Routing
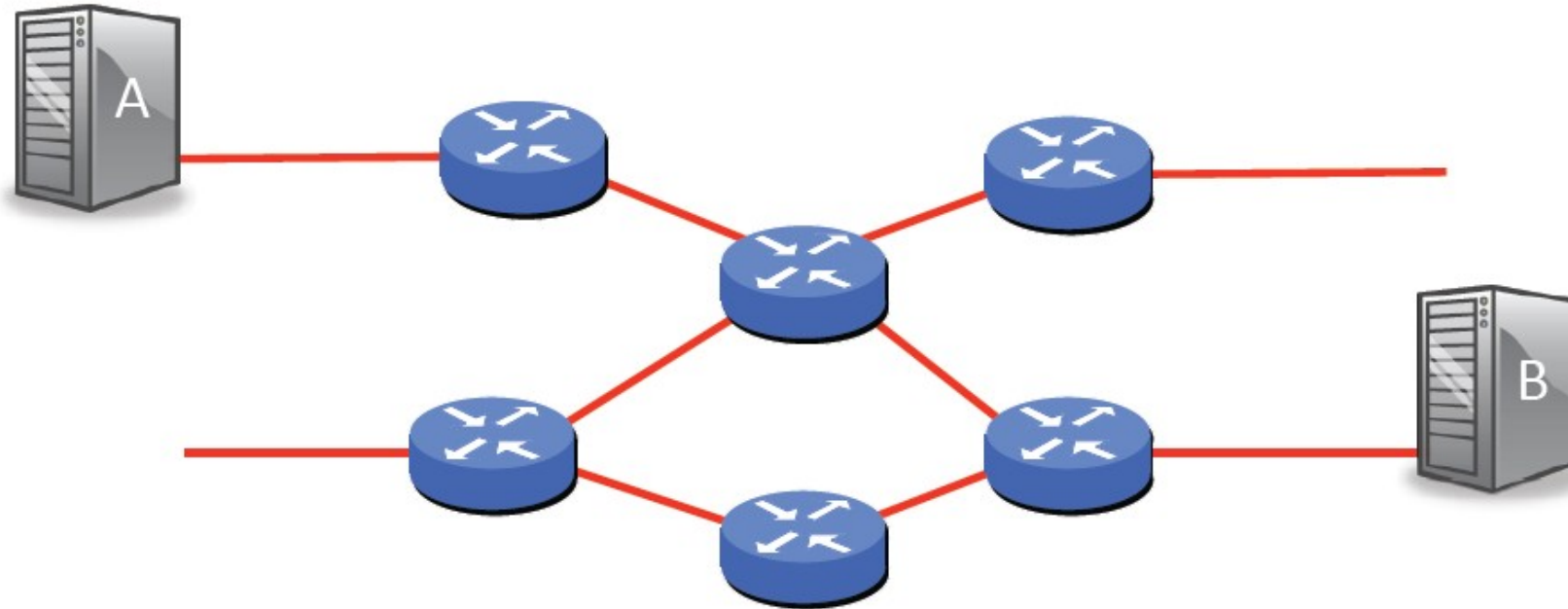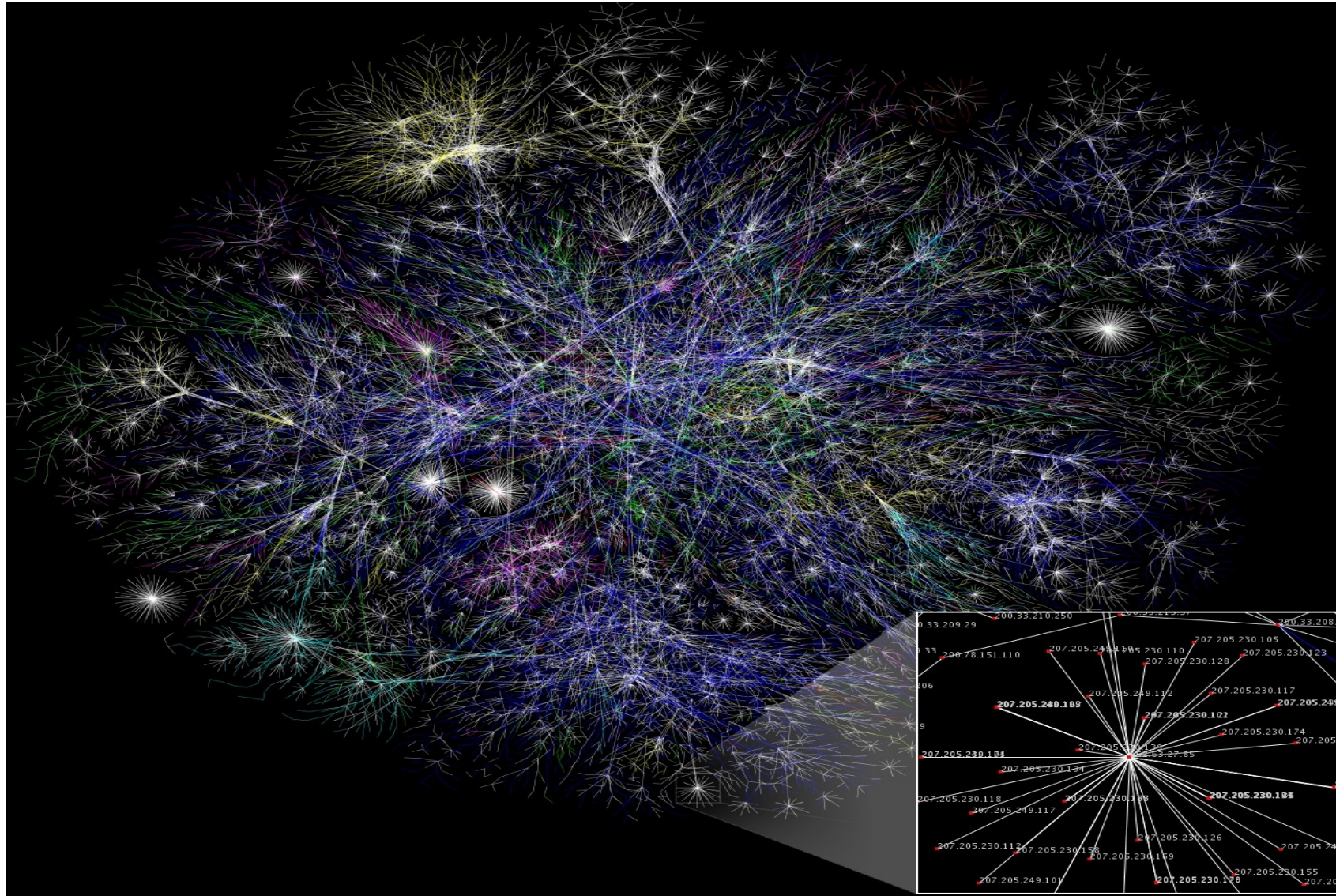
Problem

Who should determine how packets from A can reache B are ?

Route selection criteria?

What we can do with such a network ?

Autonomous system – AS

AS — *it is a system of IP networks and routers managed by one or more operators with a single routing policy.*

ASN — Autonomous system number

Globally, this is done by the IANA (Internet Assigned Numbers Authority), which delegates its tasks to the RIR (Regional Internet Registry) – i.e. regional organizations, each of which is responsible for a certain part of the planet (for Europe and Russia – it is RIPE NCC), which in turn delegates its tasks to the LIR (Local Internet Registry).
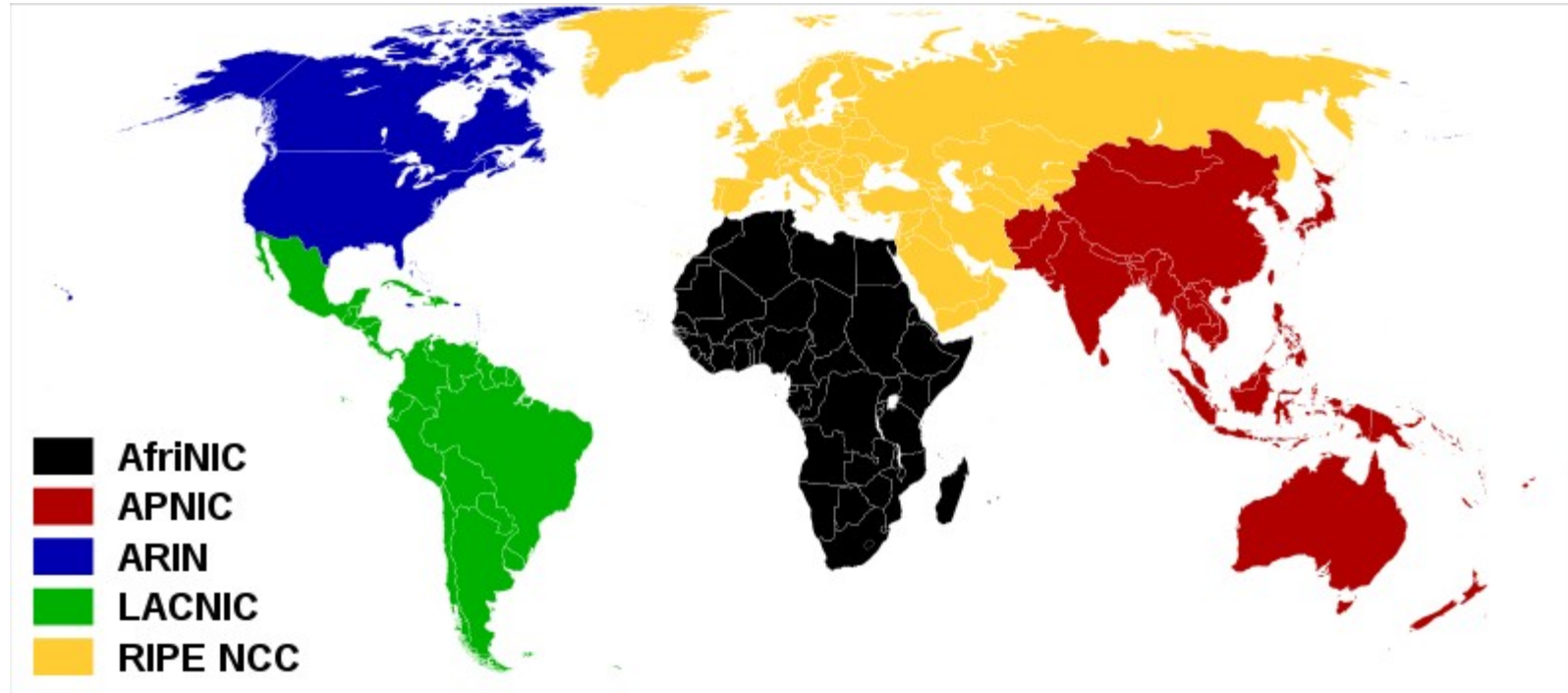
Autonomous system

Unit of hierarchy in the Internet

Inside the AS its owner decides how to route data flows

A Protocol BGP-4 (Border Gateway Protocol v4 RFC 1771) must be used between the speakers

How to find AS number?

http://whatismyipaddress.com/ip-lookup

AfriNIC
APNIC
ARIN
LACNIC
RIPE NCC

RIR & LIR

Almost any organization can become a LIR if it has the necessary documents.

At LIR'a we would take ASN (as number) - for example 64500. And he himself could have ASN 64501.

Until 2007, only 16-bit AS numbers were possible, that is, a total of 65536 numbers were available. 0 and 65535 are reserved.
Numbers 64512 through 65534 are for private AS that are not routed globally
Rooms 64496-64511 – for use in examples and documentation.

Now it is possible to use 32-bit AS numbers.

It is impossible to speak about Autonomous systems without binding to blocks of IP addresses. In practice, each AS must have some block of addresses associated with it.

Protocol OSPF

Routing algorithms

Routing involves two parallel process:
- **preparation of routing table**
- **forwarding of the datagrams (with using this table).**

Formation of the route table produces:
- **through the use routing protocol (dynamic routing)**
- **under the influence of instructions network administrator – (static routing.)**

Routing algorithms

Routing algorithms solve the problem of path selection.

- *unicast*
- *multicast.*
- *source routing*
- *«hot potato».*
- *.........*

Types of routing protocols

- *IGP (internal to your Autonomous system) - ISIS/OSPF/RIP/EIGRP.*
- *EGP (внешние) - BGP – Border Gateway Protocol.*

According to the types of algorithms used are divided into:
- *DV (Distance Vector)*
- *LS (Link State).*

Metrics

- *Typically, each segment that makes up a route, some value is assigned - to estimate this segment's.*
- *Each routing Protocol uses its own route evaluation system.*

Examples:
- *RIP – hop count*
- *OSPF – based on the link bandwidth*

IP Routing Table

```
<Quidway> display ip routing-table
Route Flags: R - relay, D - download to fib

-----------------------------------------------------------------------------------------

Routing Tables: Public
        Destinations : 6        Routes : 6

  Destination/Mask        Proto        Pre   Cost   Flags     NextHop          Interface

         1.1.1.1/32       Static        60    0        D        0.0.0.0          NULL0
         2.2.2.2/32              Static   60     0     D        100.0.0.2
Vlanif100
       100.0.0.0/24       Direct        0     0        D        100.0.0.1        Vlanif100
       100.0.0.1/32       Direct        0     0        D        127.0.0.1        Vlanif100
       127.0.0.0/8        Direct        0     0        D        127.0.0.1
InLoopBack0
       127.0.0.1/32       Direct        0     0        D        127.0.0.1
InLoopBack0
```
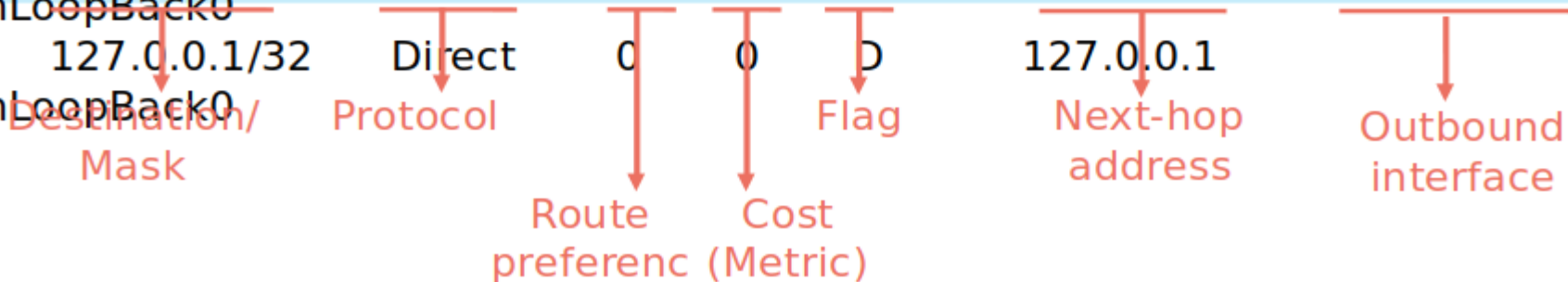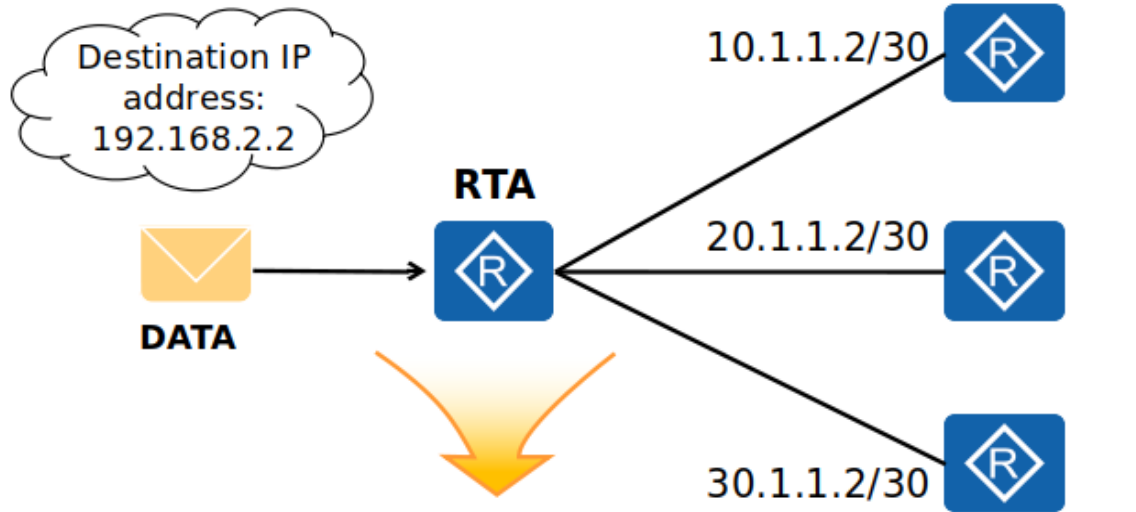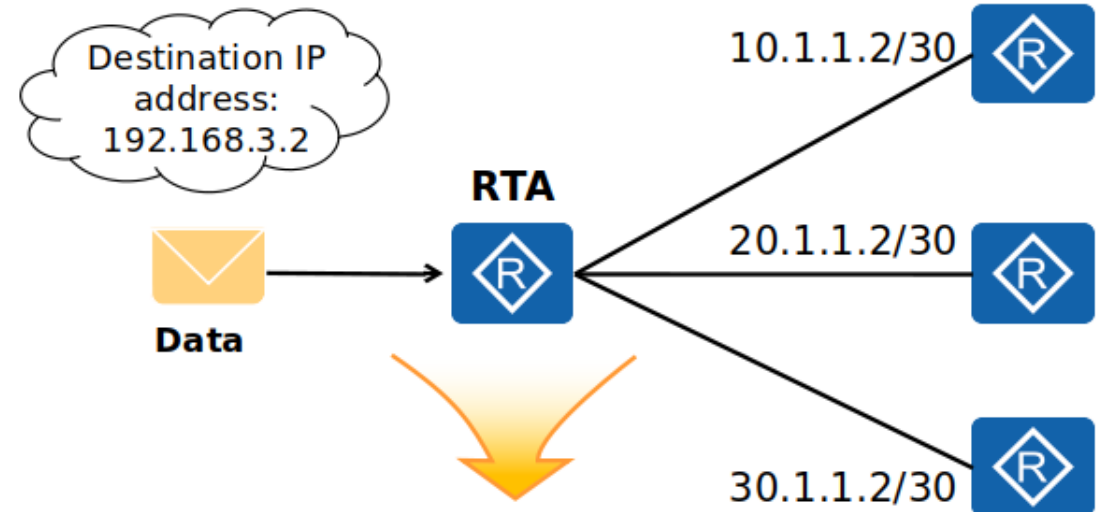
Destination/
Mask

Protocol

Flag

Next-hop
address

Outbound
interface

Route
preferenc

Cost
(Metric)

Example of Longest Matching

Destination IP address: 192.168.2.2

DATA

RTA

10.1.1.2/30

20.1.1.2/30

30.1.1.2/30

IP routing table of RTA

Destination/Mask	Next Hop
192.168.0.0/16	10.1.1.2
192.168.2.0/24	**20.1.1.2**
192.168.3.0/24	30.1.1.2

← Match

Destination IP address: 192.168.3.2

Data

RTA

10.1.1.2/30

20.1.1.2/30

30.1.1.2/30

IP routing table of RTA

Destination/Mask	Next Hop
192.168.0.0/16	10.1.1.2
192.168.2.0/24	20.1.1.2
192.168.3.0/24	**30.1.1.2**

← Match

Route-based Forwarding Process

Destination IP address: 40.0.1.2

Data

Destination/ Mask	Next Hop	Outbound Interface
20.0.1.0/24	20.0.1.2	GE0/0
30.0.1.0/24	30.0.1.1	GE0/1
10.0.1.0/24	20.0.1.1	GE0/0
40.0.1.0/24	**30.0.1.2**	**GE0/1**

IP routing table of R2

Gateway

10.0.1.0/24	GE0/1	GE0/0	20.0.1.0/24	GE0/0	GE0/1	30.0.1.0/24	GE0/0	GE0/1	40.0.1.0/24

10.0.1.1   20.0.1.1   20.0.1.2   30.0.1.1   30.0.1.2   40.0.1.1

R1   **R2**   **R3**

IP routing table of R1

Destination/ Mask	Next Hop	Outbound Interface
10.0.1.0/24	10.0.1.1	GE0/1
20.0.1.0/24	20.0.1.1	GE0/0
30.0.1.0/24	20.0.1.2	GE0/0
40.0.1.0/24	**20.0.1.2**	**GE0/0**

IP routing table of R3

Destination/ Mask	Next Hop	Outbound Interface
40.0.1.0/24	**40.0.1.1**	**GE0/1**
30.0.1.0/24	30.0.1.2	GE0/0
10.0.1.0/24	30.0.1.1	GE0/0
20.0.1.0/24	30.0.1.1	GE0/0

OSPF

- *OSPF (RFC 2328)*
- *line state changes are sent as needed*
- *each router uses Dijkstra's algorithm*
- *Can use authentication*
- *AS (AS in OSPF differs from AS in the Internet) can be divided in to areas*
- *analogue - IS-IS (RFC 1142)*

OSPF
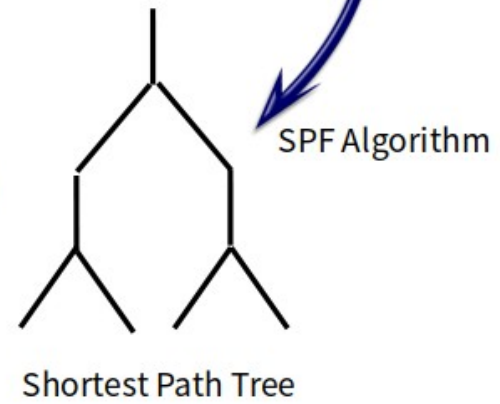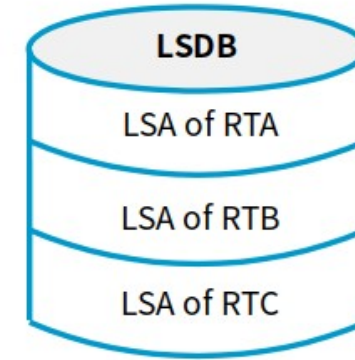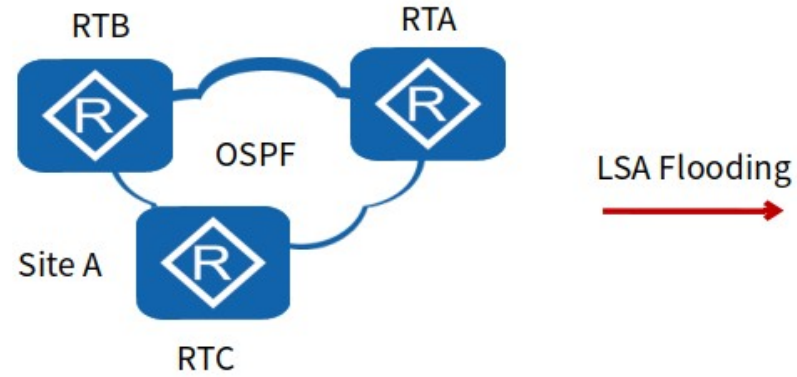
Stages of the protocol OSPF

1) Routers exchange hello packets across all interfaces on which OSPF is activated. Routers that share a common data link become neighbors when they agree on certain parameters specified in their hello packets.

2) In the next phase of the Protocol, the routers will attempt to move into an adjacency state with their neighbors (for the shared bus networks like Ethernet).

3) Each router sends the so called link state advertisements (LSA) to the routers with which it is in the adjacency state

Stages of the protocol OSPF (continuation)

4) *Each router that receives an advertisement from an adjacent router writes the information it transmits to the router's link state database and sends a copy of the advertisement to all other adjacent routers.*

5) *By sending advertisements within the same OSPF zone, all routers build an identical router link state database.*

6) *Each router independently runs Dijkstra's shortest path algorithm*

1) *Each router builds a routing table from its shortest path tree*

OSPF roughter types
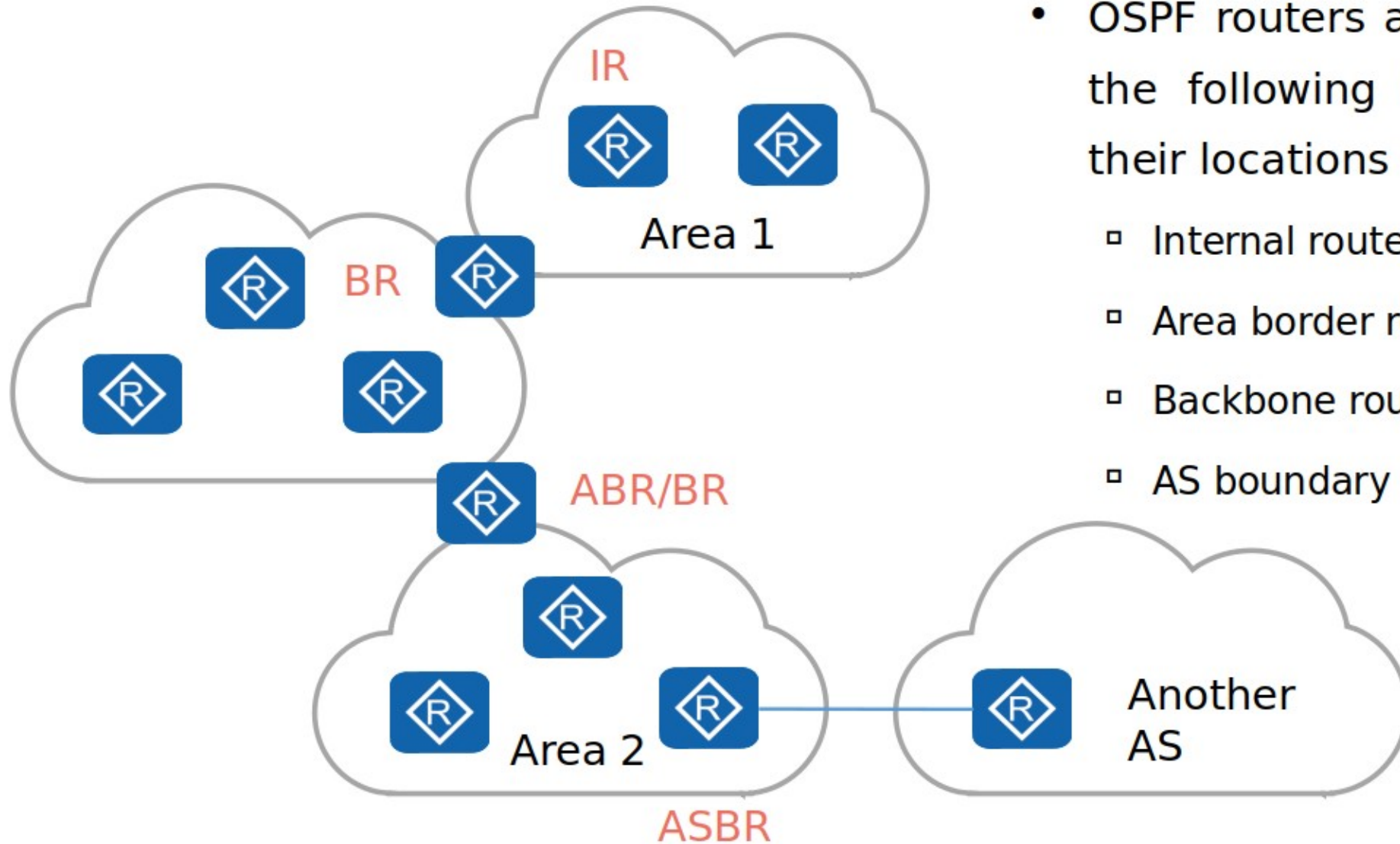
- OSPF routers are classified into the following types based on their locations or functions:
  - Internal router
  - Area border router (ABR)
  - Backbone router
  - AS boundary router (ASBR)

OSPF roughter types

- ***Internal router*** *— a router whose interfaces all belong to the same zone. These routers have only one link-state database.*
- ***Area border router, ABR*** *— connects one or more zones to the backbone zone and acts as a gateway for cross-zone traffic.*
- ***AS boundary router, ASBR*** *— a router that has one port in the OSPF domain and another port in the domain of any of the internal gateway protocols (for example, RIP or EIGRP).*
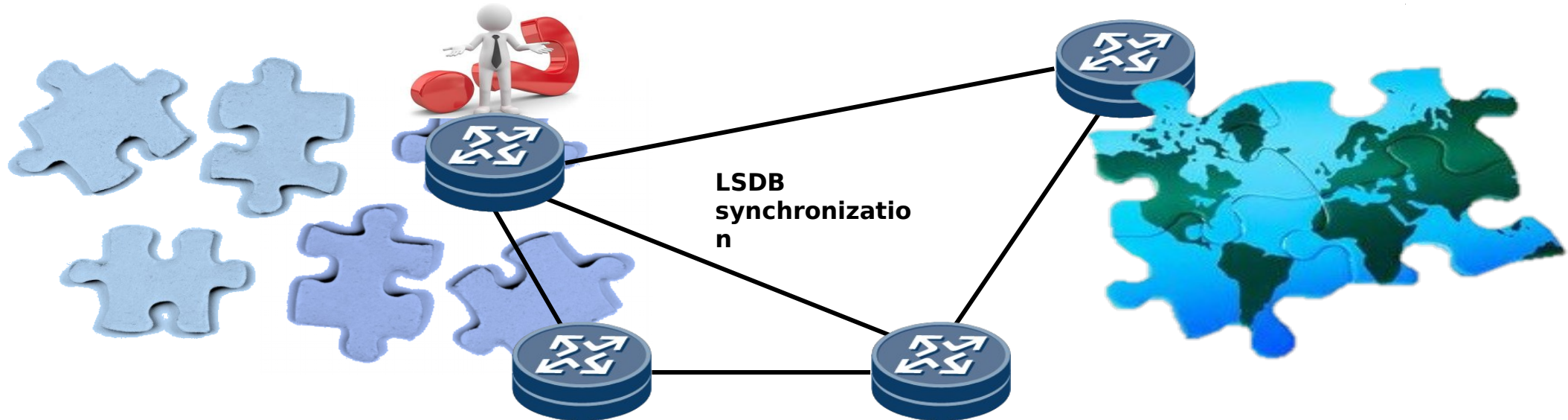
OSPF Supported Network Types

2.2.2.2
RTB

1.1.1.1
RTA

Broadcast

2.2.2.2
RTB

1.1.1.1
RTA

Point-to-Point

DR

Virtual Circuit

Virtual Circuit

ATM/FR

Non-Broadcast Multi-Access (NBMA)

Virtual Circuit

Virtual Circuit

ATM/FR

Point to Multi-Point
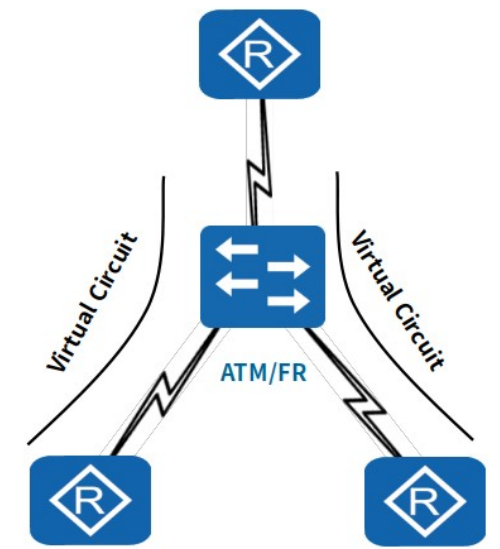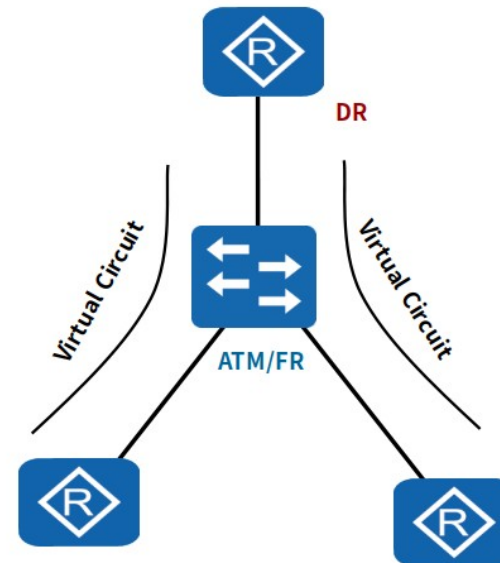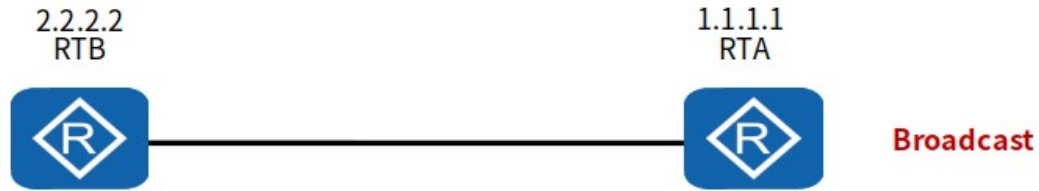
Roughter ID

Every OSPF router must have unique router ID.

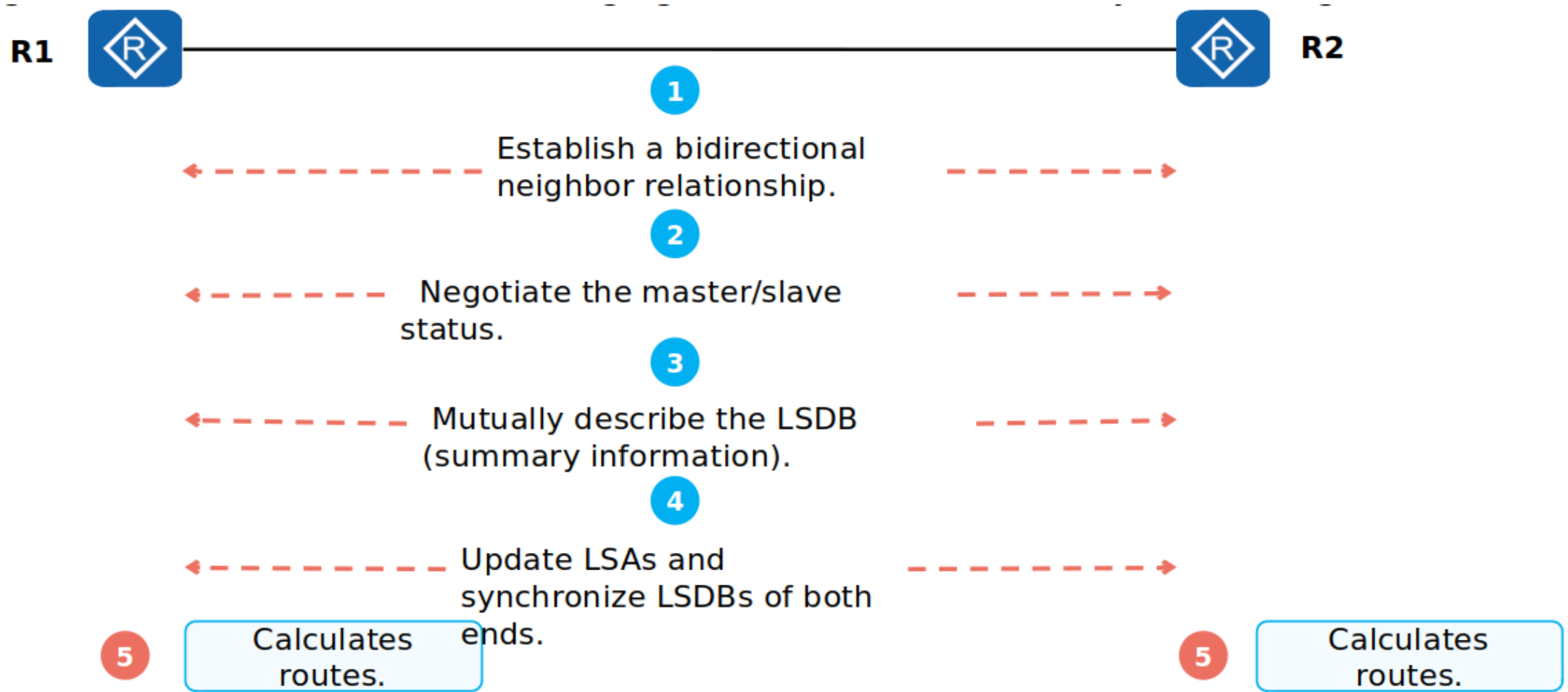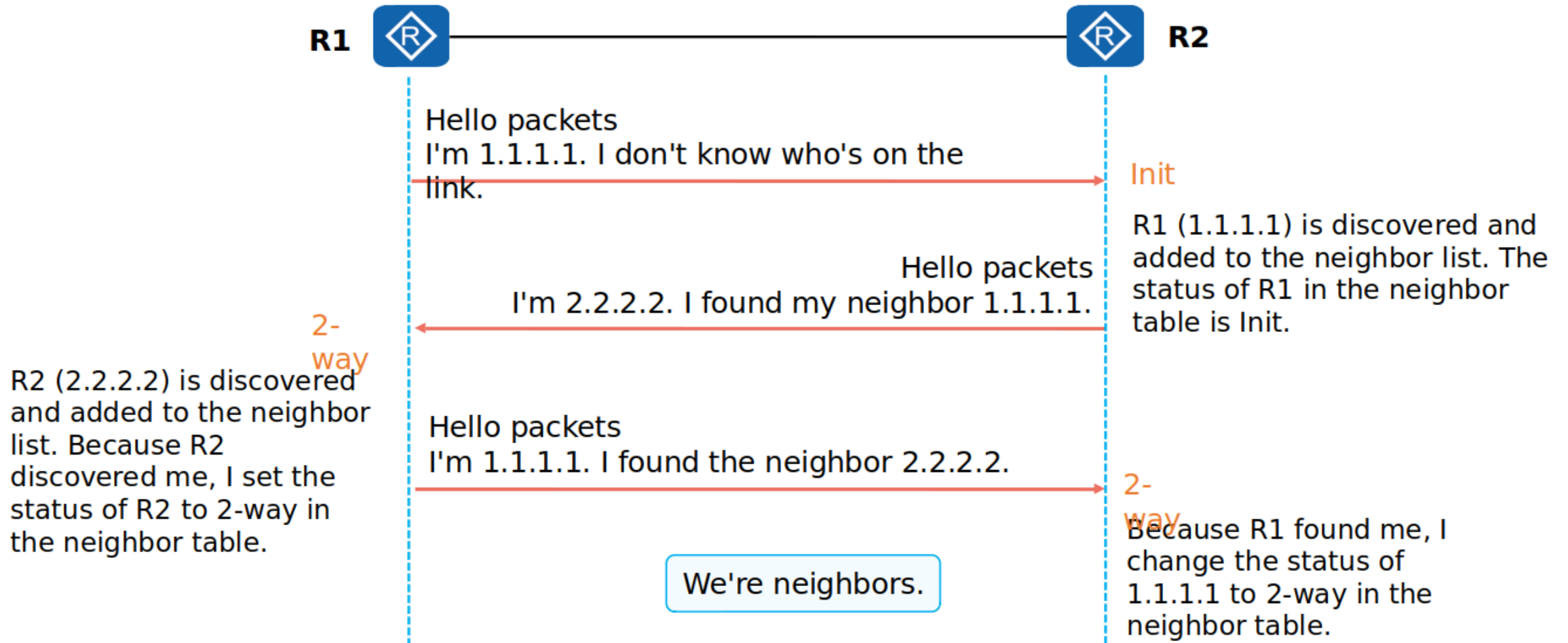Who am I and who are others? Each router must be identified in an LSDB.

LSDB synchronization

OSPF status

Process of Establishing an OSPF Adjacency Relationship

R1 — **R2**

1. Establish a bidirectional neighbor relationship.
2. Negotiate the master/slave status.
3. Mutually describe the LSDB (summary information).
4. Update LSAs and synchronize LSDBs of both ends.
5. Calculates routes.
5. Calculates routes.
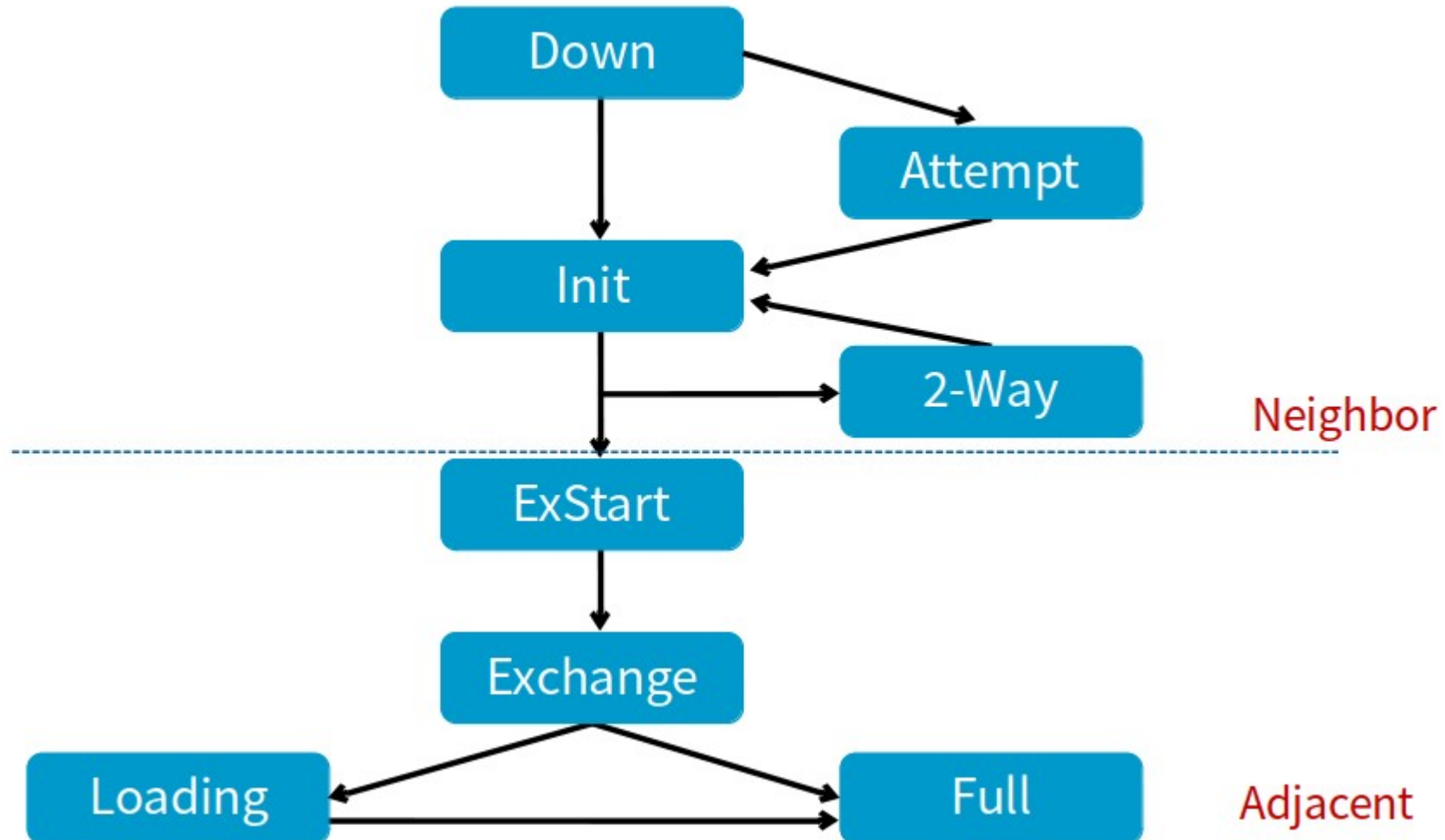
Steps 1 to 4 involve interaction between both ends, and Step 5 is performed separately on each device.

Process of Establishing an OSPF Adjacency Relationship

R1 ▢————————————————▢ **R2**

Hello packets
I'm 1.1.1.1. I don't know who's on the link.

Init

R1 (1.1.1.1) is discovered and added to the neighbor list. The status of R1 in the neighbor table is Init.

Hello packets
I'm 2.2.2.2. I found my neighbor 1.1.1.1.

2-way

R2 (2.2.2.2) is discovered and added to the neighbor list. Because R2 discovered me, I set the status of R2 to 2-way in the neighbor table.

Hello packets
I'm 1.1.1.1. I found the neighbor 2.2.2.2.

2-way

Because R1 found me, I change the status of 1.1.1.1 to 2-way in the neighbor table.

We're neighbors.

Process of Establishing an OSPF Adjacency Relationship

Process of Establishing an OSPF Adjacency Relationship

Router ID:
1.1.1.1
R1

Router ID:
2.2.2.2
R2

Loading

Loading

LSR
I want to request the complete
information about the xx LSA.

LSU
This is the complete information about the requested
LSA.

LS ACK
Confirms the reception of the LSU and the xx
LSA carried in the LSU.

LSR
I want to request the complete
information about the yyy LSA.

Full

Full

LSDBs of R1 and R2
are synchronized.

LSDBs of R1 and R2
are synchronized.
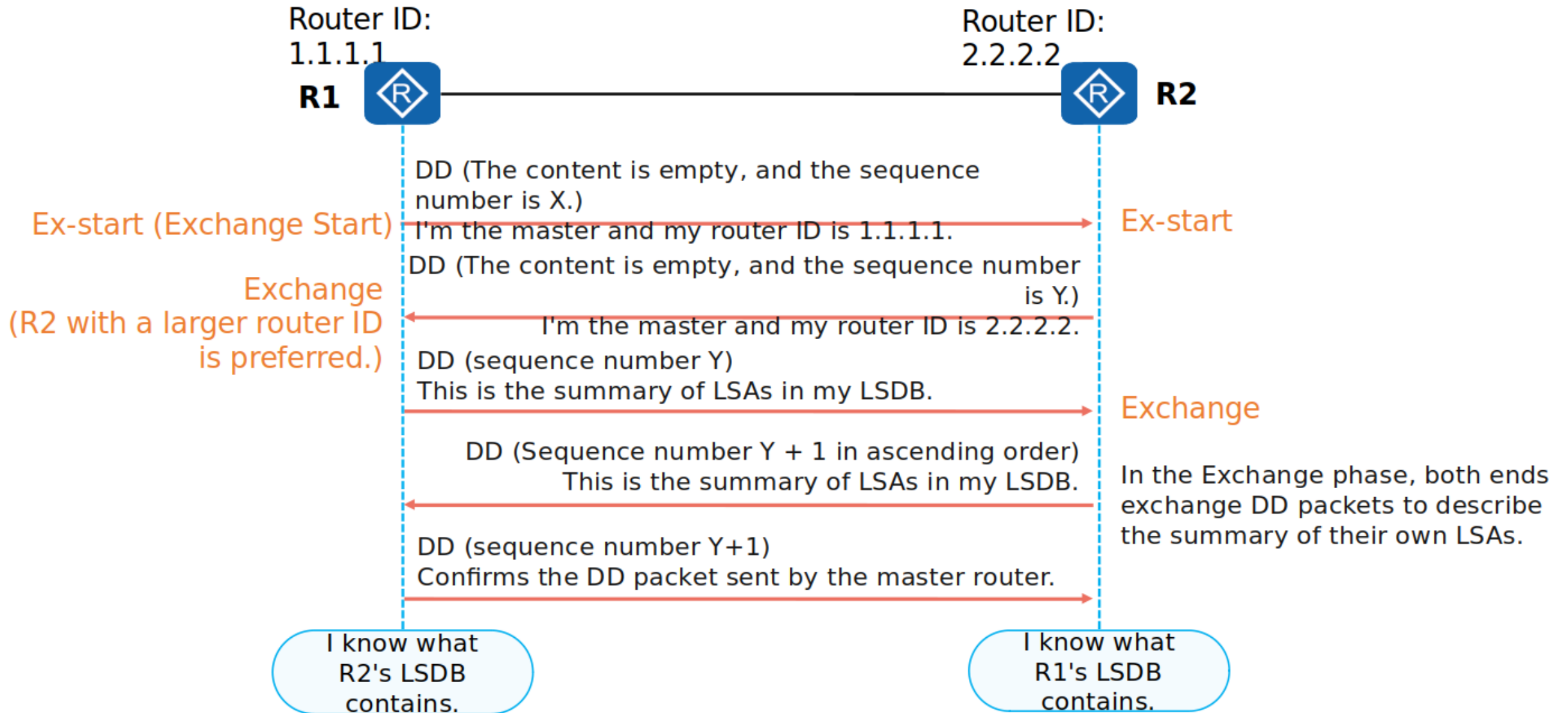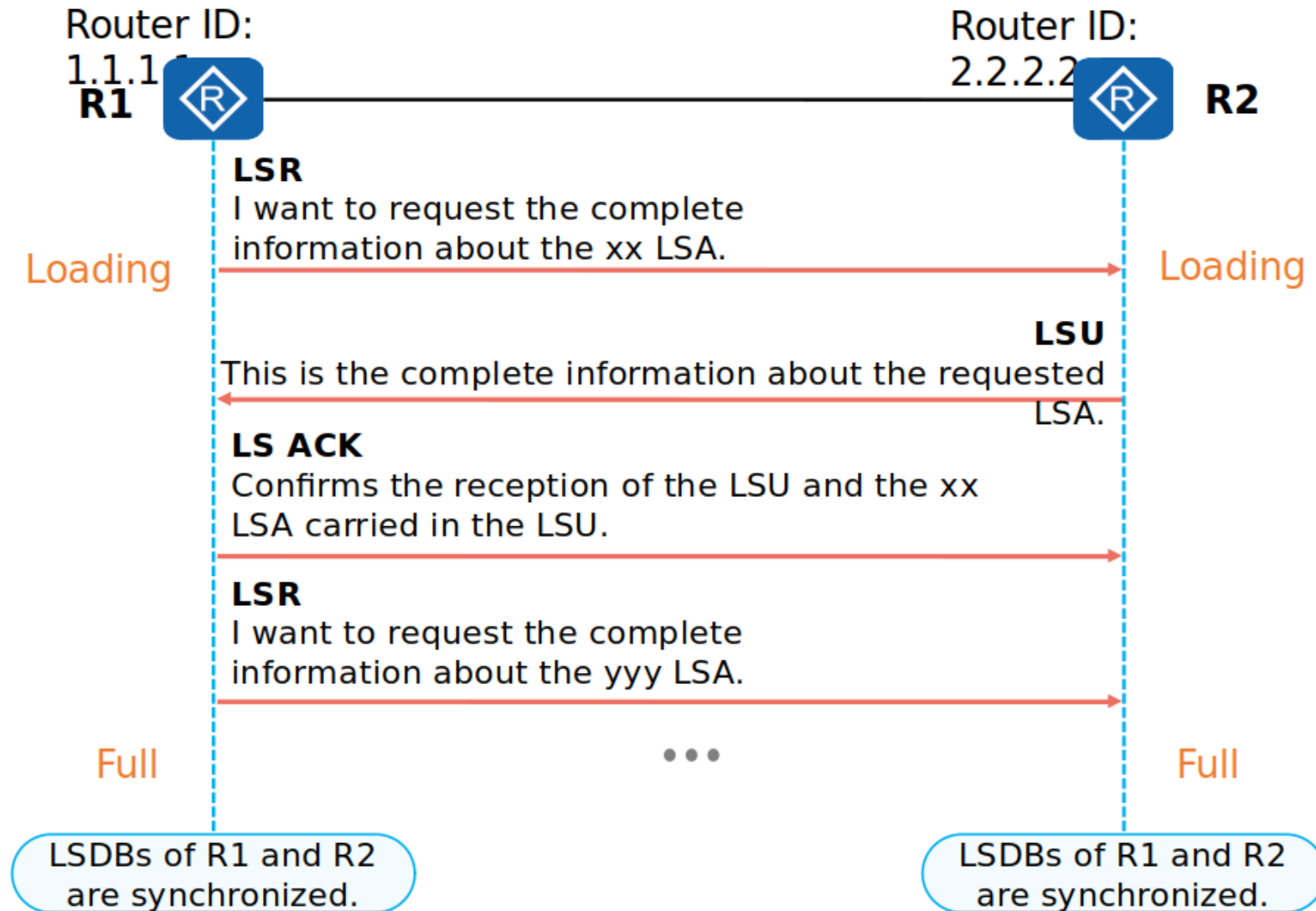
OSPF pakets types

Type	Packet Type	Packet Function
1	Hello	To discover and maintain OSPF neighbor relationships.
2	Database Description (DD)	To exchange LSDB summary.
3	Link State Request (LSR)	To request specific link state information.
4	Link State Update (LSU)	To send requested link state information.
5	Link State Ack (LSAck)	To acknowledge receipt of an LSA.

OSPF packet header format

Октет	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
0—3	Version								Type = 1								Packet length															
4—7	Router ID																															
8—11	Area ID																															
12—15	Checksum																Authentication type															
16—19	Authentication																															
20—23																																
24—27	Network mask																															
28—31	Hello interval																Options								Router priority							
32—35	Router dead interval																															
36—39	Designated router																															
40—43	Backup designated router																															
44—47	Neighbor ID																															
…	…																															

OSPF packet header format

All five types of OSPF packets are encapsulated directly in an IP packet. The OSPF Protocol number in the IP header is 89.

All OSPF packets have the same header

OSPF packet header fields:

Version number *— the version of the OSPF Protocol. OSPFv2 is for IPv4-2 and OSPFv3 is for IPv6.*

Packet type*-specifies which type of OSPF packet is sent:*

1 — Hello

2 — Database Description

3 — Link State Request

4 — Link State Update

5 — Link State Acknowledgment

Packet length *— the length of the OSPF packet in bytes. The length includes the header.*

Router ID*-determines which router sent the packet.*

Area ID*-determines in which zone the packet is generated.*

Checksum*-used to check the integrity of the OSPF packet, to detect errors during transmission.*

OSPF packet header format (c)

Authentication type — *the type of authentication that is used between routers:*

0 - authentication is not used,

1- clear text authentication,

2- MD5 authentication.

Authentication data - *used when authenticating routers.*

The **Data** *field is different for different types of OSPF packets:*

Hello-*list of famous neighbors*

DD - *contains the summary information of the channel state database, which includes all known router IDs and their latest sequence number and other information.*

LSR - *contains the type of LSU required and the ID of the router that has this LSU.*

LSU - *contains a complete record of announcements about the state of the channel. Multiple LSAs can be transferred in a single service pack.*

LSAck - *field empty*

Hello package format
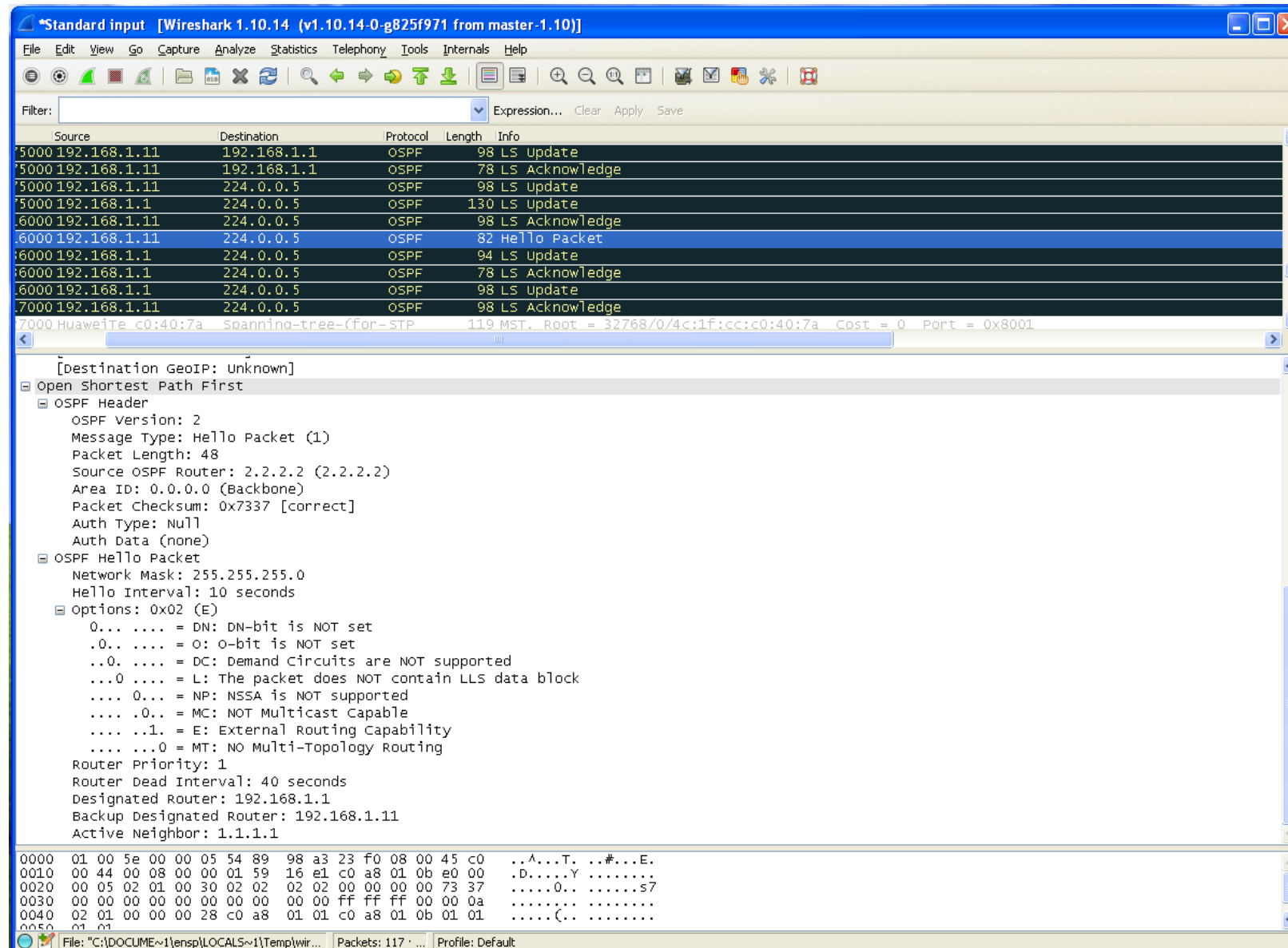
With the Hello packet, the router detects its neighbors;

- *It sends parameters that routers need to negotiate before they become neighbors;*

- *Hello packets act as keepalive packets between neighbors;*

- *Responsible for establishing two-way communications between neighboring routers (two-way communication is established when the router sees itself in the list of neighbors hello-packet received from a neighboring router);*

- *It is used when selecting DR and BDR in broadcast and non-broadcast multiple access networks.*
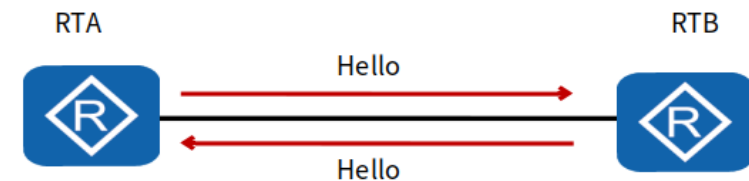
Hello Interval	Options	Router Priority
Router Dead Interval		
Designated Router		
Backup Designated Router		
Neighbor		

OSPF packet header format

OSPF packet header format

OSPF packet header format

OSPF packet header format

Source	Destination	Protocol	Length	Info
5000 192.168.1.11	192.168.1.1	OSPF	70	LS Request
5000 192.168.1.11	192.168.1.1	OSPF	86	DB Description
5000 192.168.1.1	224.0.0.5	OSPF	98	LS Update
5000 192.168.1.1	192.168.1.11	OSPF	98	LS Update
5000 192.168.1.1	192.168.1.11	OSPF	70	LS Request
5000 192.168.1.1	192.168.1.11	OSPF	66	DB Description
5000 192.168.1.11	192.168.1.1	OSPF	98	LS Update
5000 192.168.1.11	192.168.1.1	OSPF	78	LS Acknowledge
5000 192.168.1.11	224.0.0.5	OSPF	98	LS Update

```
Destination: 224.0.0.5 (224.0.0.5)
[Source GeoIP: Unknown]
[Destination GeoIP: Unknown]
Open Shortest Path First
  OSPF Header
      OSPF Version: 2
      Message Type: LS Update (4)
      Packet Length: 64
      Source OSPF Router: 1.1.1.1 (1.1.1.1)
      Area ID: 0.0.0.0 (Backbone)
      Packet Checksum: 0x9122 [correct]
      Auth Type: Null
      Auth Data (none)
  LS Update Packet
      Number of LSAs: 1
      LS Type: Router-LSA
          LS Age: 1 seconds
          Do Not Age: False
          Options: 0x02 (E)
          LS Type: Router-LSA (1)
          Link State ID: 1.1.1.1
          Advertising Router: 1.1.1.1 (1.1.1.1)
          LS Sequence Number: 0x80000003
          LS Checksum: 0x20bd
          Length: 36
          Flags: 0x00
          Number of Links: 1
          Type: Stub     ID: 192.168.1.0    Data: 255.255.255.0   Metric: 1
              IP network/subnet number: 192.168.1.0
              Link Data: 255.255.255.0
              Link Type: 3 - Connection to a stub network
              Number of TOS metrics: 0
              TOS 0 metric: 1
```

OSPF packet header format

OSPF packet header format

Source	Destination	Protocol	Length	Info
5000 192.168.1.11	192.168.1.1	OSPF	70	LS Request
5000 192.168.1.11	192.168.1.1	OSPF	86	DB Description
5000 192.168.1.1	224.0.0.5	OSPF	98	LS Update
5000 192.168.1.1	192.168.1.11	OSPF	98	LS Update
5000 192.168.1.1	192.168.1.11	OSPF	70	LS Request
5000 192.168.1.1	192.168.1.11	OSPF	66	DB Description
5000 192.168.1.11	192.168.1.1	OSPF	98	LS Update
5000 192.168.1.11	192.168.1.1	OSPF	78	LS Acknowledge
5000 192.168.1.11	224.0.0.5	OSPF	98	LS Update

```
Header length: 20 bytes
Differentiated Services Field: 0xc0 (DSCP 0x30: Class Selector 6; ECN: 0x00: Not-ECT (Not ECN-Capable Transport))
Total Length: 64
Identification: 0x0005 (5)
Flags: 0x00
Fragment offset: 0
Time to live: 1
Protocol: OSPF IGP (89)
Header checksum: 0x3544 [validation disabled]
Source: 192.168.1.11 (192.168.1.11)
Destination: 192.168.1.1 (192.168.1.1)
[Source GeoIP: Unknown]
[Destination GeoIP: Unknown]
Open Shortest Path First
    OSPF Header
        OSPF Version: 2
        Message Type: LS Acknowledge (5)
        Packet Length: 44
        Source OSPF Router: 2.2.2.2 (2.2.2.2)
        Area ID: 0.0.0.0 (Backbone)
        Packet Checksum: 0x52e0 [correct]
        Auth Type: Null
        Auth Data (none)
    LSA Header
        LS Age: 1 seconds
        Do Not Age: False
        Options: 0x02 (E)
        LS Type: Router-LSA (1)
        Link State ID: 1.1.1.1
        Advertising Router: 1.1.1.1 (1.1.1.1)
        LS Sequence Number: 0x80000003
        LS Checksum: 0x20bd
        Length: 36
```

OSPF types of LSA

Type 1 LSA — Router LSA — router channel status announcement.
Type 2 LSA — Network LSA — the announcement about the state of the network. Is distributed by DR in multiple-access networks.
Type 3 LSA — Network Summary LSA — summary announcement of the status of the network channels. The ad is distributed by edge routers.
Type 4 LSA — ASBR Summary LSA — summary announcement of the state of the Autonomous system edge router channels.
Type 5 LSA — AS External LSA — announcement of the status of the external channels of the Autonomous system.
Type 6 LSA — Multicast OSPF LSA — specialized LSA that use multicast OSPF applications.
Type 7 LSA — AS External LSA for NSSA — announcements about the status of the external channels of the Autonomous system in the NSSA area.

OSPF types of LSA

Type 8 LSA — Link LSA — announces link-local address and prefix(s)
router to all routers sharing a channel (link).
Sent only if more than one channel is present router. Distributed only within the channel (link).
Type 9 LSA — Intra-Area-Prefix LSA puts in compliance: list
IPv6 prefixes the router pointing to the Router LSA, list
IPv6 prefixes and the transit network pointing to the Network LSA.
Distributed only within one zone.

OSPF areas types

- *backbone area*
- *standard area*
- *stub area*
- *Totally stubby area*
- *Not-so-stubby area (NSSA)*

Designated Router & Backup Designated Router
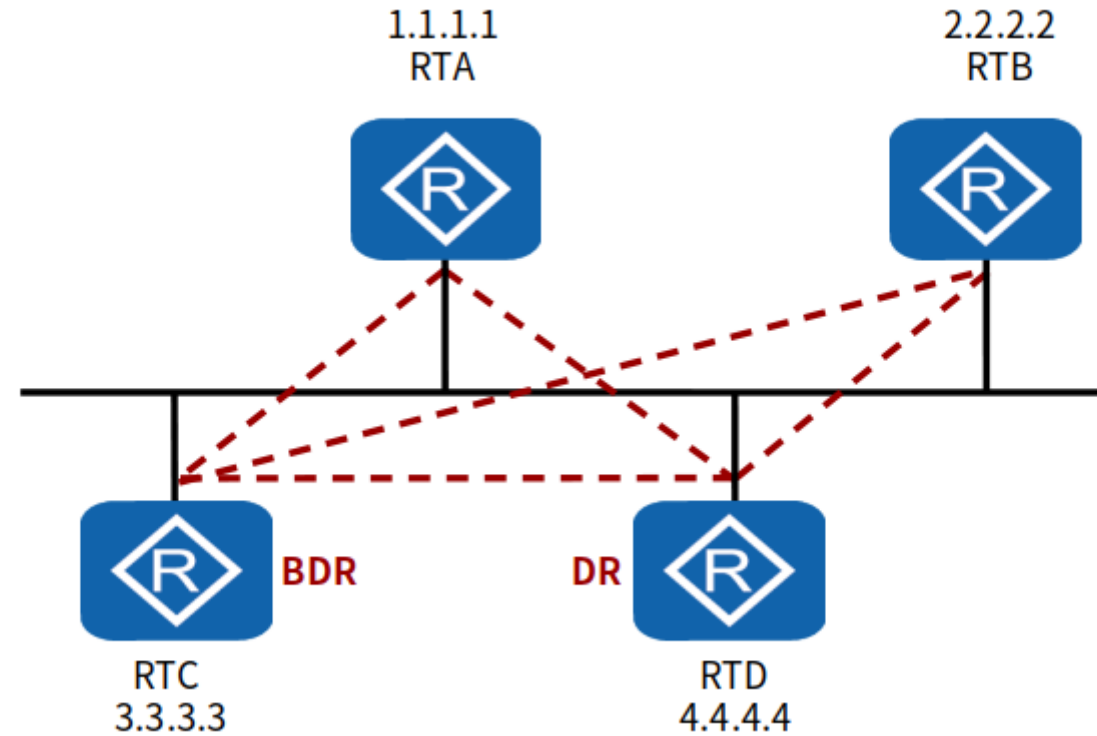
OSPF DR и BDR
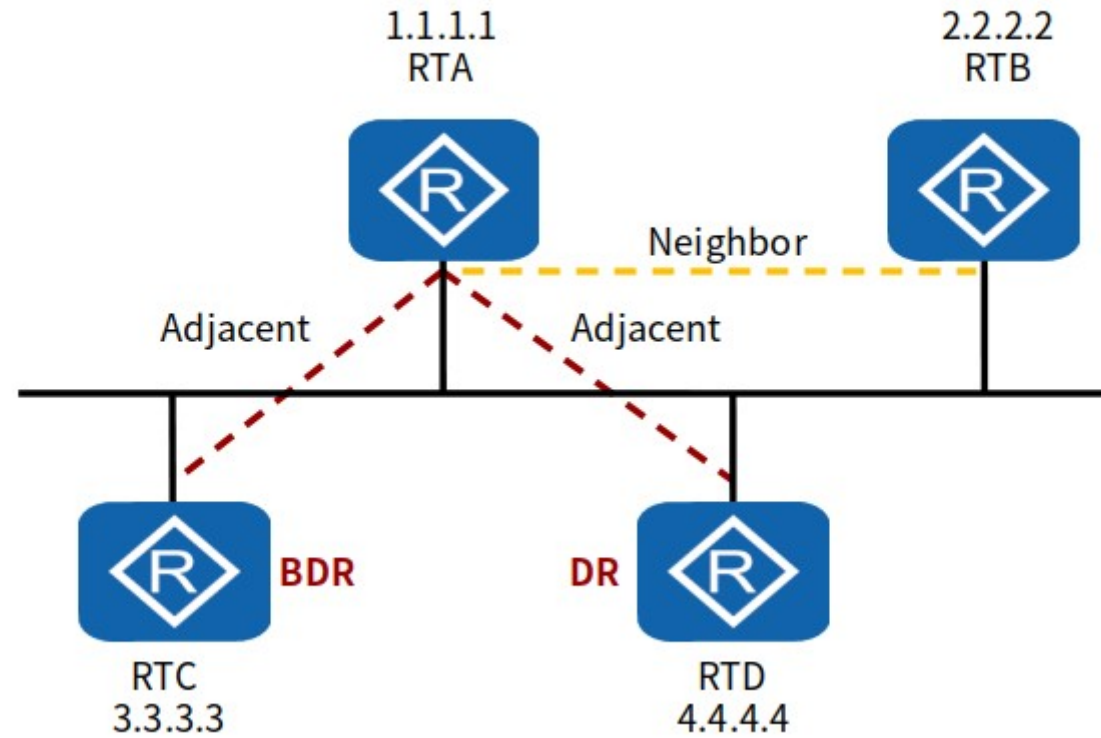
OSPF Example

ospf 1 router-id 1.1.1.1
area 0
network 10.1.12.0 0.0.0.255
network 10.1.13.0 0.0.0.255

ospf 1 router-id 2.2.2.2
area 0
network 10.1.12.0 0.0.0.255
network 10.1.24.0 0.0.0.255
network 10.1.235.0 0.0.0.255

RTA

RTB

10.1.12.0/24

10.1.24.0/24

10.1.13.0/24

10.1.235.0/24

RTD
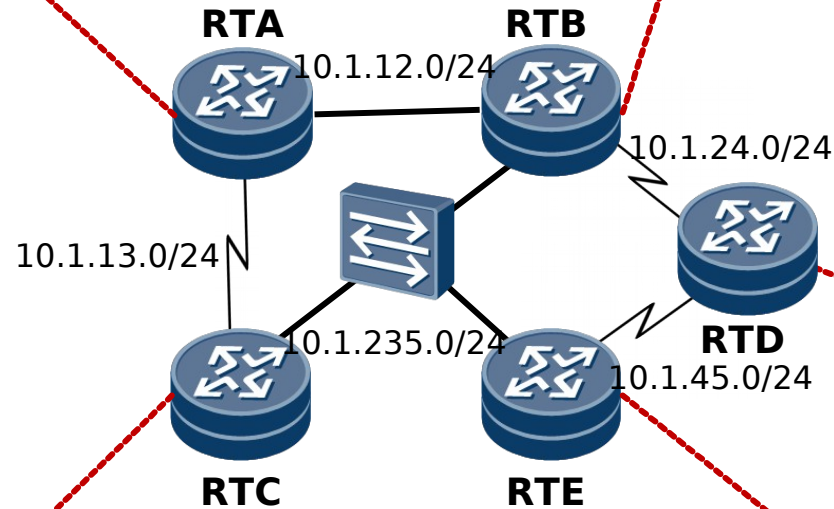
10.1.45.0/24

ospf 1 router-id 4.4.4.4
area 0
network 10.1.24.0 0.0.0.255
network 10.1.45.0 0.0.0.255

RTC

RTE

ospf 1 router-id 3.3.3.3
area 0
network 10.1.13.0 0.0.0.255
network 10.1.235.0 0.0.0.255

ospf 1 router-id 5.5.5.5
area 0
network 10.1.45.0 0.0.0.255
network 10.1.235.0 0.0.0.255

BGP

Border Gateway Protocol (BGP-4) Main terms

- *autonomous system, AS*
- *transit AS*
- *path*
- *path attributes, PA*
- *BGP speaker*
- *neighbor, peer*
- *Network Layer Reachability Information, NLRI*

Border Gateway Protocol (BGP-4)

- BGP use «path vector»
- Each BGP router sends out a list of paths (path-AS list)
  - AS_PATH
  - The 171.64/16 network can be accessed by following the route {AS7,AS52,AS13}
- The presence of a loop in a route is determined locally and such routes are ignored
- From the variety of available routes, the one that most corresponds to the AS policy is selected
- If the router/lines fail, the route is removed from the list

AS list

Next hop

4.4.0.0/16	11.11.1.2
3.3.0.0/16	11.11.1.2
2.2.0.0/16	11.11.1.2

AS 100
1.1.0.0/16

11.1.1.1

11.1.1.2

1.1.0.0/16	11.11.1.1

AS 200
2.2.0.0/16

AS 400
4.4.0.0/16

AS 300
3.3.0.0/16
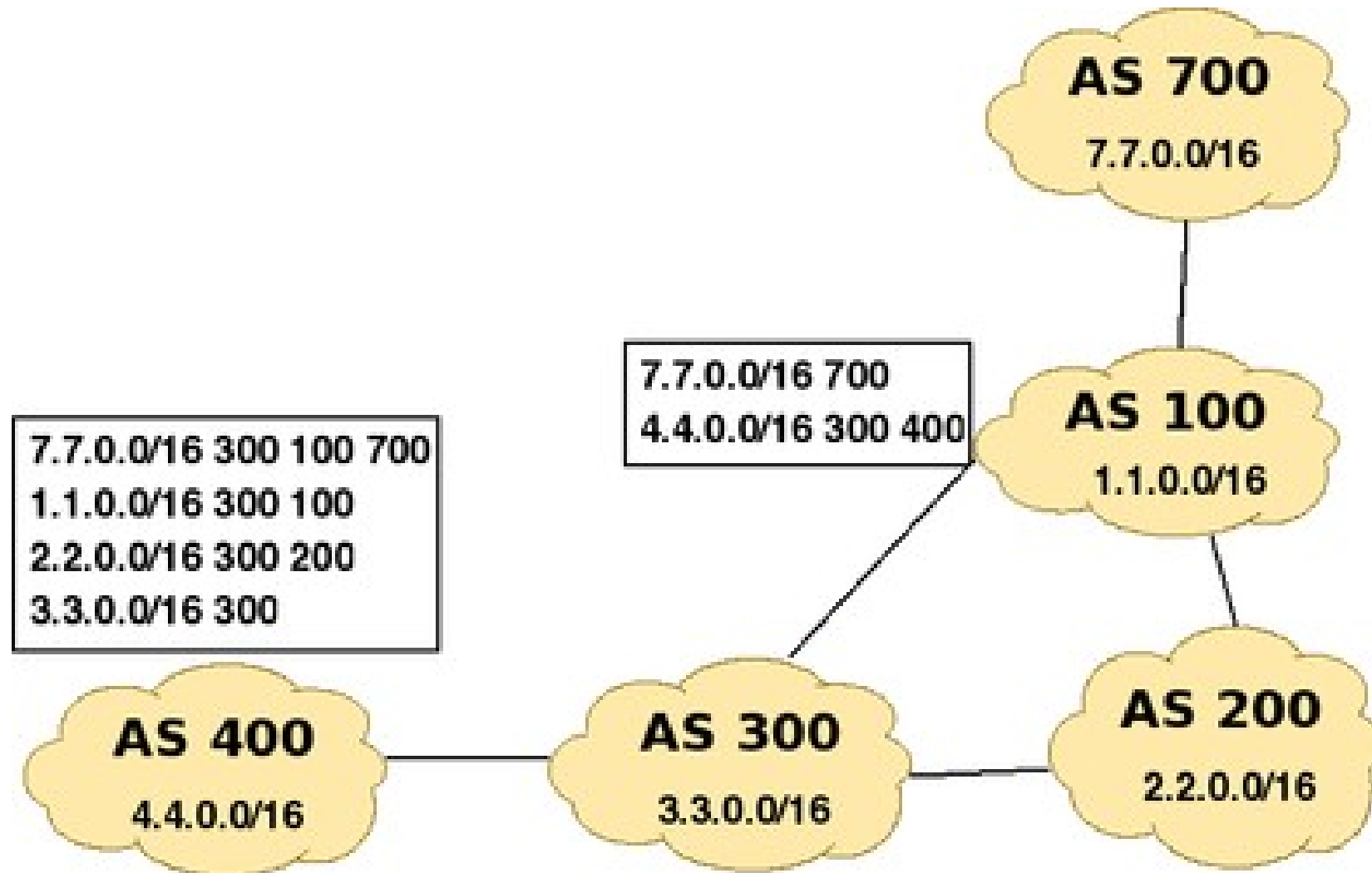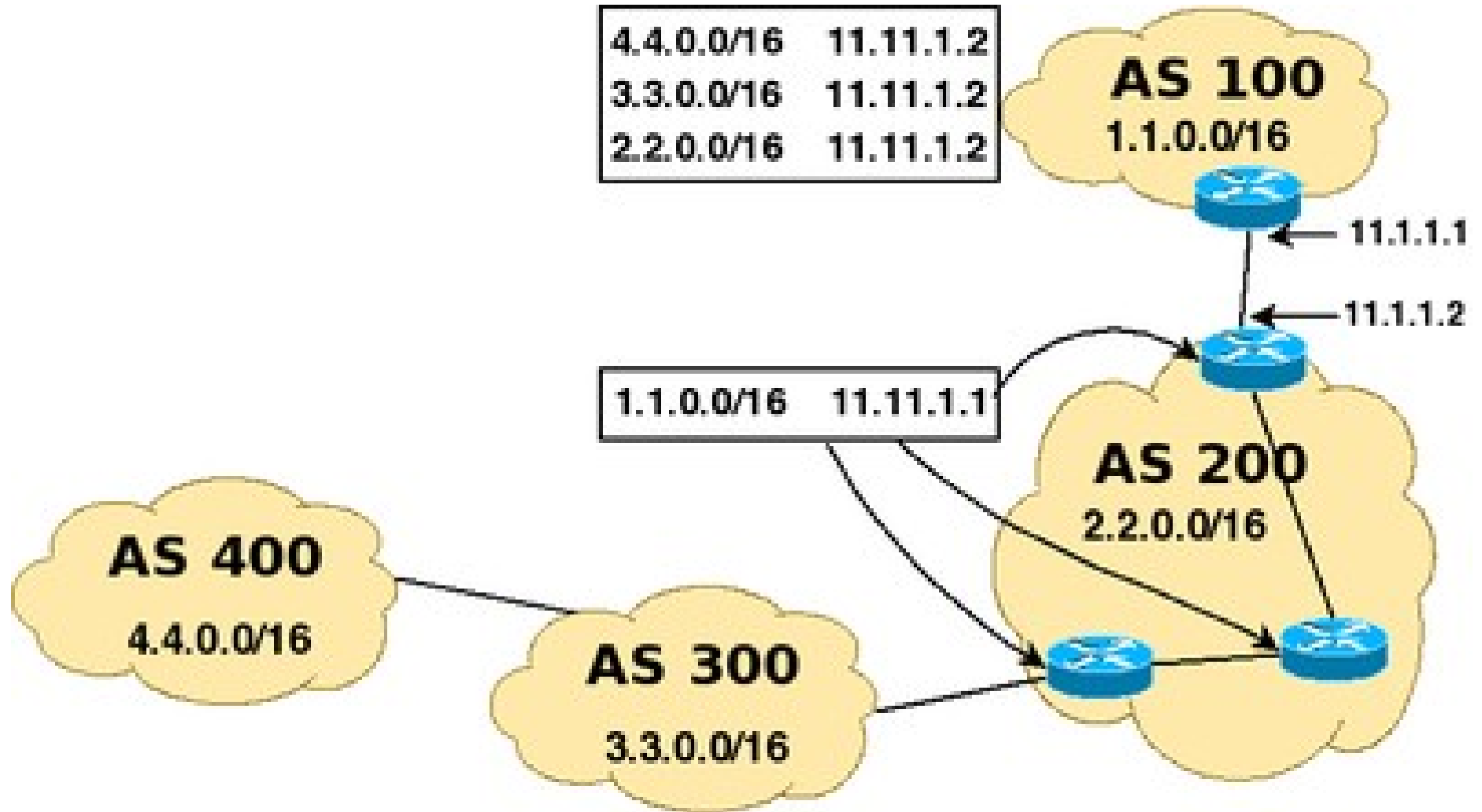
BGP packets

- ***Open*** **-** BGP session setup (TCP port 179)
- ***Keep Alive -*** Health check at regular intervals
- ***Notification*** **-** Closing peering session
- ***Update*** **-** Advertising a new route or withdrawing a previously advertised route

BGP advertisement = prefix + path attributes

- **Path attributes**
  - Origin - who suggested this route
  - AS path - preferences, the rule for determining the output
  - next hop - IP address of BGP gateway of the next AS,
  - multi_exit_disc – how to select one of several gateways advertizing the rought
  - Local pref - prefixes of local speakers
  - Aggregate – combining several routes with a common prefix

Open message

Update message

272 18:44:45.729710	198.51.100.1	198.51.100.2	BGP	106 UPDATE Message	
273 18:44:45.759710	198.51.100.2	198.51.100.1	BGP	106 UPDATE Message	
274 18:44:45.799710	198.51.100.1	198.51.100.2	BGP	92 KEEPALIVE Message, KEEPALIVE Message	
275 18:44:45.819710	198.51.100.2	198.51.100.1	BGP	92 KEEPALIVE Message, KEEPALIVE Message	
276 18:44:46.029710	198.51.100.1	198.51.100.2	TCP	60 60882 > bgp [ACK] Seq=155 Ack=155 Win=16230 Len=0	

⊞ Frame 272: 106 bytes on wire (848 bits), 106 bytes captured (848 bits)
⊞ Ethernet II, Src: c0:00:14:74:00:00 (c0:00:14:74:00:00), Dst: c0:01:14:74:00:00 (c0:01:14:74:00:00)
⊞ Internet Protocol Version 4, Src: 198.51.100.1 (198.51.100.1), Dst: 198.51.100.2 (198.51.100.2)
⊞ Transmission Control Protocol, Src Port: 60882 (60882), Dst Port: bgp (179), Seq: 65, Ack: 65, Len: 52
⊟ Border Gateway Protocol – UPDATE Message
    Marker: ffffffffffffffffffffffffffffffff
    Length: 52
    Type: UPDATE Message (2)
    Unfeasible routes length: 0 bytes
    Total path attribute length: 25 bytes
⊟ Path attributes
   ⊞ ORIGIN: IGP (4 bytes)
   ⊞ AS_PATH: 100 (7 bytes) | Атрибуты пути
   ⊞ NEXT_HOP: 198.51.100.1 (7 bytes)
   ⊞ MULTI_EXIT_DISC: 0 (7 bytes)
⊟ Network layer reachability information: 4 bytes
   ⊟ 100.0.0.0/23 | Информация о новых или удалённых маршрутах
      NLRI prefix length: 23
      NLRI prefix: 100.0.0.0 (100.0.0.0)

Border Gateway Protocol (BGP-4)
Выбор пути

If path selection policies are not configured, selection occurs as follows :
1. Maximum importance weight (local to router Cisco).
2. The maximum value of local preference (for the entire AS).
3. Prefer local router route (next hop = 0.0.0.0).
4. The shortest path through Autonomous systems. (the short AS_PATH)
5. The minimum value of the origin code (IGP < EGP < incomplete).
6. Minimum MED value (spread between Autonomous system).
7. The eBGP path is better than the iBGP path.
8. Choose a path through the nearest IGP neighbor.
9. Select the oldest route for the eBGP path.
10. Choose a path through the neighbor with the smallest BGP router ID.
11. Select the path through the neighbor with the lowest IP address.

Full View и Default Route

Full View router learns all the routes of the Internet.
There are more than 600 000 such routes.
If you use 2 providers, you can get more than a million routes.
- ***Full View**. Full information about the structure of the Internet. To any address on the Internet you can view the path from the current device on the network*
- ***Load distribution**. This is achieved by setting, for example, route priorities for specific prefixes*
- *Full View is **required** if your AS is transit*

***Default Route**-only one default route comes from each provider*
- *Saving equipment resources*
- *Easy maintenance*

BGP

- *All users to interact on the Internet must use BGP-4*
- *BGP-4 uses a path vector routing algorithm that easily recognizes loops*
- *BGP-4 has a complex interface that allows each AS to set its own local routing policy*
- *Each AS sets its own policies for routing, security, and local features*

Доп. главы Компьютерных сетей чл.-корр. РАН
Смелянский Р.Л.

What is IBGP

Let's start with what Internal BGP is. In fact, it is the same BGP, but inside the AS.

Reservation. *When there are several links to providers and it is undesirable to have all of them on one border router, several routers are put, and IBGP rises between them in order that on them there was always up-to-date information on all routes.*

BGP client connection. If the task is to connect the client via BGP, and you have more than one router, IBGP is used .

What is IBGP
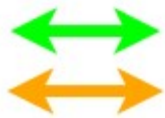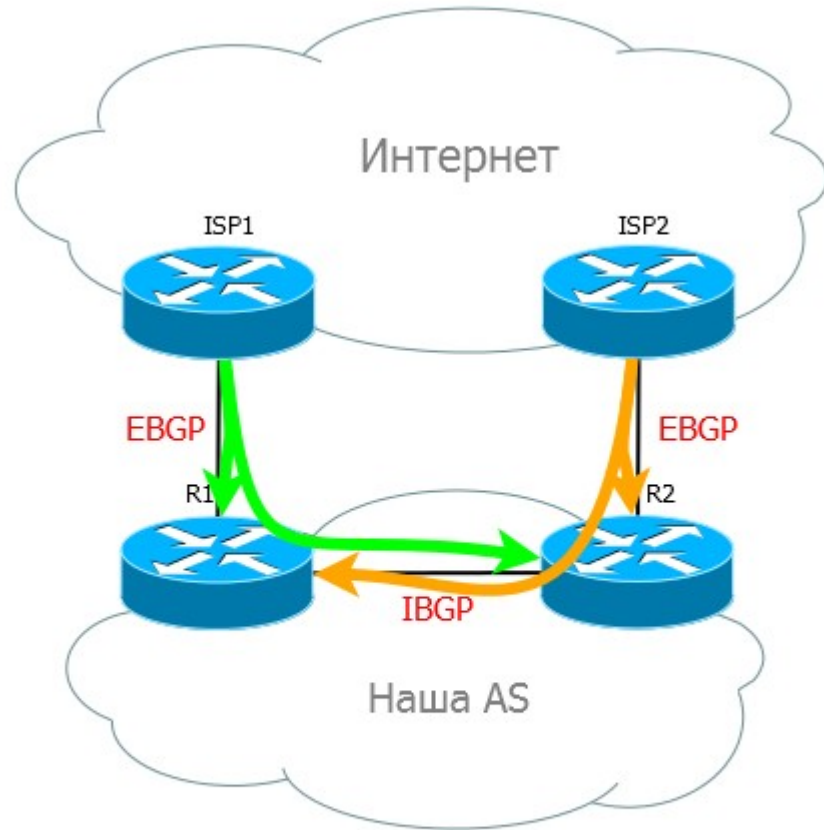
IBGP & EBGP differences

1) Resistance to loops generation

- *EBGP handles loops with AS-Path. If there was already an AS number of the local AS in the list, this route is discarded.*

- *When you advertize a route inside an Autonomous System, the AS-Path does not change. IBGP uses a fully connected Full Mesh.*

- *In this case, the route received from the IBGP neighbor is not announced to other IBGP neighbors.*

- *This allows all routers to have all routes and still avoid loops.*

IBGP & EBGP differences

2) Next Hop address.

- *In External BGP, when a router sends an announcement to its EBGP neighbor, it first changes the Next-Hop address to his own one, and then sends.*

- *If the router sends an IBGP announcement to a neighbor, the next-Hop address does not change.*

- *The concept of Next-Hop is different from that used in IGP. In IBGP, it reports the exit point from the local AS.*

- *It is important that the recipient of such an announcement has a route to Next-Hop. If it is not, the route will be placed in the BGP table, but it will not be in the routing table.*

IBGP & EBGP differences

3) Connecting neighbors

- *While EBGP typically involves two neighbors connecting directly to each other, in Internal BGP, neighbors can be connected through multiple intermediate devices.*

- *in EBGP, you can also configure neighbors that are several intermediate devices away from each other, but for IBGP, this works by default.*

- *allows IBGP partnership between Loopback addresses, EBGP does not use this*

What is IBGP
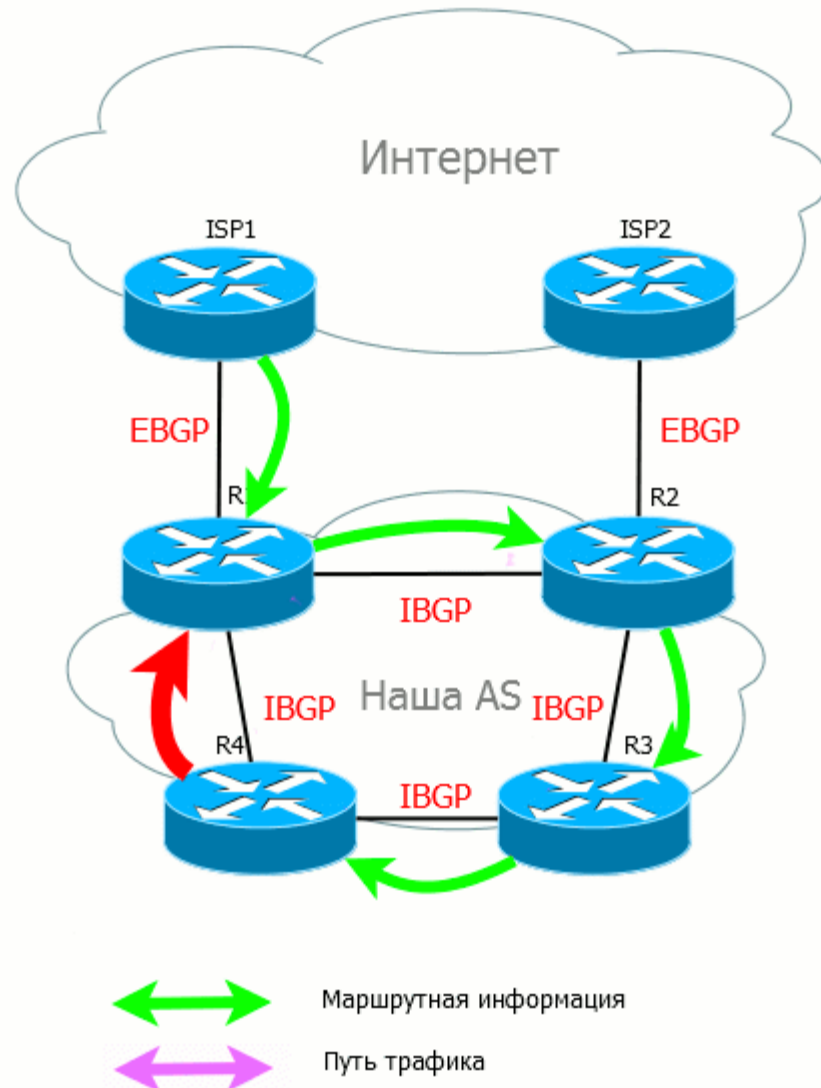
- *In the case of a fully connected topology and the Split Horizon rule, a situation is excluded when even a new route from R4 can get out as a priority, this is inefficient, since the routes will be studied incorrectly, a loop may form and the traffic will not get to the destination point.*

- *R1 received the announcement from ISP1, transmits it immediately to all its neighbors: R2, R3, R4. And those, in turn, keep these announcements, but transmit only to EBGP-partners, but not IBGP, precisely because they are received from the IBGP-partner. That is, all BGP routers have up-to-date information and loops are excluded.*

Typically, IBGP is configured using IGP existing on the network.

IGP provides*:*
- *IP connectivity of all routers,*
- *quick response to changes in the topology*
- *transfer route information about internal networks.*

1. Use Loopback interfaces :
These settings will be used to determine the Router ID for both OSPF and BGP.
2. Configure internal OSPF routing

BGP - Setting up interaction with neighbors

AS 200

RTD

4.4.4.4

.4

EBGP **AS 2**

10.1.24.0/24

.2

RTB

2.2.2.2

.2

10.1.12.0/24

IBGP

AS 300

RTE

5.5.5.5

.5

EBGP

10.1.35.0/24

.3

router id 5.5.5.5
bgp 300
peer 10.1.35.3 as-number 100

router id 3.3.3.3
bgp 100
peer 10.1.35.5 as-number 300
peer 10.1.12.2 as-number 100

RTC

3.3.3.3

.3

10.1.13.0/24

RTA

.1   .1

AS 100

1.1.1.1

OSPF

BGP - Setting up interaction with neighbors

AS 200

RTD
4.4.4.4

.4

AS 2

EBGP
10.1.24.0/24

.2

RTB
2.2.2.2

IBGP

.2
10.1.12.0/24

RTA

.1   .1

AS 100

1.1.1.1

AS 300

RTE
5.5.5.5

.5

EBGP
10.1.35.0/24

.3

RTC
3.3.3.3

.3
10.1.13.0/24

OSPF

```
router id 5.5.5.5
bgp 300
peer 10.1.35.3 as-number 100
```

```
router id 3.3.3.3
bgp 100
peer 10.1.35.5 as-number 300
peer 2.2.2.2 as-number 100
Peer 2.2.2.2 connect-interface loopback 0
```

BGP - Configure the advertisement of routes

```
router id 3.3.3.3
bgp 200
peer 10.1.23.2 as-number 100
```

RTC
3.3.3.3

AS 200

.3

EBGP
10.1.23.0/24

.2

```
router id 2.2.2.2
bgp 100
peer 10.1.23.2 as-number 200
 ipv4-family unicast
   network 10.1.12.0 255.255.255.0
   network 100.0.0.0 255.255.255.0
   network 100.0.1.0 255.255.255.0
```

AS 100

100.0.0.0/24
100.0.1.0/24

.2

10.1.12.0/24

RTB
2.2.2.2

.1

OSPF
Area 0

1.1.1.1   RTA
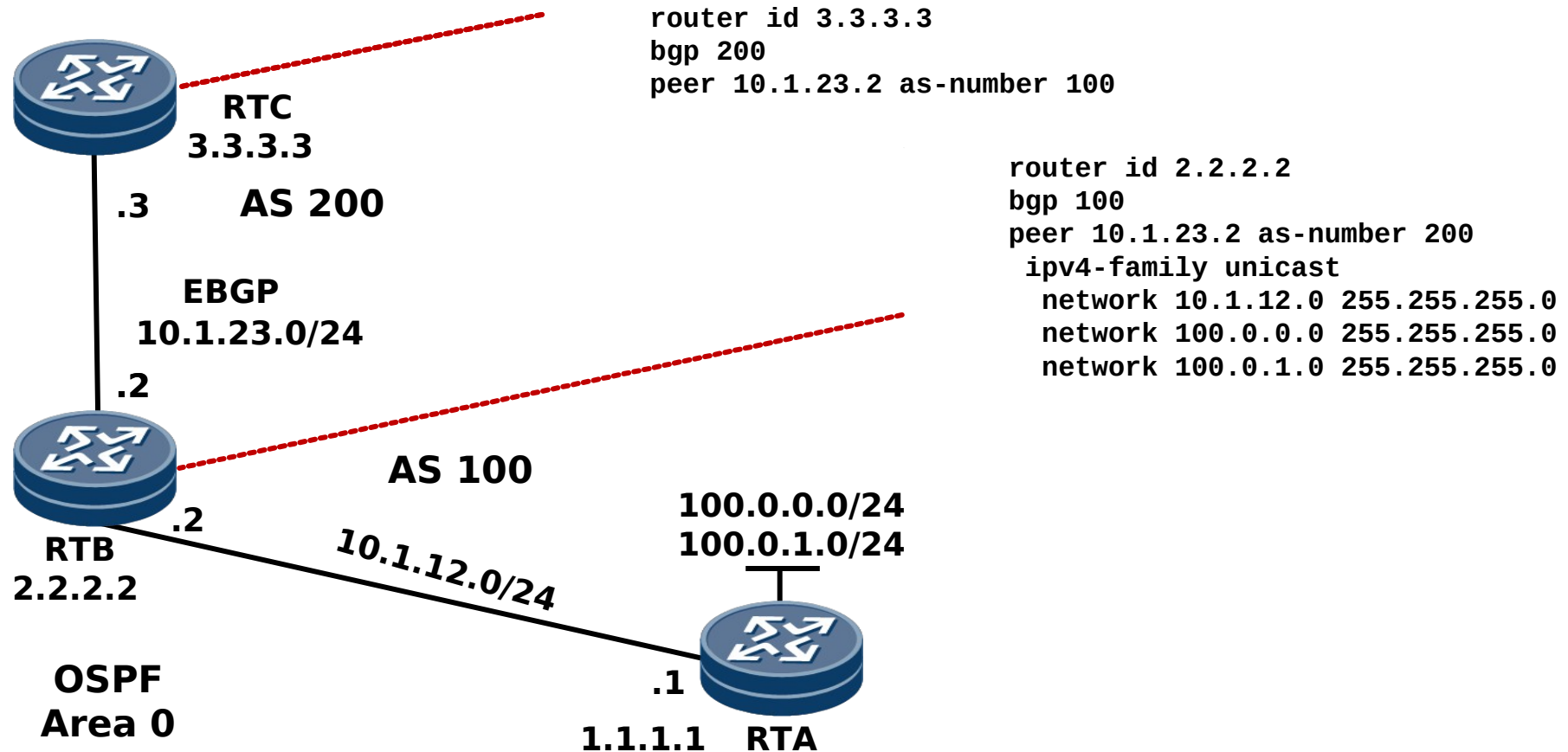
The *network* command is used to import routes that exist in the routing table into the BGP route table.

BGP — routes advertisement check

<RTC>display bgp routing-table

```
           Network           NextHop      MED    LocPrf    PrefVal Path/Ogn
*>    10.1.12.0/24       10.1.23.2     0       0         100i
*>    100.0.0.0/24       10.1.23.2     0       0         100i
*>    100.0.1.0/24       10.1.23.2     1       0         100i
```

RTC
3.3.3.3

.3   **AS 200**

EBGP

10.1.23.0/24

.2

<RTB>display ip routing-table

```
Destination/Mask    Proto    Pre   Cost   Flags    NextHop       Interface
10.1.12.0/24        Direct   0     0      D        10.1.12.2     GigabitEthernet0/0/0
100.0.0.0/24        Static   60    0      RD       10.1.12.1     GigabitEthernet0/0/0
100.0.1.0/24        OSPF     10    1      D        10.1.12.1     GigabitEthernet0/0/0
```

AS 100

.2

RTB
2.2.2.2

10.1.12.0/24

100.0.0.0/24
100.0.1.0/24

OSPF
Area 0

.1

1.1.1.1   RTA

Can BGP work without IGP

- *IGP and IBGP work together, each doing their job.*
- *IGP provides internal IP connectivity, response to changes in the network, notification of all nodes. He knows about the public addresses of the AS.*
- *IBGP handles external routes in the AS and their transit to and from customers. Usually he knows nothing about the structure of the internal network.*

Additional features of IBGP

Route Reflector

Route Reflector is a special IBGP router that performs the function of "reflecting " routes — he gets a route from one neighbor, and sends it to all others. That is, on IBGP routers, you need to configure a session with only one neighbor — with Route Reflector. Direct analogy - DR in OSPF.

Additional features of BGP

MP-BGP and Address Families

The Border Gateway Protocol 4 (BGP-4) transmits only IPv4 unicast routing information and cannot transmit routing information for other network layer protocols, such as IPv6 and multicast protocols.

To support multiple types of network layer protocols, the IETF extended BGP-4 to Multiprotocol Extensions for BGP-4 (MP-BGP). The current MP-BGP standard is RFC 4760.

As an enhancement of BGP-4, MP-BGP provides routing information for various protocols, such as IPv6 (BGP4+) and multicast

Additional features of BGP

MP-BGP and Address Families

Multiprotocol Reachable NLRI (MP_REACH_NLRI) is used to advertise reachable routes and information about the next hop. MP_REACH_NLRI is coded as one or more 3-tuples of the form <Address Family Information, Next Hop Information, Network Layer Reachability Information>.

Address Family Information (3 bytes)
Next Hop Network Address Information (variable length)
Network Layer Reachable Information (variable length)

Address Family Information: consists of a 2-byte Address Family Identifier (AFI) and a 1-byte Subsequent Address Family Identifier (SAFI):

The AFI identifies a network layer protocol. Defined values for this field are specified in RFC 1700 (Address Family Number). For example, 1 indicates IPv4; 2 indicates IPv6.

The SAFI indicates the type of the NLRI field.

If the AFI is 1 and the SAFI is 128, the address in the NLRI field is a BGP-VPNv4 address.

Additional features of BGP

MP-BGP and Address Families

Network Layer Reachable Information: is a variable-length field that lists NLRI for the routes being advertised.

Length (1 byte)
Label (variable length)
Prefix (variable length)

Descriptions of each part of the NLRI field are as follows:

Length: indicates the total bits of the label and prefix.

Label: consists of one or more labels. The length of a label is 3 bytes.

Prefix: In a BGP/Multiprotocol Label Switching (MPLS) IP VPN, the prefix field consists of a route distinguisher (RD) and IPv4 address prefix.

VPNv4 update messages exchanged between provider edges (PEs) or autonomous system boundary routers (ASBRs) carry MP_REACH_NLRI. An Update message can carry multiple reachable routes with the same routing attributes.

Additional features of BGP

BGP-LS

BGP-link state (LS) enables BGP to report topology information collected by IGPs to the controller.

Without BGP-LS, the router uses an IGP (OSPF or IS-IS) to collect network topology. This method has the following disadvantages:

- The controller must have high computing capabilities and support the IGP.
- The controller cannot gain the complete inter-domain topology information and therefore is unable to calculate optimal E2E paths.
- Different IGPs report topology information separately to the controller, which complicates the controller's analysis and processing.

BGP-LS has the following advantages:

- Reduces computing capability requirements and spares the necessity of IGPs on the controller.
- Facilitates route selection and calculation on the controller by using BGP to summarize process or AS topology information and report the complete information to the controller.
- Requires only one routing protocol (BGP) to report topology information to the controller.