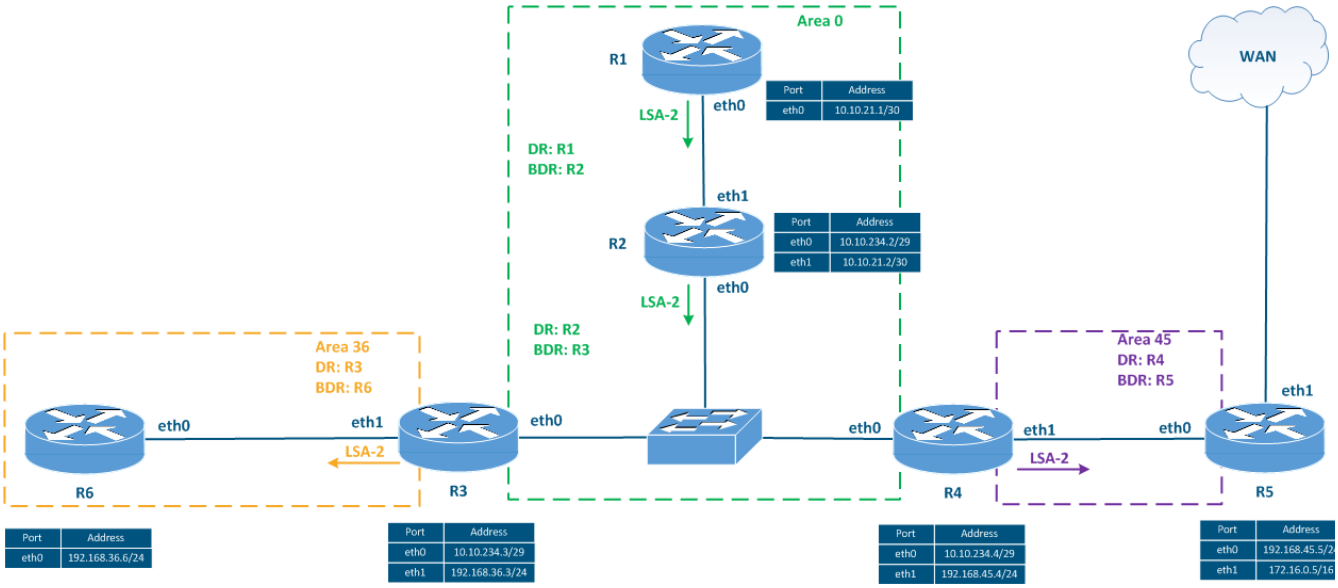

Figure 5c - Distribution of the LSA type 2

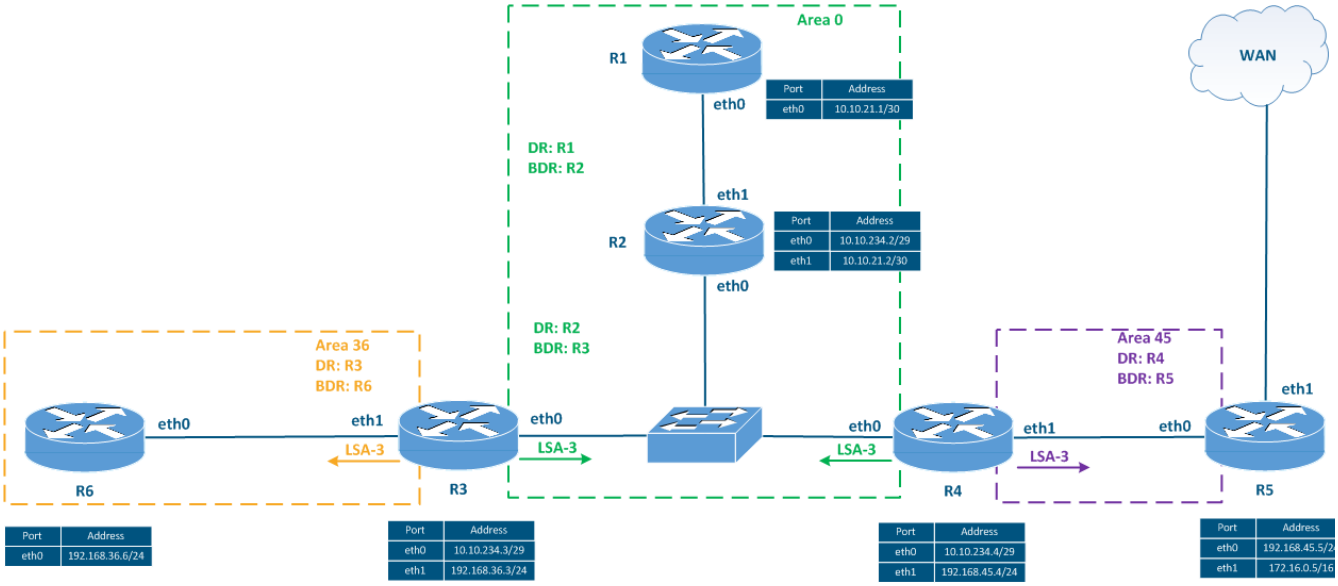

Figure 5d - Distribution of the LSA type 3

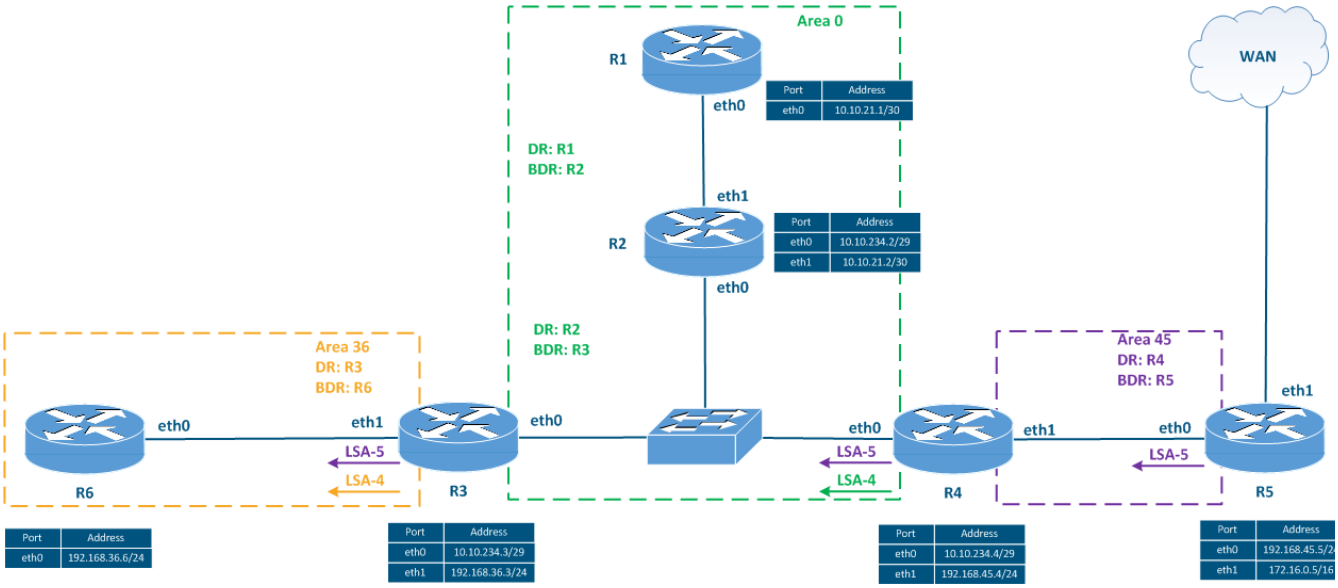

Figure 5e - Distribution of the LSA type 4 and of the LSA type 5

Building the shortest path tree

After the LSDB synchronization, each router performs a shortest path tree calculation using Dijkstra's algorithm.

In the networks having redundancy implemented, the LSDB contains announcements 1about the same network received from different sources. Such routes are transmitted to the RIB in the following order:

- **Intra-area routes:** routes distributed within the same area using LSA types 1 and 2.
- **Inter-area routes:** routes received from neighboring areas using LSA type 3.
- **Type 1 external routes:** routes to external networks received from the ASBR. The route metric for this type of routes is counted as the sum of the metric set by the ASBR during the announcement plus the metric of the path to the ASBR.
- **Type 2 external routes:** are similar to the type 1 external routes, with a different method for the metric calculation. The metric is equal to the value set by the ASBR during the announcement and does not include the path to the ASBR.
- **Metric value:** for two routes to the same network received from sources of the same type, the metric values are compared. The route with the lower metric value will be added to the RIB.

Area types

The way to reduce the OSPF service traffic volume is to use different types of areas. The OSPF protocol provides the following types of areas:

- Normal;
- Stub;
- Totally Stub;
- NSSA;
- Totally NSSA.

Let's look at the main features of different area types using the example in (Figure 6): routers R1, R2, R3 and R4 are connected in chain with each other, forming three OSPF areas. Routers R3 and R4 have external links. In each example, we will change the type of area 4 and analyze the LSA types associated with that area. In these examples, the LSA's not related with area 4 and the type 1 and 2 LSAs will be omitted because they are distributed within any area type.

Figure 6 - Network scheme used for the description of the area types

Normal

Normal areas do not change the propagation of the LSAs and the processing logic described above (Figure 7a). This area type is used by default. The backbone area is a special case of the Normal area.

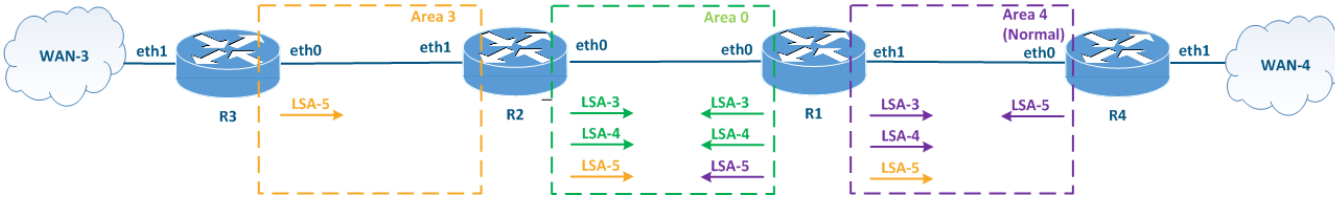

Figure 7a - LSA distribution in the Normal area

Stub

The Stub area is characterized by the following features (Figure 7b):

- The Stub area cannot have external links. Thus LSA types 5 and 4 are prohibited in the Stub area.
- The stub area's routing information is distributed to the neighboring areas using LSA type 3.
- LSA Type 3 messages about the networks in different areas are distributed in the Stub area, similarly to the Normal areas.
- When an LSA type 5 from a different area enters the Stub area, it is converted to LSA type 3 with the default route information.

Stub areas are used in LAN segments that have no connection with the external networks, but the routers in this area must receive full routing information from the neighboring areas. Using Stub areas allows to obtain a small increase in performance by reducing the LSA number and to protect the network from attacks from the external network segment.

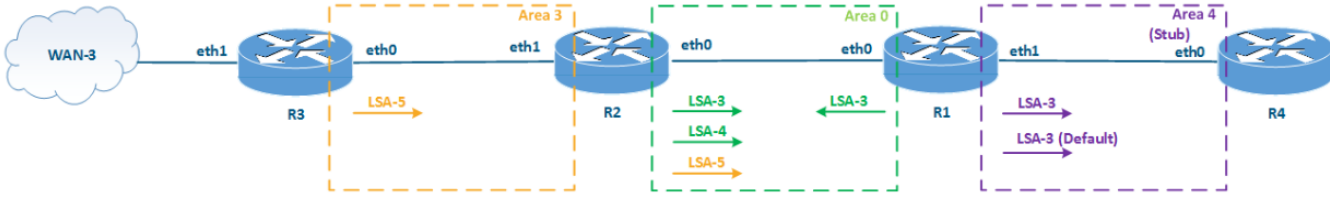

Figure 7b - LSA distribution in the Stub area

Totally Stub

The Totally Stub area behaves similarly to the Stub area with one exception: LSA's of types 3 and 5 from the neighboring areas are replaced with one LSA type 3 with a default route (Figure 7c).

Totally Stub area applications are similar to the ones of the Stub area, but the routers in a totally stub area will not have all the routing information about the neighboring areas. This offers a significant performance increase, as Totally Stub area routers will use a single default route to transmit data to the neighboring areas.

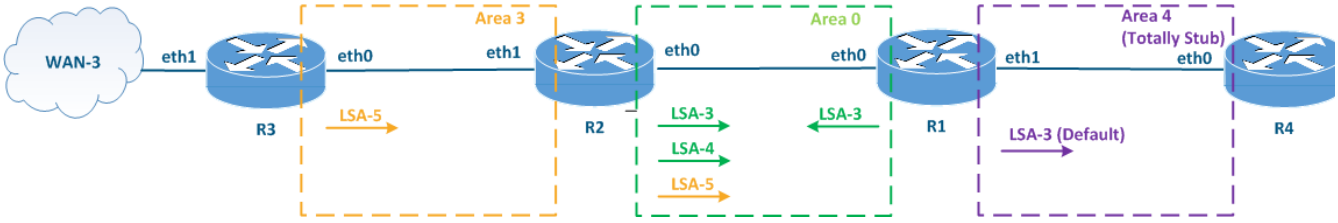

Figure 7c - LSA distribution in the Totally Stub area

NSSA

The NSSA area has characteristics similar to the Stub area with one exception: the NSSA can have an external link (Figure 7d). Since LSA type 5 which is used to distribute routing information about the external links, is prohibited in Stub areas, NSSA areas use LSA type 7 for this purpose. This LSA type has the same structure as LSA type 5, but it is permitted in NSSA areas. At the area border, the ABR converts the LSA type 7 to LSA type 5, setting itself as the routing source. Since the ABR performing the LSA conversion becomes the source, there is no need to generate an additional type 4 LSA.

Usually, the NSSA area usage is a result of the network's development: connecting an external communication channel to the Stub area requires changing its type to NSSA.

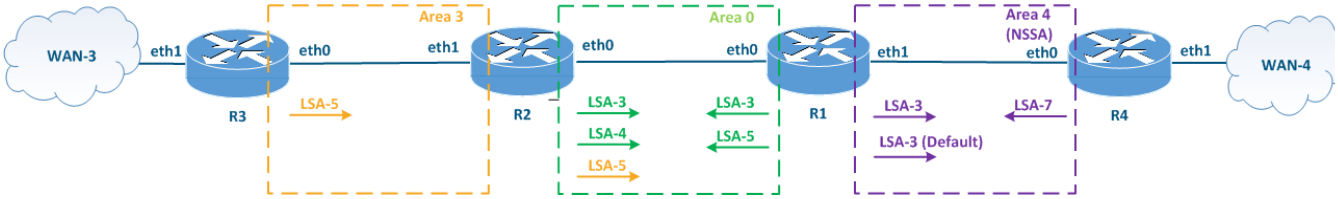

Figure 7d - LSA distribution in the NSSA area

Totally NSSA

Totally NSSA areas behave similarly to the NSSA areas with one exception: only one type 3 LSA with a default route is exported to the Totally NSSA area (Figure 7e).

Totally NSSA areas are a result of the network development: connecting an external link to a Totally Stub area requires changing the type of the area to Totally NSSA.

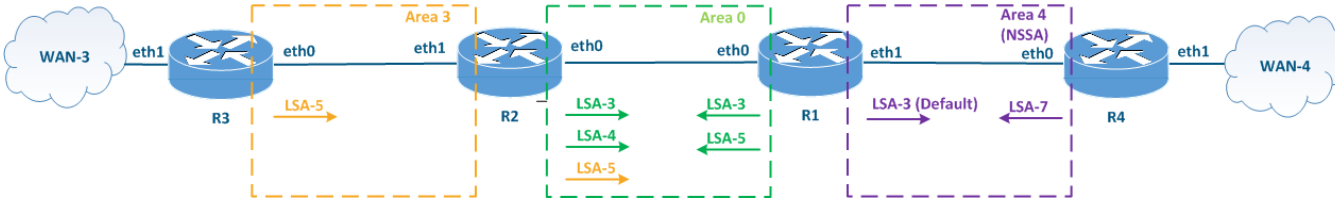

Figure 7e - LSA distribution in the Totally NSSA area

Virtual link

One of the OSPF principles is to always connect two non-backbone areas only through the backbone area. Despite that, as a result of the historical development, the structure of some networks does not match with this principle. Bringing such networks up to the OSPF backbone can be costly, so OSPF has been extended with the virtual link concept.

The virtual link has the following features (Figure 8):

- A virtual link is a logical connection configured on two ABRs, one of which is connected to the backbone area. Routers R1 and R2 are ABRs on which a virtual network interface is created and R2 is connected to the backbone area via the eth1 interface.
- The virtual link is the interface used by R2 to connect area 4. All LSA types are distributed over the virtual link like through a normal interface.
- The area that is common for two ABRs sharing a virtual link is called a transit area. In the example below, area 7 is the transit area.
- The transit area should have the Normal type. It is not possible to establish a virtual link through Stub or NSSA areas.

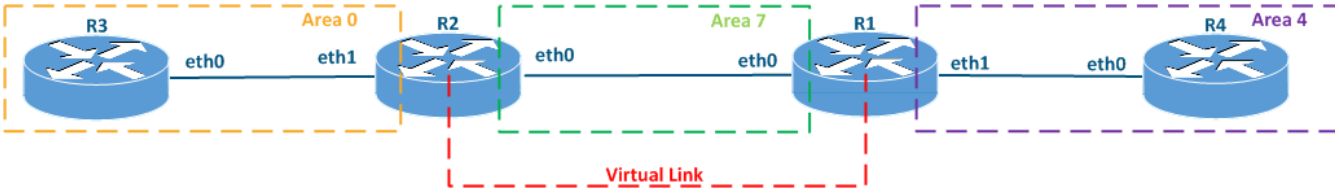

Figure 8 - Network scheme with a virtual link

OSPF's protocol features

The OSPF protocol features can be summarized as follows:

- **Open implementation:** OSPF is an open protocol, so it can be used by equipment from different manufacturers.
- **Easy configuration:** in small networks, the protocol can be started with only two commands.
- **Flexible configuration:** the wide protocol tool set allows to implement many network schemes.
- **Scalability, fault tolerance, balancing, efficiency:** similar to ODR, OSPF has the advantages of a dynamic routing protocol.
- **High entry threshold:** understanding the OSPF terminology and logic is time consuming.

OSPF practice

The examples on how to configure OSPF are present in the child page: OSPF protocol's configuration.

Additional materials

Webinars

1. Typical scenario of routing setting using Infinet Wireless devices. Part II

Other

1. Ifconfig command (interfaces configuration)
2. ARDA (Aqua Router Daemon)
3. OSPF command
4. arip command
5. rip command
6. RFC 2328