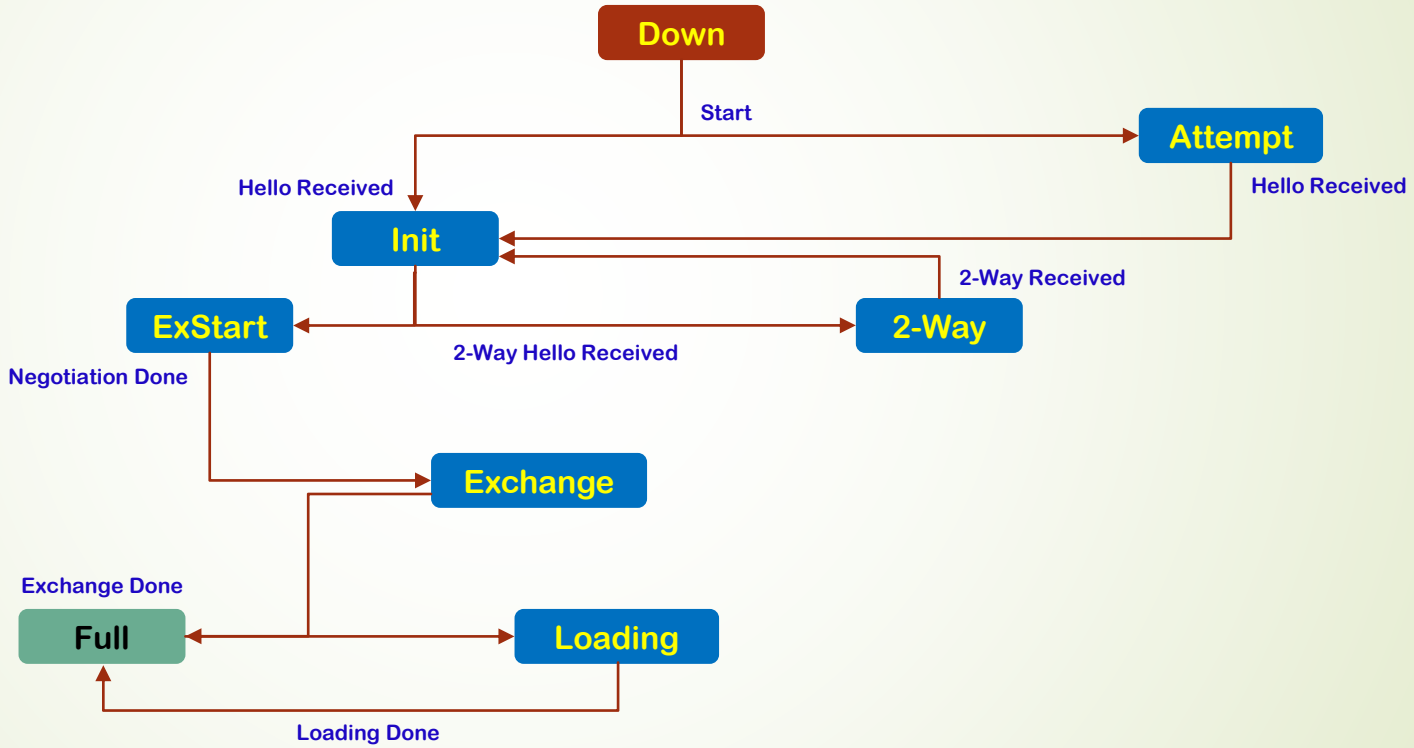
Implementing Enterprise Routing and Switching Network

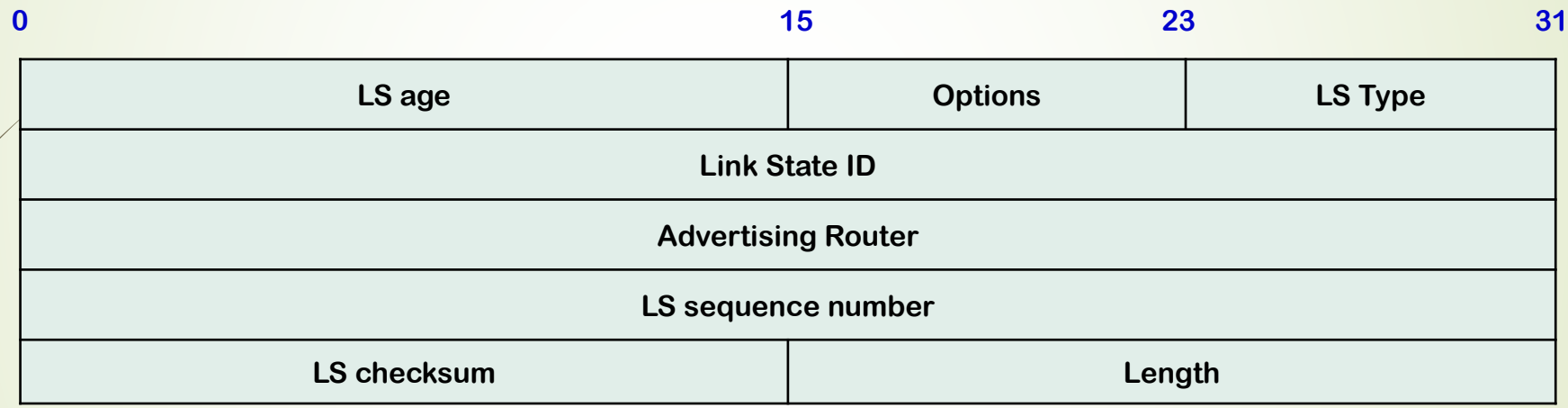
OSPF
Protocol Basics

OSPF Neighbor State Mechanism



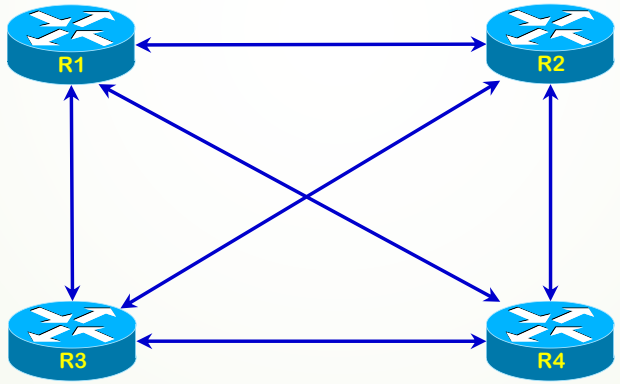
LSA Header

- LSA carry OSPF link state information.



Problems in MA Networks

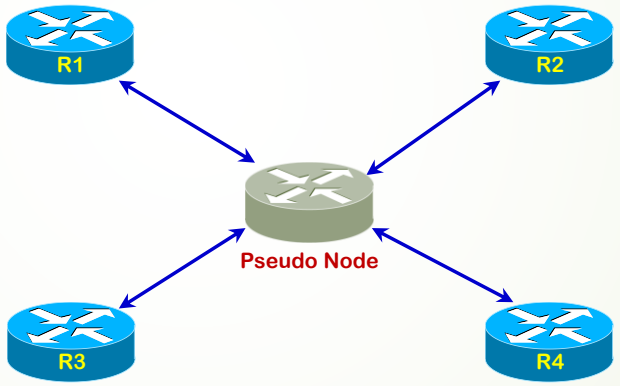
- Managing N x (N-1) / 2 adjacencies is complicated.
- Unnecessarily flooding repeated LSAs is a waste of resources.



Routers (N)	Adjacencies N(N-1) / 2
4	6
5	10
10	45
20	190
50	1225

Responsibilities of DR and BDR

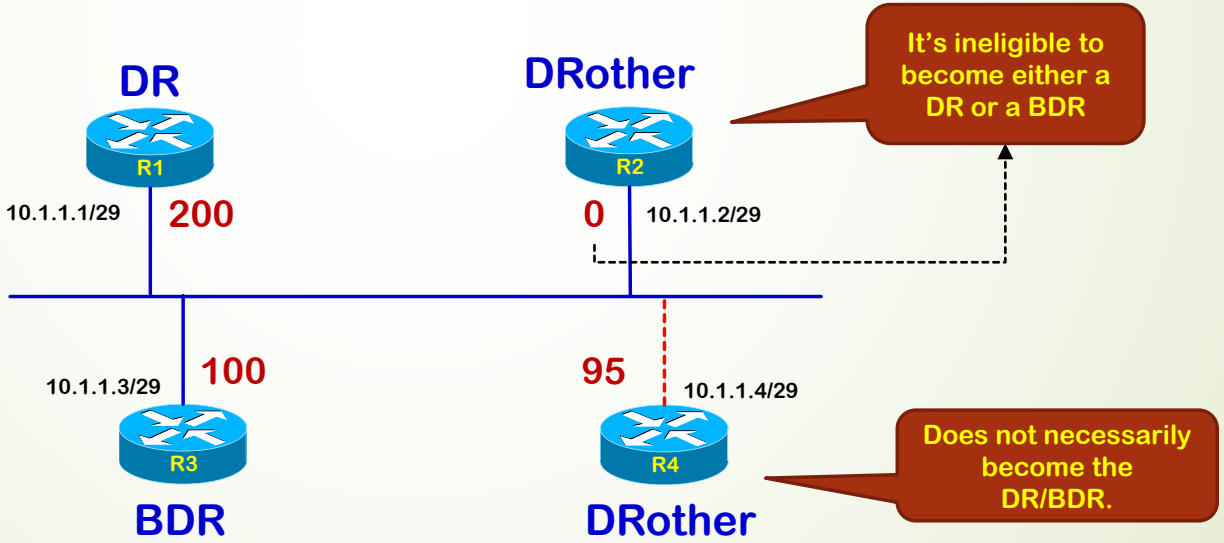
- Reduces the number of adjacencies.
- Reduces OSPF traffic.



- Question : How to avoid DR as the single point of failure?

DR and BDR Election

- Election rules: DR/BDR election occurs among interfaces.
  - An interface with the highest router priority takes precedence.
  - If interfaces have the same router priority, the interface with a larger router ID takes precedence.

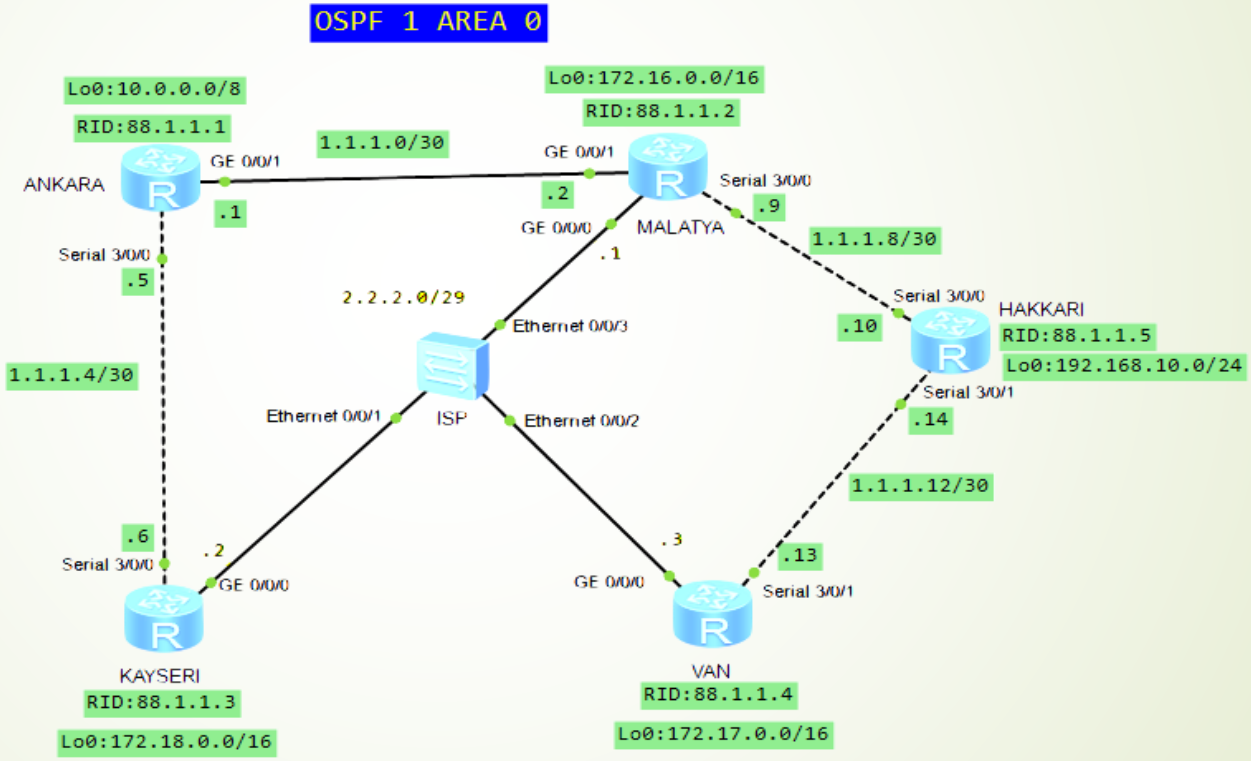


Neighbor Relationship and Adjacency

Network Type	Whether to Establish an Adjacency
Point-to-Point	Yes
Broadcast	The DR establishes adjacencies with the BDR and Drothers. The BDR establishes adjacencies with the DR and Drothers. DRothers establish only neighbor relationships with each other.
NBMA	
Point-to-Multipoint	Yes

OSPF
Intra-Area Routing

A Router-LSA Describing a Point-to-Point Interface

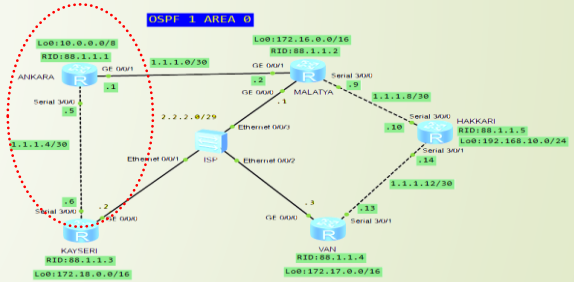


A Router-LSA Describing a Point-to-Point Interface

```
<ANKARA>display ospf lsdb router self-originate

        OSPF Process 1 with Router ID 88.1.1.1
                    Area: 0.0.0.0
                Link State Database

Type       : Router          // LSA Type
Ls id      : 88.1.1.1        // Link State ID
Adv rtr    : 88.1.1.1        // ID of the router that generate the LSA
Ls age     : 83
Len        : 72
Options    :  E
seq#       : 80000009
chksum     : 0xcdba
Link count : 4
 * Link ID  : 10.0.0.1       // IP Address of the STUB Network
   Data     : 255.255.255.255 // Subnet mask of this STUB network
   Link Type: StubNet
   Metric   : 0
   Priority : Medium
 * Link ID  : 88.1.1.3       // ID of the Neighbor Router
   Data     : 1.1.1.5         // IP Address of the Router Interface that advertises the LSA
   Link Type: P-2-P
   Metric   : 48              // Cost value
```

Routing Information

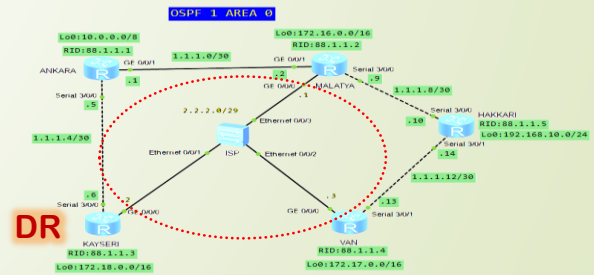
Topology Information



A Router-LSA Describing a Broadcast or an NBMA Interface

```
<MALATYA>display ospf lsdb router self-originate

        OSPF Process 1 with Router ID 88.1.1.2
                    Area: 0.0.0.0
                Link State Database


  Type          : Router          // LSA Type
  Ls id         : 88.1.1.2        // Link State ID
  Adv rtr       : 88.1.1.2        // ID of the router that generates the LSA
  Ls age        : 900
  Len           : 84
  Options       :  E
  seq#          : 80000010
  chksum        : 0x534f
  Link count    : 5
   * Link ID    : 2.2.2.2         // Interface IP Address of a Designated Router
     Data       : 2.2.2.1         // IP Address of the router interface that advertises the LSA
     Link Type  : TransNet
     Metric     : 1
   ......
```

Topology Information

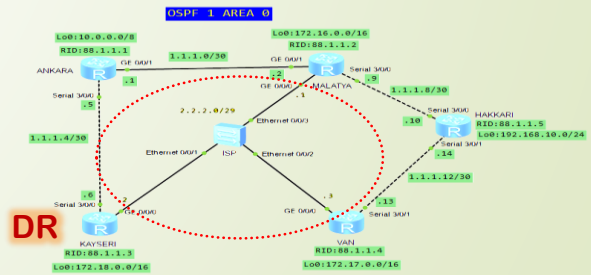


- Q: Where is the network ID or Subnet Mask?

A Network-LSA Describing a Broadcast or an NBMA Interface

DR

```
<KAYSERI>display ospf lsdb network self-originate

        OSPF Process 1 with Router ID 88.1.1.3
                    Area: 0.0.0.0
                Link State Database

Topology and Routing Information

Type          : Network                  // LSA Type
Ls id         : 2.2.2.2                  // Interface of IP Address of the DR
Adv rtr       : 88.1.1.3                 // Router ID of the DR
Ls age        : 457
Len           : 36
Options       :  E
seq#          : 80000008
chksum        : 0xe9ee
Net mask      : 255.255.255.248          // Subnet Mask
Priority      : Low
  Attached Router    88.1.1.3            // A list of routers connected to this network
  Attached Router    88.1.1.2
  Attached Router    88.1.1.4
```



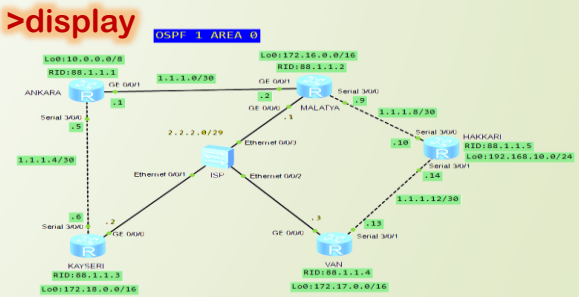
LSDB in an OSPF Area

<ANKARA>display ospf lsdb

OSPF Process 1 with Router ID 88.1.1.1
Link State Database

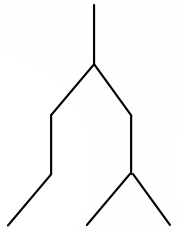
Area: 0.0.0.0

Type	LinkState ID	AdvRouter	Age	Len	Sequence	Metric
Router	88.1.1.3	88.1.1.3	1058	72	8000000D	1
Router	88.1.1.2	88.1.1.2	1044	84	8000000E	1
Router	88.1.1.1	88.1.1.1	1052	72	80000009	1
Router	88.1.1.5	88.1.1.5	1088	84	80000008	0
Router	88.1.1.4	88.1.1.4	1059	72	8000000B	1
Network	2.2.2.2	88.1.1.3	1058	36	80000005	0
Network	1.1.1.2	88.1.1.2	1044	32	80000002	0

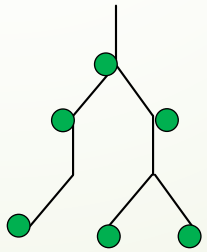


SPF Algorithm

- Phase 1: Create a shortest path tree (SPT).
  - Create an SPT on the basis of topology information in Router-LSAs and Network-LSAs.



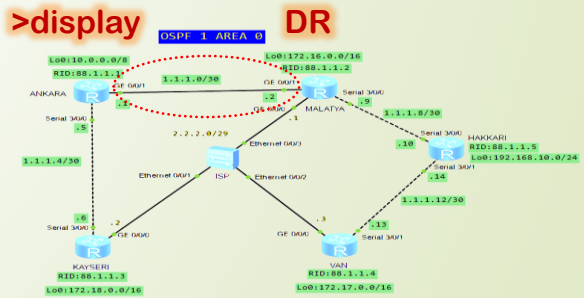
- Phase 2: Calculate an optimal route.
  - Calculate an optimal route on the basis of the SPT and the routing information in Router-LSAs and Network-LSAs.



Creating an SPT (1)

```
<ANKARA>display ospf lsdb router self-originate

        OSPF Process 1 with Router ID 88.1.1.1
                    Area: 0.0.0.0
                Link State Database

Type       : Router
Ls id      : 88.1.1.1
Adv rtr    : 88.1.1.1
Ls age     : 83
Len        : 72
Options    :  E
seq#       : 80000009
chksum     : 0xcdba
Link count : 4
 * Link ID: 1.1.1.2
   Data     : 1.1.1.1
   Link Type: TransNet
   Metric   : 1
  ......

 * Link ID: 88.1.1.3
   Data     : 1.1.1.5
   Link Type: P-2-P
   Metric   : 48
  ......
```

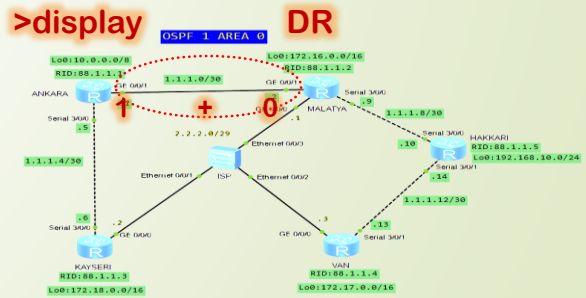
Candidate	Total Cost for Candidate	Parent Node
1.1.1.2	1	88.1.1.1
88.1.1.3	48	88.1.1.1



Creating an SPT (2)

```
<ANKARA>display ospf lsdb network 1.1.1.2

        OSPF Process 1 with Router ID 88.1.1.1
                Area: 0.0.0.0
                Link State Database


Type            : Network
Ls id           : 1.1.1.2
Adv rtr         : 88.1.1.2
Ls age          : 955
Len             : 32
Options         :  E
seq#            : 80000007
chksum          : 0x77c6
Net mask        : 255.255.255.252
Priority        : Low
  Attached Router    88.1.1.2
  Attached Router    88.1.1.1
```

Candidate	Total Cost for Candidate	Parent Node
88.1.1.3	48	88.1.1.1
88.1.1.2	1 + 0	1.1.1.2



Creating an SPT (3)

<ANKARA>dis ospf lsdb router 88.1.1.2

OSPF Process 1 with Router ID 88.1.1.1
Area: 0.0.0.0
Link State Database

Type : Router
Ls id : 88.1.1.2
Adv rtr : 88.1.1.2
Ls age : 188
Len : 84
Options : E
seq# : 80000014
chksum : 0x4b53
Link count : 5

* Link ID : 2.2.2.2
Data : 2.2.2.1
Link Type : TransNet
Metric : 1

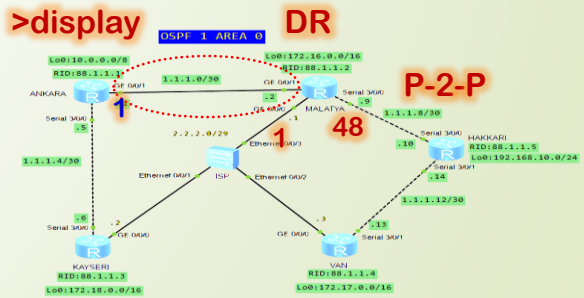
* Link ID : 1.1.1.2
Data : 1.1.1.2
Link Type : TransNet
Metric : 1

.....

* Link ID : 88.1.1.5
Data : 1.1.1.9
Link Type : P-2-P
Metric : 48

.....

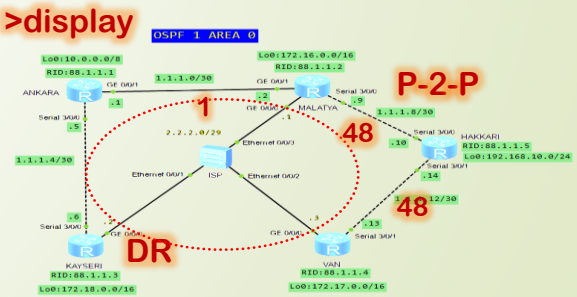
Candidate	Total Cost for Candidate	Parent Node
88.1.1.2	1 + 1	88.1.1.2
88.1.1.5	48	88.1.1.2
88.1.1.5	1 + 48	88.1.1.1



Creating an SPT (4)

```
<ANKARA>display ospf lsdb network 2.2.2.2

        OSPF Process 1 with Router ID 88.1.1.1
                Area: 0.0.0.0
             Link State Database

Type       : Network
Ls id      : 2.2.2.2
Adv rtr    : 88.1.1.3
Ls age     : 108
Len        : 36
Options    :  E
seq#       : 8000000c
chksum     : 0xe1f2
Net mask   : 255.255.255.248
Priority   : Low
  Attached Router    88.1.1.3
  Attached Router    88.1.1.2
  Attached Router    88.1.1.4
```

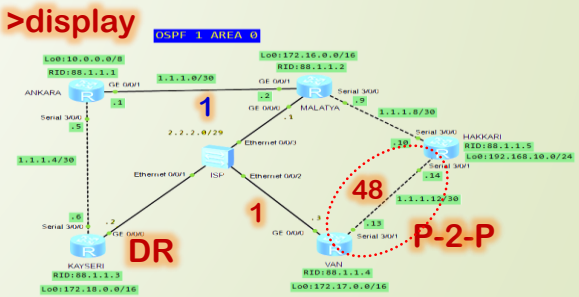
Candidate	Total Cost for Candidate	Parent Node
88.1.1.2	1 + 1	88.1.1.2
88.1.1.4	1 + 48 + 48	88.1.1.4



Creating an SPT (5)

```
<ANKARA>display ospf lsdb  router 88.1.1.4

        OSPF Process 1 with Router ID 88.1.1.1
                    Area: 0.0.0.0
                Link State Database


Type            : Router
Ls id           : 88.1.1.4
Adv rtr         : 88.1.1.4
Ls age          : 1091
Len             : 72
Options         :  E
seq#            : 80000012
chksum          : 0x3279
Link count      : 4
* Link ID: 2.2.2.2
  Data    : 2.2.2.3
  Link Type: TransNet
  Metric : 1
......
* Link ID: 1.1.1.12
  Data    : 255.255.255.252
  Link Type: StubNet
  Metric : 48
  Priority : Low
```

Candidate	Total Cost for Candidate	Parent Node
1.1.1.12	1 + 1 + 48	88.1.1.4

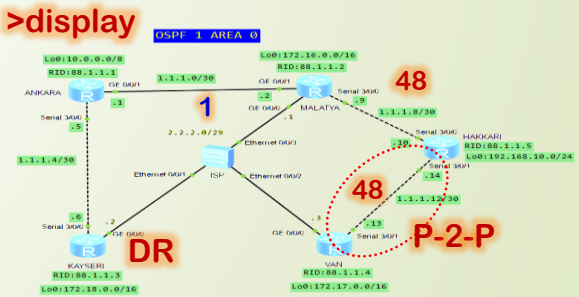


Creating an SPT (6)

```
<ANKARA>display ospf lsdb router 88.1.1.5

        OSPF Process 1 with Router ID 88.1.1.1
                Area: 0.0.0.0
                Link State Database


Type          : Router
Ls id         : 88.1.1.5
Adv rtr       : 88.1.1.5
Ls age        : 1167
Len           : 84
Options       :  E
seq#          : 8000000f
chksum        : 0xbe69
Link count    : 5
.....
* Link ID     : 88.1.1.2
  Data        : 1.1.1.10
  Link Type   : P-2-P
  Metric      : 48
* Link ID     : 1.1.1.8
  Data        : 255.255.255.252
  Link Type   : StubNet
  Metric      : 48
  Priority    : Low
```

Candidate	Total Cost for Candidate	Parent Node
1.1.1.12	~~1 + 48 + 48~~	88.1.1.5

When you compare with 88.1.1.4, it is not preferable



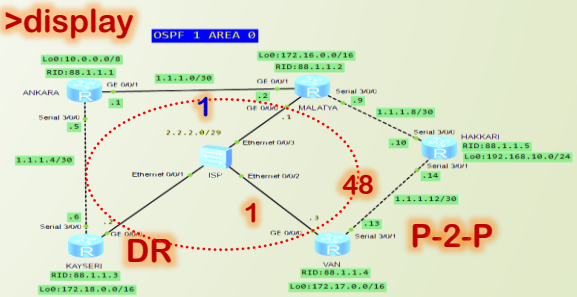
Viewing the OSPF Routing Table

```
<ANKARA>display ospf routing

    OSPF Process 1 with Router ID 88.1.1.1
          Routing Tables
```

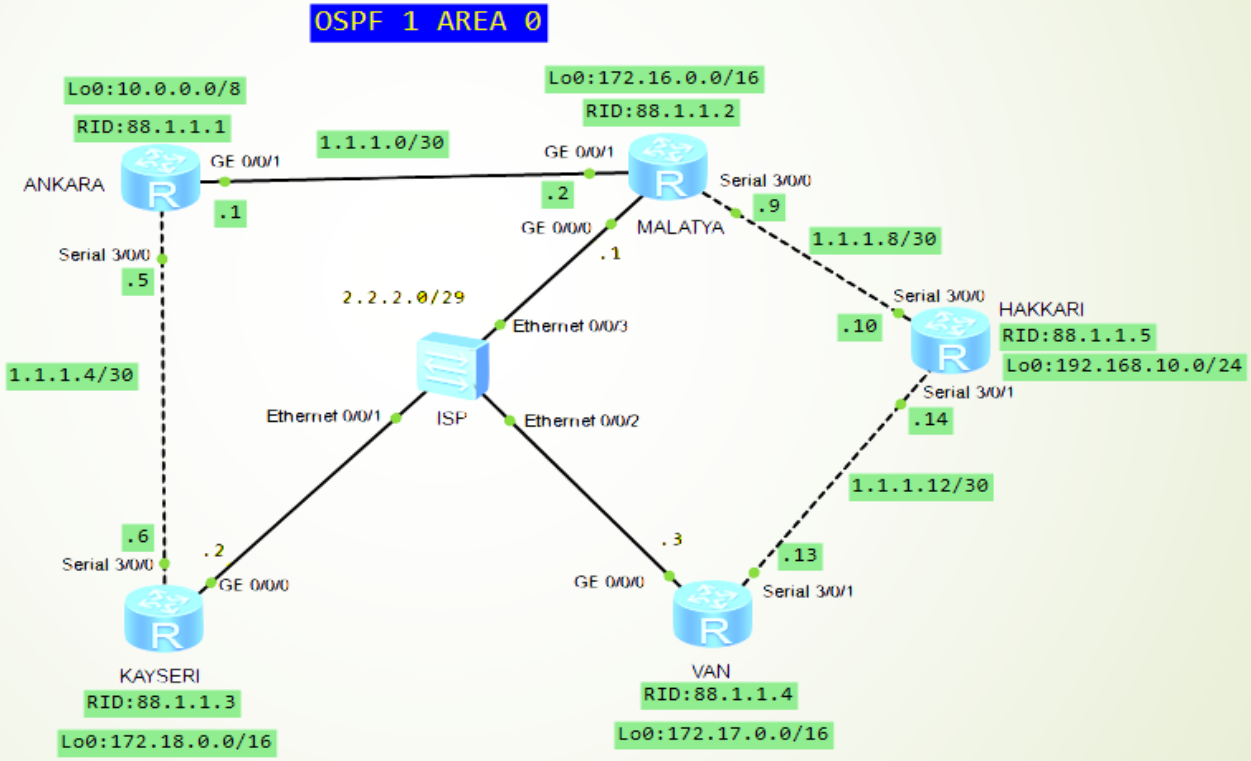
Routing for Network

Destination	Cost	Type	NextHop	AdvRouter	Area
1.1.1.0/30	1	Transit	1.1.1.1	88.1.1.1	0.0.0.0
1.1.1.4/30	48	Stub	1.1.1.5	88.1.1.1	0.0.0.0
10.0.0.1/32	0	Stub	10.0.0.1	88.1.1.1	0.0.0.0
1.1.1.8/30	49	Stub	1.1.1.2	88.1.1.2	0.0.0.0
1.1.1.12/30	50	Stub	1.1.1.2	88.1.1.4	0.0.0.0
2.2.2.0/29	2	Transit	1.1.1.2	88.1.1.3	0.0.0.0
172.16.0.1/32	1	Stub	1.1.1.2	88.1.1.2	0.0.0.0
172.17.0.1/32	2	Stub	1.1.1.2	88.1.1.4	0.0.0.0
172.18.0.1/32	2	Stub	1.1.1.2	88.1.1.3	0.0.0.0
192.168.10.1/32	49	Stub	1.1.1.2	88.1.1.5	0.0.0.0

Total Nets: 10
Intra Area: 10 Inter Area: 0 ASE: 0 NSSA: 0



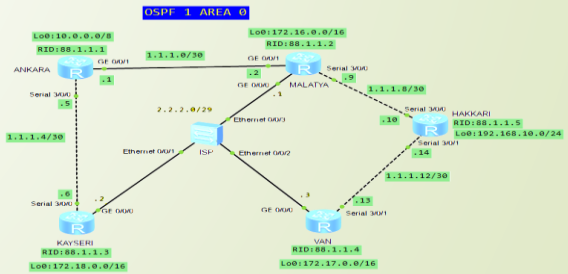
Example of Configuring a Single Area OSPF Network



Example of Configuring a Single Area OSPF Network

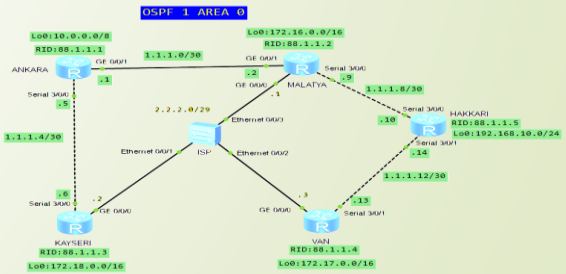
```
[ANKARA] dis curr | begin ospf
ospf 1 router-id 88.1.1.1
 silent-interface LoopBack0
 area 0.0.0.0
  network 1.1.1.0 0.0.0.3
  network 1.1.1.4 0.0.0.3
  Network 10.0.0.0 0.255.255.255
```

```
[ANKARA]display ospf peer brief

        OSPF Process 1 with Router ID 88.1.1.1
                Peer Statistic Information
 ----------------------------------------------------------------------------
 Area Id          Interface                     Neighbor id      State
 0.0.0.0          GigabitEthernet0/0/1          88.1.1.2         Full
 0.0.0.0          Serial3/0/0                   88.1.1.3         Full
 ----------------------------------------------------------------------------
```



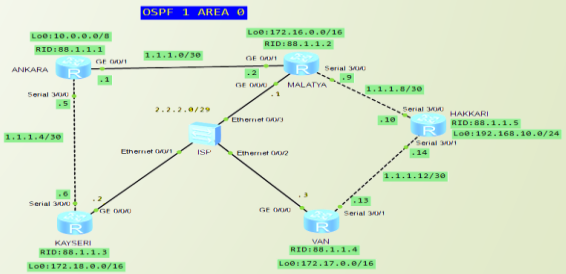
Example of Configuring a Single Area OSPF Network

```
[ANKARA] display ospf peer

          OSPF Process 1 with Router ID 88.1.1.1
                  Neighbors

 Area 0.0.0.0 interface 1.1.1.1(GigabitEthernet0/0/1)'s neighbors
 Router ID: 88.1.1.2         Address: 1.1.1.2
   State: Full  Mode: Nbr is  Master  Priority: 1
   DR: 1.1.1.2  BDR: 1.1.1.1  MTU: 0
   Dead timer due in 36  sec
   Retrans timer interval: 0
   Neighbor is up for 00:03:25
   Authentication Sequence: [ 0 ]

                  Neighbors

 Area 0.0.0.0 interface 1.1.1.5 (Serial3/0/0)'s neighbors
 Router ID: 88.1.1.3         Address: 1.1.1.6
   ……
```



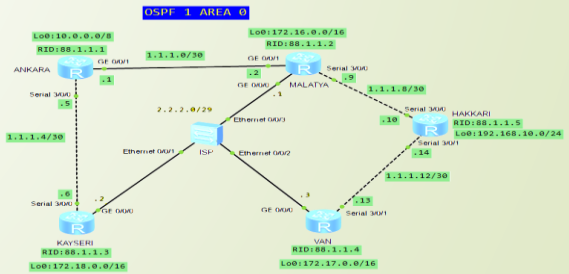
Example of Configuring a Single Area OSPF Network

```
<ANKARA> display ospf lsdb brief

         OSPF Process 1 with Router ID 88.1.1.1
                 LS Database Statistics

Area ID         Stub    Router   Network    S-Net   S-ASBR      Type-7  | Subtotal
0.0.0.0         0       5        2          0       0           0       | 7
Total           0       5        2          0       0           0       |
-------------------------------------------------------------------------+--------
Area ID                 Opq-9    Opq-10                                  | Subtotal
0.0.0.0         0       0                                                | 0
Total           0       0                                                |
-------------------------------------------------------------------------+--------
                ASE     Opq-11                                           | Subtotal
Total           0       0                                                | 0
-------------------------------------------------------------------------+--------
                                                                         | Total
                                                                         | 7
```

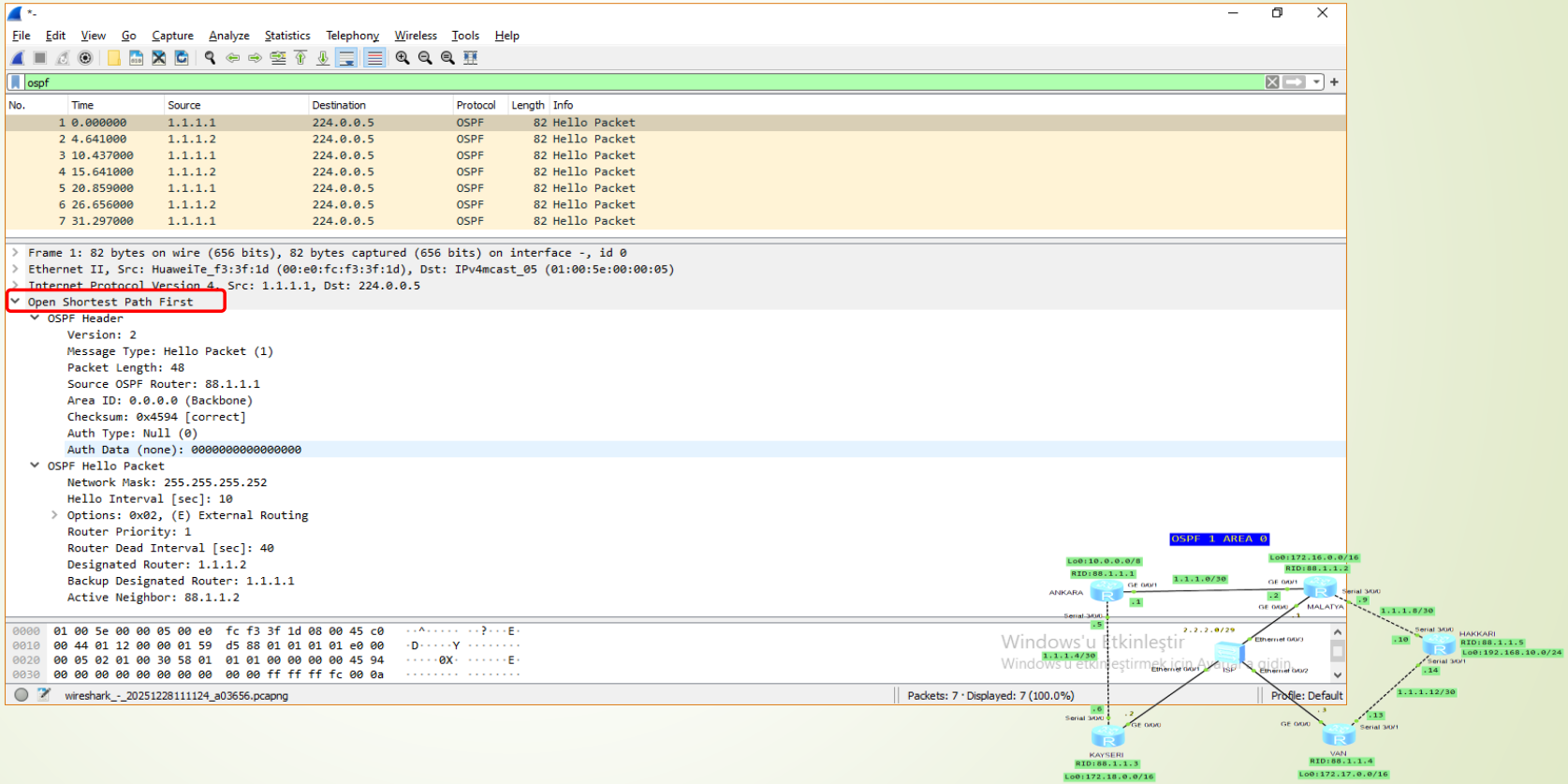


Example of Configuring a Single Area OSPF Network

```
<ANKARA> display ip routing-table protocol ospf
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Public routing table : OSPF
         Destinations : 7        Routes : 7

OSPF routing table status : <Active>
         Destinations : 7        Routes : 7

Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

        1.1.1.8/30  OSPF    10   49        D   1.1.1.2         GigabitEthernet 0/0/1
       1.1.1.12/30  OSPF    10   50        D   1.1.1.2         GigabitEthernet 0/0/1
        2.2.2.0/29  OSPF    10   2         D   1.1.1.2         GigabitEthernet 0/0/1
     172.16.0.1/32  OSPF    10   1         D   1.1.1.2         GigabitEthernet 0/0/1
     172.17.0.1/32  OSPF    10   2         D   1.1.1.2         GigabitEthernet 0/0/1
     172.18.0.1/32  OSPF    10   2         D   1.1.1.2         GigabitEthernet 0/0/1
   192.168.10.1/32  OSPF    10   49        D   1.1.1.2         GigabitEthernet 0/0/1

OSPF routing table status : <Inactive>
         Destinations : 0        Routes : 0
```



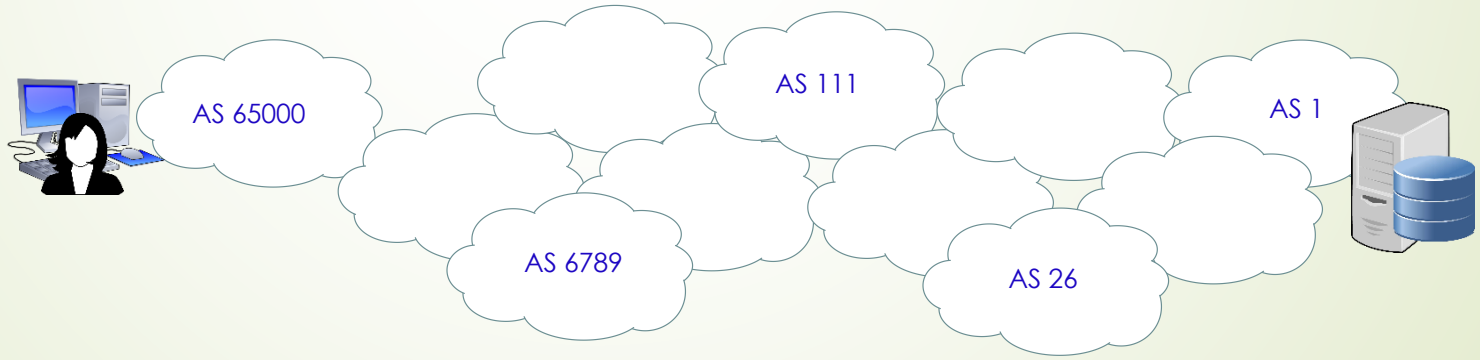
Example of Configuring a Single Area OSPF Network



BGP
Principles and Configurations

Border Gateway Protocol

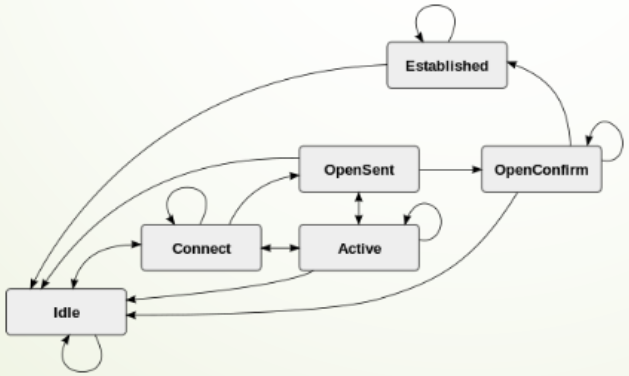
- Autonomous System (AS) is introduced into the Exterior Gateway Protocol (EGP). An AS is a set of routers under a single technical administration and using the same routing policies.
- The İnterior Gateway Protocol (IGP) is used within an AS to calculate and discover routes. Routers within the same AS trust each other, so IGP route calculation and information flooding are completely open and require little manual intervention.
- Inter-AS connection requirements promote the development of EGP. BGP is an EGP and used to control routes and select optimal routes between ASs.

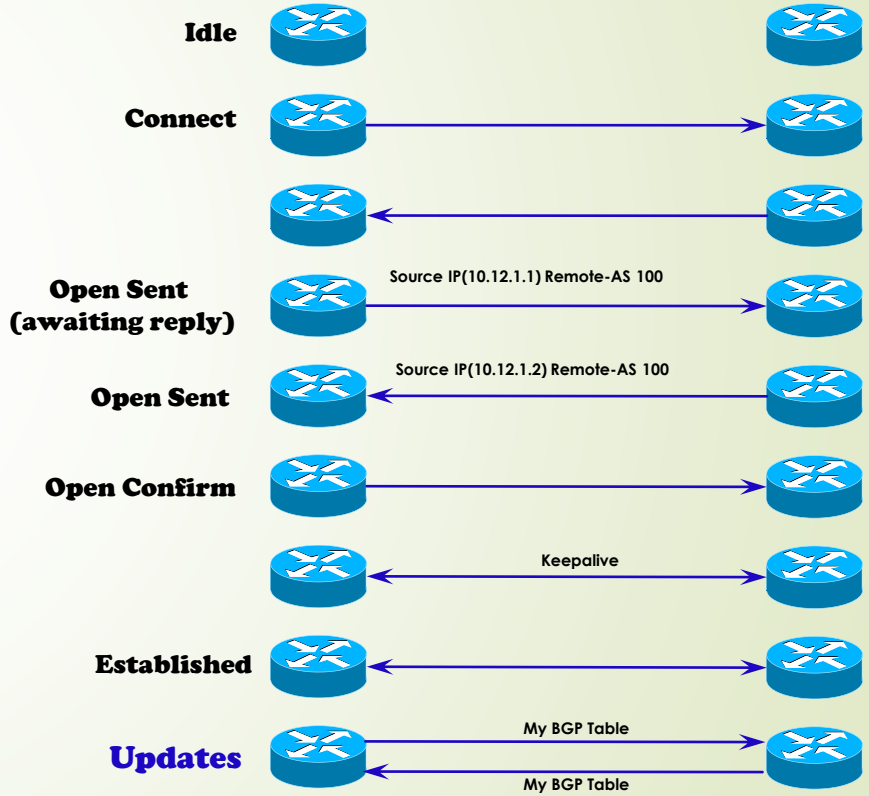


BGP Neighbor Relationship Establishment

The BGP session may report the following states:

- **Idle**
- **Connect**
- **Active**
- **OpenSent**
- **OpenConfirm**
- **Established**





BGP Neighbor Relationship Establishment

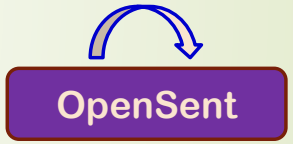
- BGP routers exchange BGP messages to establish neighbor relationships and update routing information. BGP messages are classified into Open, Update, Notification, Keepalive, and Route-refresh messages.
  - Open message: is the first message sent after a TCP connection is established. It is used to establish a BGP connection between neighbors. After a BGP neighbor receives an Open message and negotiation succeeds, the neighbor sends a Keepalive message to confirm and retain the connection. Then BGP neighbors can exchange Update, Notification, Keepalive, and Route-refresh messages.
  - Update message: is used to exchange routing information between BGP neighbors. It can advertise multiple reachable routes with the same route attributes and withdraw multiple unreachable routes.
  - Notification message: is sent when a BGP router detects an error. Then a BGP connection is terminated immediately.
  - Keepalive message: is sent periodically from a BGP router to its neighbor to retain their connection.
  - Route-refresh message: is sent by a BGP router to request its neighbor to send routing information again after this router changes its routing policy.

BGP Neighbor Relationship Establishment

- During message exchange, the Idle state is the initial state of a BGP router. In Idle state, the BGP router rejects the connection request from its neighbor. Only after receiving the Start event of itself, the BGP router tries to establish a TCP connection with its neighbor and transitions to the Connect state.
- In Connect state, the BGP router starts the Connect Retry timer, waiting for a TCP connection to be established.
  - If a TCP connection is established, the BGP router sends an Open message to its neighbor and transitions to the OpenSent state.
  - If a TCP connection fails to be established, the BGP router transitions to the Active state.
  - If the BGP router does not receive any response from its neighbor until the Connect Retry timer expires, the BGP router continues to try to establish a TCP connection with its neighbor and stays in the Connect state.

BGP Neighbor Relationship Establishment

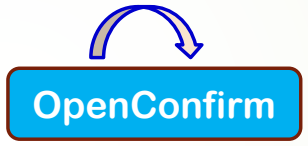
- In Active state, the BGP router always tries to establish a TCP connection.
  - If a TCP connection is established, the BGP router sends an Open message to its neighbor, turns off the Connect Retry timer, and transitions **to the OpenSent state.**
  - If a TCP connection fails to be established, the BGP router stays in **the Active state.**
  - If the BGP router does not receive any response from its neighbor until the Connect Retry timer expires, the BGP router transitions **to the Connect state.**



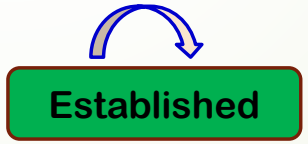
- In OpenSent state, the BGP router waits for an Open message from its neighbor and checks information carried in the Open message, including AS number, version number, and authentication password.
  - If the received Open message is correct, the BGP router sends a Keepalive message to its neighbor and transitions **to the OpenConfirm state.**
  - If the received Open message is incorrect, the BGP router sends a Notification message to its neighbor and transitions **to the Idle state.**

BGP Neighbor Relationship Establishment

- In OpenConfirm state, the BGP router waits for a Keepalive or Notification message from its neighbor. If it receives a Keepalive message, it transitions to the Established state. If it receives a Notification message, it transitions to the Idle state.



- In Established state, the BGP router can exchange Update, Keepalive, Route-refresh, and Notification messages with its neighbor.



BGP Route Generation Mode - Network

- Two BGP route generation modes are available: using the network command and using the import command.
- The network command imports existing routes in an IP routing table into a BGP routing table one by one.

```
#
 sysname LONDON
#
bgp 65500
 peer 10.0.14.1 as-number 65000
 #
 ipv4-family unicast
  undo synchronization
  network 50.1.1.0 255.255.255.0
  network 200.0.1.0
  network 200.0.2.0
  network 200.0.3.0
  peer 10.0.14.1 enable
#
```

BGP Route Generation Mode - Network

```
[LONDON]display bgp routing-table

BGP Local router ID is 10.0.14.2
Status codes: * - valid, > - best, d - damped,
              h - history,  i - internal, s - suppressed, S - Stale
              Origin : i - IGP, e - EGP, ? - incomplete


Total Number of Routes: 9
     Network            NextHop        MED LocPrf    PrefVal Path/Ogn

*>   10.0.12.0/30       10.0.14.1      0             0       65000?
*>   10.0.13.0/30       10.0.14.1      0             0       65000?
     10.0.14.0/30       10.0.14.1      0             0       65000?
*>   10.0.23.0/30       10.0.14.1                    0       65000?
*>   50.1.1.0/24        0.0.0.0        0             0       i
*>   100.3.3.0/24       10.0.14.1                    0       65000i
*>   200.0.1.0          0.0.0.0        0             0       i
*>   200.0.2.0          0.0.0.0        0             0       i
*>   200.0.3.0          0.0.0.0        0             0       i
```

BGP Route Generation Mode - Network

```
[ISTANBUL]display ip routing-table protocol bgp
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Public routing table : BGP
         Destinations : 6        Routes : 6

BGP routing table status : <Active>
         Destinations : 6        Routes : 6

Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

        10.0.23.0/30  IBGP   255  0           RD  10.0.12.2       GigabitEthernet 0/0/1
         50.1.1.0/24  EBGP   255  0           D   10.0.14.2       GigabitEthernet 0/0/0
        100.3.3.0/24  IBGP   255  0           RD  10.0.12.2       GigabitEthernet 0/0/1
        200.0.1.0/24  EBGP   255  0           D   10.0.14.2       GigabitEthernet 0/0/0
        200.0.2.0/24  EBGP   255  0           D   10.0.14.2       GigabitEthernet 0/0/0
        200.0.3.0/24  EBGP   255  0           D   10.0.14.2       GigabitEthernet 0/0/0
```

BGP Route Generation Mode - Import

```
[LONDON]ip ip-prefix IMPORT index 10 permit 200.0.1.0 24
[LONDON]ip ip-prefix IMPORT index 20 permit 200.0.2.0 24
[LONDON]ip ip-prefix IMPORT index 30 permit 200.0.3.0 24

[LONDON]route-policy IMPORT permit node 10
Info: New Sequence of this List.

[LONDON-route-policy]if-match ip-prefix IMPORT
[LONDON-route-policy]q

[LONDON]bgp 65500
[LONDON-bgp]ipv4-family unicast
[LONDON-bgp-af-ipv4]import-route direct route-policy IMPORT
```

BGP Route Generation Mode - Import

```
[LONDON]display bgp routing-table

BGP Local router ID is 10.0.14.2
Status codes: * - valid, > - best, d - damped,
              h - history,  i - internal, s - suppressed, S - Stale
              Origin : i - IGP, e - EGP, ? - incomplete


Total Number of Routes: 9
     Network            NextHop        MED        LocPrf     PrefVal Path/Ogn

*>   10.0.12.0/30       10.0.14.1      0                     0       65000?
*>   10.0.13.0/30       10.0.14.1      0                     0       65000?
     10.0.14.0/30       10.0.14.1      0                     0       65000?
*>   10.0.23.0/30       10.0.14.1                            0       65000?
*>   50.1.1.0/24        0.0.0.0        0                     0       i
*>   100.3.3.0/24       10.0.14.1                            0       65000i
*>   200.0.1.0          0.0.0.0        0                     0       ?
*>   200.0.2.0          0.0.0.0        0                     0       ?
*>   200.0.3.0          0.0.0.0        0                     0       ?
```

BGP Route Generation Mode - Import

```
[ISTANBUL]display ip routing-table protocol bgp
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Public routing table : BGP
         Destinations : 6        Routes : 6

BGP routing table status : <Active>
         Destinations : 6        Routes : 6

Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

        10.0.23.0/30  IBGP   255  0           RD  10.0.12.2       GigabitEthernet 0/0/1
         50.1.1.0/24  EBGP   255  0           D   10.0.14.2       GigabitEthernet 0/0/0
        100.3.3.0/24  IBGP   255  0           RD  10.0.12.2       GigabitEthernet 0/0/1
        200.0.1.0/24  EBGP   255  0           D   10.0.14.2       GigabitEthernet 0/0/0
        200.0.2.0/24  EBGP   255  0           D   10.0.14.2       GigabitEthernet 0/0/0
        200.0.3.0/24  EBGP   255  0           D   10.0.14.2       GigabitEthernet 0/0/0

BGP routing table status : <Inactive>
         Destinations : 0        Routes : 0
```

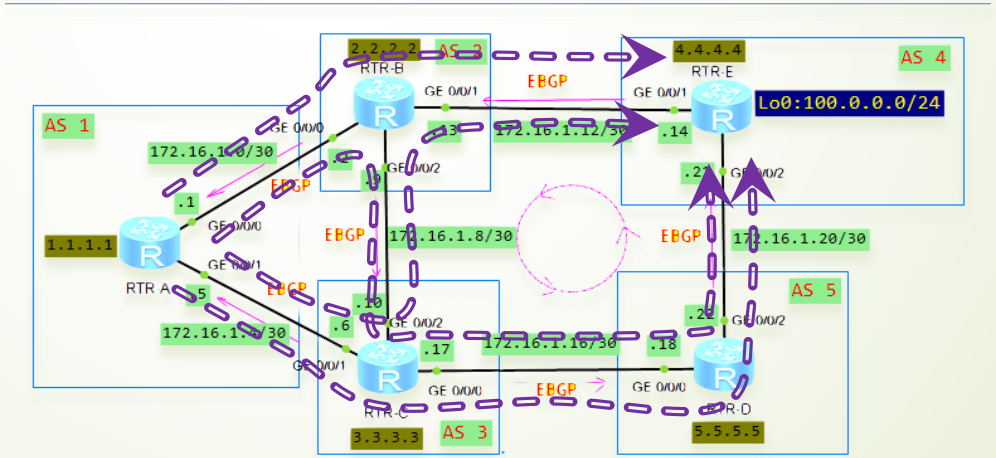
BGP Update Message

- BGP routes are generated in either network or import mode. They are encapsulated in UPDATE messages and advertised to neighbors. BGP advertises routing information only after a neighbor relationship is established.
- Update messages are used to advertise reachable routes and withdraw unreachable routes. An Update message contains the following information:
  - Network Layer Reachability Information (NLRI): Advertises the IP Prefix and prefix length.
  - Path Attribute: Provides loop detection and controls optimal route selection.
  - Withdrawn Route: Describes the prefix and prefix length of the unreachable withdrawn route.
- BGP route advertisement must follow specific rules to prevent potential problems.

BGP
Route Selection Rules

BGP Route Selection Rules

- After a BGP router advertises a route to neighbors, each neighbor selects the optimal route:
  - If this route is the only route to the destination, the neighbor selects it as the optimal route.
  - If there are multiple routes to the same destination, the neighbor selects the route with the highest priority as the optimal route.
  - If there are multiple routes that are destined for the same destination and have the same priority the neighbor compares attributes of these routes to select the optimal route.

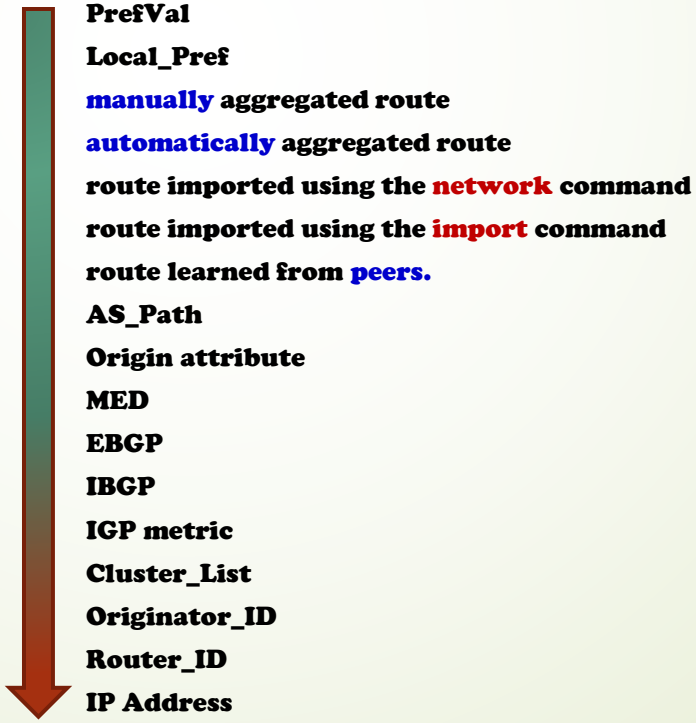


BGP Route Selection Rules

- Generally, BGP router compares the following attributes in sequence to select the optimal route:
  - Discards routes with unreachable next hops.
  - Prefers the route with the largest **PrefVal** value. The PrefVal attribute is a Huawei property attribute and is valid only on the device where it is configured.
  - Prefers the route with the highest **Local_Pref.**
  - Prefers the following routes in descending order of priority: manually aggregated route, automatically aggregated route, route imported using the network command, route imported using the import command, and route learned from peers.
  - Prefers the route with the shortest **AS_Path.**
  - Prefers the route with the **Origin attribute** IGP, EGP, and incomplete in sequence.
  - Prefers the route with the **smallest MED** if routes are received from the same AS.
  - Prefers an **EBGP** route to an **IBGP** route.
  - Prefers the route with the **smallest IGP metric** within an AS.

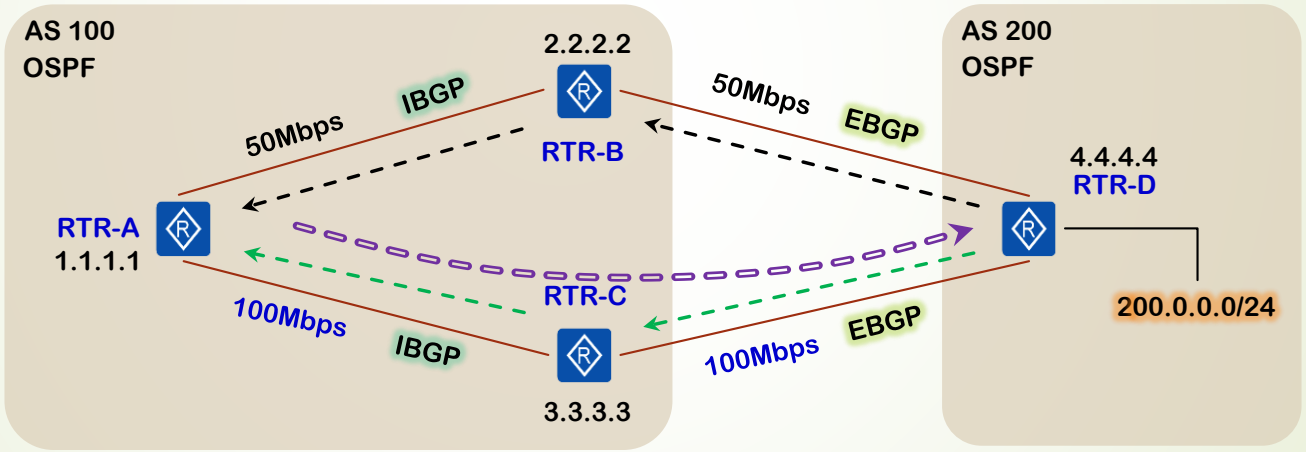
BGP Route Selection Rules

- Prefers the route with the shortest **Cluster_List.**
- Prefers the route with the smallest **Originator_ID.**
- Prefers the route advertised by the router with the smallest **Router_ID.**
- Prefers the route learned from the neighbor with the lowest **IP Address.**

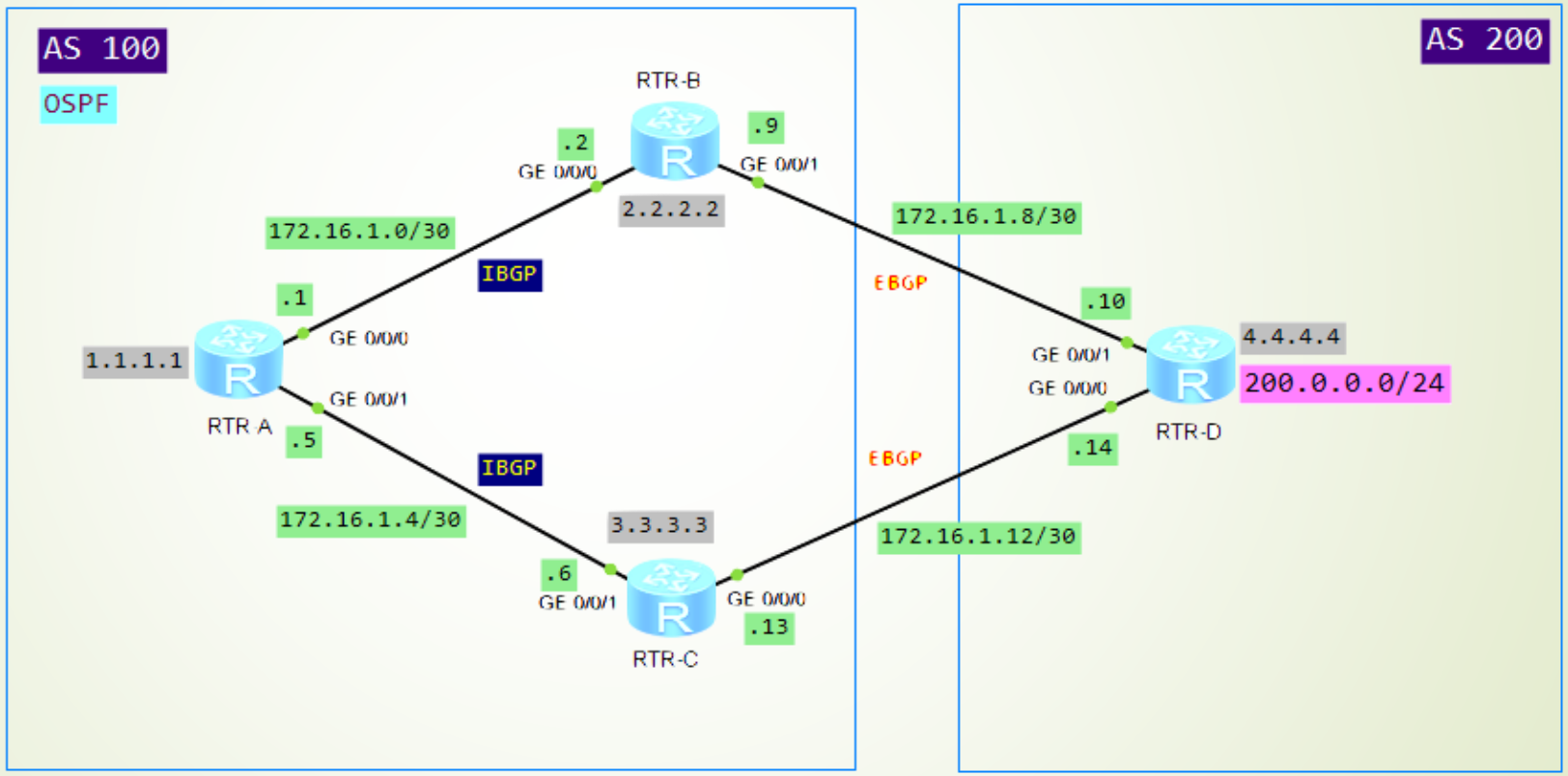


Impact of PrefVal on BGP Route Selection

- The PrefVal attribute is a Huawei property attribute and is valid only on the device where it is configured. It corresponds to the weight in BGP route selection rules. A larger **PrefVal value** a higher priority.



BGP Attribute – PrefVal



BGP Attribute – PrefVal

```
<RTR-A>display current-configuration interface lo0
[V200R003C00]
#
interface LoopBack0
 ip address 1.1.1.1 255.255.255.255
#
return
<RTR-A>display current-configuration interface g0/0/0
[V200R003C00]
#
interface GigabitEthernet0/0/0
 ip address 172.16.1.1 255.255.255.252
#
return
<RTR-A>display current-configuration interface g0/0/1
[V200R003C00]
#
interface GigabitEthernet0/0/1
 ip address 172.16.1.5 255.255.255.252
#
```

BGP Attribute – PrefVal

```
<RTR-A>display current-configuration section bgp
[V200R003C00]
#
bgp 100
 router-id 1.1.1.1
 peer 172.16.1.2 as-number 100
 peer 172.16.1.6 as-number 100
 #
 ipv4-family unicast
  undo synchronization
  network 1.1.1.1 255.255.255.255
  peer 172.16.1.2 enable
  peer 172.16.1.6 enable
#
return
```

BGP Attribute – PrefVal

```
<RTR-A>display bgp routing-table

BGP Local router ID is 1.1.1.1
Status codes: * - valid, > - best, d - damped,
              h - history,  i - internal, s - suppressed, S - Stale
              Origin : i - IGP, e - EGP, ? - incomplete


Total Number of Routes: 7
     Network            NextHop         MED LocPrf    PrefVal Path/Ogn

*>   1.1.1.1/32         0.0.0.0         0             0       i
*>i  2.2.2.2/32         172.16.1.2      0    100      0       i
*>i  3.3.3.3/32         172.16.1.6      0    100      0       i
*>i  4.4.4.4/32         172.16.1.2      0    100      0       200i
* i                     172.16.1.6      0    100      0       200i
*>i  200.0.0.0          172.16.1.6      0    100      0       200i
* i                     172.16.1.2      0    100      0       200i
```

BGP Attribute – PrefVal

```
<RTR-A>display ip routing-table protocol bgp
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Public routing table : BGP
         Destinations : 4        Routes : 4

BGP routing table status : <Active>
         Destinations : 4        Routes : 4

Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

        2.2.2.2/32  IBGP    255  0         RD    172.16.1.2      GigabitEthernet 0/0/0
        3.3.3.3/32  IBGP    255  0         RD    172.16.1.6      GigabitEthernet 0/0/1
        4.4.4.4/32  IBGP    255  0         RD    172.16.1.2      GigabitEthernet 0/0/0
      200.0.0.0/24  IBGP    255  0         RD    172.16.1.6      GigabitEthernet 0/0/1

BGP routing table status : <Inactive>
         Destinations : 0        Routes : 0
```

BGP Attribute – PrefVal

```
<RTR-B>display current-configuration interface lo0
[V200R003C00]
#
interface LoopBack0
 ip address 2.2.2.2 255.255.255.255
#
return
<RTR-B>display current-configuration interface g0/0/0
[V200R003C00]
#
interface GigabitEthernet0/0/0
 ip address 172.16.1.2 255.255.255.252
#
return
<RTR-B>display current-configuration interface g0/0/1
[V200R003C00]
#
interface GigabitEthernet0/0/1
 ip address 172.16.1.9 255.255.255.252
#
```

BGP Attribute – PrefVal

```
<RTR-B>display current-configuration section bgp
[V200R003C00]
#
bgp 100
 router-id 2.2.2.2
 peer 172.16.1.1 as-number 100
 peer 172.16.1.10 as-number 200
 #
 ipv4-family unicast
  undo synchronization
  network 2.2.2.2 255.255.255.255
  peer 172.16.1.1 enable
  peer 172.16.1.1 next-hop-local
  peer 172.16.1.10 enable
#
return
```

BGP Attribute – PrefVal

```
<RTR-B>display bgp routing-table

BGP Local router ID is 2.2.2.2
Status codes: * - valid, > - best, d - damped,
              h - history,  i - internal, s - suppressed, S - Stale
              Origin : i - IGP, e - EGP, ? - incomplete


Total Number of Routes: 4
     Network            NextHop        MED    LocPrf    PrefVa  Path/Ogn

*>i  1.1.1.1/32         172.16.1.1     0      100       0       i
*>   2.2.2.2/32         0.0.0.0        0                0       i
*>   4.4.4.4/32         172.16.1.10    0                0       200i
*>   200.0.0.0          172.16.1.10    0                0       200i
```

BGP Attribute – PrefVal

```
<RTR-B>display ip routing-table protocol bgp
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Public routing table : BGP
         Destinations : 3        Routes : 3

BGP routing table status : <Active>
         Destinations : 3        Routes : 3

Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

        1.1.1.1/32  IBGP    255  0         RD    172.16.1.1      GigabitEthernet 0/0/0
        4.4.4.4/32  EBGP    255  0         D     172.16.1.10     GigabitEthernet 0/0/1
      200.0.0.0/24  EBGP    255  0         D     172.16.1.10     GigabitEthernet 0/0/1

BGP routing table status : <Inactive>
         Destinations : 0        Routes : 0
```

BGP Attribute – PrefVal

```
<RTR-C>display current-configuration interface lo0
[V200R003C00]
#
interface LoopBack0
 ip address 3.3.3.3 255.255.255.255
#
return
<RTR-C>display current-configuration interface g0/0/0
[V200R003C00]
#
interface GigabitEthernet0/0/0
 ip address 172.16.1.13 255.255.255.252
#
return
<RTR-C>display current-configuration interface g0/0/1
[V200R003C00]
#
interface GigabitEthernet0/0/1
 ip address 172.16.1.6 255.255.255.252
#
```

BGP Attribute – PrefVal

```
<RTR-C>display current-configuration section bgp
[V200R003C00]
#
bgp 100
 router-id 3.3.3.3
 peer 172.16.1.5 as-number 100
 peer 172.16.1.14 as-number 200
 #
 ipv4-family unicast
  undo synchronization
  network 3.3.3.3 255.255.255.255
  peer 172.16.1.5 enable
  peer 172.16.1.5 next-hop-local
  peer 172.16.1.14 enable
#
return
```

BGP Attribute – PrefVal

```
<RTR-C>display bgp routing-table

BGP Local router ID is 3.3.3.3
Status codes: * - valid, > - best, d - damped,
              h - history,  i - internal, s - suppressed, S - Stale
              Origin : i - IGP, e - EGP, ? - incomplete


Total Number of Routes: 4
     Network            NextHop        MED    LocPrf    PrefVal Path/Ogn

*>i  1.1.1.1/32         172.16.1.5     0      100       0       i
*>   3.3.3.3/32         0.0.0.0        0                0       i
*>   4.4.4.4/32         172.16.1.14    0                0       200i
*>   200.0.0.0          172.16.1.14    0                0       200i
```

BGP Attribute – PrefVal

```
<RTR-C>display ip routing-table protocol bgp
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Public routing table : BGP
         Destinations : 3        Routes : 3

BGP routing table status : <Active>
         Destinations : 3        Routes : 3

Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

        1.1.1.1/32  IBGP    255  0         RD    172.16.1.5      GigabitEthernet 0/0/1
        4.4.4.4/32  EBGP    255  0         D     172.16.1.14     GigabitEthernet 0/0/0
      200.0.0.0/24  EBGP    255  0         D     172.16.1.14     GigabitEthernet 0/0/0

BGP routing table status : <Inactive>
         Destinations : 0        Routes : 0
```

BGP Attribute – PrefVal

```
<RTR-D>display current-configuration interface  lo0
#
interface LoopBack0
 ip address 4.4.4.4 255.255.255.255
#

<RTR-D>display current-configuration interface lo1
#
interface LoopBack1
 ip address 200.0.0.1 255.255.255.0
#

<RTR-D>display current-configuration interface g0/0/0
#
interface GigabitEthernet0/0/0
 ip address 172.16.1.14 255.255.255.252
#

<RTR-D>display current-configuration interface g0/0/1
#
interface GigabitEthernet0/0/1
 ip address 172.16.1.10 255.255.255.252
#
```

BGP Attribute – PrefVal

```
<RTR-D>display current-configuration section bgp
[V200R003C00]
#
bgp 200
 router-id 4.4.4.4
 peer 172.16.1.9 as-number 100
 peer 172.16.1.13 as-number 100
 #
 ipv4-family unicast
  undo synchronization
  network 4.4.4.4 255.255.255.255
  network 200.0.0.0
  peer 172.16.1.9 enable
  peer 172.16.1.13 enable
#
return
```

BGP Attribute – PrefVal

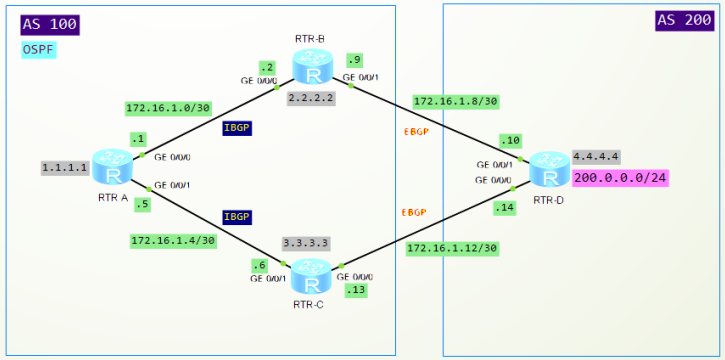
```
<RTR-D>display bgp routing-table

BGP Local router ID is 4.4.4.4
Status codes: * - valid, > - best, d - damped,
              h - history,  i - internal, s - suppressed, S - Stale
              Origin : i - IGP, e - EGP, ? - incomplete


Total Number of Routes: 6
     Network            NextHop        MED LocPrf    PrefVal Path/Ogn

*>   1.1.1.1/32         172.16.1.9                   0       100i
*                       172.16.1.13                  0       100i
*>   2.2.2.2/32         172.16.1.9     0             0       100i
*>   3.3.3.3/32         172.16.1.13    0             0       100i
*>   4.4.4.4/32         0.0.0.0        0             0       i
*>   200.0.0.0          0.0.0.0        0             0       i
```

BGP Attribute – PrefVal

```
<RTR-D>display ip routing-table protocol bgp
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Public routing table : BGP
         Destinations : 3        Routes : 3

BGP routing table status : <Active>
         Destinations : 3        Routes : 3

Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

        1.1.1.1/32  EBGP    255  0         D     172.16.1.9      GigabitEthernet 0/0/1
        2.2.2.2/32  EBGP    255  0         D     172.16.1.9      GigabitEthernet 0/0/1
        3.3.3.3/32  EBGP    255  0         D     172.16.1.13     GigabitEthernet 0/0/0

BGP routing table status : <Inactive>
         Destinations : 0        Routes : 0
```

BGP Attribute – PrefVal

- In the topology, AS 200 has a user network segment 200.0.0.0/24. The administrator of AS 100 wants to access this network segment over the high-bandwidth links and expects that the policy configured on RTA affects only its route selection.
- Method to meet this requirement:
  - On RTR-A, configure ip-prefix to match the route 200.0.0.0/24 received from RTR-C, configure a route-policy to invoke this ip-prefix, set the PrefVal value of the route to 100, and apply the route-policy in the import direction.
- Verification: Run the tracert command on RTR-C to check the routers that the route 200.0.0.0/24 passes through.



BGP Attribute – PrefVal

```
[RTR-A]ip ip-prefix PREFVAL_200 permit 200.0.0.0 24

[RTR-A]route-policy SET_PREFVAL permit node 10
Info: New Sequence of this List.
[RTR-A-route-policy]if-match ip-prefix PREFVAL_200
[RTR-A-route-policy]apply preferred-value 100
[RTR-A-route-policy]quit

[RTR-A]route-policy SET_PREFVAL permit node 20
Info: New Sequence of this List.
[RTR-A-route-policy]quit

[RTR-A]bgp 100
[RTR-A-bgp]ipv4-family unicast
[RTR-A-bgp-af-ipv4]peer 172.16.1.6 route-policy SET_PREFVAL import
[RTR-A-bgp-af-ipv4]q
[RTR-A-bgp]q
[RTR-A]q
```

BGP Attribute – PrefVal

```
<RTR-A>refresh bgp all import

<RTR-A>display bgp routing-table 200.0.0.0

 BGP local router ID : 1.1.1.1
 Local AS number : 100
 Paths:   2 available, 1 best, 1 select
 BGP routing table entry information of 200.0.0.0/24:
 From: 172.16.1.6 (3.3.3.3)
 Route Duration: 00h00m20s
 Relay IP Nexthop: 0.0.0.0
 Relay IP Out-Interface: GigabitEthernet0/0/1
 Original nexthop: 172.16.1.6
 Qos information : 0x0
 AS-path 200, origin igp, MED 0, localpref 100, pref-val 100, valid, internal, best, select, active, pre
255
 Not advertised to any peer yet
```

BGP Attribute – PrefVal

```
<RTR-A>display bgp routing-table 200.0.0.0
......

 BGP routing table entry information of 200.0.0.0/24:
 From: 172.16.1.2 (2.2.2.2)
 Route Duration: 00h00m20s
 Relay IP Nexthop: 0.0.0.0
 Relay IP Out-Interface: GigabitEthernet0/0/0
 Original nexthop: 172.16.1.2
 Qos information : 0x0
 AS-path 200, origin igp, MED 0, localpref 100, pref-val 0, valid, internal, pre 255, not preferred for
PreVal
 Not advertised to any peer yet
```

BGP Attribute – PrefVal

```
<RTR-A>display bgp routing-table

BGP Local router ID is 1.1.1.1
Status codes: * - valid, > - best, d - damped,
              h - history,  i - internal, s - suppressed, S - Stale
              Origin : i - IGP, e - EGP, ? - incomplete


Total Number of Routes: 7
     Network            NextHop        MED LocPrf    PrefVal Path/Ogn

*>   1.1.1.1/32         0.0.0.0        0             0       i
*>i  2.2.2.2/32         172.16.1.2     0    100      0       i
*>i  3.3.3.3/32         172.16.1.6     0    100      0       i
*>i  4.4.4.4/32         172.16.1.2     0    100      0       200i
* i                     172.16.1.6     0    100      0       200i
*>i  200.0.0.0          172.16.1.6     0    100      100     200i
* i                     172.16.1.2     0    100      0       200i
```

BGP Attribute – PrefVal

```
<RTR-A>tracert -a 1.1.1.1 200.0.0.1

traceroute to  200.0.0.1(200.0.0.1), max hops: 30 ,packet length: 40,press CTRL_C to break

1 172.16.1.6   30 ms      20 ms      20 ms
2 172.16.1.14  50 ms      20 ms      20 ms
```

QUESTIONS?

Prepared by Ayhan KOŞAR
ayhankosar@gmail.com
ayhan.k@inetworkeskisehir.com
www.inetworkeskisehir.com

linkedin.com/in/ayhankosar