OSPF Virtual Links: Transit capability

Contents

Introduction

The purpose of this document is to demonstrate the Open Shortest Path First (OSPF) behavior when the V-bit (Virtual-link bit) is present in a non-backbone area. The V-bit is signaled in Type-1 LSA only if the router is the endpoint of one or more fully adjacent virtual links. When the V-bit is set this could change path calculation preference between intra-area and inter-area routes.

Prerequisites

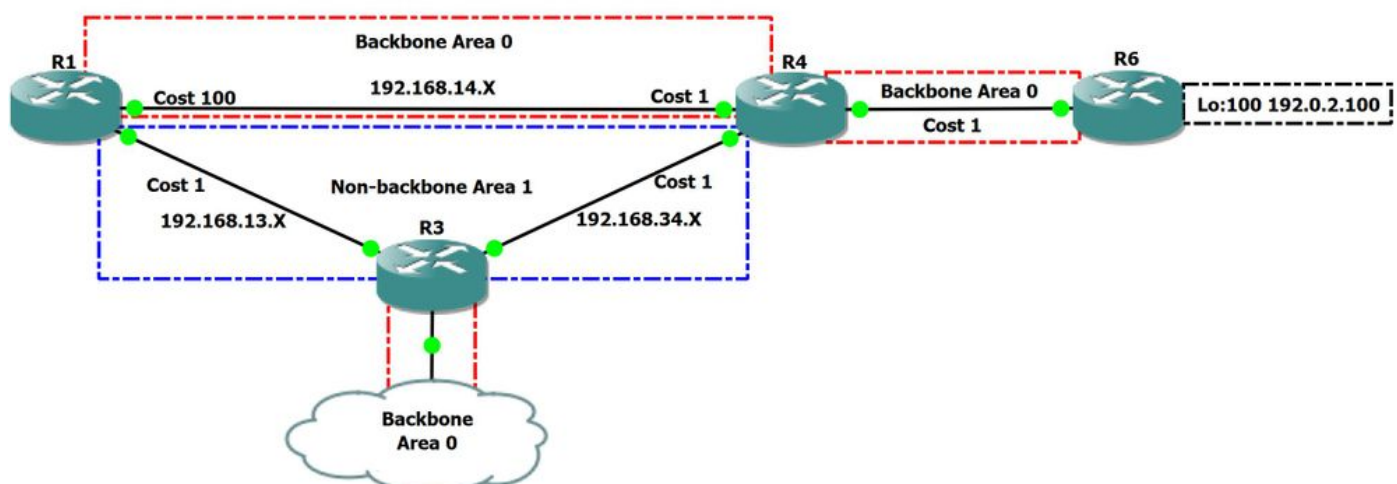Refer to the network diagram in Figure 1 as you use this document:

Figure 1

In the network diagram above, we have both backbone area 0 and non-backbone area 1. R1 is an Area Border Router (ABR) connecting both area 0 and area 1, R4 and R3 have a similar role in this network. In this topology area 0 is discontiguous since R3 and R4 are not connected via area 0.

Background Information

All areas in an OSPF autonomous system must be connected to the backbone area (area 0). In some cases where you have a non-backbone area between your backbone area, this could cause some areas of the autonomous system to become unreachable and results in your network being discontiguous. When it is not possible to have a contiguous backbone area, you may use a virtual link to connect your backbone through a non-backbone area. The area through which you configure the virtual link is known as a transit area.
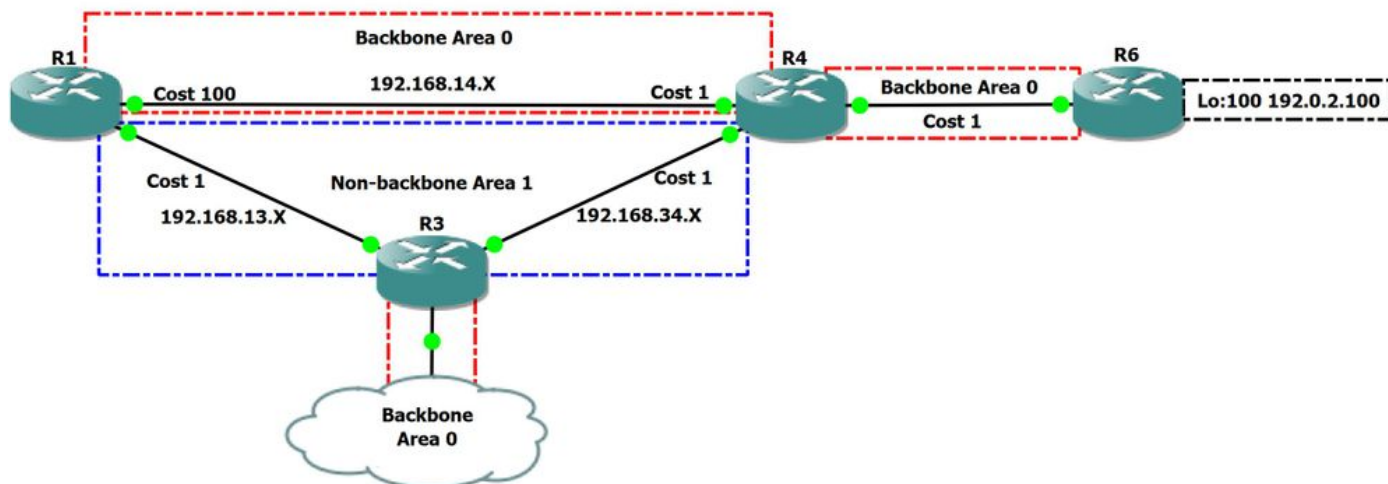
Scenario 1

Network Diagram:

Figure 2

In this scenario, we will be going over expected path calculation in the above network topology. We will be investigating what path is preferred when routing from R1 towards R6 loopback 100 which has an ip address of 192.0.2.100/32

Lets have a look at the OSPF database on R1 to further undestand the topology:

```
R1#show ip ospf database

            OSPF Router with ID (1.1.1.1) (Process ID 1)

            Router Link States (Area 0)

Link ID         ADV Router      Age         Seq#        Checksum Link count
1.1.1.1         1.1.1.1         22          0x8000000C 0x00CD7A 2
4.4.4.4         4.4.4.4         289         0x8000000F 0x00434E 4
6.6.6.6         6.6.6.6         374         0x80000009 0x00630A 3

            Summary Net Link States (Area 0)

Link ID         ADV Router      Age         Seq#        Checksum
192.168.13.0    1.1.1.1         18          0x80000001 0x00348D
192.168.13.0    4.4.4.4         207         0x80000001 0x00E3D0
192.168.34.0    1.1.1.1         8           0x80000001 0x005655
192.168.34.0    4.4.4.4         683         0x80000001 0x00F1AE

            Router Link States (Area 1)

Link ID         ADV Router      Age         Seq#        Checksum Link count
1.1.1.1         1.1.1.1         17          0x80000009 0x00EC2B 2
3.3.3.3         3.3.3.3         18          0x8000000E 0x005A64 4
```

```
4.4.4.4          4.4.4.4          544          0x80000005 0x0007CF 2

                    Summary Net Link States (Area 1)

Link ID          ADV Router       Age          Seq#       Checksum
155.1.37.0       3.3.3.3          1558         0x80000004 0x00A7C3
192.0.2.100      1.1.1.1          23           0x80000001 0x009F0C    <- R6 Loopback
192.0.2.100      4.4.4.4          370          0x80000001 0x0059AA    <- R6 Loopback
192.168.14.0     1.1.1.1          23           0x80000001 0x000B52
192.168.14.0     4.4.4.4          331          0x80000001 0x00CEE5
192.168.34.0     1.1.1.1          3608         0x80000002 0x00406C
192.168.46.0     1.1.1.1          23           0x80000001 0x00B388
192.168.46.0     4.4.4.4          484          0x80000001 0x006D27
```

From the above output we can see that R1 learns R6 Lo100:192.0.2.100 via R4 as a Type-3 Summary LSA, R1 is also originating itself a Type-3 Summary LSA since it knows R6 Lo100:192.0.2.100 via intra-area backbone. In the below output we can see that R6 has 192.0.2.100 directly connected.

```
R1#show ip ospf da router 6.6.6.6

 OSPF Router with ID (1.1.1.1) (Process ID 1)

            Router Link States (Area 0)

 LS age: 614
 Options: (No TOS-capability, DC)
 LS Type: Router Links
 Link State ID: 6.6.6.6
 Advertising Router: 6.6.6.6
 LS Seq Number: 8000000D
 Checksum: 0x5B0E
 Length: 60
 Number of Links: 3

 Link connected to: a Stub Network
 (Link ID) Network/subnet number: 192.0.2.100       <-- Loopback 100 directly connected
 (Link Data) Network Mask: 255.255.255.255
 Number of MTID metrics: 0
 TOS 0 Metrics: 1

 Link connected to: another Router (point-to-point)
 (Link ID) Neighboring Router ID: 4.4.4.4
 (Link Data) Router Interface address: 192.168.46.6
 Number of MTID metrics: 0
 TOS 0 Metrics: 1

 Link connected to: a Stub Network
 (Link ID) Network/subnet number: 192.168.46.0
 (Link Data) Network Mask: 255.255.255.0
 Number of MTID metrics: 0
 TOS 0 Metrics: 1
```

Abstract from RFC 2328 Section 16.2

```
16.2.  Calculating the inter-area routes


        (5) Next, look up the routing table entry for the destination N.
```

(If N is an AS boundary router, look up the "router" routing
table entry associated with Area A). If no entry exists for
N or if the entry's path type is "type 1 external" or "type
2 external", then install the inter-area path to N, with
associated area Area A, cost IAC, next hop equal to the list
of next hops to router BR, and Advertising router equal to
BR.

(6) Else, if the paths present in the table are intra-area
paths, do nothing with the LSA **(intra-area paths are always
preferred).**

(7) **Else, the paths present in the routing table are also
inter-area paths. Install the new path through BR if it is
cheaper,** overriding the paths in the routing table.
Otherwise, if the new path is the same cost, add it to the
list of paths that appear in the routing table entry.

In the above output we can see that it is stated intra-area routes are prefered over inter-area routes. So in our scenario R1 should prefer going via intra-area backbone per RFC 2328.

Lets check if this behaviour is observed in our topology:

```
R1#show ip ospf rib 192.0.2.100

   OSPF Router with ID (1.1.1.1) (Process ID 1)

 Base Topology (MTID 0)

OSPF local RIB
Codes: * - Best, > - Installed in global RIB
LSA: type/LSID/originator

*> 192.0.2.100/32, Intra, cost 102, area 0
 SPF Instance 9, age 02:19:34
 Flags: RIB, HiPrio
 via 192.168.14.4, GigabitEthernet3 label 1048578
 Flags: RIB
 LSA: 1/6.6.6.6/6.6.6.6

R1#show ip route 192.0.2.100
Routing entry for 192.0.2.100/32
 Known via "ospf 1", distance 110, metric 102, type intra area
 Last update from 192.168.14.4 on GigabitEthernet3, 02:26:29 ago
 Routing Descriptor Blocks:
 * 192.168.14.4, from 6.6.6.6, 02:26:29 ago, via GigabitEthernet3
 Route metric is 102, traffic share count is 1
```

As you can see from the outputs above we prefer going over backbone area 0 towards R6 loopback100. In our Link State Database we are also aware of an inter-area path through R3 then R4. The summary LSA which is learned via R4 with a cost of 2 can be seen below:

```
R1#show ip ospf database summary 192.0.2.100

            OSPF Router with ID (1.1.1.1) (Process ID 1)

                Summary Net Link States (Area 1)

  LS age: 523
  Options: (No TOS-capability, DC, Upward)
  LS Type: Summary Links(Network)
```

```
Link State ID: 192.0.2.100 (summary Network Number)
Advertising Router: 1.1.1.1
LS Seq Number: 80000005
Checksum: 0x9710
Length: 28
Network Mask: /32
      MTID: 0          Metric: 102


LS age: 973
Options: (No TOS-capability, DC, Upward)
LS Type: Summary Links(Network)
Link State ID: 192.0.2.100 (summary Network Number)
Advertising Router: 4.4.4.4                        <- This is Type-3 LSA injected by ABR R4
LS Seq Number: 80000005
Checksum: 0x51AE
Length: 28
Network Mask: /32
      MTID: 0          Metric: 2
```

Please take into account that this cost of 2 reflects the cost which the ABR has towards the destination prefix. Type-3 LSAs are flooded from area 0 into non-backbone areas and vice-versa, it describes ABR's reachability towards links in other areas. It Includes the cost from the ABRs perspective who enjected the Type-3 LSA, but hides full cost from the router who received the Type-3 LSA.

From the output above we now know we have two paths we could take to reach R6 loopback from R1:

1. Intra-area which has a cost of 102
2. Inter-area which has a cost of 2 known via Type-3 LSA + R1 cost towards R4 which is also 2. This gives us a total cost of 4

In this scenario we have already observed that we are preferring a higher cost intra-area path since it is defined in RFC 2328 that intra-area is prefered over inter-area.

Before proceding with scenario 2 here is an example of how OSPF interprets Type-3 LSAs:

• ABR R4 can reach link A intra-area with cost of X
• R1 can reach ABR R4 with a cost Y
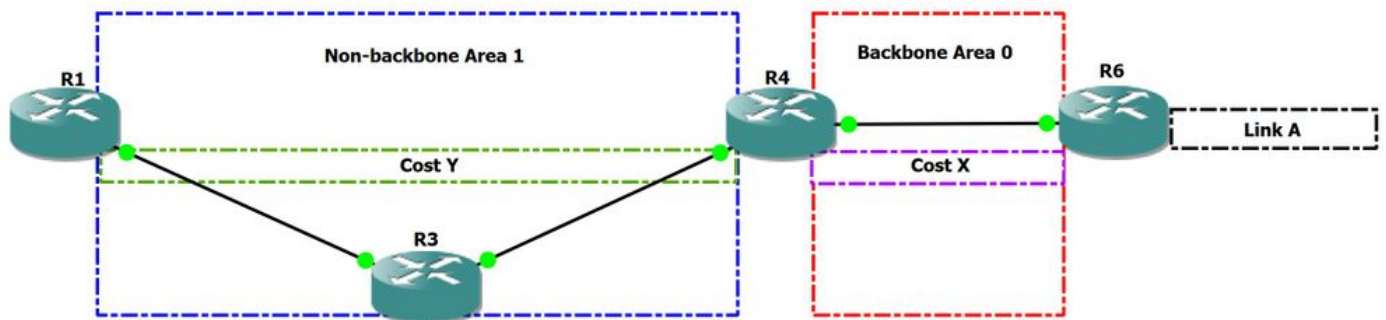• Implies R1 can reach link A via SPT with a cost of X + Y

Figure 3

This is why inter-area routing is usually compared with distance vector protocols, since information between areas is hidden.
Because inter-area OSPF is distance vector, it is vulnerable to routing loops. It avoids loops by mandating a loop-free inter-area topology, in which traffic from one area can only reach another area through area 0.
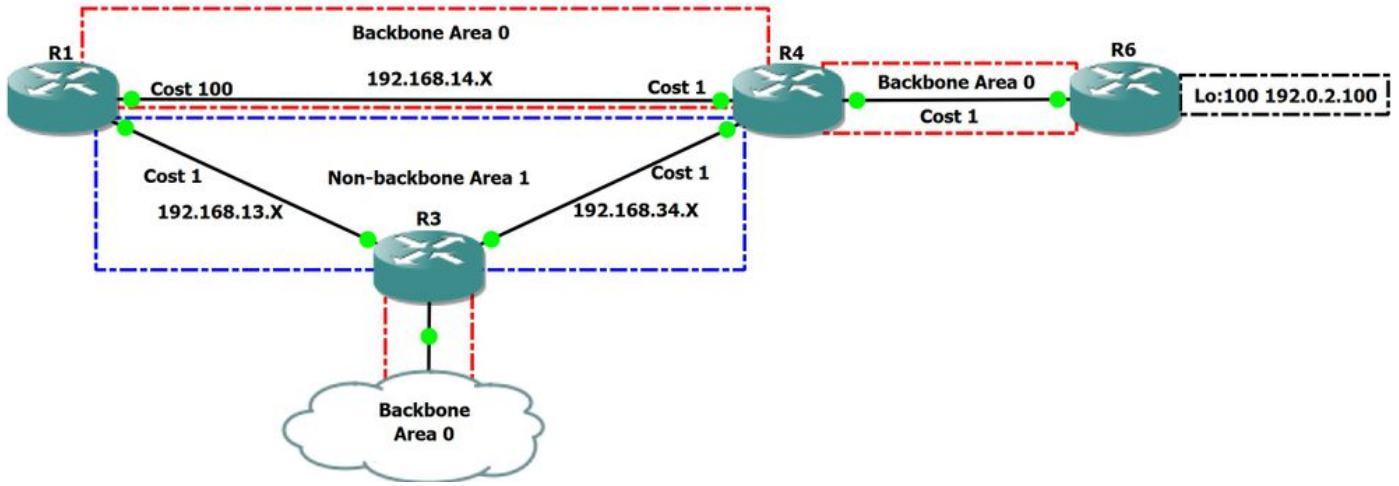
Scenario 2

Network Diagram:

Figure 4

In this scenario we set the V-bit on R3 and R4 so we could check path preference when this bit is present in Type-1 LSA of non-backbone area 1.

Abstract from RFC 2328 Section 6

```
6. The Area Data Structure

TransitCapability
        This parameter indicates whether the area can carry data traffic
        that neither originates nor terminates in the area itself. This
        parameter is calculated when the area's shortest-path tree is
        built (see Section 16.1, where TransitCapability is set to TRUE
        if and only if there are one or more fully adjacent virtual
        links using the area as Transit area), and is used as an input
        to a subsequent step of the routing table build process (see
        Section 16.3). When an area's TransitCapability is set to TRUE,
        the area is said to be a "transit area".
```

Abstract from RFC 2328 Section 16.1

```
16.1 Calculating the shortest-path tree for an area

        (2) Call the vertex just added to the tree vertex V.  Examine
          the LSA associated with vertex V.  This is a lookup in the
          Area A's link state database based on the Vertex ID.  If
          this is a router-LSA, and bit V of the router-LSA (see
          Section A.4.2) is set, set Area A's TransitCapability to
          TRUE.  In any case, each link described by the LSA gives the
          cost to an adjacent vertex.  For each described link, (say
          it joins vertex V to vertex W):
```

From the above statement in RFC we can see that when the V-bit is set in the router-LSA, we know that area in which the bit is set to be transit capable or in other words when running Dijkstra algorithm the TransitCapability is true for that area.

Once we know that an area could be considered for capability transit if there is a V-bit set, we must check if this functionality is configured: The OSPF Area Transit Capability feature is enabled by default.

```
R1#show run all | sec ospf
router ospf 1
capability opaque
capability lls
capability transit
```

To set the V-bit in area 1 we will create a virtual-link from R3 towards R4. When the virtual link is brought up, we should see the V-bit set in the Type-1 LSA.

```
R3(config)#router ospf 1
R3(config-router)#area 1 virtual-link 4.4.4.4

R3#show ip ospf interface brief
Interface   PID   Area            IP Address/Mask    Cost   State Nbrs F/C
VL0         1     0               192.168.34.3/24    1      P2P   1/1     <-- Here we have
Virtual-link present and 1 neighborship over VL0
Gi3         1     0               192.168.80.3/24    1      DR    0/0
Gi2         1     1               192.168.13.3/24    1      P2P   1/1
Gi1         1     1               192.168.34.3/24    1      P2P   1/1
R3#
```

Now lets check Type-1 LSA for R3 area 1.

```
R3#show ip ospf 1 1 database router 3.3.3.3 OSPF Router with ID (3.3.3.3) (Process ID 1) Router Link States (Area 1)
LS age: 189 Options: (No TOS-capability, DC) LS Type: Router Links Link State ID: 3.3.3.3 Advertising Router:
3.3.3.3 LS Seq Number: 80000018 Checksum: 0x525E Length: 72 Area Border Router Virtual Link Endpoint      <-
V-bit set
  Number of Links: 4

    Link connected to: another Router (point-to-point)
     (Link ID) Neighboring Router ID: 1.1.1.1
     (Link Data) Router Interface address: 192.168.13.3
      Number of MTID metrics: 0
       TOS 0 Metrics: 1

    Link connected to: a Stub Network
     (Link ID) Network/subnet number: 192.168.13.0
     (Link Data) Network Mask: 255.255.255.0
      Number of MTID metrics: 0
       TOS 0 Metrics: 1

    Link connected to: another Router (point-to-point)
     (Link ID) Neighboring Router ID: 4.4.4.4
     (Link Data) Router Interface address: 192.168.34.3
      Number of MTID metrics: 0
       TOS 0 Metrics: 1

    Link connected to: a Stub Network
     (Link ID) Network/subnet number: 192.168.34.0
     (Link Data) Network Mask: 255.255.255.0
      Number of MTID metrics: 0
       TOS 0 Metrics: 1
```

As we can see in the above output, R3 now has the V-bit set on its Type-1 LSA for area 1 and has capability transit enabled in routing process level.

We can also see that R1 has cababilty transit enabled for area 1 in the below output:

```
R1#show ip ospf
 Routing Process "ospf 1" with ID 1.1.1.1
 Start time: 00:02:48.412, Time elapsed: 01:27:00.690
 Supports only single TOS(TOS0) routes
 Supports opaque LSA
 Supports Link-local Signaling (LLS)
 Supports area transit capability
 Supports NSSA (compatible with RFC 3101)
 Supports Database Exchange Summary List Optimization (RFC 5243)
 Event-log enabled, Maximum number of events: 1000, Mode: cyclic
 It is an area border router
 Router is not originating router-LSAs with maximum metric
 Initial SPF schedule delay 5000 msecs
 Minimum hold time between two consecutive SPFs 10000 msecs
 Maximum wait time between two consecutive SPFs 10000 msecs
 Incremental-SPF disabled
 Minimum LSA interval 5 secs
 Minimum LSA arrival 1000 msecs
 LSA group pacing timer 240 secs
 Interface flood pacing timer 33 msecs
 Retransmission pacing timer 66 msecs
 EXCHANGE/LOADING adjacency limit: initial 300, process maximum 300
 Number of external LSA 0. Checksum Sum 0x000000
 Number of opaque AS LSA 0. Checksum Sum 0x000000
 Number of DCbitless external and opaque AS LSA 0
 Number of DoNotAge external and opaque AS LSA 0
 Number of areas in this router is 2. 2 normal 0 stub 0 nssa
 Number of areas transit capable is 1
 External flood list length 0
 IETF NSF helper support enabled
 Cisco NSF helper support enabled
 Reference bandwidth unit is 100 mbps
    Area BACKBONE(0)
        Number of interfaces in this area is 1
        Area has no authentication
        SPF algorithm last executed 00:00:33.554 ago
        SPF algorithm executed 11 times
        Area ranges are
        Number of LSA 10. Checksum Sum 0x05EB7B
        Number of opaque link LSA 0. Checksum Sum 0x000000
        Number of DCbitless LSA 0
        Number of indication LSA 0
        Number of DoNotAge LSA 3
        Flood list length 0
    Area 1
        Number of interfaces in this area is 1
        This area has transit capability         <-- This area is transit capabile
        Area has no authentication
        SPF algorithm last executed 00:00:04.259 ago
        SPF algorithm executed 8 times
        Area ranges are
        Number of LSA 10. Checksum Sum 0x0517AA
        Number of opaque link LSA 0. Checksum Sum 0x000000
        Number of DCbitless LSA 0
        Number of indication LSA 0
        Number of DoNotAge LSA 0
        Flood list length 0
```

Since area 1 now passes all criteria to become a transit area we should now observe a different path calculation/preference then seen before in our first scenario.

It is stated RFC 2328 if an area is considered as transit area it should be examinded differently than non-transit areas

Abstract from RFC 2328 Section 16.1

```
16.3. Examining transit areas' summary-LSAs
```

This step is only performed by area border routers attached to one or more non-backbone areas that are capable of carrying transit traffic (i.e., "transit areas", or those areas whose TransitCapability parameter has been set to TRUE in Step 2 of the Dijkstra algorithm (see Section 16.1).

```
The purpose of the calculation below is to examine the transit
areas to see whether they provide any better (shorter) paths
than the paths previously calculated in Sections 16.1 and 16.2.
Any paths found that are better than or equal to previously
discovered paths are installed in the routing table.
```

According to the RFC, if the area is transit-capable, it is subject to the path calculation described in section 16.3 of RFC 2328

Note: that in this example the virtual link enables transit data traffic to be forwarded through Area 1, but the actual path the transit data traffic takes does not need follow the virtual link. In other words, virtual links allow transit traffic to be forwarded through an area, but do not dictate the precise path that the traffic will take.

Let's assume capability transit was disabled on R1. Let's check the path towards the destination R6 loopback:100 192.0.2.100 with a traceroute.

```
R1#traceroute 192.0.2.100
Tracing the route to 192.0.2.100
VRF info: (vrf in name/id, vrf out name/id)
1 192.168.14.4 2 msec 2 msec 2 msec    <--R4
2 192.168.46.6 3 msec 3 msec *         <--R6
```

Once we turn this functionality on with the V-bit set in area 1 we observe the following logs:

```
R1#debug ip ospf spf intra
OSPF SPF intra debugging is on
R1#debug ip ospf spf inter OSPF SPF inter debugging is on R1#conf Enter configuration commands,
one per line. End with CNTL/Z. R1(config)#router ospf 1 R1(config-router)#capability transit
R1(config-router)#
```
*Aug 14 15:28:07.934: OSPF-1 INTER: Running spf for summaries in transit area 1
```
*Aug 14 15:28:07.934: OSPF-1 INTER: Summary transit processing lsid 192.0.2.100 adv_rtr 4.4.4.4
type 3 seq 0x8000000B
*Aug 14 15:28:07.934: OSPF-1 INTER: Summary metric 2
*Aug 14 15:28:07.934: OSPF-1 INTER: found best path to adv_rtr:
i,ABR [2] via 192.168.13.3, GigabitEthernet1, Area 1 orp_txit_adv_rtr 0.0.0.0 pathflag 0x0
*Aug 14 15:28:07.934: OSPF-1 INTER: Add transit path via area 1
*Aug 14 15:28:07.934: OSPF-1 SPF : Exist path: next-hop 192.168.13.3, interface GigabitEthernet1
```
*Aug 14 15:28:07.934: OSPF-1 INTRA: Route update succeeded for 192.0.2.100/255.255.255.255, metric 4, Next Hop: GigabitEthernet1/192.168.13.3 area 0

Now lets check how R1 routes towards R6 loopback100

```
R1#show ip ospf rib 192.0.2.100

            OSPF Router with ID (1.1.1.1) (Process ID 1)


                Base Topology (MTID 0)

OSPF local RIB
Codes: * - Best, > - Installed in global RIB
LSA: type/LSID/originator

*>  192.0.2.100/32, Intra, cost 4, area 0
     SPF Instance 14, age 00:12:28
     Flags: RIB, HiPrio, Transit
      via 192.168.13.3, GigabitEthernet1 label 1048578
       Flags: RIB
       LSA: 1/6.6.6.6/6.6.6.6

R1#show ip route 192.0.2.100
Routing entry for 192.0.2.100/32
Known via "ospf 1", distance 110, metric 4, type intra area
Last update from 192.168.13.3 on GigabitEthernet1, 00:01:26 ago
Routing Descriptor Blocks:
 * 192.168.13.3, from 6.6.6.6, 00:01:26 ago, via GigabitEthernet1
Route metric is 4, traffic share count is 1
```

Why do we see Intra-area instead of Inter-area? In RFC 2328 section 16.3 it is mentioned that when doing path calulcation if we have a route which is of lower cost over transit area (Type-3) we should update the next-hop of the prefix. This is indeed the behaviour we are seeing in the above output. The next-hop mentioned is correct, but the type is misleading.

Abstract from RFC 2328 Section 16.3

```
16.3. Examining transit areas' summary-LSAs

(4) Look up the routing table entry for the advertising router
BR associated with the Area A. If it is unreachable, examine
the next LSA. Otherwise, the cost to destination N is the
sum of the cost in BR's Area A routing table entry and the
cost advertised in the LSA. Call this cost IAC.
```

(5) **If this cost is less than the cost occurring in N's routing**
        table entry, overwrite N's list of next hops with those used
        for BR, and set N's routing table cost to IAC. Else, if IAC
        is the same as N's current cost, add BR's list of next hops
        to N's list of next hops. In any case, the area associated
        with N's routing table entry must remain the backbone area,
        and the path type (either intra-area or inter-area) must
        also remain the same.

R1 is prefering inter-area Type-3 over Type-1 intra-area route, although it is stated as intra-area in the output. We clearly see the next-hop is not assosiated to area 0

```
R1#show ip ospf neighbor

Neighbor ID     Pri   State           Dead Time   Address         Interface
4.4.4.4           0   FULL/  -        00:00:39    192.168.14.4    GigabitEthernet3
3.3.3.3           0   FULL/  -        00:00:32    192.168.13.3    GigabitEthernet1


R1#show ip ospf neighbor detail
```

```
 Neighbor 4.4.4.4, interface address 192.168.14.4
    In the area 0 via interface GigabitEthernet3
    Neighbor priority is 0, State is FULL, 6 state changes
    DR is 0.0.0.0 BDR is 0.0.0.0
    Options is 0x12 in Hello (E-bit, L-bit)
    Options is 0x52 in DBD (E-bit, L-bit, O-bit)
    LLS Options is 0x1 (LR)
    Dead timer due in 00:00:36
    Neighbor is up for 00:30:20
    Index 1/1/1, retransmission queue length 0, number of retransmission 3
    First 0x0(0)/0x0(0)/0x0(0) Next 0x0(0)/0x0(0)/0x0(0)
    Last retransmission scan length is 1, maximum is 2
    Last retransmission scan time is 135 msec, maximum is 135 msec

Neighbor 3.3.3.3, interface address 192.168.13.3
    In the area 1 via interface GigabitEthernet1
    Neighbor priority is 0, State is FULL, 6 state changes
    DR is 0.0.0.0 BDR is 0.0.0.0
    Options is 0x12 in Hello (E-bit, L-bit)
    Options is 0x52 in DBD (E-bit, L-bit, O-bit)
    LLS Options is 0x1 (LR)
    Dead timer due in 00:00:39
    Neighbor is up for 00:30:20
    Index 1/1/2, retransmission queue length 0, number of retransmission 3
    First 0x0(0)/0x0(0)/0x0(0) Next 0x0(0)/0x0(0)/0x0(0)
    Last retransmission scan length is 4, maximum is 4
    Last retransmission scan time is 126 msec, maximum is 126 msec
```

Let's also traceroute towards the destination of R6 loopback100:

```
R1#traceroute 192.0.2.100
Tracing the route to 192.0.2.100
VRF info: (vrf in name/id, vrf out name/id)
1 192.168.13.3 2 msec 4 msec 3 msec          <-- R3
2 192.168.34.4 5 msec 3 msec 3 msec          <-- R4
3 192.168.46.6 5 msec 6 msec *               <-- R6
R1#
```
Hence in the above output we see that the non-backbone area 1 is preferred over the backbone area 0 to reach R6 loopback 100.

It is also possible to have ECMP (Equal Cost Multipath) using both intra-area and inter-area routes if the cost between them is equal. This could be done in our topology by decreasing R1s link towards R4 from 100 to 2.

When this is done we have the following output in both RIB and OSPF RIB:

```
R1#show ip ospf rib 192.0.2.100 OSPF Router with ID (1.1.1.1) (Process ID 1) Base Topology (MTID 0) OSPF local RIB
Codes: * - Best, > - Installed in global RIB LSA: type/LSID/originator *> 192.0.2.100/32, Intra, cost 4, area 0 SPF
Instance 14, age 00:13:08 Flags: RIB, HiPrio, Transit, OldTrans via 192.168.13.3, GigabitEthernet1 label 1048578
Flags: RIB LSA: 1/6.6.6.6/6.6.6.6 via 192.168.14.4, GigabitEthernet3 label 1048578 Flags: RIB LSA: 1/6.6.6.6/6.6.6.6
```

```
R1#show ip route 192.0.2.100
Routing entry for 192.0.2.100/32
Known via "ospf 1", distance 110, metric 4, type intra area
Last update from 192.168.14.4 on GigabitEthernet3, 00:12:44 ago
Routing Descriptor Blocks:
192.168.14.4, from 6.6.6.6, 00:12:44 ago, via GigabitEthernet3
Route metric is 4, traffic share count is 1
* 192.168.13.3, from 6.6.6.6, 00:12:44 ago, via GigabitEthernet1
Route metric is 4, traffic share count is 1
```