You make **possible**

Troubleshooting OSPF

Nicholas Russo
CCIE 42518 (RS/SP), CCDE 20160041
BRKRST-3310

Cisco live!
June 9-13, 2019 • San Diego, CA
#CLUS

Agenda

- Introduction
- Graph Theory
- Repairing Adjacencies
- Optimal Routing
- Mastering “Forward Address”
- Automated Troubleshooting

Cisco Webex Teams

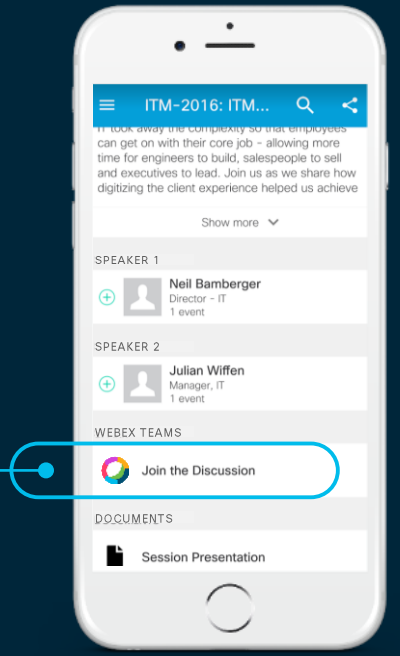
Questions?

Use Cisco Webex Teams to chat with the speaker after the session

How

1. Find this session in the Cisco Live Mobile App
2. Click “Join the Discussion”
3. Install Webex Teams or go directly to the team space
4. Enter messages/questions in the team space

Webex Teams will be moderated by the speaker until June 16, 2019.



cs.co/ciscolivebot#BRKRST-3310

What this session is all about

- Brilliance in the basics
- Deep understanding
- A set of things you'll actually remember

What this session is not about

- A fact recitation
- A break-fix Tour de Force
- An exhaustive list of every OSPF detail
  - Bits, flags, codes, and RFCs

- Reference icon in upper right hand corner

Graph Theory & Scientific Method

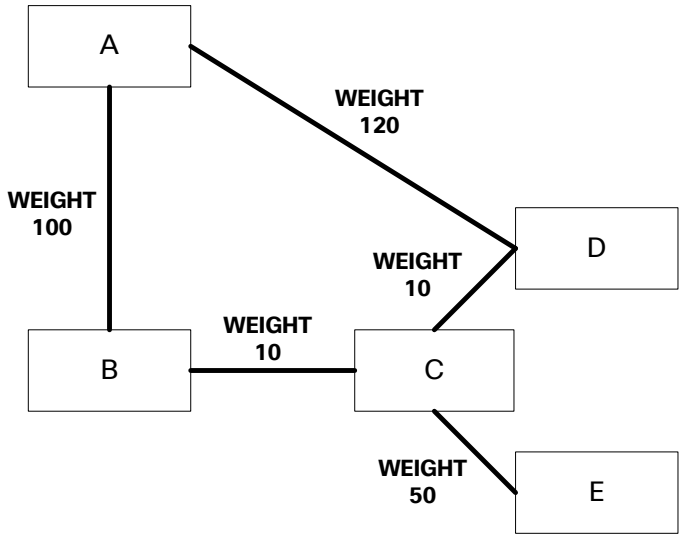
You make networking **possible**

Cisco live!

Graph Theory simplified

- A graph is a set of vertices (nodes) connected together via edges
  - Each vertex has a unique identifier
  - Each edge may have a weight and direction
  - Each edge connects exactly two vertices

- In OSPF terms
  - Each node has a unique 4 byte ID
  - Each link consists of two directed edges
  - Each link connects exactly two nodes



The Scientific Method simplified

- Structured approach to generating empirical evidence
  1. Observe the environment
  2. Ask how and/or why?
  3. Form a hypothesis
  4. Conduct an experiment (i.e., test the hypothesis)
  5. Analyze the result

“Plans are worthless,
but planning is everything.”

Dwight D. Eisenhower

34th President of the United States

The Scientific Method on Twitter

- I hypothesized the following:
  - The majority of OSPF engineers want to master the LSDB (the tool)
  - Those who encounter poor designs want to understand them (the outcome)

Nicholas Russo @nickrusso42518

Since I'm covering Troubleshooting OSPF at Cisco Live 2018, I need to validate hypotheses before building content. Which of the following is the most important OSPF skill you wish you had? Be honest.

32% Mastery of link state DB

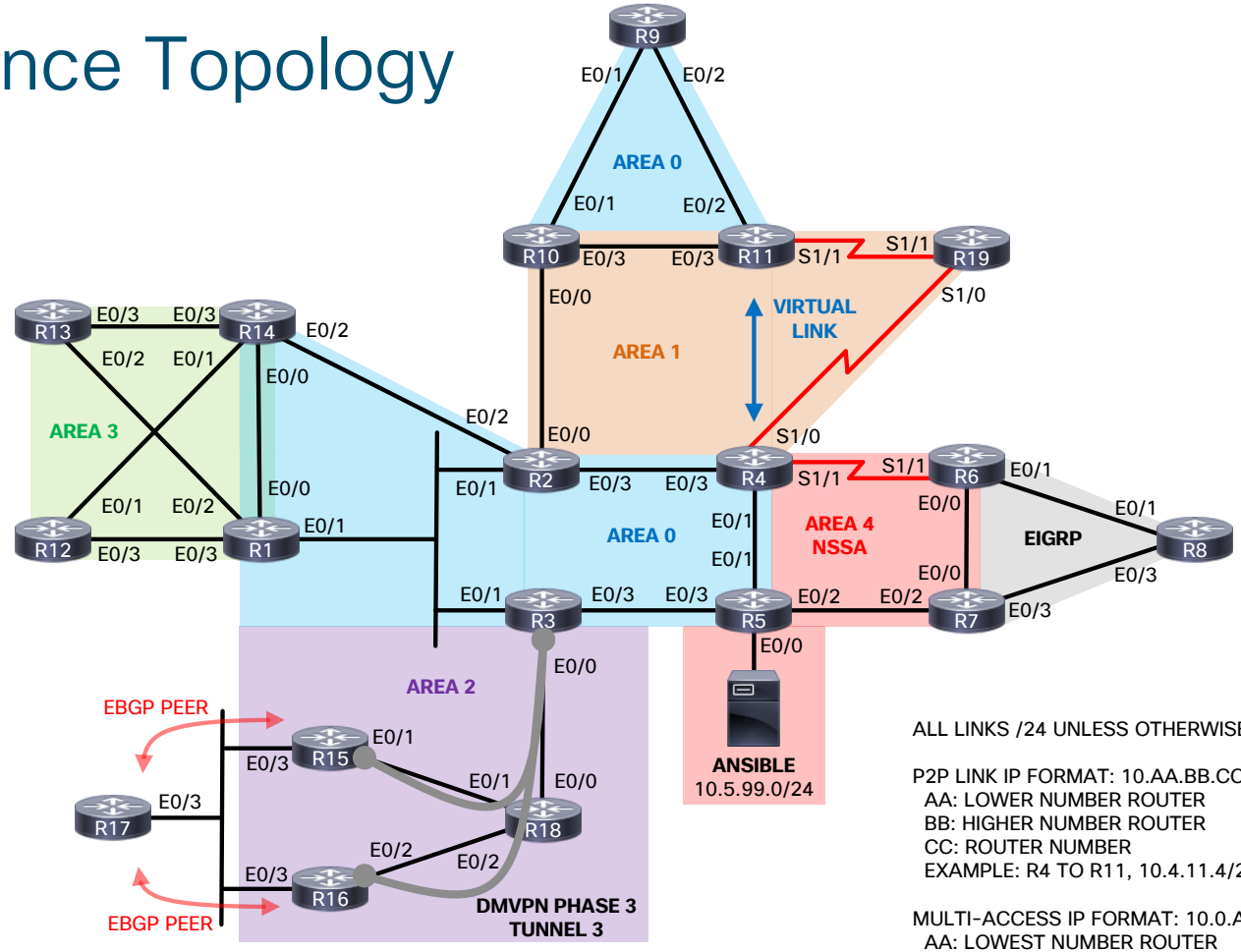
8% Fixing broken adjacencies

14% Performance: SPF, timers

46% Navigating poor designs

78% of respondents supported my hypothesis (n = 405)

Reference Topology



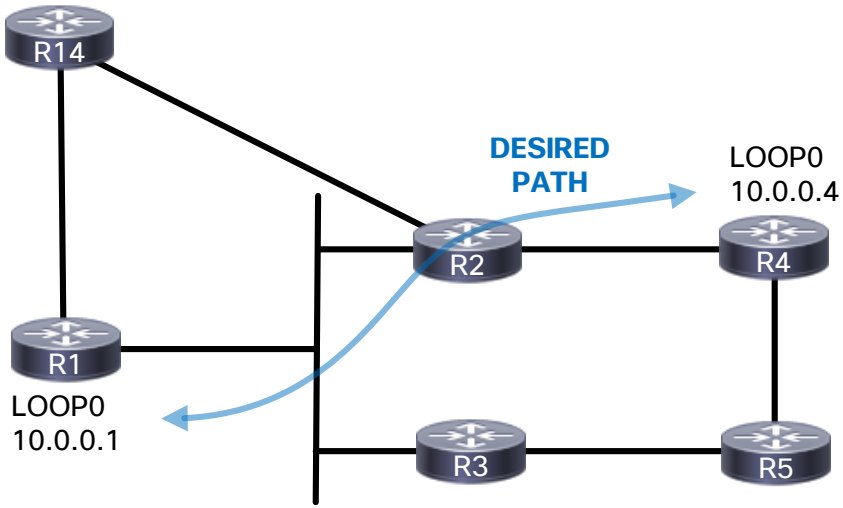
ALL LINKS /24 UNLESS OTHERWISE NOTED

P2P LINK IP FORMAT: 10.AA.BB.CC/24
- AA: LOWER NUMBER ROUTER
- BB: HIGHER NUMBER ROUTER
- CC: ROUTER NUMBER
- EXAMPLE: R4 TO R11, 10.4.11.4/24 ON R4

MULTI-ACCESS IP FORMAT: 10.0.AA.CC/24
- AA: LOWEST NUMBER ROUTER
- CC: ROUTER NUMBER
- EXAMPLE: R1/R2/R3, 10.0.1.3/24 ON R3

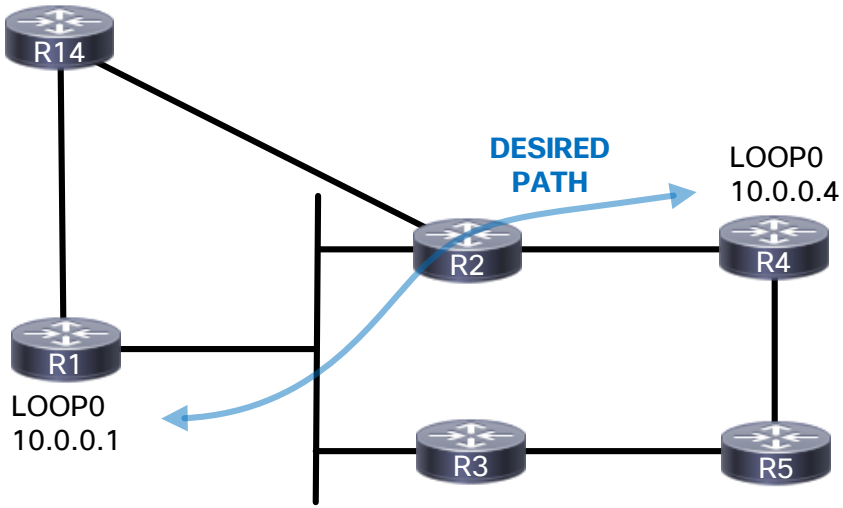
Web downloads don’t work

- R4 is trying to download R1’s startup config via HTTP
- Download keeps timing out
- R14 was recently hardened for security reasons
- Routing via R2 is desired
- The customer is only giving you access to R4



Scientific Method in action

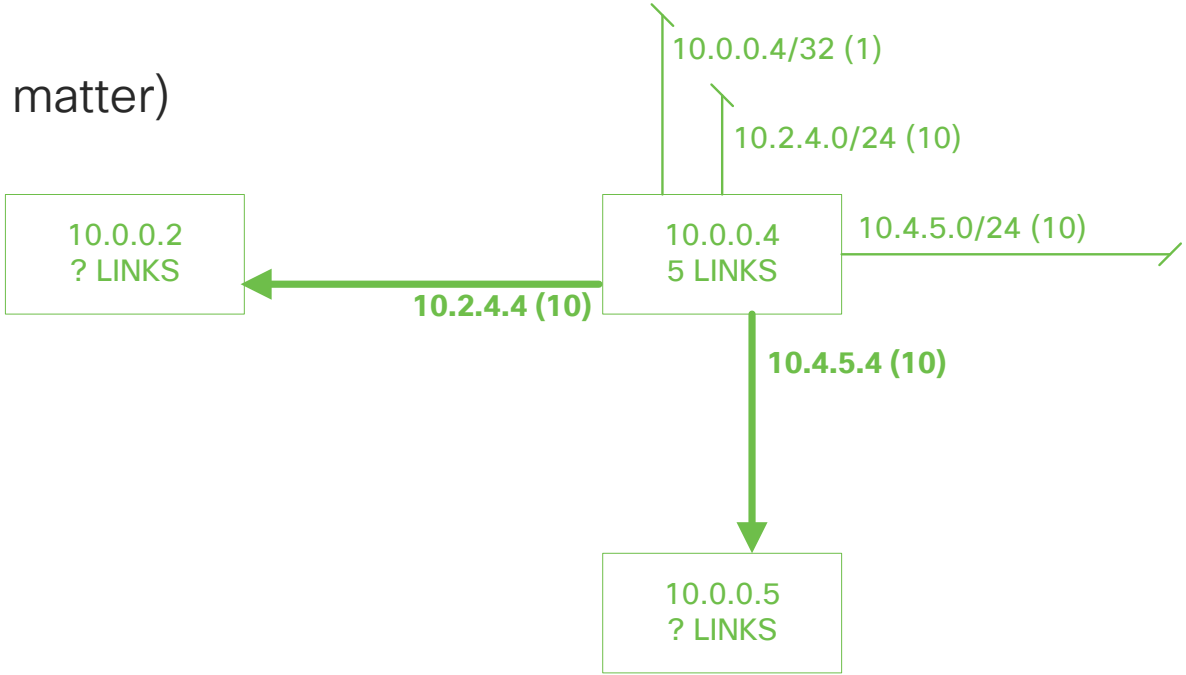
- Observe
  - R4→R1 ping/traceroute look OK
- Ask how and/or why?
  - Why are web downloads failing when ping/traceroute work?



- Form hypothesis
  - Return traffic R1→R4 is erroneously routing via R14

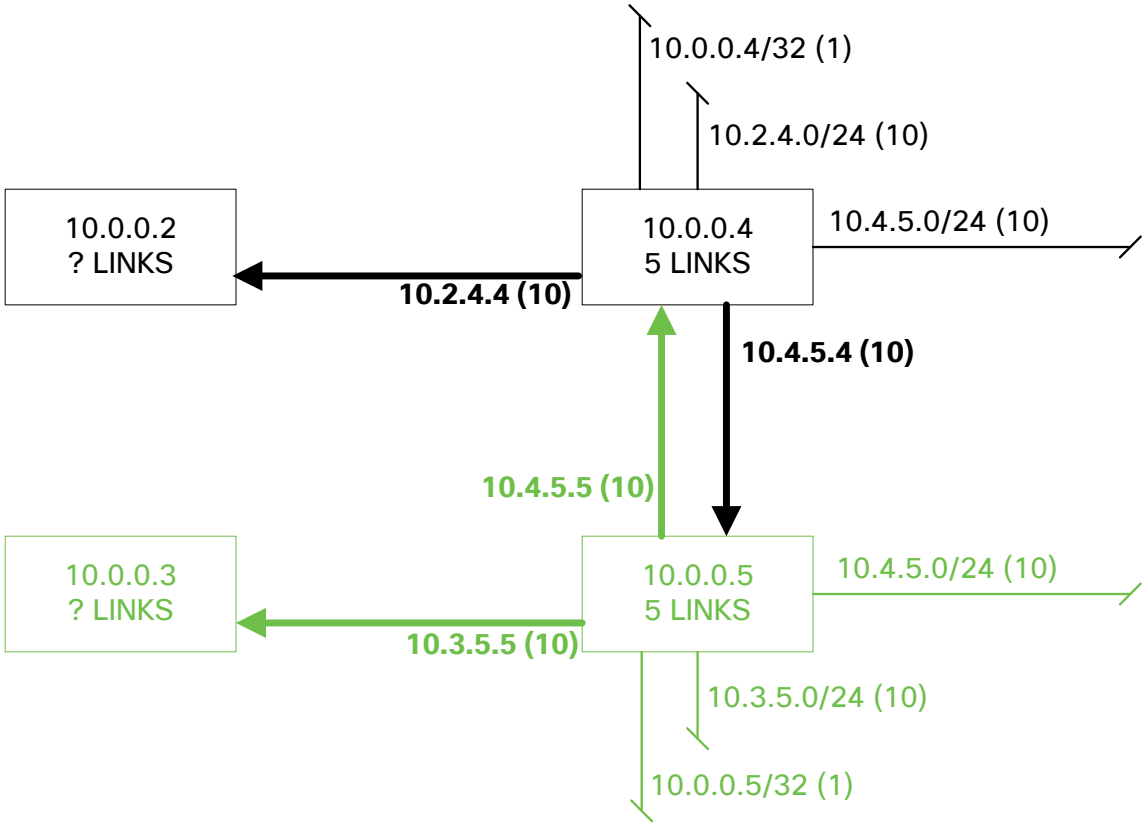
Begin Drawing Graph

- Start at R4 (doesn't matter)
- Ask it about itself

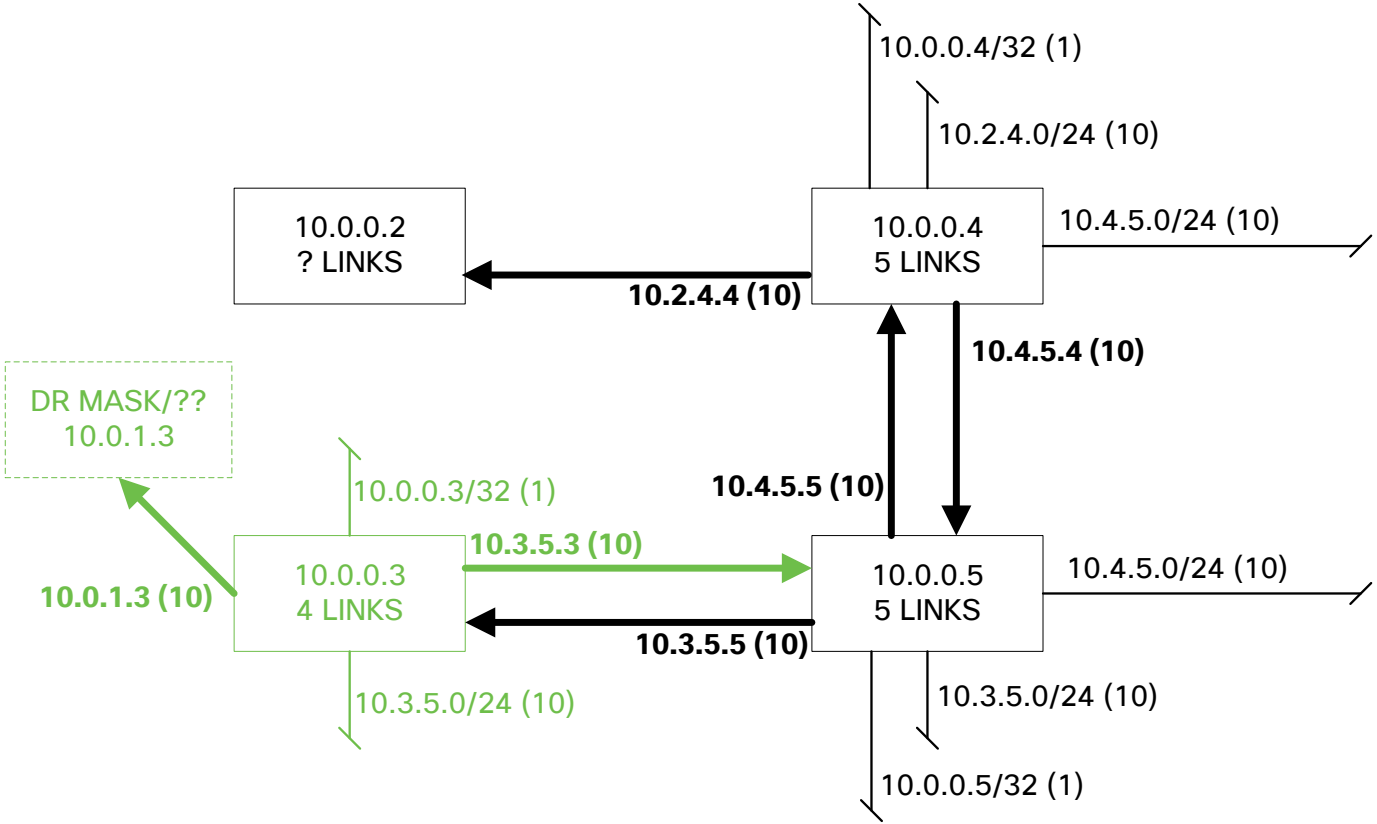


Continue Graph (R5)

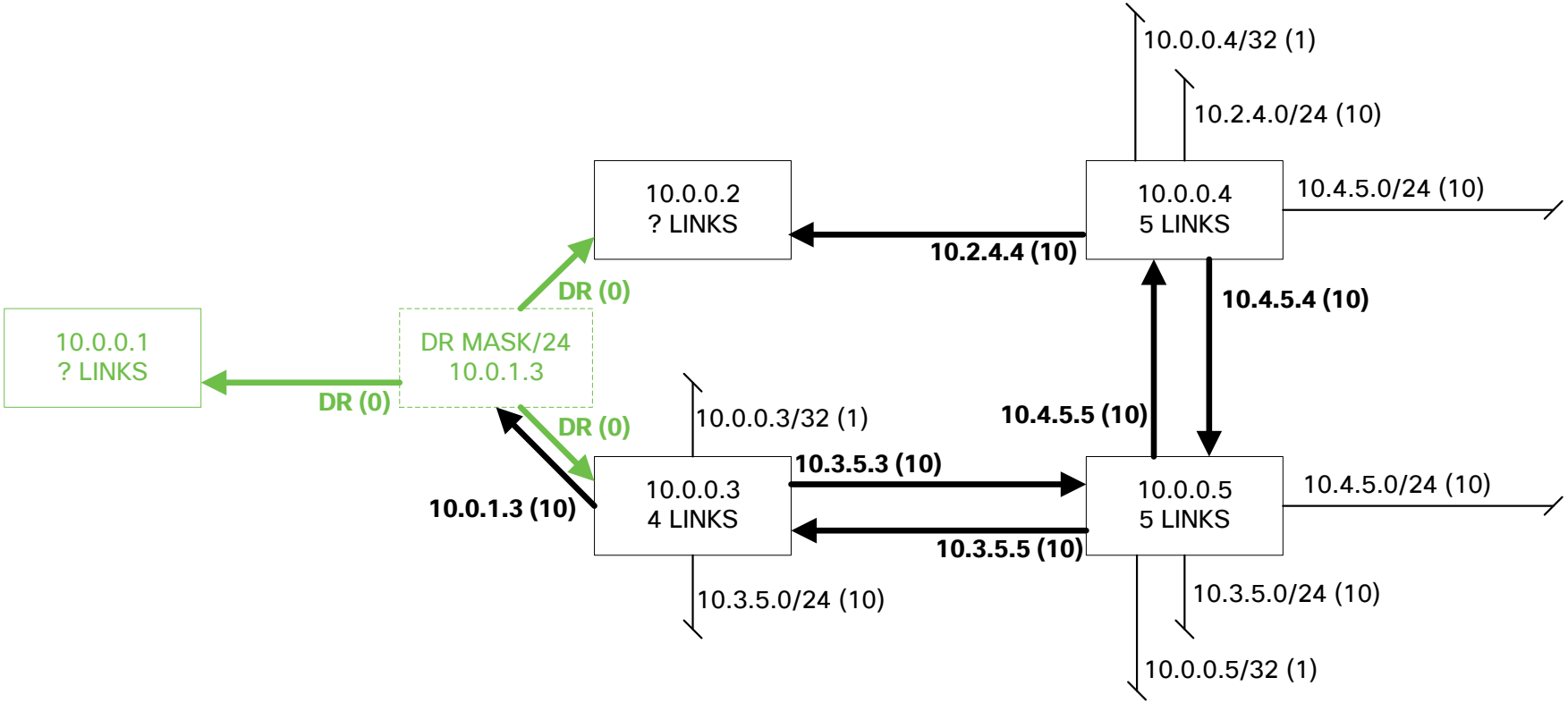
- Advance to a newly discovered node
- Does not matter which one
- I chose R5



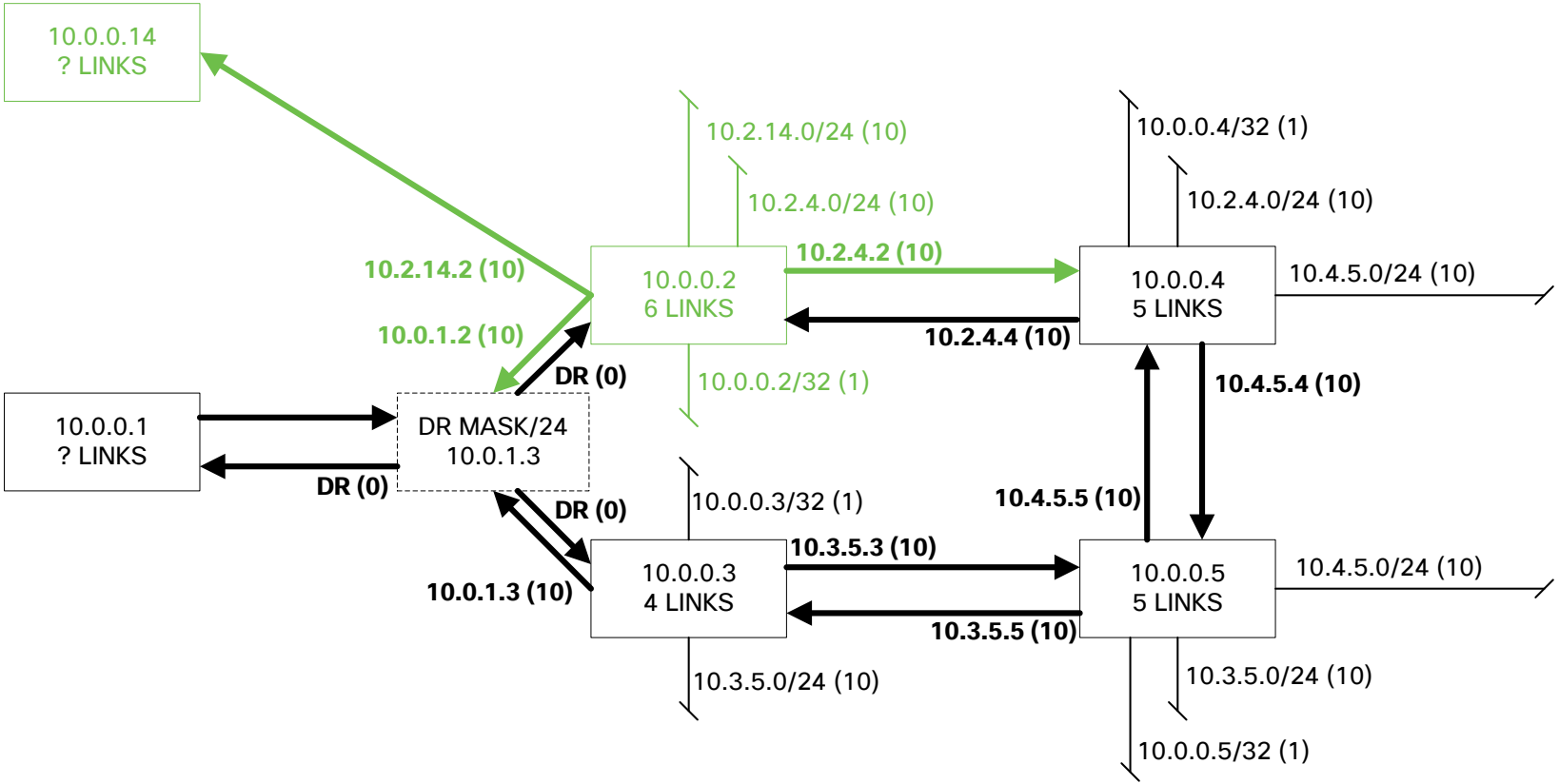
Continue Graph (R3)



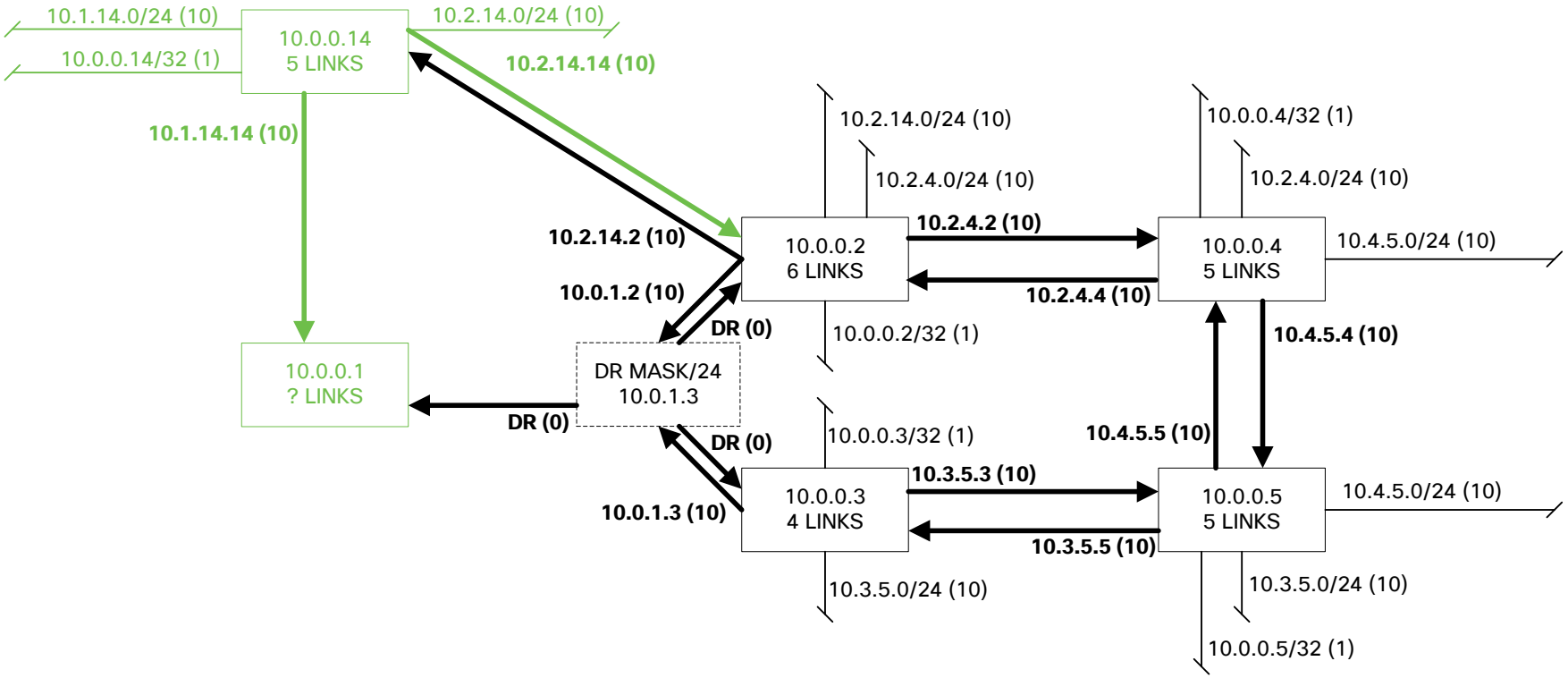
Continue Graph (DR)



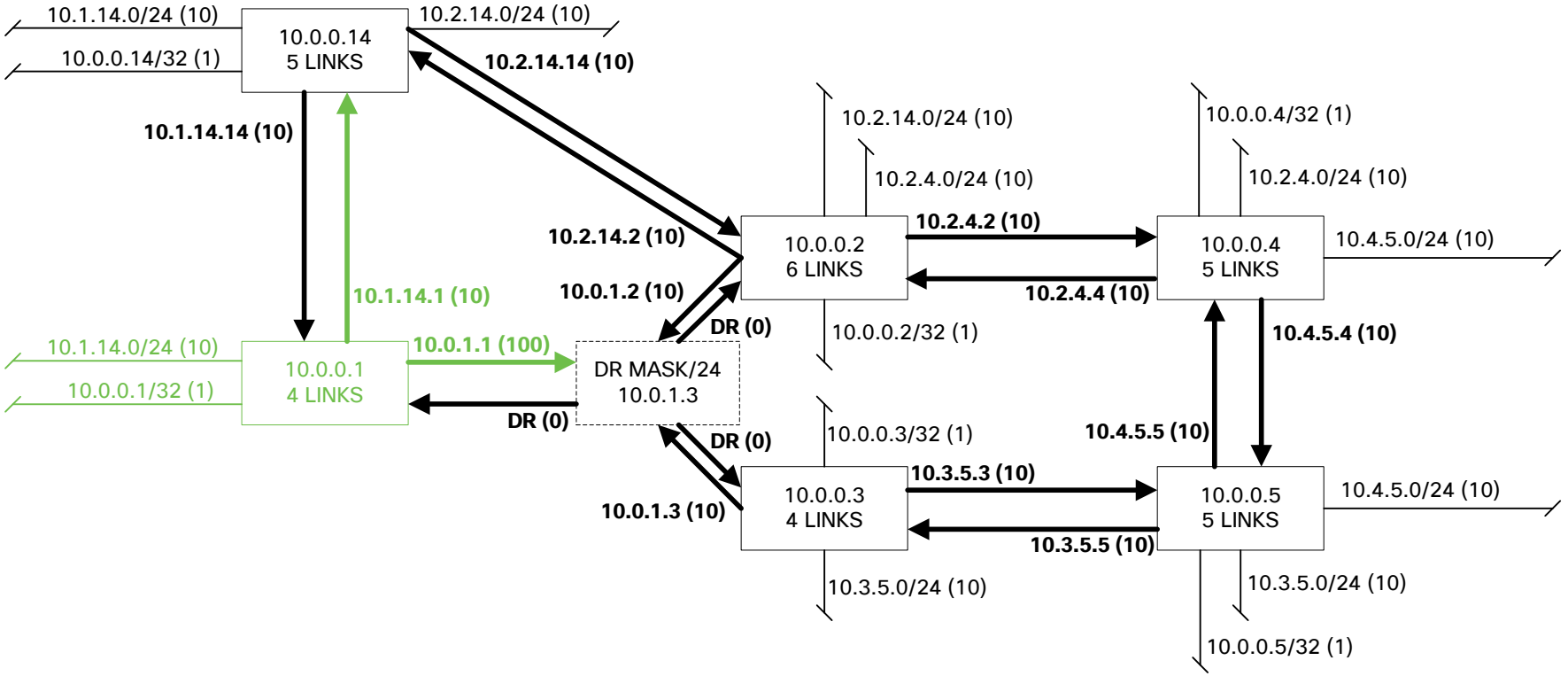
Continue Graph (R2)



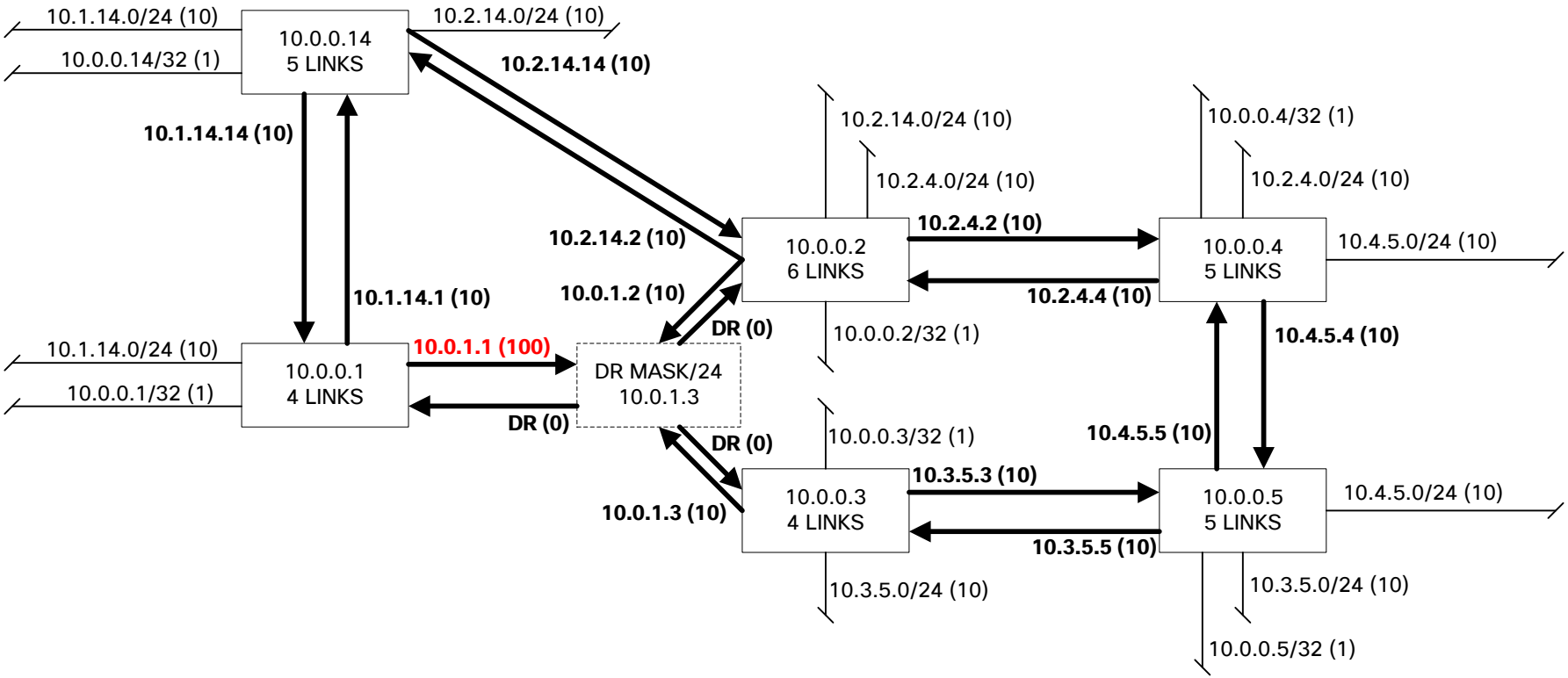
Continue Graph (R14)



Continue Graph (R1)



Final Graph



Finishing up

- Conduct experiment
  - We magically get access to R1 now ☺
- Analyze the result
- Takeaway
  - We used ONE command on ONE device!

Repairing Adjacencies

You make customer experience **possible**

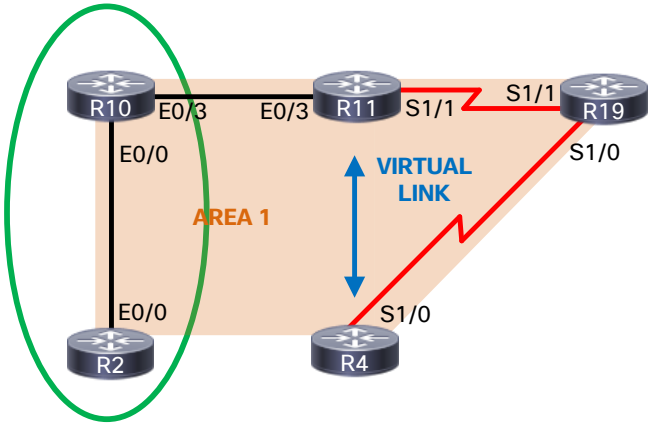
Repairing adjacencies

- Common but usually easy to fix
  - Wide range of show/debug commands
  - Plenty of online resources
  - Most engineers are experienced in it
- We will focus on one graph-related issue today
  - Other issues included in the session PDF

Issue #1: R2 to R10

```
R10#debug condition interface ethernet 0/0
Condition 1 set

R10#debug ip ospf packet
OSPF packet debugging is on

*Feb 19 01:43:27.659: OSPF-1 PAK  : Et0/0:  IN: 10.2.10.2->224.0.0.5: ver:2 type:1
len:44 rid:10.0.0.2 area:0.0.0.1 chksum:E41B auth:1

*Feb 19 01:43:29.035: OSPF-1 PAK  : Et0/0: OUT: 10.2.10.10->10.2.10.2: ver:2 type:1
len:44 rid:10.0.0.10 area:0.0.0.1 chksum:CE86 auth:1
```

- The problem is not:
  - OSPF disabled or link down
  - Passive interface
  - Authentication type mismatch
  - Area ID mismatch
  - Wrong IP configured
  - Duplicate RID
- What about the unicast reply on R10?

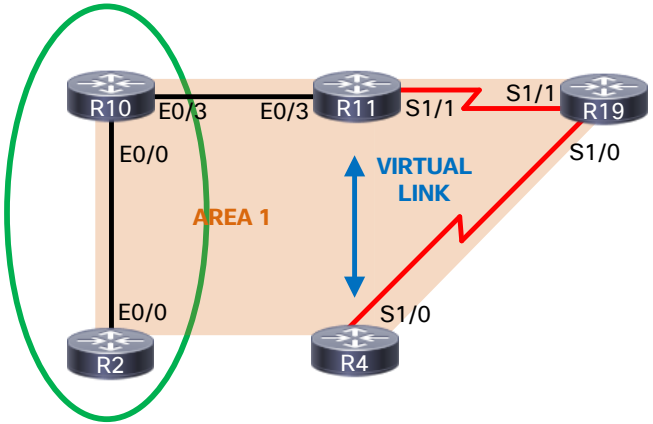


Issue #1: R2 to R10 continued

```
R10#debug ip ospf adj
OSPF adjacency debugging is on

*Feb 19 01:50:36.168: OSPF-1 ADJ   Et0/0: Drop packet from 10.2.10.2 with TTL: 1

R10#show ip ospf interface ethernet 0/0 | include TTL
  Strict TTL checking enabled
```

```
R2#show ip ospf interface ethernet 0/0 | include TTL
[no output]
```

- Hypotheses:
  - TTL security enabled on R10
  - TTL security not enabled on R2
- Experiment:
  - Enable TTL security on R2

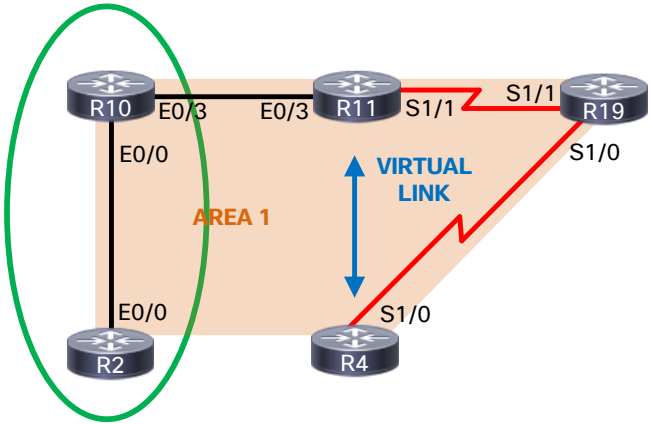


Issue #2: R2 to R10

```
R10#debug ip ospf adj
OSPF adjacency debugging is on

*Feb 19 01:52:02.058: OSPF-1 ADJ   Et0/0: Rcv pkt from 10.2.10.2,   : Mismatched
Authentication Key - Clear Text

R2#show running-config interface ethernet 0/0 | include authentication-key
 ip ospf authentication-key OSPF
```

```
R2#show running-config interface ethernet 0/0 | include authentication-key
 ip ospf authentication-key OSPF123
```

- Hypotheses:
  - Mismatched key between R2 and R10
- Experiment:
  - Change R2 key to “OSPF”

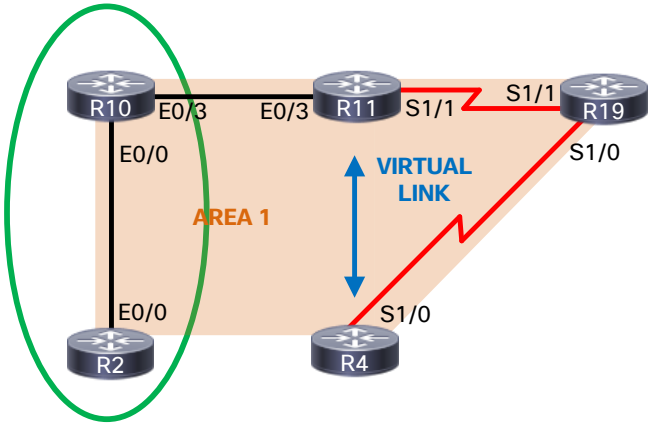


Issue #3: R2 to R10

```
R10#debug ip ospf hello
OSPF hello debugging is on

*Feb 19 01:53:45.413: OSPF-1 HELLO Et0/0: Rcv hello from 10.0.0.2 area 1 10.2.10.2
*Feb 19 01:53:45.413: OSPF-1 HELLO Et0/0: Mismatched hello parameters from 10.2.10.2
*Feb 19 01:53:45.413: OSPF-1 HELLO Et0/0: Dead R 40 C 40, Hello R 10 C 10 Mask R
255.255.255.128 C 255.255.255.0

R10#show ip interface ethernet 0/0 | include Internet
  Internet address is 10.2.10.10/24
```

```
R2#show ip interface ethernet 0/0 | include Internet
  Internet address is 10.2.10.2/25
```

- Hypotheses:
  - Given lab constraint of /24, wrong mask on R2 of /25
- Experiment:
  - Change R2 mask to /24

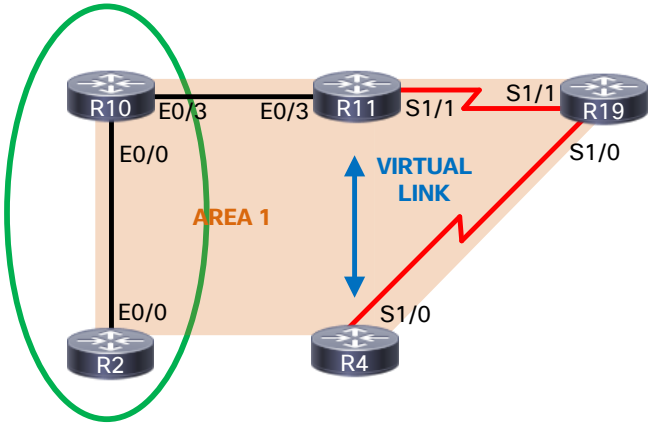


Issue #4: R2 to R10

```
R10#debug ip ospf hello
OSPF hello debugging is on

*Feb 19 01:57:26.703: OSPF-1 HELLO Et0/0: Rcv hello from 10.0.0.2 area 1 10.2.10.2
*Feb 19 01:57:26.703: OSPF-1 HELLO Et0/0: Hello from 10.2.10.2 with mismatched
Stub/Transit area option bit
```

```
R2#show ip ospf | begin Area_1
    Area 1
        It is a stub area
        [snip]
```

- Hypotheses:
  - R2 is incorrectly configured with area 1 as a stub area
- Experiment:
  - Remove stub configuration on R2 for area 1

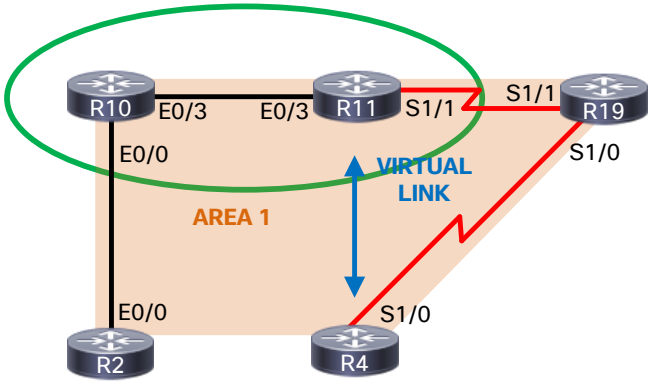


Issue #1: R10 to R11

```
R11#debug ip ospf packet
OSPF packet debugging is on

*Feb 19 02:03:56.346: OSPF-1 PAK  : Et0/3: OUT: 10.10.11.11->224.0.0.5: ver:2 type:1
len:44 rid:10.0.0.11 area:0.0.0.1 chksum:E292 auth:0

*Feb 19 02:04:05.390: OSPF-1 PAK  : Et0/3: OUT: 10.10.11.11->224.0.0.5: ver:2 type:1
len:44 rid:10.0.0.11 area:0.0.0.1 chksum:E292 auth:0
```

```
R10#show ip ospf interface ethernet 0/3
%OSPF: OSPF not enabled on Ethernet0/3
```

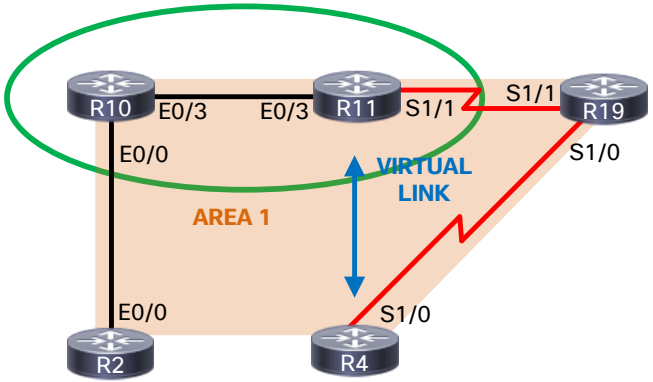
- Hypotheses:
  - OSPF not correctly enabled on R10
- Experiment:
  - Enable OSPF on R10



Issue #2: R10 to R11

```
R11#debug ip ospf packet
OSPF packet debugging is on

*Feb 19 02:05:10.707: OSPF-1 PAK  : Et0/3:  IN: 11.10.11.10->224.0.0.5: ver:2 type:1
len:44 rid:10.0.0.10 area:0.0.0.1 chksum:E293 auth:0

*Feb 19 02:05:12.290: OSPF-1 PAK  : Et0/3: OUT: 10.10.11.11->224.0.0.5: ver:2 type:1
len:44 rid:10.0.0.11 area:0.0.0.1 chksum:E292 auth:0
```

- The problem is not:
  - OSPF disabled or link down
  - Passive interface
  - Authentication type mismatch
  - Area ID mismatch
  - Duplicate RID
- ... but source IP from R10 is suspicious

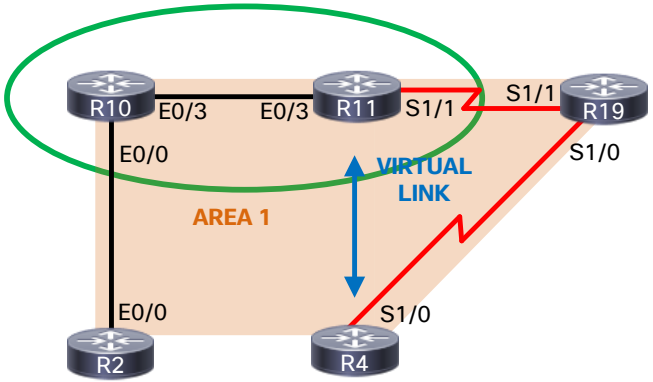


Issue #2: R10 to R11 continued

```
R11#debug ip ospf adj
OSPF adjacency debugging is on

*Feb 19 02:06:36.768: OSPF-1 ADJ   Et0/3: Rcv pkt from 11.10.11.10, area 0.0.0.1 : src
not on the same network
```

```
R10#show ip interface ethernet 0/3 | include Internet
  Internet address is 11.10.11.10/24
```

- Hypothesis:
  - Wrong IP configured on R10
- Experiment:
  - Use IP 10.10.11.10/24 on R10

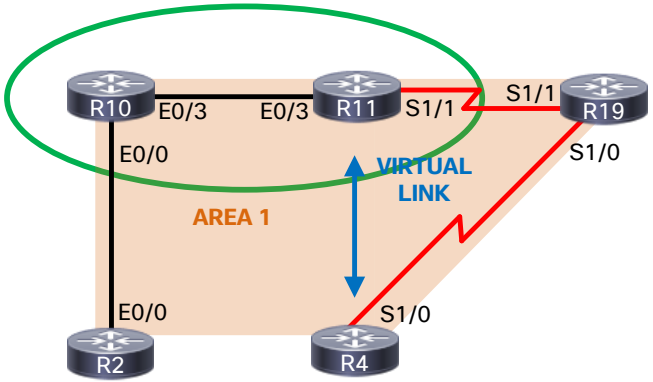


Issue #3: R10 to R11

```
R11#
*Feb 19 02:08:10.166: %OSPF-4-NET_TYPE_MISMATCH: Received Hello from 10.0.0.10 on
Ethernet0/3 indicating a potential network type mismatch

R11#show ip ospf interface ethernet 0/3 | include Network
  Process ID 1, Router ID 10.0.0.11, Network Type POINT_TO_POINT, Cost: 10
```

```
R10#show ip ospf interface ethernet 0/3 | include Network
  Process ID 1, Router ID 10.0.0.10, Network Type BROADCAST, Cost: 10
```

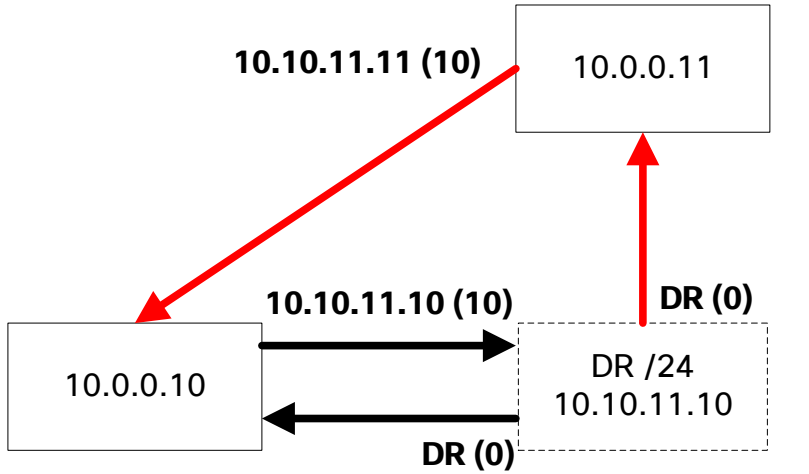
- Hypothesis:
  - Neighbor is up, but graph is broken
- Experiment:
  - Draw the graph first
  - Change R10 network to P2P



Examining the R10 to R11 link

```
R10#show ip ospf 1 1 database router self-originate
[snip]
  Link State ID: 10.0.0.10
  Advertising Router: 10.0.0.10

  [snip]
    Link connected to: a Transit Network
     (Link ID) Designated Router address: 10.10.11.10
     (Link Data) Router Interface address: 10.10.11.10
      Number of MTID metrics: 0
       TOS 0 Metrics: 10

R10#show ip ospf 1 1 database network 10.10.11.10
[snip]
  Link State ID: 10.10.11.10 (address of DR)
  Advertising Router: 10.0.0.10
  [snip]
  Network Mask: /24
        Attached Router: 10.0.0.10
        Attached Router: 10.0.0.11
```

```
R10#show ip ospf 1 1 database router 10.0.0.11
[snip]
  Adv Router is not-reachable in topology Base ...
  Link State ID: 10.0.0.11
  Advertising Router: 10.0.0.11
  [snip]
    Link connected to: another Router (point-to-point)
     (Link ID) Neighboring Router ID: 10.0.0.10
     (Link Data) Router Interface address: 10.10.11.11
      Number of MTID metrics: 0
       TOS 0 Metrics: 10
```



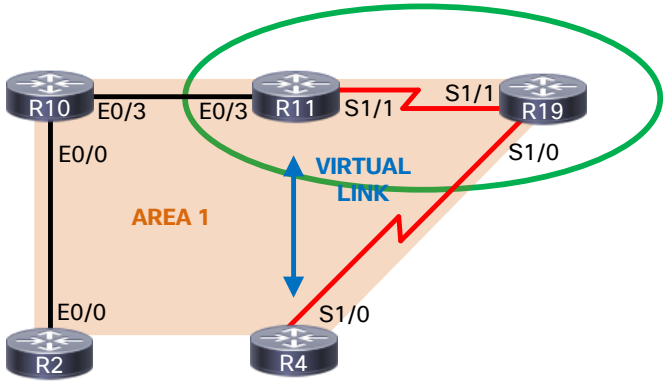
Issue #1: R11 to R19

```
R19#debug ip ospf packet
OSPF packet debugging is on
[no output]

R19#show ip ospf interface serial1/1
Serial1/1 is up, line protocol is down
  Internet Address 10.11.19.19/24, Area 1, Attached via Interface Enable
  [snip]
```

```
R19#show interfaces serial 1/1 | include protocol|Encap
Serial1/1 is up, line protocol is down
  Encapsulation HDLC, crc 16, loopback not set
```

```
R11#show interfaces serial 1/1 | include protocol|Encap
Serial1/1 is up, line protocol is down
  Encapsulation PPP, LCP Closed, crc 16, loopback not set
```

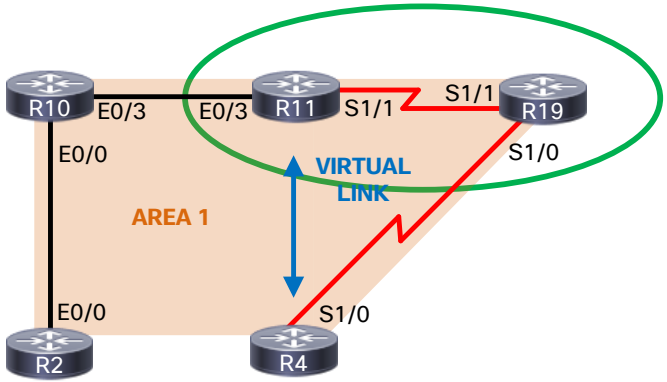
- Hypothesis:
  - Serial encapsulation mismatch
- Experiment:
  - Change R11 to HDLC



Issue #2: R11 to R19

```
R19#debug ip ospf packet
OSPF packet debugging is on

*Feb 19 02:23:30.225: OSPF-1 PAK  : Se1/1:  IN: 10.0.0.11->224.0.0.5: ver:2 type:1
len:44 rid:10.0.0.11 area:0.0.0.1 chksum:E193 auth:0

*Feb 19 02:23:31.208: OSPF-1 PAK  : Se1/1: OUT: 10.11.19.19->224.0.0.5: ver:2 type:1
len:44 rid:10.0.0.19 area:0.0.0.1 chksum:E28F auth:0
```

- The problem is not:
  - OSPF disabled or link down
  - Passive interface
  - Authentication type mismatch
  - Area ID mismatch
  - Duplicate RID
- … but source IP from R19 is suspicious

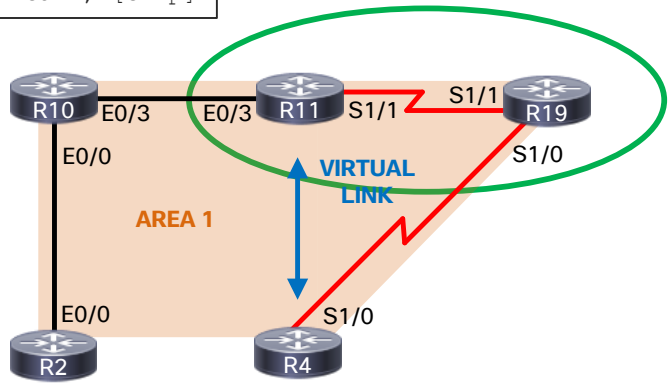


Issue #2: R11 to R19 continued

```
R19#debug ip ospf adj
OSPF adjacency debugging is on

*Feb 19 02:24:36.864: OSPF-1 ADJ   Se1/1: Rcv pkt from 10.0.0.11, area 0.0.0.1 : src not on the same network

R19#show ip ospf interface serial 1/1
Serial1/1 is up, line protocol is up
  Internet Address 10.11.19.19/24, Area 1, Attached via Interface Enable
```

```
R11#show ip ospf interface serial 1/1
Serial1/1 is up, line protocol is up
  Interface is unnumbered. Using address of Loopback0 (10.0.0.11), Area 1, [snip]
```

- Hypothesis:
  - IP unnumbered mismatch
- Experiment:
  - Enabled IP unnumbered on R19
  - It's more fun this way ☺

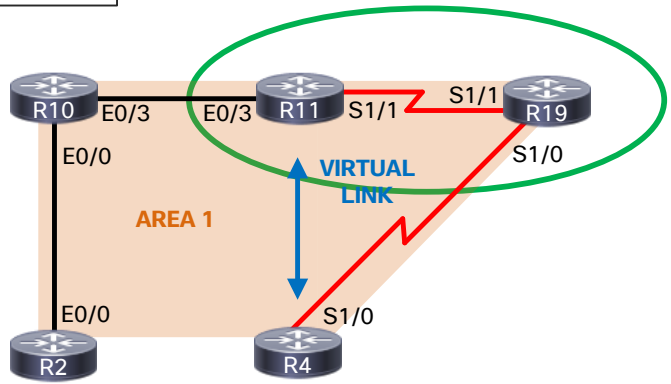


Issue #3: R11 to R19

```
R19#debug ip ospf hello
OSPF hello debugging is on

*Feb 19 02:26:40.552: OSPF-1 HELLO Se1/1: Rcv hello from 10.0.0.11 area 1 10.0.0.11
*Feb 19 02:26:40.553: OSPF-1 HELLO Se1/1: Mismatched hello parameters from 10.0.0.11
*Feb 19 02:26:40.553: OSPF-1 HELLO Se1/1: Dead R 40 C 36, Hello R 10 C 9

R19#show ip ospf interface serial 1/1 | include Timer
  Timer intervals configured, Hello 9, Dead 36, Wait 36, Retransmit 5
```

```
R11#show ip ospf interface serial 1/1 | include Timer
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
```

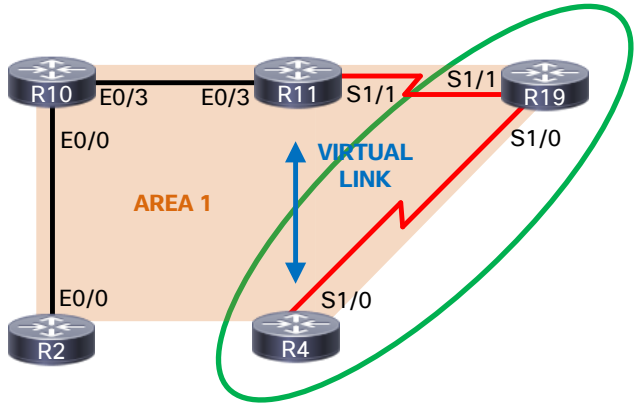
- Hypothesis:
  - Hello/dead intervals are misconfigured on R19 as 9/36, versus 10/40
- Experiment:
  - Use 10/40 timers on R19



Issue #1: R4 to R19

```
R4#debug ip ospf packet
OSPF packet debugging is on

*Feb 19 02:31:42.848: OSPF-1 PAK   Se1/0: Drop packet, OSPF not running or passive

R4#show ip ospf interface serial 1/0
Serial1/0 is up, line protocol is up
[snip]
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
    oob-resync timeout 40
    No Hellos (Passive interface)
    [snip]
```

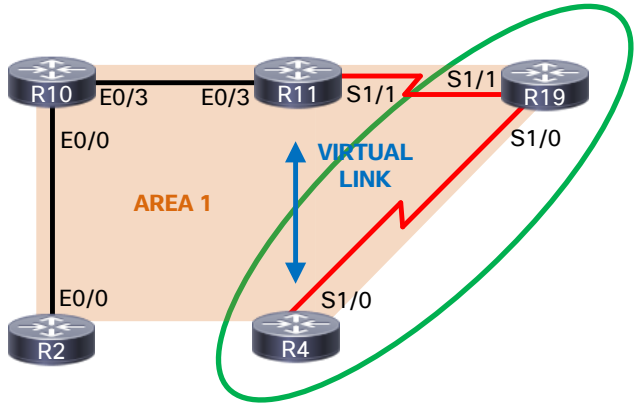
- Hypothesis:
  - Incorrect passive-interface configuration on R4
- Experiment:
  - Remove passive-interface on R4



Issue #2: R4 to R19

```
R4#debug ip ospf packet
OSPF packet debugging is on

*Feb 19 02:32:39.883: OSPF-1 PAK  : Se1/0:  IN: 10.4.19.19->224.0.0.5: ver:2 type:1
len:44 rid:10.0.0.19 area:0.0.0.1 chksum:0 auth:2 keyid:1 seq:0x5A8A

*Feb 19 02:32:44.996: OSPF-1 PAK  : Se1/0: OUT: 10.4.19.4->224.0.0.5: ver:2 type:1
len:44 rid:10.0.0.4 area:0.0.0.2 chksum:E298 auth:0
```

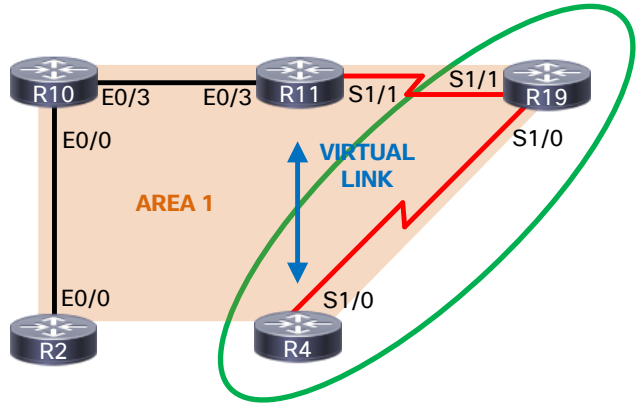
- The problem is not:
  - OSPF disabled or link down
  - Passive interface
  - Wrong IP addressing
  - Duplicate RID
- ... but a few suspicions
  - Area ID mismatch?
  - Auth type mismatch?



Issue #2: R4 to R19 continued

```
R4#debug ip ospf adj
OSPF adjacency debugging is on

*Feb 19 02:33:18.101: OSPF-1 ADJ   Se1/0: Rcv pkt from 10.4.19.19, area 0.0.0.2, mismatched
area 0.0.0.1 in the header

R4#show ip ospf interface brief | include ^Interface|^Se1/0
Interface    PID   Area            IP Address/Mask    Cost  State Nbrs F/C
Se1/0        1     2               10.4.19.4/24       32768 P2P   0/0
```

- Hypothesis:
  - Incorrect area assignment on R4
- Experiment:
  - Configure R4 in area 1

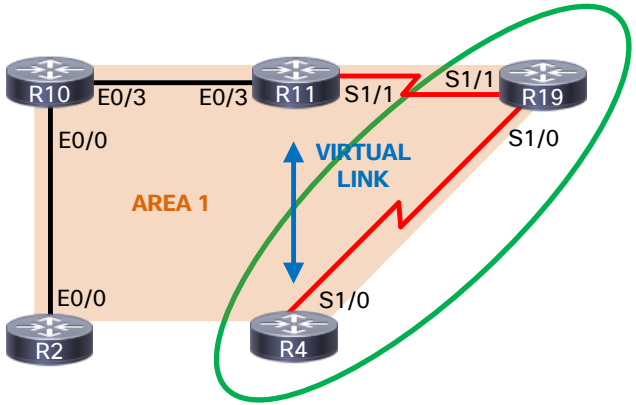


Issue #3: R4 to R19

```
R4#debug ip ospf adj
OSPF adjacency debugging is on

*Feb 19 02:35:19.997: OSPF-1 ADJ   Se1/0: Rcv pkt from 10.4.19.19 : Mismatched
Authentication type. Input packet specified type 2, we use type 0

Hypothesis: R4 is not configured for authentication. Need to check R19 first

R4#show ip ospf interface serial 1/0 | begin Crypto
[no output]
```

```
R19#show ip ospf interface serial 1/0 | begin Crypto
Cryptographic authentication enabled
Sending SA: Key 1, Algorithm HMAC-SHA-256 - key chain KC_OSPF_AUTH
```

- Hypothesis:
  - Authentication applied on R19, not R4
- Experiment:
  - Add authentication to R4

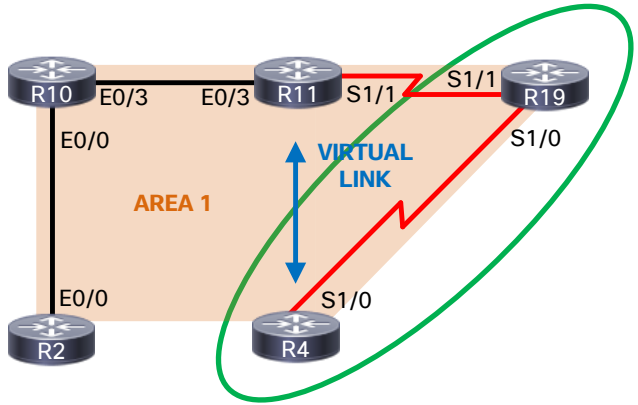


Issue #4: R4 to R19

```
R4#debug ip ospf adj
OSPF adjacency debugging is on

*Feb 19 02:39:31.957: OSPF-1 ADJ   Se1/0: Rcv DBD from 10.0.0.19 seq 0x5F9 opt 0x52
flag 0x7 len 32  mtu 1500 state EXSTART
*Feb 19 02:39:31.958: OSPF-1 ADJ   Se1/0: Nbr 10.0.0.19 has larger interface MTU

R4#show ip ospf neighbor 10.0.0.19
[snip]
    Neighbor priority is 0, State is EXSTART, 3 state changes
    [snip]
    Number of retransmissions for last database description packet 20
```

```
R4#show interfaces serial 1/0 | include MTU
  MTU 1400 bytes, BW 1544 Kbit/sec, DLY 20000 usec,
```

- Hypothesis:
  - MTU mismatch between R4 and R19
- Experiment:
  - Increase R4 MTU to 1500



Things that cause adjacencies to fail

- R2 to R10
  - TTL-security mismatch
  - Authentication key mismatch
  - Area type (stub/NSSA) mismatch
  - Subnet mask mismatch
- R10 to R11
  - OSPF disabled
  - IP not in correct subnet
  - Network type mismatch
- R11 to R19
  - Interface layer-1 or layer-2 down
  - IP unnumbered mismatch
  - Hello/dead interval mismatch
- R19 to R4
  - Passive interface
  - Area ID mismatch
  - Authentication type mismatch
  - MTU mismatch

Fun with Virtual Links

You make the power of data **possible**

Cisco live!

Virtual Links

- Three main uses
  - Repair noncontiguous area 0 (our case)
  - Create new ABRs between nonzero area boundaries
  - Traffic engineering across nonzero areas

- That's nice, but our VL is down
  - Observe
  - Ask how/why
  - Form a hypothesis

Verify current state

```
R4#show ip ospf virtual-links
Virtual Link OSPF_VL0 to router 10.0.0.11 is down
  Run as demand circuit
  DoNotAge LSA allowed.
  Transit area 1
 Topology-MTID    Cost    Disabled     Shutdown      Topology Name
        0          65535     no          no            Base
  Transmit Delay is 1 sec, State DOWN,
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
```

```
R11#show ip ospf virtual-links
Virtual Link OSPF_VL0 to router 10.0.0.4 is down
  Run as demand circuit
  DoNotAge LSA allowed.
  Transit area 1
 Topology-MTID    Cost    Disabled     Shutdown      Topology Name
        0          65535     no          no            Base
  Transmit Delay is 1 sec, State DOWN,
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
```

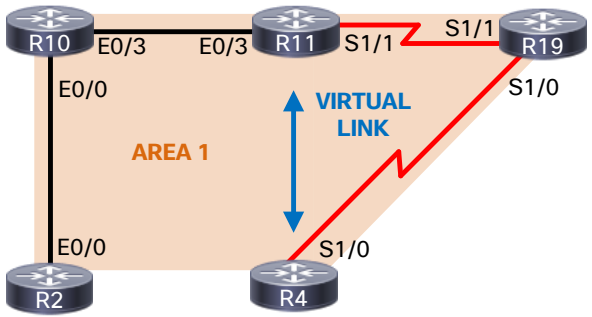
- A cost of 65535 between two VL endpoints should work!

Looking Deeper

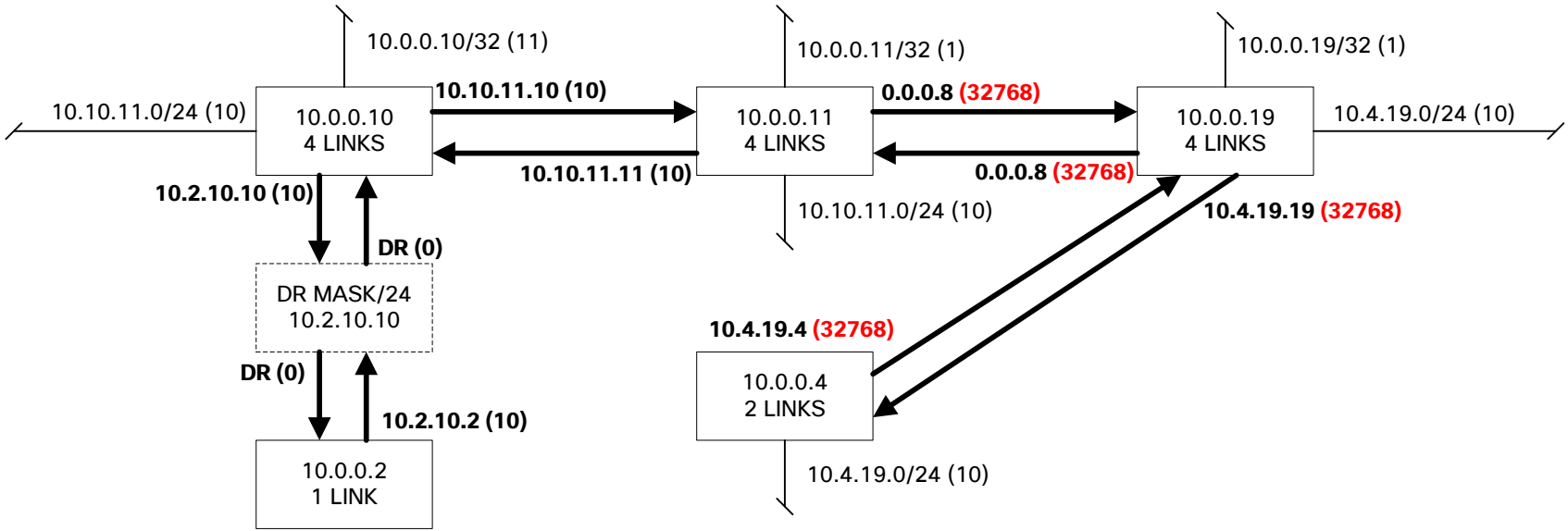
```
R4#show ip ospf border-routers | include 10.0.0.11
i 10.0.0.11 [65536] via 10.4.19.19, Serial1/0, ABR, Area 1, SPF 11
```

```
R11#show ip ospf border-routers | include 10.0.0.4
i 10.0.0.4 [65536] via 10.0.0.19, Serial1/1, ABR/ASBR, Area 1, SPF 8
```

- So what's the problem?
  - Cost of 65535 is cosmetic issue!
  - Actual cost between these nodes is 65536
- Hypothesis
  - Path cost between R4 and R19 is too high



Area 1 graph



VL is “up”, but not really

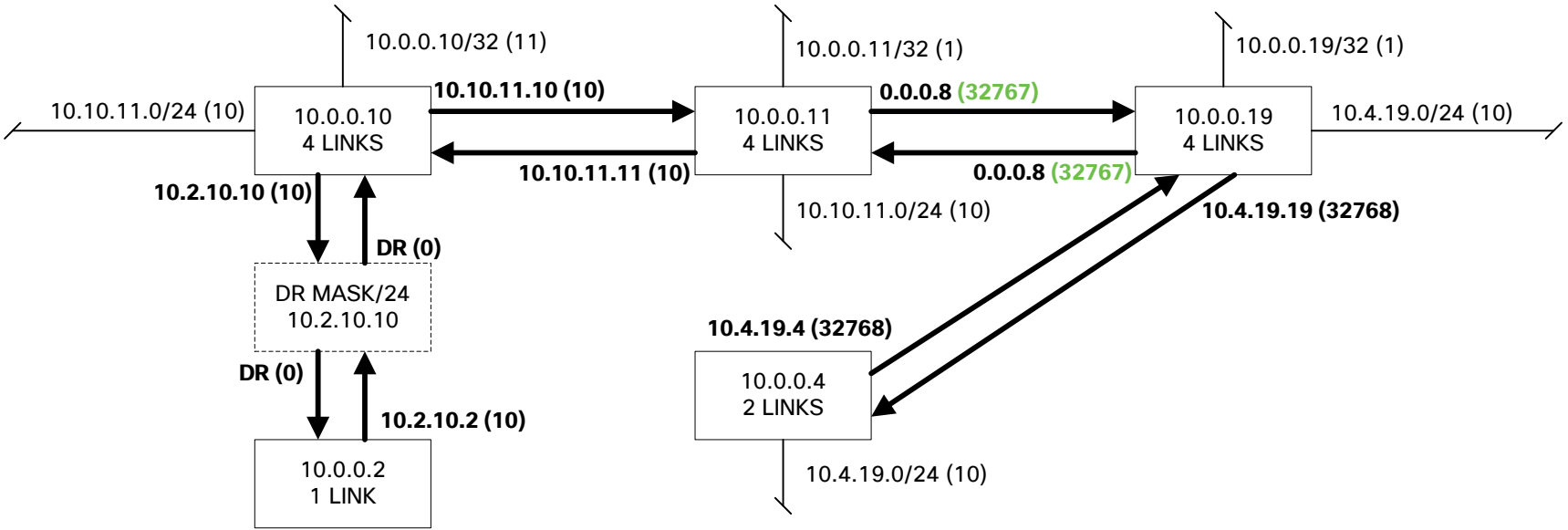
```
! Config applied to R11
interface Serial1/1
 ip ospf cost 32767

R11#show ip ospf border-routers | include 10.0.0.4
i 10.0.0.4 [65535] via 10.0.0.19, Serial1/1, ABR/ASBR, Area 1, SPF 9

R11#show ip ospf virtual-links
Virtual Link OSPF_VL0 to router 10.0.0.4 is up
  Run as demand circuit
  DoNotAge LSA allowed.
  Transit area 1, via interface Serial1/1
 Topology-MTID    Cost    Disabled     Shutdown      Topology Name
        0           65535     no          no            Base
  Transmit Delay is 1 sec, State POINT_TO_POINT,
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
    Hello due in 00:00:07
```

- Must differentiate between two “up” states:
  - VL being “up” only means that path to VL endpoint is valid
  - Having an OSPF neighbor is the real end-state

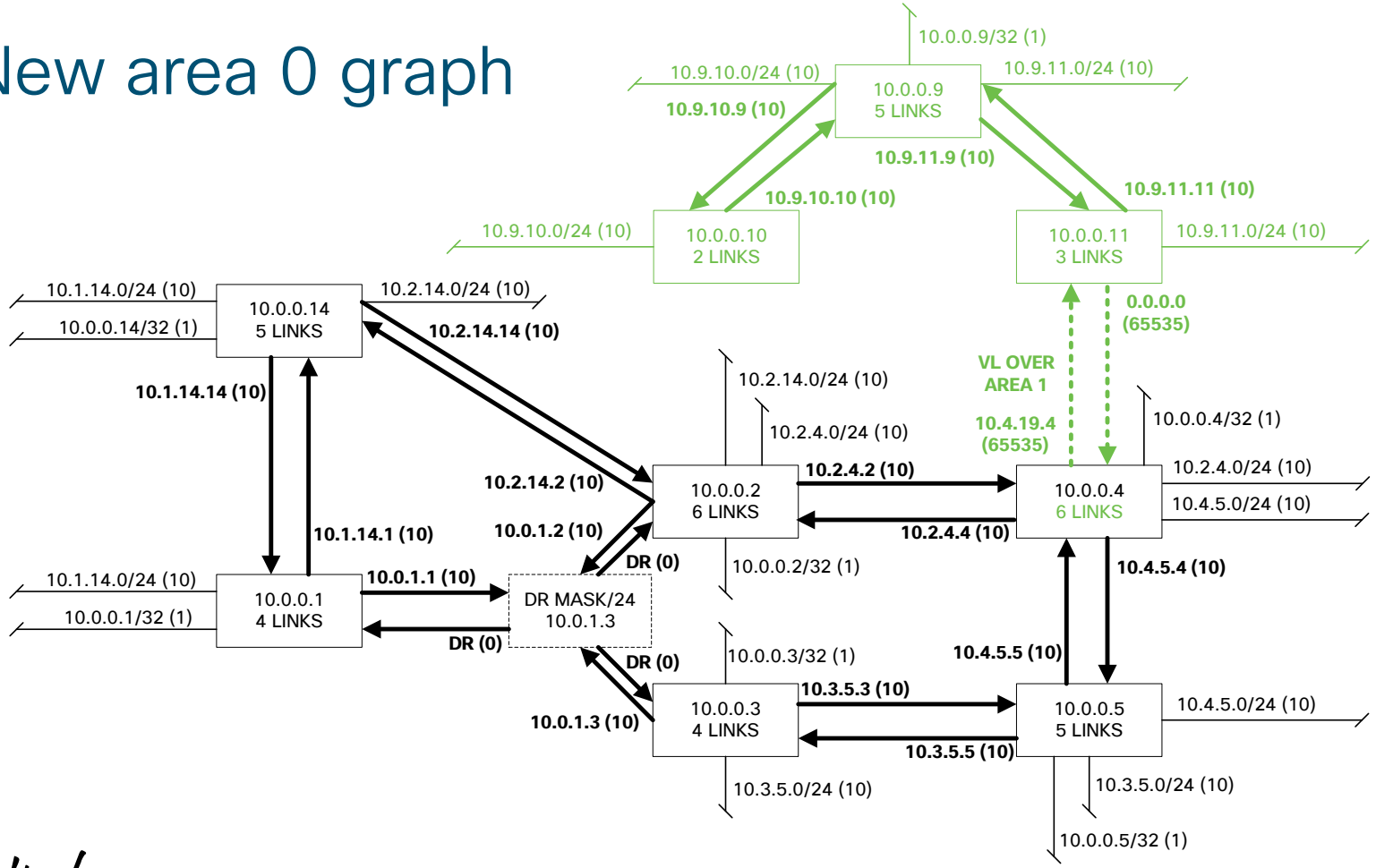
OK, seriously this time

```
! Config applied to R19
interface Serial1/1
 ip ospf cost 32767

R4#show ip ospf border-routers | include 10.0.0.11
i 10.0.0.11 [65535] via 10.4.19.19, Serial1/0, ABR, Area 1, SPF 12

R4#show ip ospf virtual-links
Virtual Link OSPF_VL0 to router 10.0.0.11 is up
  Run as demand circuit
  DoNotAge LSA allowed.
  Transit area 1, via interface Serial1/0
 Topology-MTID    Cost    Disabled     Shutdown      Topology Name
        0           65535     no          no            Base
  Transmit Delay is 1 sec, State POINT_TO_POINT,
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
    Hello due in 00:00:00
    Adjacency State FULL (Hello suppressed)
    Index 1/3/5, retransmission queue length 0, number of retransmission 0
    First 0x0(0)/0x0(0)/0x0(0) Next 0x0(0)/0x0(0)/0x0(0)
    Last retransmission scan length is 0, maximum is 0
    Last retransmission scan time is 0 msec, maximum is 0 msec
```

New area 1 graph

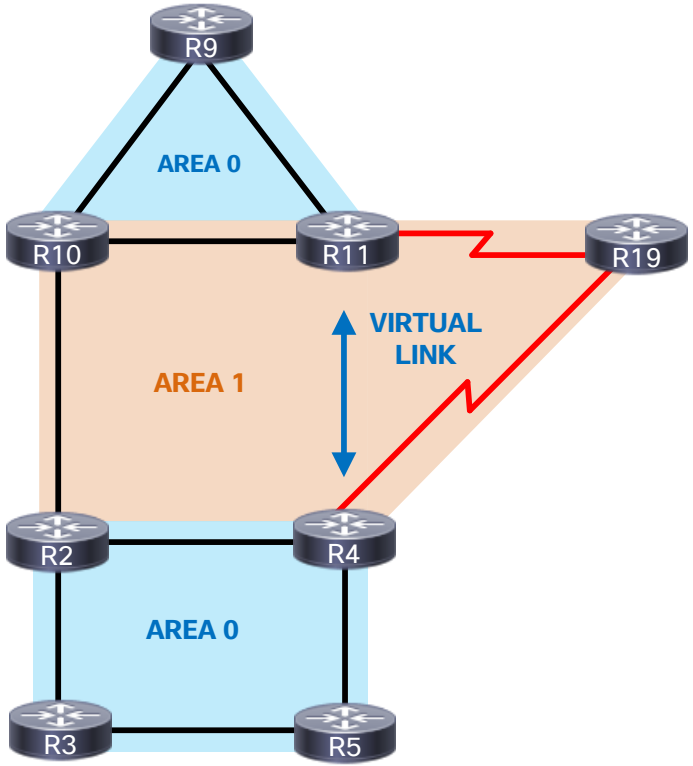


New area 0 graph



Let's test some flows!

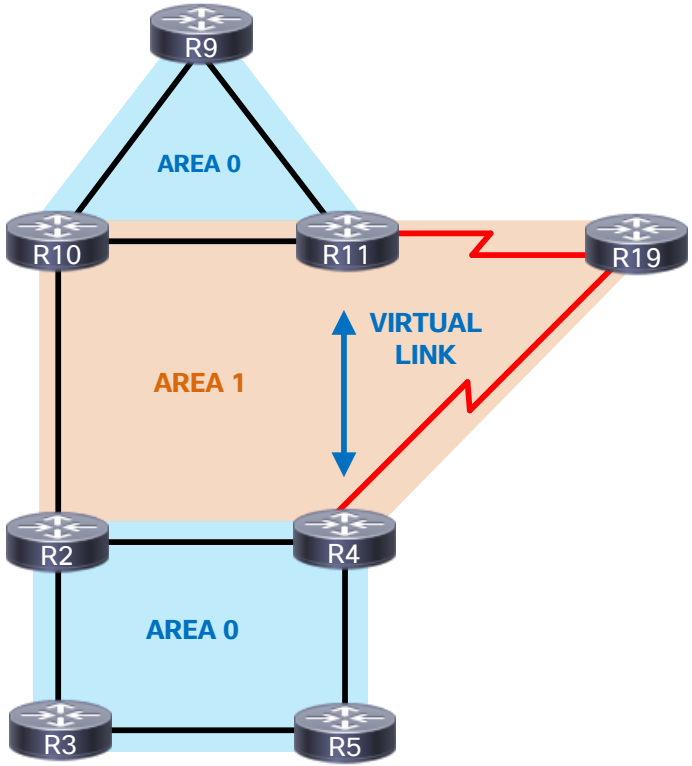
```
R9#show ip route 10.0.0.5
Routing entry for 10.0.0.5/32
  Known via "ospf 1", distance 110, metric 65556, type intra area
  Last update from 10.9.11.11 on Ethernet0/2, 00:00:35 ago
  Routing Descriptor Blocks:
  * 10.9.11.11, from 10.0.0.5, 00:00:35 ago, via Ethernet0/2
      Route metric is 65556, traffic share count is 1

R9#traceroute 10.0.0.5 source loopback 0
Type escape sequence to abort.
Tracing the route to 10.0.0.5
VRF info: (vrf in name/id, vrf out name/id)
  1 10.9.11.11 1 msec 5 msec 2 msec
  2 10.10.11.10 2 msec 2 msec 1 msec
  3 10.2.10.2 2 msec 1 msec 1 msec
  4 10.2.4.4 11 msec 7 msec 12 msec
  5 10.4.5.5 12 msec 12 msec 11 msec
```



- We just looked at the area 0 graph
  - R11 routes via R10 and R2 using area 1 links
  - How is this possible?

Smarter paths across area 1

```
R11#show ip route 10.0.0.5
Routing entry for 10.0.0.5/32
  Known via "ospf 1", distance 110, metric 41, type intra area
  Last update from 10.10.11.10 on Ethernet0/3, 00:01:39 ago
  Routing Descriptor Blocks:
  * 10.10.11.10, from 10.0.0.5, 00:01:39 ago, via Ethernet0/3
      Route metric is 41, traffic share count is 1
```



- Analysis
  - R11 preferred R10 over R19 for routing towards R5
  - Control-plane / data-plane separation

Introducing Capability Transit

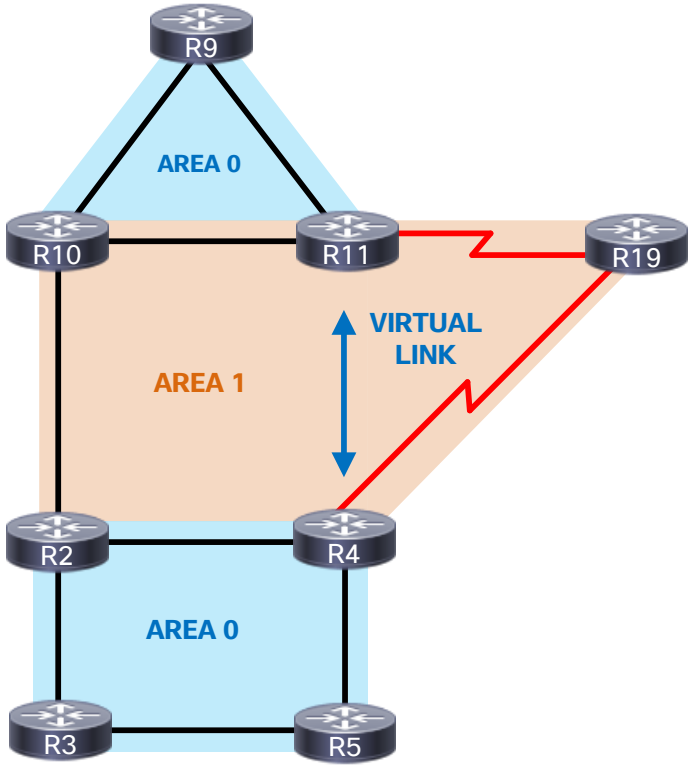
```
R11#show ip ospf | include transit
 Supports area transit capability
 Number of areas transit capable is 1
        This area has transit capability: Virtual Link Endpoint

R11#show ip ospf border-routers | include 10.0.0.[24]
i 10.0.0.2 [65545] via 10.0.0.19, Serial1/1, ABR, Area 0, SPF 12
i 10.0.0.2 [20] via 10.10.11.10, Ethernet0/3, ABR, Area 1, SPF 14
i 10.0.0.4 [30] via 10.10.11.10, Ethernet0/3, ASBR, Area 0, SPF 12 (transit)
i 10.0.0.4 [65535] via 10.0.0.19, Serial1/1, ABR, Area 0, SPF 12
i 10.0.0.4 [65535] via 10.0.0.19, Serial1/1, ABR/ASBR, Area 1, SPF 14
```

- Analysis
  - R11's path cost to R4 is 30 across area 0 …
  - This was derived via the "transit" capability

Let's disable it!

```
! Config applied to R11
router ospf 1
 no capability transit

R11#show ip ospf | include transit
 Does not support area transit capability

R11#show ip ospf border-routers | include 10.0.0.[24]
i 10.0.0.2 [65545] via 10.0.0.19, Serial1/1, ABR, Area 0, SPF 11
i 10.0.0.2 [20] via 10.10.11.10, Ethernet0/3, ABR, Area 1, SPF 13
i 10.0.0.4 [65535] via 10.0.0.19, Serial1/1, ABR/ASBR, Area 0, SPF 11
i 10.0.0.4 [65535] via 10.0.0.19, Serial1/1, ABR/ASBR, Area 1, SPF 13
```

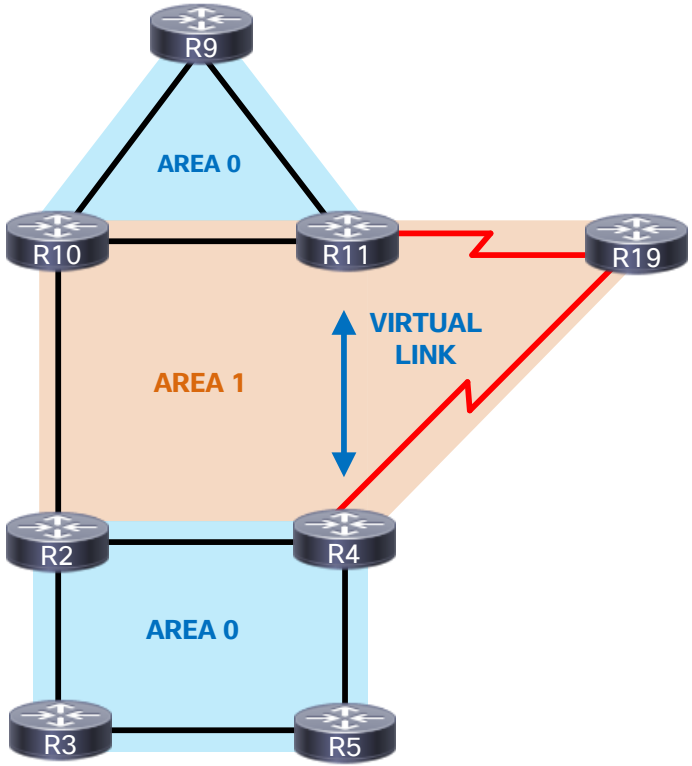
- Analysis
  - R11's path to R4 is 65535 across area 0
  - This is the exact link cost of the VL
  - R11 cannot use area 1 for optimal transit



Going the long way

```
R9#traceroute 10.0.0.5 source loopback 0
Type escape sequence to abort.
Tracing the route to 10.0.0.5
VRF info: (vrf in name/id, vrf out name/id)
  1 10.9.11.11 1 msec 1 msec 1 msec
  2 10.0.0.19 10 msec 10 msec 8 msec
  3 10.4.19.4 19 msec 21 msec 18 msec
  4 10.4.5.5 16 msec 21 msec 20 msec
```

```
R11#show ip route 10.0.0.5
Routing entry for 10.0.0.5/32
  Known via "ospf 1", distance 110, metric 65546, type intra area
  Last update from 10.0.0.19 on Serial1/1, 00:01:09 ago
  Routing Descriptor Blocks:
  * 10.0.0.19, from 10.0.0.5, 00:01:09 ago, via Serial1/1
      Route metric is 65546, traffic share count is 1
```



- R11 decides to follow the virtual link
  - Control/data plane flows route similarly

Capability Transit in Review

- Separates VL control plane from packet forwarding
  - Allows shortest path forwarding across transit area
  - Local decision on VL endpoint (R11)
  - Sometimes doesn't matter (R4)

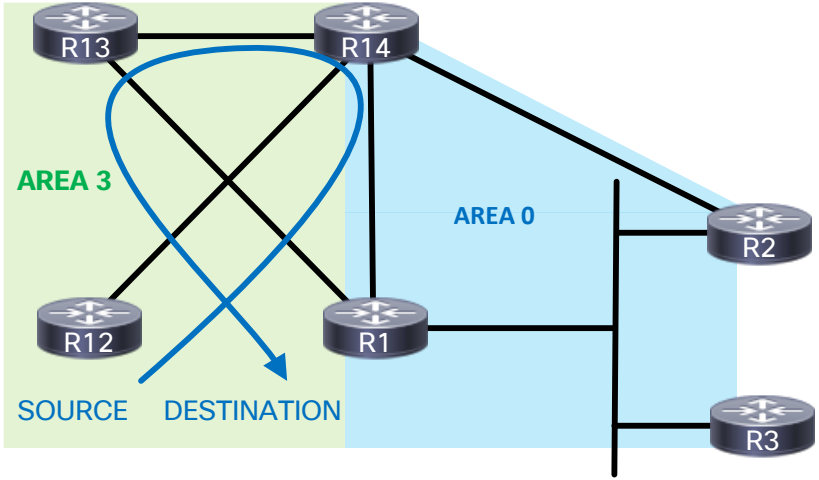
- When disabled:
  - Traffic follows the path of the virtual link itself
  - Advantage: symmetric routing (pretend R19 was a firewall)
  - Disadvantage: suboptimal routing, possibly a routing loop

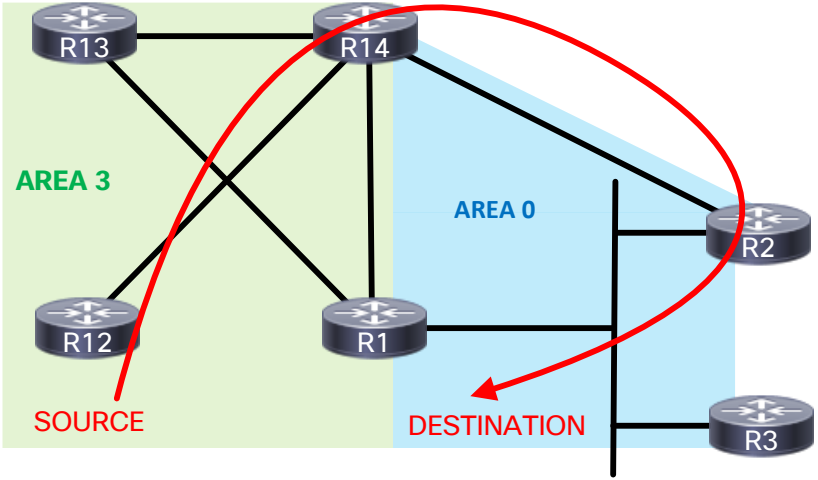
Optimal Routing

You make multi-cloud **possible**

Suboptimal Routing

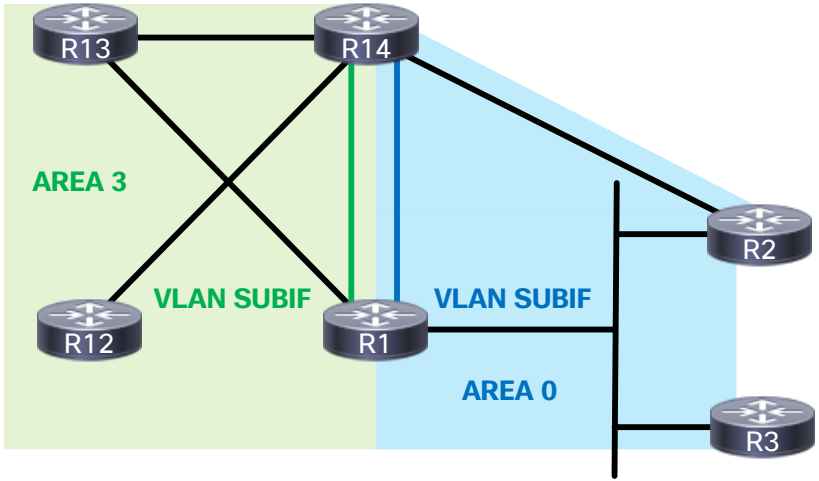
- Links can only belong to one area
- Suppose the R1–R12 link fails





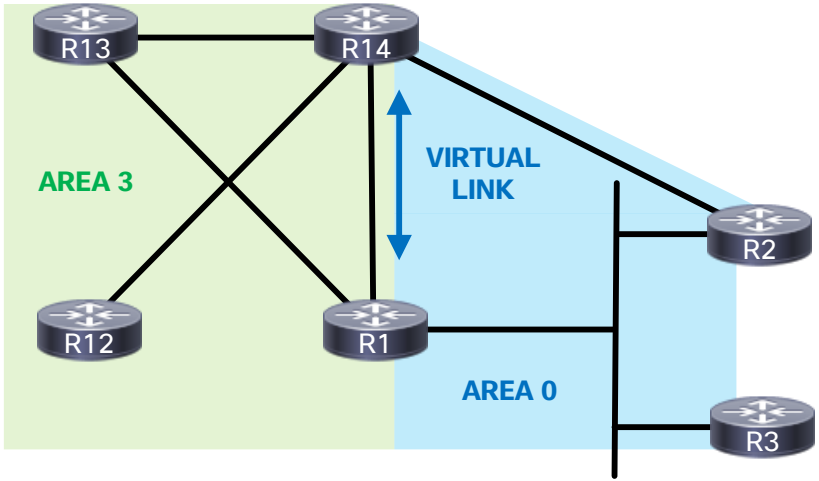
Solution 1: Link multiplexing

- Layer-2 tech needs to support it (Ethernet VLAN, FR DLCI, etc.)
- Configuration intensive, may need new IPs



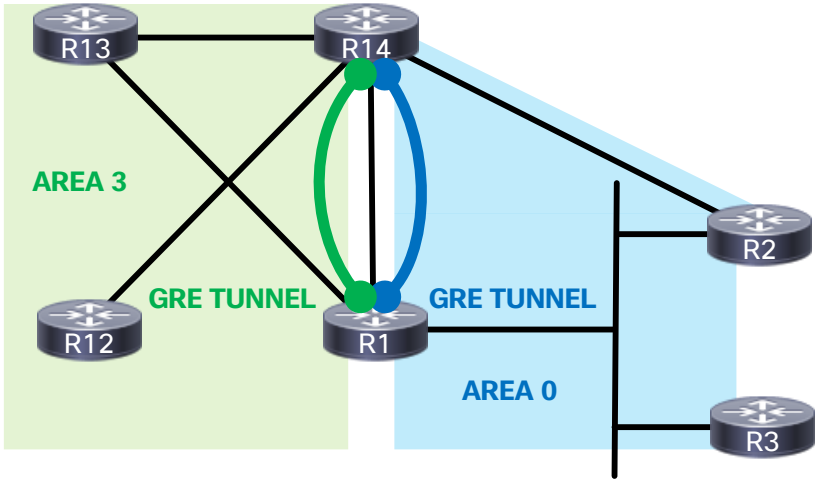
Solution 2: OSPF Virtual Link

- Use nonzero area as the base, run VL over top
- Only works with one nonzero area



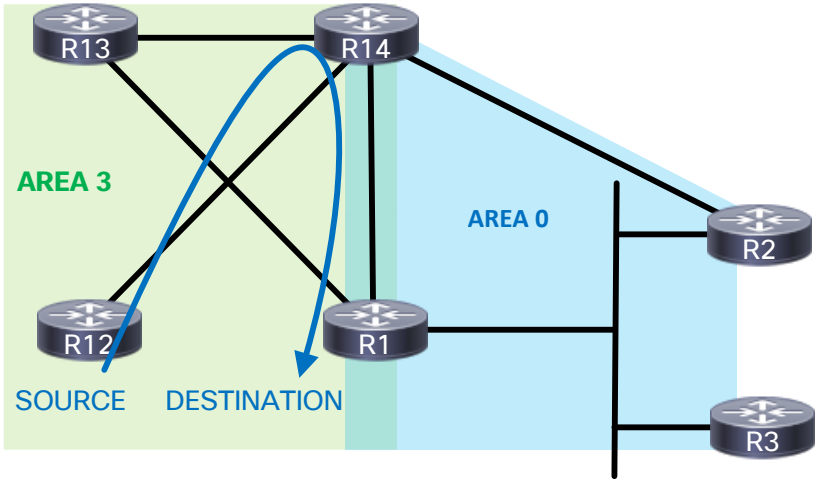
Solution 3: GRE tunnels

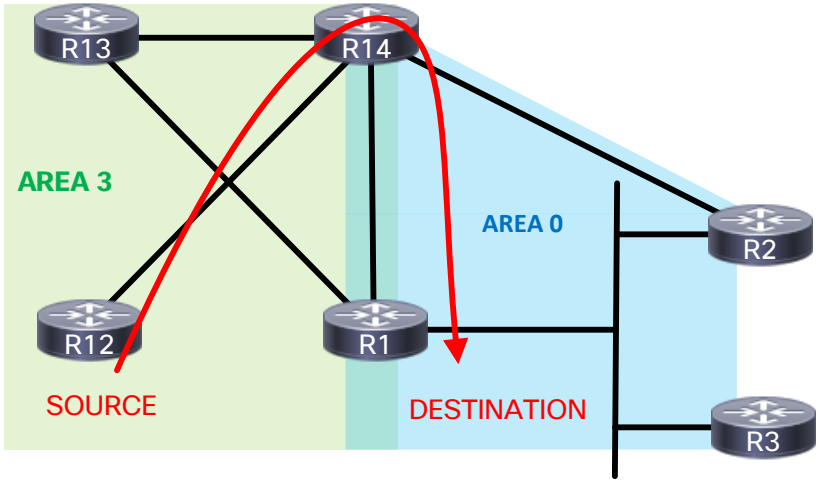
- Multiple P2P tunnels with varying keys over a non-OSPF link
- Config intensive, additional encap, FW/IPS challenges



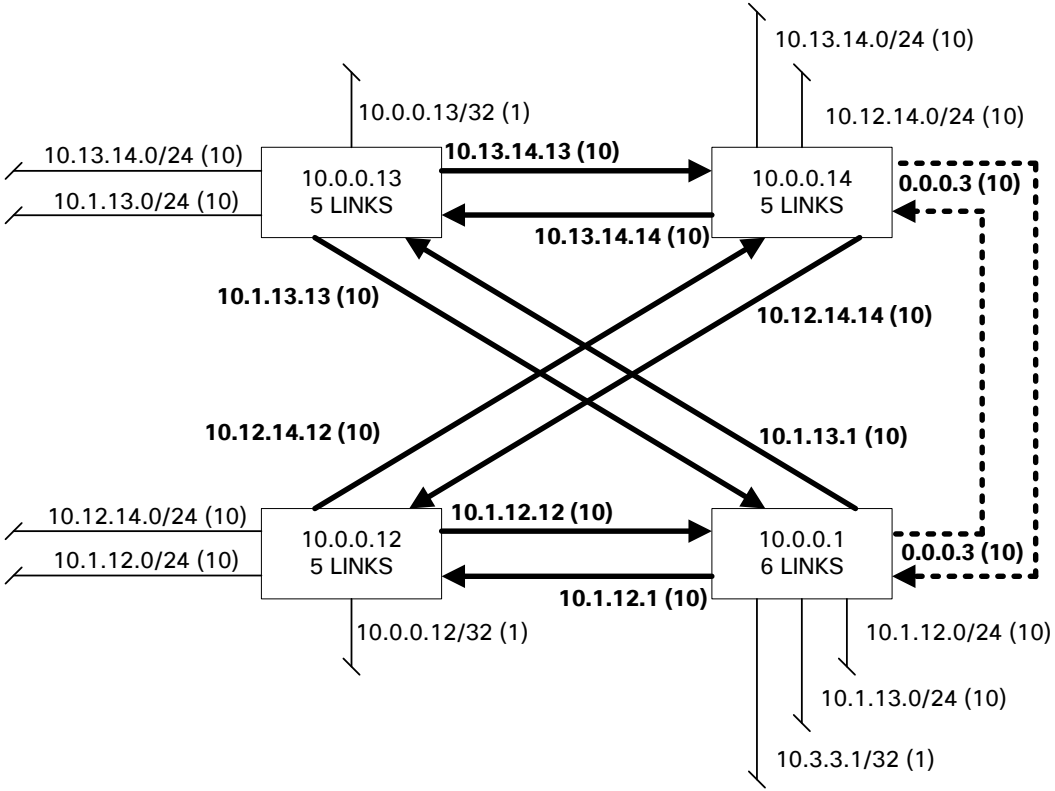
Better way: Multi-area Adjacency

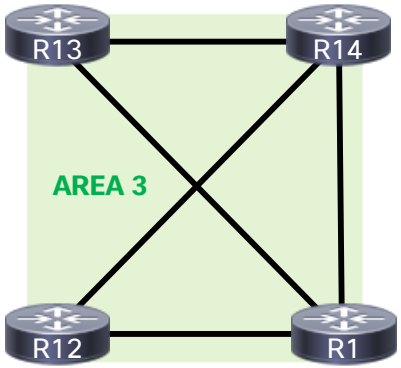
- Similar to IS-IS level-1/level-2 on the same link
- Creates new logical connection for a given area





Area 3 Graph





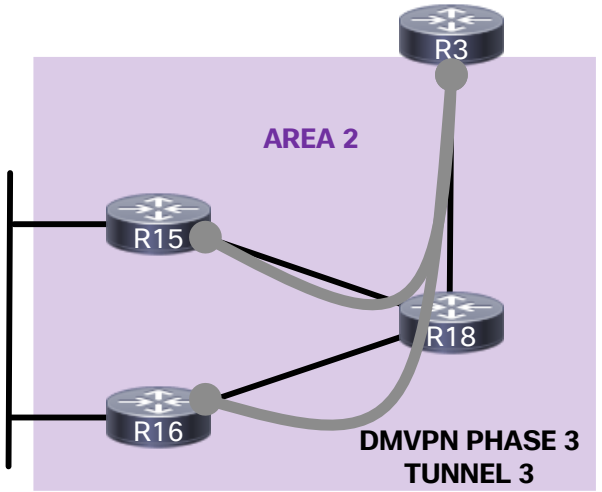
OSPF running over Hub/Spoke Networks

You make networking **possible**

OSPF over DMVPN

- A suboptimal design choice
  - Spokes share topology information
  - Scalability is poor
  - Not commonly used

- Can we improve it?



Nothing fishy from R3's view

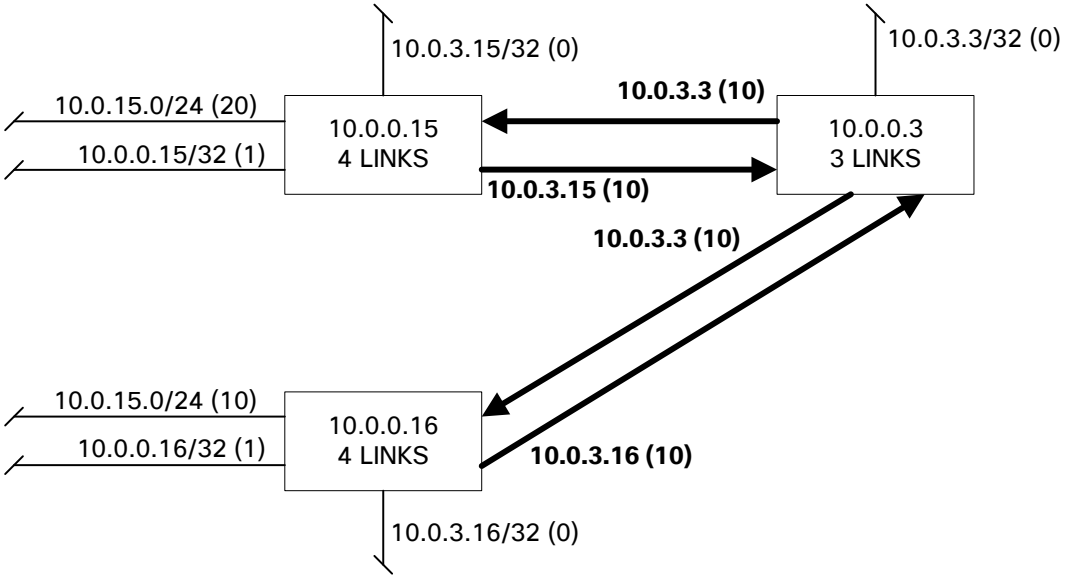
```
R3#show ip ospf 1 2 database router 10.0.0.15
[snip]
  Advertising Router: 10.0.0.15
  [snip]
  AS Boundary Router
  Number of Links: 4

    Link connected to: a Stub Network
     (Link ID) Network/subnet number: 10.0.0.15
     (Link Data) Network Mask: 255.255.255.255
      Number of MTID metrics: 0
       TOS 0 Metrics: 1
[snip]
```

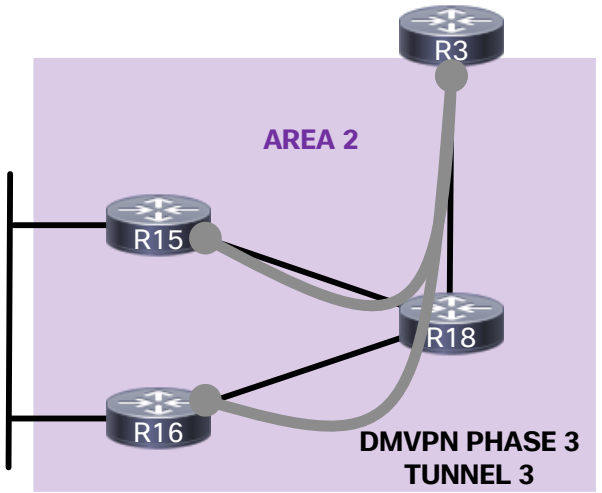
```
R3#show ip ospf 1 2 database router 10.0.0.16
[snip]
  Advertising Router: 10.0.0.16
  [snip]

  Number of Links: 4

    Link connected to: a Stub Network
     (Link ID) Network/subnet number: 10.0.0.16
     (Link Data) Network Mask: 255.255.255.255
      Number of MTID metrics: 0
       TOS 0 Metrics: 1
[snip]
```

```
R3#show ip route ospf | include Tunnel3
O        10.0.0.15/32 [110/11] via 10.0.3.15, 01:17:05, Tunnel3
O        10.0.0.16/32 [110/11] via 10.0.3.16, 01:17:05, Tunnel3
[snip]
```

- R3 sees R15 and R16 within area 2

Examine the Graph from R3's view



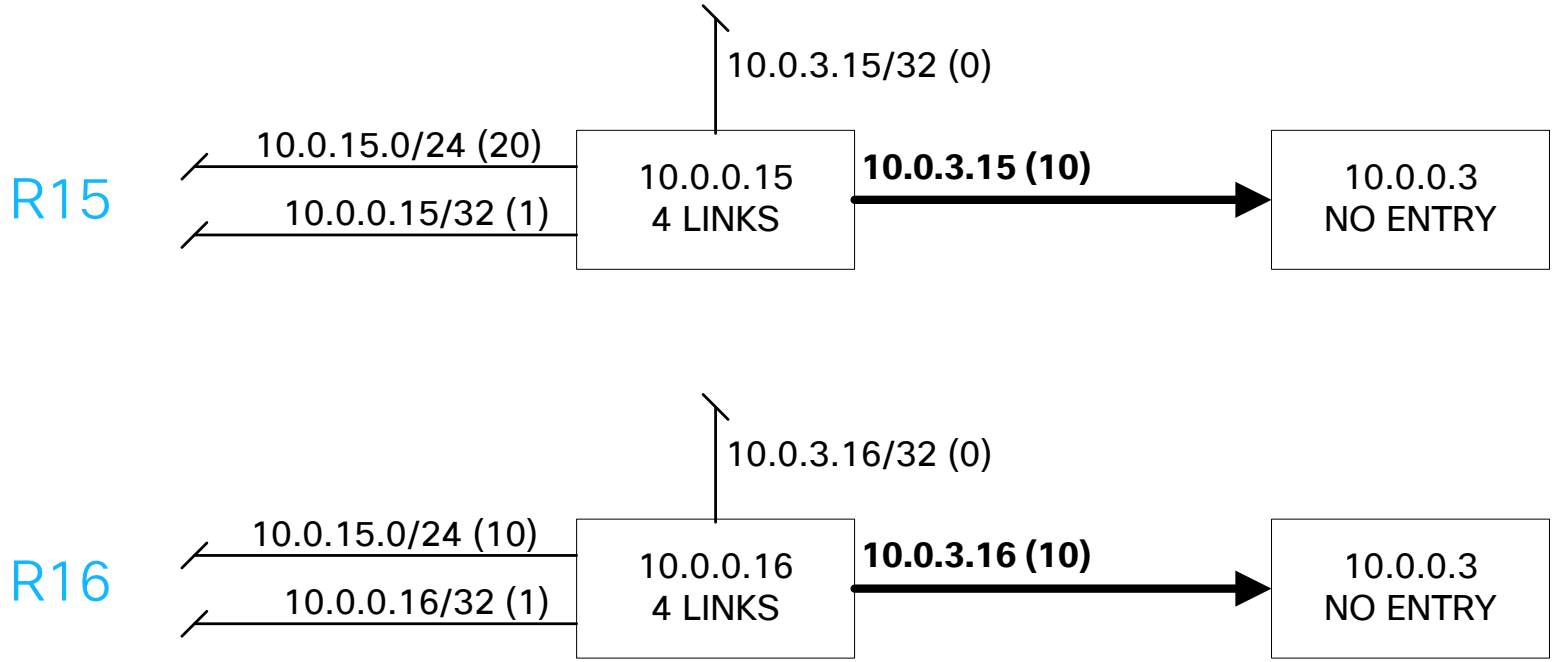


Definitely fishy from R16's view

```
R16#show ip ospf database
[snip]
Link ID         ADV Router      Age     Seq#       Checksum Link count
10.0.0.16       10.0.0.16       225     0x8000000A 0x00A88F 4

R16#show ip route ospf | include ^O
[no output]
```

- Analysis
  - R3 has OSPF routes, R16 does not
  - Check R15 as well; same behavior
- Hypothesis
  - R3 is performing intra-area LSA filtering

Examine the spokes Graphs



More Analysis

- Routers in an area have an inconsistent view
  - But this is not necessarily a problem
- Hub learns OSPF routes from spokes
- Spokes have a set of static routes upstream
  - Equivalent of “hub summarization” used with EIGRP/BGP
  - Spoke-to-spoke dynamic tunnels still work ☺

So how does it work?

```
R16#show ip route 10.0.0.15
Routing entry for 10.0.0.0/8
  Known via "static", distance 1, metric 0
  Advertised by bgp 1
  Routing Descriptor Blocks:
  * 10.0.3.3
      Route metric is 0, traffic share count is 1

R16#traceroute 10.0.0.15 source 10.0.0.16
Type escape sequence to abort.
Tracing the route to 10.0.0.15
VRF info: (vrf in name/id, vrf out name/id)
  1 10.0.3.3 6 msec 2 msec 4 msec
  2 10.0.3.15 9 msec 3 msec 3 msec
```

```
R16#show ip route 10.0.0.15
Routing entry for 10.0.0.15/32
  Known via "nhrp", distance 250, metric 255
  Last update from 10.0.3.15 on Tunnel3, 00:00:05 ago
  Routing Descriptor Blocks:
  * 10.0.3.15, from 10.0.3.15, 00:00:05 ago, via Tun3
      Route metric is 255, traffic share count is 1
      MPLS label: none

R16#traceroute 10.0.0.15 source 10.0.0.16
Type escape sequence to abort.
Tracing the route to 10.0.0.15
VRF info: (vrf in name/id, vrf out name/id)
  1 10.0.3.15 9 msec 6 msec 5 msec
```

- The spokes have upstream static routes to cover all destinations
- DMVPN still handles spoke-to-spoke dynamic NHRP routes
- Works easily with one OSPF hub; more than one is challenging

Understanding “Forward Address”

You make the power of data **possible**

Cisco live!

No surprises so far ...

```
R15#show bgp ipv4 unicast summary | begin ^Neighbor
Neighbor        V           AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.0.15.17      4        42518      90      91        3    0    0 01:18:08        1

R15#show ip route bgp | include ^B
B        10.0.0.17/32 [20/0] via 10.0.15.17, 01:18:02
```

```
R16#show bgp ipv4 unicast summary | begin ^Neighbor
Neighbor        V           AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.0.15.17      4        42518      90      91        3    0    0 01:18:15        1

R16#show ip route bgp | include ^B
B        10.0.0.17/32 [20/0] via 10.0.15.17, 01:17:42
```

- R15 and R16 have an eBGP peering to R17
- R15 and R16 learn only R17's loopback

... but now, find the difference!

```
R15#show ip route 10.0.0.17
Routing entry for 10.0.0.17/32
  Known via "bgp 1", distance 20, metric 0
  Tag 42518, type external
  Redistributing via ospf 1
  Advertised by ospf 1 metric-type 2 subnets
  Last update from 10.0.15.17 01:18:25 ago
  Routing Descriptor Blocks:
  * 10.0.15.17, from 10.0.15.17, 01:18:25 ago
      Route metric is 0, traffic share count is 1
      AS Hops 1
      Route tag 42518
      MPLS label: none
```

```
R16#show ip route 10.0.0.17
Routing entry for 10.0.0.17/32
  Known via "bgp 1", distance 20, metric 0
  Tag 42518, type external
  Last update from 10.0.15.17 01:18:29 ago
  Routing Descriptor Blocks:
  * 10.0.15.17, from 10.0.15.17, 01:18:29 ago
      Route metric is 0, traffic share count is 1
      AS Hops 1
      Route tag 42518
      MPLS label: none
```

- Only R15 is redistributing it
  - This suggests R16 would never be used for forwarding ...
  - Sounds like a good hypothesis

A simple experiment; Verify!

```
R3#show ip ospf border-routers | include Area 2
i 10.0.0.15 [10] via 10.0.3.15, Tun3, ASBR, Area 2, SPF 5

R3#show ip ospf database external 10.0.0.17
[snip]
  Link State ID: 10.0.0.17 (External Network Number )
  Advertising Router: 10.0.0.15
  [snip]
  Network Mask: /32
        Metric Type: 2 (Larger than any link state path)
        MTID: 0
        Metric: 1
        Forward Address: 10.0.15.17
        External Route Tag: 42518
```

```
R3#show ip route 10.0.0.17
Routing entry for 10.0.0.17/32
  Known via "ospf 1", distance 110, metric 1
  Tag 42518, type extern 2, forward metric 20
  Last update from 10.0.3.16 on Tun3, 00:08:00 ago
  Routing Descriptor Blocks:
  * 10.0.3.16, from 10.0.0.15, 00:08:00 ago, via Tun3
      Route metric is 1, traffic share count is 1
      Route tag 42518
```

- Our hypothesis was wrong
  - The route (via LSA5) came from R15
  - The next-hop is via R16

Shortest path to FA

```
R3#show ip route 10.0.15.17
Routing entry for 10.0.15.0/24
  Known via "ospf 1", distance 110, metric 20, type intra area
  Last update from 10.0.3.16 on Tunnel3, 01:23:17 ago
  Routing Descriptor Blocks:
  * 10.0.3.16, from 10.0.0.16, 01:23:17 ago, via Tunnel3
      Route metric is 20, traffic share count is 1

R3#show ip ospf interface tunnel 3 | include Cost:
  Process ID 1, Router ID 10.0.0.3, Network Type POINT_TO_MULTIPOINT, Cost: 10
```

```
R3#show ip ospf 1 2 database router 10.0.0.15
[snip]
    Link connected to: a Stub Network
     (Link ID) Network/subnet number: 10.0.15.0
     (Link Data) Network Mask: 255.255.255.0
      Number of MTID metrics: 0
       TOS 0 Metrics: 20
```

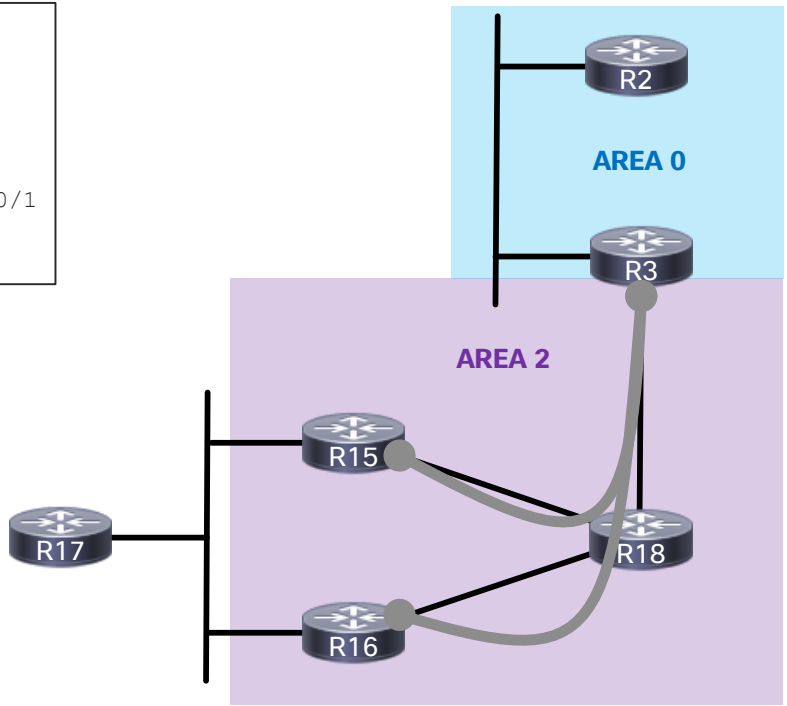
```
R3#show ip ospf 1 2 database router 10.0.0.16
[snip]
    Link connected to: a Stub Network
     (Link ID) Network/subnet number: 10.0.15.0
     (Link Data) Network Mask: 255.255.255.0
      Number of MTID metrics: 0
       TOS 0 Metrics: 10
```

- Cost to FA from R3 is 20 (all intra area)
  - R3 cost to R16 is 10
  - R16 FA LAN cost is 10

What about beyond area 2?

```
R2#show ip route 10.0.0.17
Routing entry for 10.0.0.17/32
  Known via "ospf 1", distance 110, metric 1
  Tag 42518, type extern 2, forward metric 30
  Last update from 10.0.1.3 on Ethernet0/1, 00:00:10 ago
  Routing Descriptor Blocks:
  * 10.0.1.3, from 10.0.0.15, 00:00:10 ago, via Ethernet0/1
      Route metric is 1, traffic share count is 1
      Route tag 42518
```

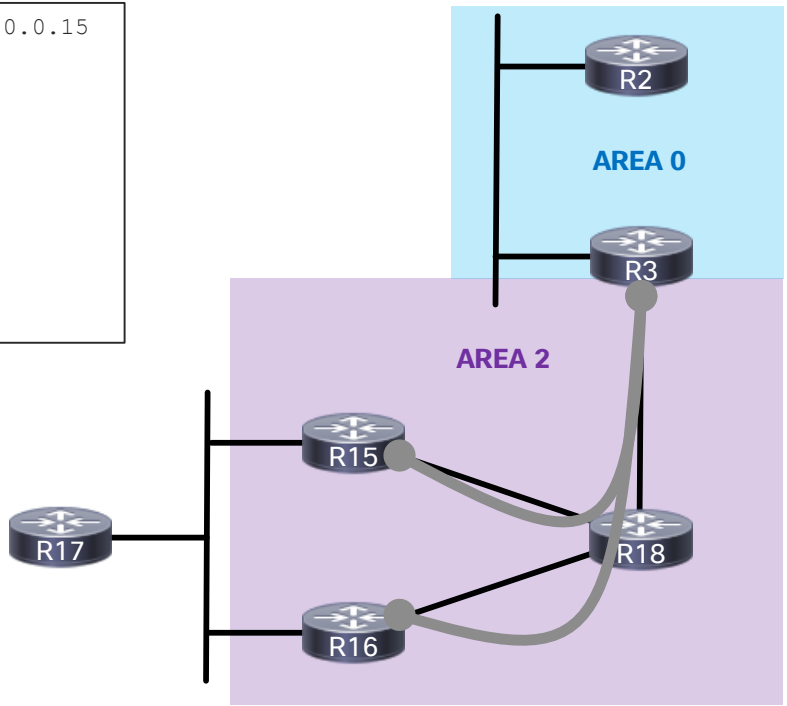
- Three key navigation questions
  1. Where did “metric 1” come from?
  2. Where did “forward metric 30” come from?
  3. How did R2 know about R15 at all?



Where did “metric 1” come from?

```
R2#show ip ospf database external 10.0.0.17 adv-router 10.0.0.15
[snip]
  Link State ID: 10.0.0.17 (External Network Number )
  Advertising Router: 10.0.0.15
  [snip]
  Network Mask: /32
        Metric Type: 2 (Larger than any link state path)
        MTID: 0
        Metric: 1
        Forward Address: 10.0.15.17
        External Route Tag: 42518
```

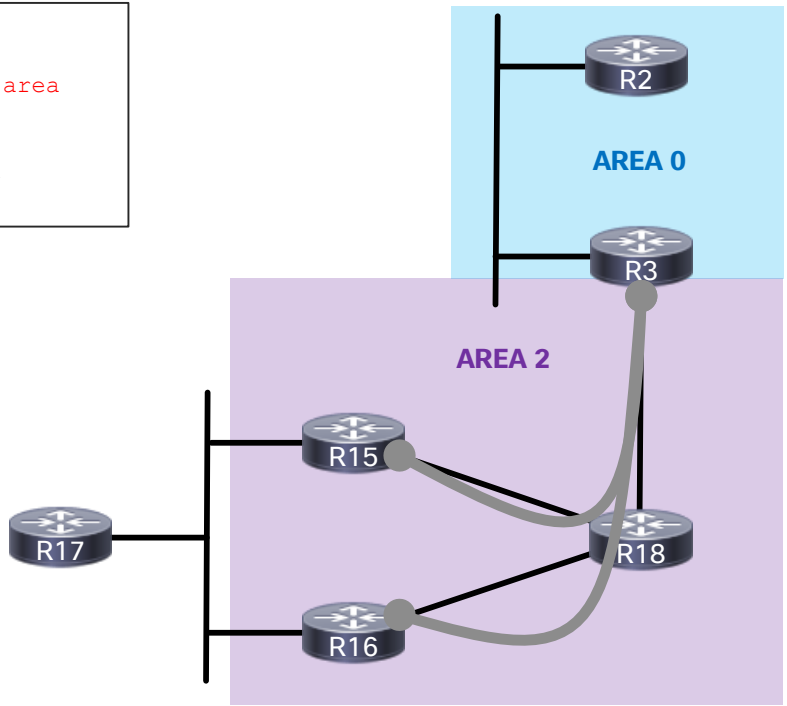
- Answer: the LSA5 seed metric
  - External-2 not summed with forward metric
  - BGP → OSPF redistributed default metric is 1



Where did “forward metric 30” come from?

```
R2#show ip route 10.0.15.0
Routing entry for 10.0.15.0/24
  Known via "ospf 1", distance 110, metric 30, type inter area
  Last update from 10.0.1.3 on Ethernet0/1, 01:35:38 ago
  Routing Descriptor Blocks:
  * 10.0.1.3, from 10.0.0.3, 01:35:38 ago, via Ethernet0/1
      Route metric is 30, traffic share count is 1
```

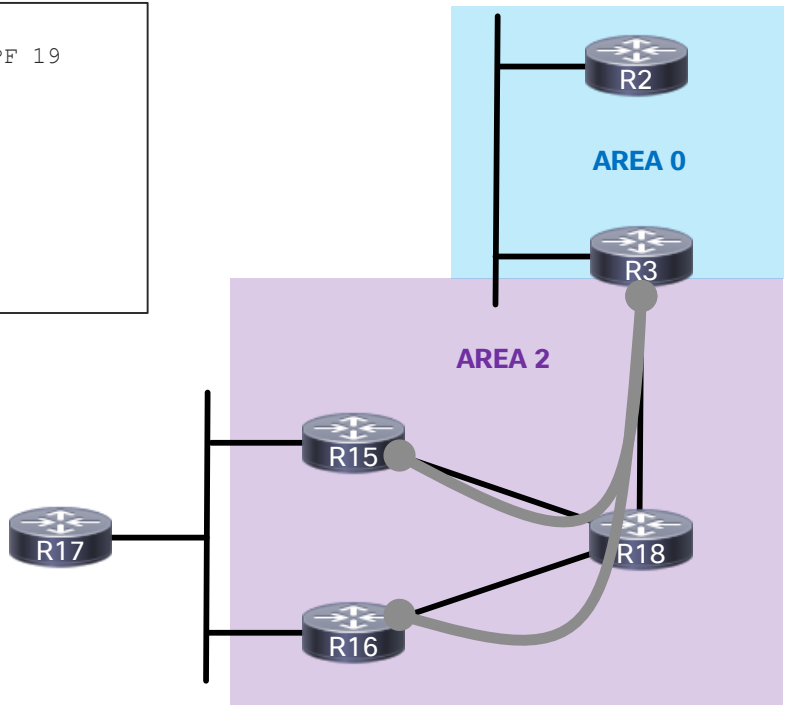
- Superficial answer: the cost to the FA
  - Better to answer where 30 came from
  - Before we do, note that this is an inter-area route



Really, where did 30 come from?

```
R2#show ip ospf border-routers | include 10.0.0.3
i 10.0.0.3 [10] via 10.0.1.3, Ethernet0/1, ABR, Area 0, SPF 19

R2#show ip ospf 1 0 database summary 10.0.15.0
[snip]
  Link State ID: 10.0.15.0 (summary Network Number)
  Advertising Router: 10.0.0.3
  [snip]
  Network Mask: /24
        MTID: 0         Metric: 20
```

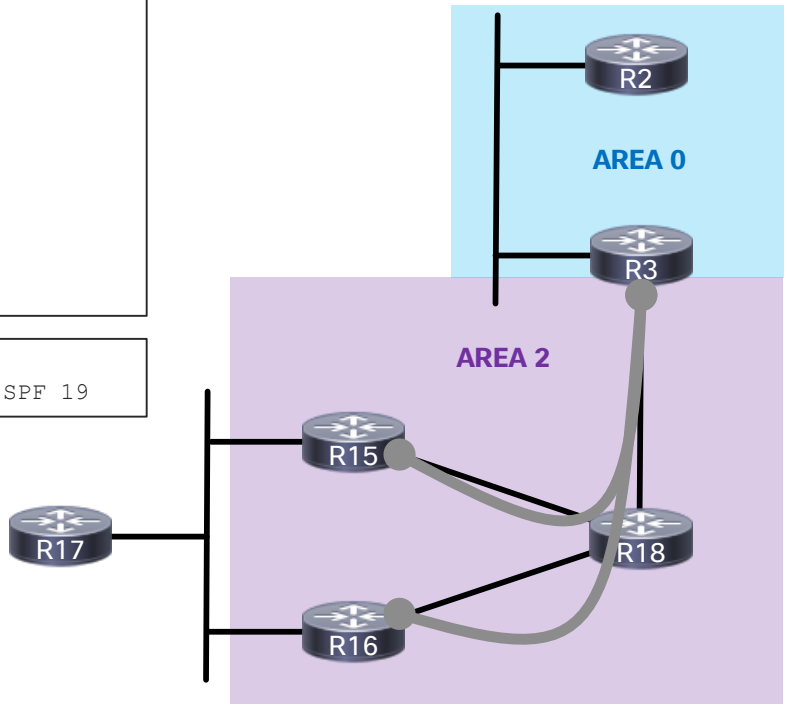
- Better answer: the cost to ABR plus the ABR's cost to FA
  - Cost from us to ABR is 10
  - Cost from ABR to FA is 20 from LSA3



How did R2 know about R15 at all?

```
R2#show ip ospf 1 0 database router 10.0.0.15
[no relevant output]

R2#show ip ospf 1 0 database asbr-summary 10.0.0.15
[snip]
  LS Type: Summary Links(AS Boundary Router)
  Link State ID: 10.0.0.15 (AS Boundary Router address)
  Advertising Router: 10.0.0.3
  [snip]
  Network Mask: /0
        MTID: 0         Metric: 10
```

```
R2#show ip ospf border-routers | include 10.0.0.15
I 10.0.0.15 [20] via 10.0.1.3, Ethernet0/1, ASBR, Area 0, SPF 19
```

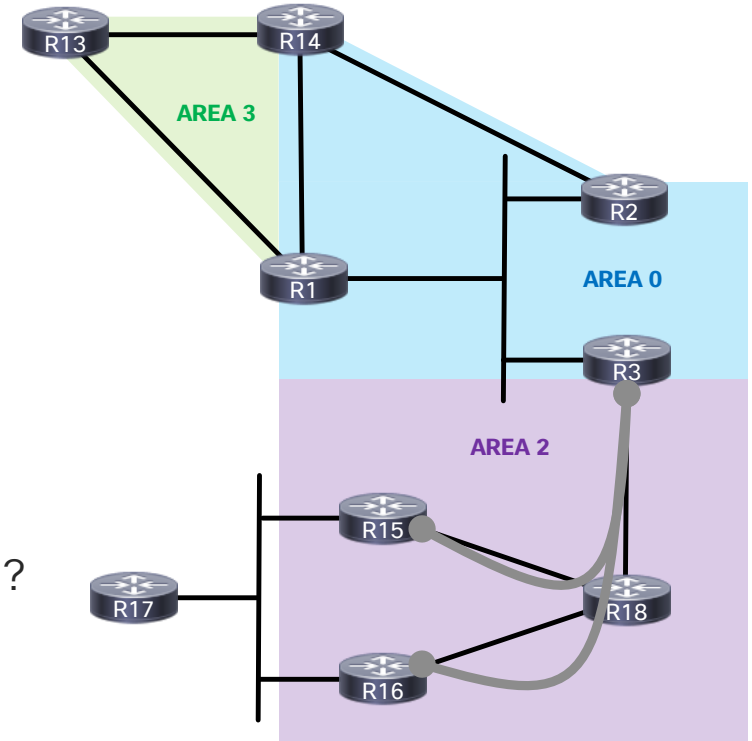


- Answer: The LSA4
  - Used when performing an inter-area lookup on an external route
  - Originated by an ABR to advertise its cost to an ASBR

What about beyond area 0?

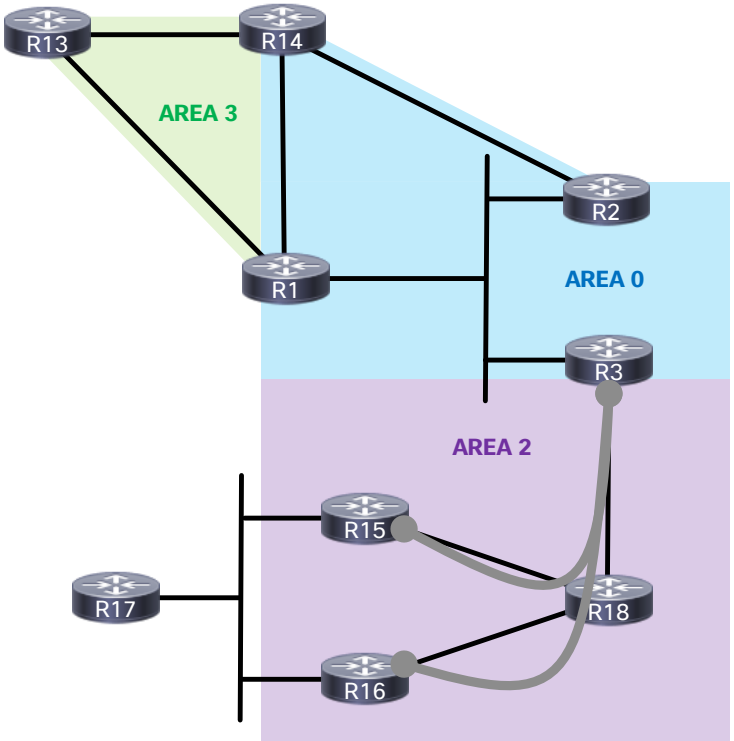
```
R13#show ip route 10.0.0.17
Routing entry for 10.0.0.17/32
  Known via "ospf 1", distance 110, metric 1
  Tag 42518, type extern 2, forward metric 40
  Last update from 10.1.13.1 on Ethernet0/2, 00:20:06 ago
  Routing Descriptor Blocks:
  * 10.1.13.1, from 10.0.0.15, 00:20:06 ago, via Ethernet0/2
      Route metric is 1, traffic share count is 1
      Route tag 42518
```

- Three key navigation questions (again)
  1. Where did “metric 1” come from?
  2. Where did “forward metric 40” come from?
  3. How did R13 know about R15 at all?



Where did “metric 1” come from?

```
R13#show ip ospf database external 10.0.0.17 adv-router 10.0.0.15
[snip]
  Link State ID: 10.0.0.17 (External Network Number )
  Advertising Router: 10.0.0.15
  [snip]
  Network Mask: /32
        Metric Type: 2 (Larger than any link state path)
        MTID: 0
        Metric: 1
        Forward Address: 10.0.15.17
        External Route Tag: 42518
```



- Answer: the LSA5 seed metric
  - Same as before!
  - Intermediate OSPF routers can’t modify this

Where did “forward metric 40” come from?

```
R13#show ip route 10.0.15.0
Routing entry for 10.0.15.0/24
  Known via "ospf 1", distance 110, metric 40, type inter area
  Last update from 10.1.13.1 on Ethernet0/2, 01:54:49 ago
  Routing Descriptor Blocks:
  * 10.1.13.1, from 10.0.0.1, 01:54:49 ago, via Ethernet0/2
      Route metric is 40, traffic share count is 1
```

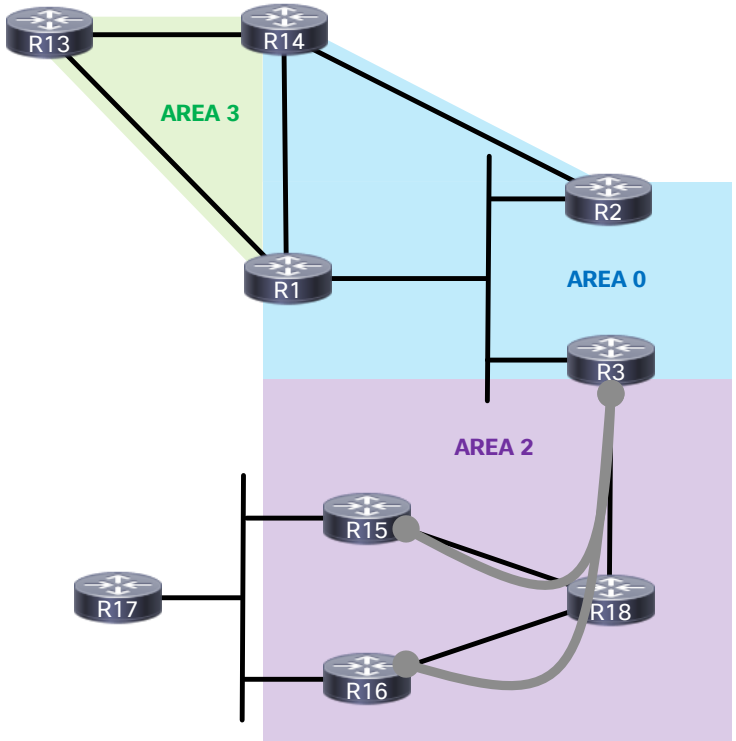
```
R13#show ip ospf 1 3 database summary 10.0.15.0
  [snip]
  Advertising Router: 10.0.0.14
  [snip]
  Network Mask: /24
        MTID: 0         Metric: 40
```

```
... continued from left ...
[snip]
  Advertising Router: 10.0.0.1
  [snip]
  Network Mask: /24
        MTID: 0         Metric: 30
```

- Answer: the cost to the FA (40)
  - Two LSA3s are received from two ABRs with varying metrics

If R1's LSA3 was used, why is cost 40?

```
R13#show ip ospf 1 3 database summary 10.0.15.0 adv-router 10.0.0.1
[snip]
  Link State ID: 10.0.15.0 (summary Network Number)
  Advertising Router: 10.0.0.1
  [snip]
  Network Mask: /24
        MTID: 0         Metric: 30


R13#show ip ospf border-routers | include ABR
i 10.0.0.1 [10] via 10.1.13.1, Ethernet0/2, ABR, Area 3, SPF 18
i 10.0.0.14 [10] via 10.13.14.14, Ethernet0/3, ABR, Area 3, SPF 18
```



- Answer: ABR cost + cost to ABR ☺
  - Use the border-routers command!
  - Cost to both ABRs is 10, so 10 + 30 = 40

How did the R13 know about R15 at all?

```
R13#show ip ospf database asbr-summary 10.0.0.15
[snip]
  Advertising Router: 10.0.0.1
  [snip]
  Network Mask: /0
        MTID: 0         Metric: 20
```

```
... continued from left ...
[snip]
  Advertising Router: 10.0.0.14
  [snip]
  Network Mask: /0
        MTID: 0         Metric: 30
```

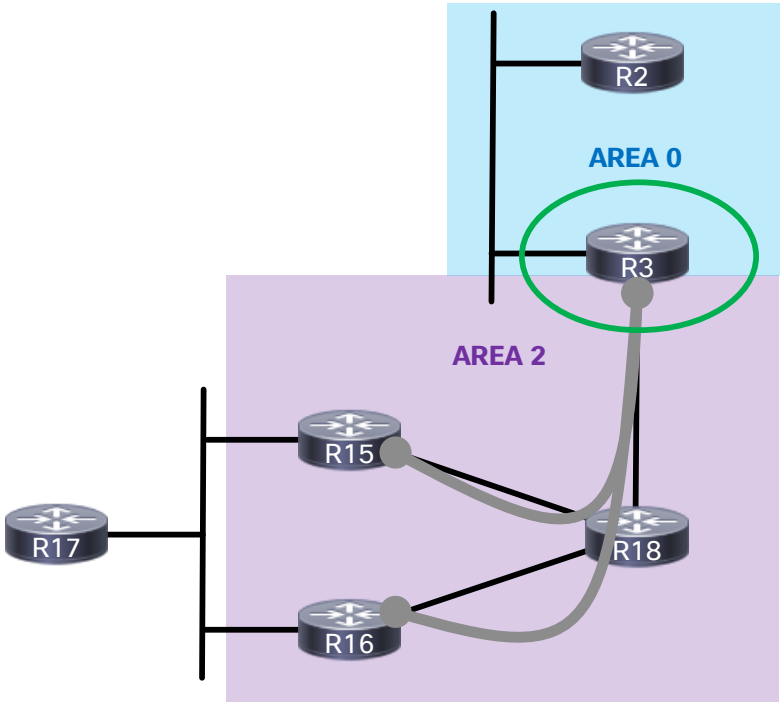
```
R13#show ip ospf border-routers | include 10.0.0.15
I 10.0.0.15 [30] via 10.1.13.1, Ethernet0/2, ASBR, Area 3, SPF 18
```

- Answer: the LSA4
  - Originated by R1 and R14 (the ABRs) with their respective costs
  - As usual, “border-routers” is your friend!

FA problems

```
R3#show ip prefix-list PL_AREA3_OUT
ip prefix-list PL_AREA3_OUT: 2 entries
   seq 5 deny 10.0.15.0/24
   seq 10 permit 0.0.0.0/0 le 32


R3#show ip ospf 1 | begin Area 2
    Area 2
        [snip]
        Area-filter PL_AREA3_OUT out
        [snip]
```

- Before we observe, discuss the impact



FA problems

```
R13#show ip route 10.0.15.0
% Subnet not in table

R13#show ip ospf 1 3 database summary 10.0.15.0
[no relevant output]
```

```
R13#show ip route 10.0.0.17
% Subnet not in table

R13#show ip ospf database external 10.0.0.17
[snip]
  Advertising Router: 10.0.0.15
  [snip]
  Network Mask: /32
        Metric Type: 2 (...)
        MTID: 0
        Metric: 1
        Forward Address: 10.0.15.17
        External Route Tag: 42518
```

- Observe
  - We expected 10.0.15.0/24 to disappear, and it did
  - Why did 10.0.0.17/32 also disappear?
- Hypothesis
  - We need a route to the FA

A valiant attempt

```
! Config applied to R13
ip route 10.0.15.0 255.255.255.0 10.1.13.1


R13#show ip route 10.0.15.0
Routing entry for 10.0.15.0/24
  Known via "static", distance 1, metric 0
  Routing Descriptor Blocks:
  * 10.1.13.1
      Route metric is 0, traffic share count is 1

R13#show ip ospf 1 3 database summary 10.0.15.0
[no relevant output]
```

```
R13#show ip route 10.0.0.17
% Subnet not in table

R13#show ip ospf database external 10.0.0.17
[snip]
  Advertising Router: 10.0.0.15
  [snip]
  Network Mask: /32
        Metric Type: 2 (...)
        MTID: 0
        Metric: 1
        Forward Address: 10.0.15.17
        External Route Tag: 42518
```

- Experiment: let's toss a static route on R13
- Analysis
  - Didn't solve our problem ☹
  - Route to FA must be OSPF intra or inter-area

Another solution: Disable the FA

```
R15#show ip ospf interface ethernet 0/3
%OSPF: OSPF not enabled on Ethernet0/3

... or ...

R15#show ip ospf interface ethernet 0/3 | include Passive
    No Hellos (Passive interface)

... or ...

R15#show ip ospf interface ethernet 0/3 | include Network
 Process ID 1, Router ID 10.0.0.15, Network Type POINT_TO_POINT, Cost: 10
```

- Can be accomplished by any of the following on R15 E0/3
  - Disabling OSPF entirely
  - Passive interface
  - Non-DR network type (point-to-anything)

Disable the FA: the aftermath

```
R13#show ip ospf database external 10.0.0.17
[snip]
  Advertising Router: 10.0.0.15
  [snip]
  Network Mask: /32
        Metric Type: 2 (...)
        MTID: 0
        Metric: 1
        Forward Address: 0.0.0.0
        External Route Tag: 42518
```

```
R13#show ip route 10.0.0.17
Routing entry for 10.0.0.17/32
  Known via "ospf 1", distance 110, metric 1
  Tag 42518, type extern 2, forward metric 30
  Last update from 10.1.13.1 on Eth0/2, 00:00:41 ago
  Routing Descriptor Blocks:
  * 10.1.13.1, from 10.0.0.15, 00:00:41 ago, via Eth0/2
      Route metric is 1, traffic share count is 1
      Route tag 42518
```

```
R13#show ip ospf border-routers | include 10.0.0.15
I 10.0.0.15 [30] via 10.1.13.1, Ethernet0/2, ASBR, Area 3, SPF 18
```

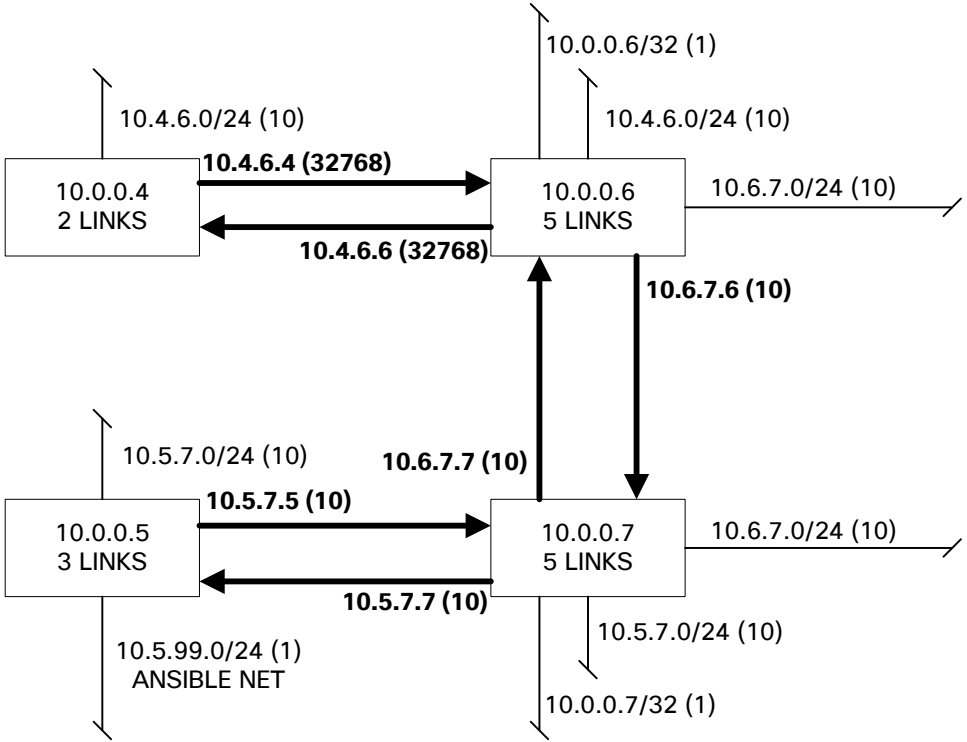
- FA set to 0.0.0.0
  - Use the cost to the ASBR (R15) instead
  - “border-routers” command reveals the forward metric

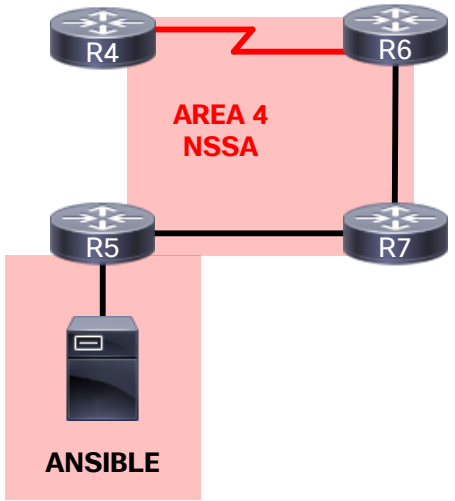
Exploring the NSSA

You make customer experience **possible**

Cisco live!

Area 4 Graph





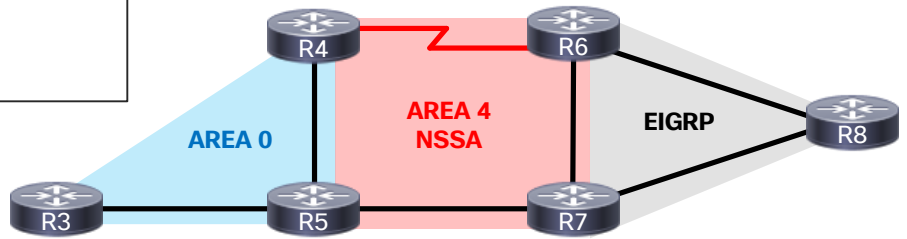
NSSA: Interesting from the outside

```
R3#show ip route 10.0.0.8
Routing entry for 10.0.0.8/32
  Known via "ospf 1", distance 110, metric 20, type extern 2, forward metric 21
  Last update from 10.3.5.5 on Ethernet0/3, 00:03:07 ago
  Routing Descriptor Blocks:
  * 10.3.5.5, from 10.0.0.5, 00:03:07 ago, via Ethernet0/3
      Route metric is 20, traffic share count is 1

R3#show ip ospf database | begin -5
                Type-5 AS External Link States

Link ID         ADV Router      Age         Seq#       Checksum Tag
10.0.0.8        10.0.0.5        101         0x80000001 0x006D01 0
10.0.0.17       10.0.0.15       1567        0x80000001 0x000791 42518
```

- Let's ask our 3 questions with a small modification:
  1. Where did "metric 20" come from?
  2. Where did "forward metric 21" come from?
  3. Since R3 doesn't know about R6/R7, what is R5's role?

Where did “metric 20” come from?

```
R3#show ip ospf database external 10.0.0.8 adv-router 10.0.0.5
[snip]
  Link State ID: 10.0.0.8 (External Network Number )
  Advertising Router: 10.0.0.5
  [snip]
  Network Mask: /32
        Metric Type: 2 (Larger than any link state path)
        MTID: 0
        Metric: 20
        Forward Address: 10.0.0.7
        External Route Tag: 0
```

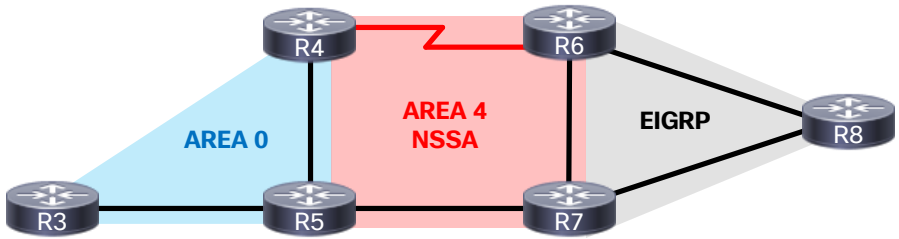


- Answer: the LSA5 seed metric
  - External-2 not summed with forward metric
  - Non-BGP → OSPF default seed metric is 20

Where did “forward metric 21” come from?

```
R3#show ip route 10.0.0.7
Routing entry for 10.0.0.7/32
  Known via "ospf 1", distance 110, metric 21, type inter area
  Last update from 10.3.5.5 on Ethernet0/3, 00:29:35 ago
  Routing Descriptor Blocks:
  * 10.3.5.5, from 10.0.0.5, 00:29:35 ago, via Ethernet0/3
      Route metric is 21, traffic share count is 1

R3#show ip ospf 1 0 database summary 10.0.0.7 adv-router 10.0.0.5
[snip]
  Network Mask: /32
        MTID: 0         Metric: 11

R3#show ip ospf border-routers | include 10.0.0.5
i 10.0.0.5 [10] via 10.3.5.5, Ethernet0/3, ABR/ASBR, Area 0, SPF 12
```

- Answer: the cost to the FA
  - R5 advertises a cost of 11
  - R3’s cost to R5 is 10



What is R5's role?

```
R5#show ip ospf | begin Area_4
    Area 4
        Number of interfaces in this area is 1
        It is a NSSA area
        Perform type-7/type-5 LSA translation
```

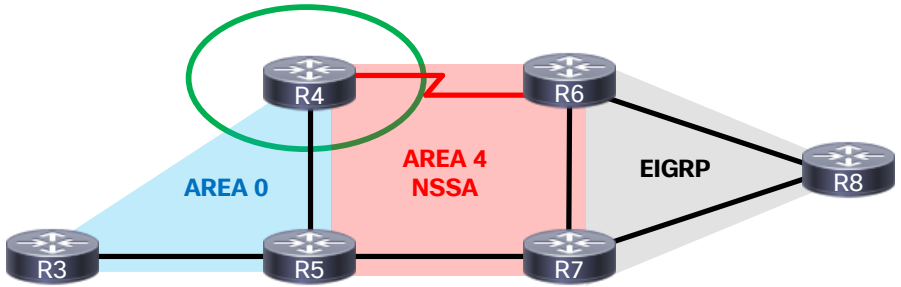
```
R5#show ip ospf database external 10.0.0.8 self-orig
[snip]
  Link State ID: 10.0.0.8 (External Network Number )
  Advertising Router: 10.0.0.5
  [snip]
  Network Mask: /32
        Metric Type: 2 (Larger than any ...)
        MTID: 0
        Metric: 20
        Forward Address: 10.0.0.7
        External Route Tag: 0
```

- Answer: 7to5 translator
  - Obviates need for LSA4 from area 4 to area 0
  - R5 re-originates the LSA7 as an LSA5

What happens if R4 is the 7to5 translator?

```
! Config applied to R4
router ospf 1
 area 4 nssa translate type7 always

R4#show ip ospf | begin Area_4
    Area 4
        Number of interfaces in this area is 1
        It is a NSSA area
        Configured to translate Type-7 LSAs
        Perform type-7/type-5 LSA translation
        [snip]
```

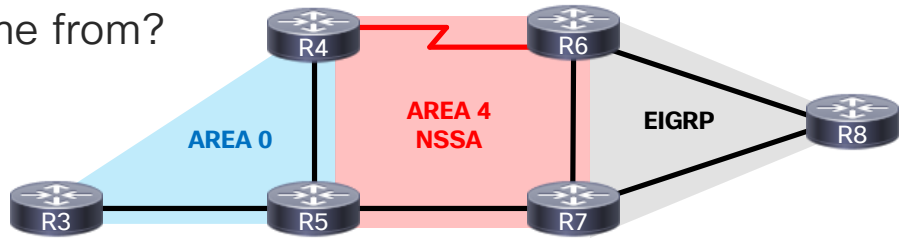
- Two mutually exclusive hypotheses
  - Traffic R3→R8 will flow through R4
  - Traffic R3→R8 will flow through R5



Analysis

```
R3#show ip route 10.0.0.8
Routing entry for 10.0.0.8/32
  Known via "ospf 1", distance 110, metric 20, type extern 2, forward metric 31
  Last update from 10.3.5.5 on Ethernet0/3, 00:00:59 ago
  Routing Descriptor Blocks:
  * 10.3.5.5, from 10.0.0.4, 00:00:59 ago, via Ethernet0/3
      Route metric is 20, traffic share count is 1
```

- Route from R4, traffic through R5
  1. Where did “metric 20” come from?
  2. Where did “forward metric 31” come from?



Where did “metric 20” come from?

```
R3#show ip ospf database external 10.0.0.8 adv-router 10.0.0.4
[snip]
  Link State ID: 10.0.0.8 (External Network Number )
  Advertising Router: 10.0.0.4
  [snip]
  Network Mask: /32
        Metric Type: 2 (Larger than any link state path)
        MTID: 0
        Metric: 20
        Forward Address: 10.0.0.6
        External Route Tag: 0
```

- Answer: the seed metric
  - This is old news
  - … but the FA has changed from 10.0.0.7 to 10.0.0.6

Where did “forward metric 31” come from?

```
R3#show ip route 10.0.0.6
Routing entry for 10.0.0.6/32
  Known via "ospf 1", distance 110, metric 31, type inter area
  Last update from 10.3.5.5 on Ethernet0/3, 00:38:52 ago
  Routing Descriptor Blocks:
  * 10.3.5.5, from 10.0.0.5, 00:38:52 ago, via Ethernet0/3
      Route metric is 31, traffic share count is 1

R3#show ip ospf 1 0 database summary 10.0.0.6 adv-router 10.0.0.5
[snip]
  Link State ID: 10.0.0.6 (summary Network Number)
  Advertising Router: 10.0.0.5
  [snip]
  Network Mask: /32
        MTID: 0         Metric: 21

R3#show ip ospf border-routers | include 10.0.0.5
i 10.0.0.5 [10] via 10.3.5.5, Ethernet0/3, ABR/ASBR, Area 0, SPF 12
```

- Answer: the cost to the FA
  - The forward metric went up, but the path didn't change!

Why did the FA change from R7 to R6?

```
R4#show ip ospf database nssa-external 10.0.0.8
[snip]
  Advertising Router: 10.0.0.6
  [snip]
  Network Mask: /32
        Metric Type: 2 (Larger than any link state path)
        MTID: 0
        Metric: 20
        Forward Address: 10.0.0.6
        External Route Tag: 0

  [snip]
  Advertising Router: 10.0.0.7
  [snip]
  Network Mask: /32
        Metric Type: 2 (Larger than any link state path)
        MTID: 0
        Metric: 20
        Forward Address: 10.0.0.7
        External Route Tag: 0
```

- Nothing suspicious here, just two LSA7s from each ASBR

Really, why did the FA change?

```
R4#show ip route 10.0.0.8
Routing entry for 10.0.0.8/32
  Known via "ospf 1", distance 110, metric 20, type NSSA extern 2, forward metric 32769
  Last update from 10.4.6.6 on Serial1/1, 00:13:37 ago
  Routing Descriptor Blocks:
  * 10.4.6.6, from 10.0.0.6, 00:13:37 ago, via Serial1/1
      Route metric is 20, traffic share count is 1

R4#show ip ospf border-routers | include 10.0.0.[67]
i 10.0.0.7 [32778] via 10.4.6.6, Serial1/1, ASBR, Area 4, SPF 11
i 10.0.0.6 [32768] via 10.4.6.6, Serial1/1, ASBR, Area 4, SPF 11
```

- Answer: Because R6 is closer to R4 than R7 is
  - From R4's view, it's sensible to translate the "better" LSA7
  - The FA helps area 0 (and beyond) account for the high cost R4-R6 link

Another Idea in 2019

- I hypothesized the following:
  - People don't want identical CL presentations year after year
  - People are interested in automating networks with reliable tools

Nicholas Russo @nickrusso42518

This is an extremely important question and I'll ask everyone to RT this. My troubleshooting OSPF session was approved for CLUS 2019! I'm debating replacing the last 15-20 minutes of NSSA deep dive with a production-grade OSPF troubleshooter Ansible demo. Please vote!

27% Leave NSSA deep dive

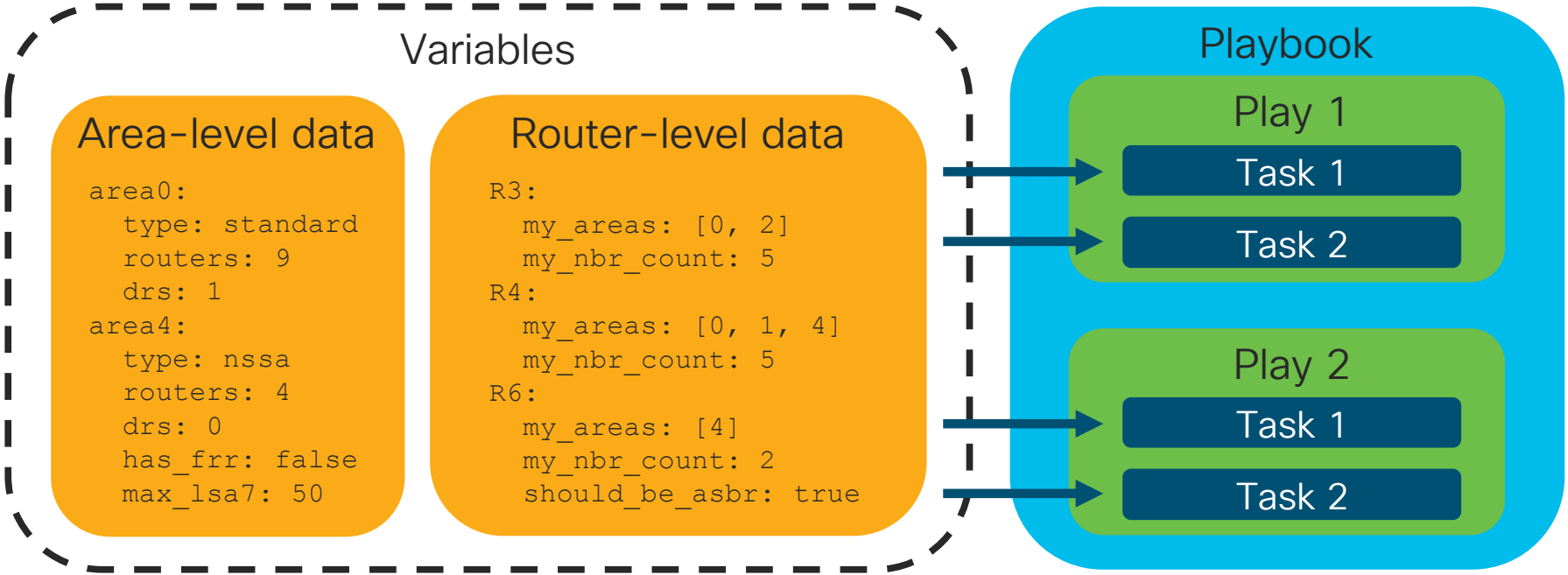
73% Demo OSPF Ansible tool

73% of respondents supported my hypothesis (n = 269)

Why Automate Troubleshooting?

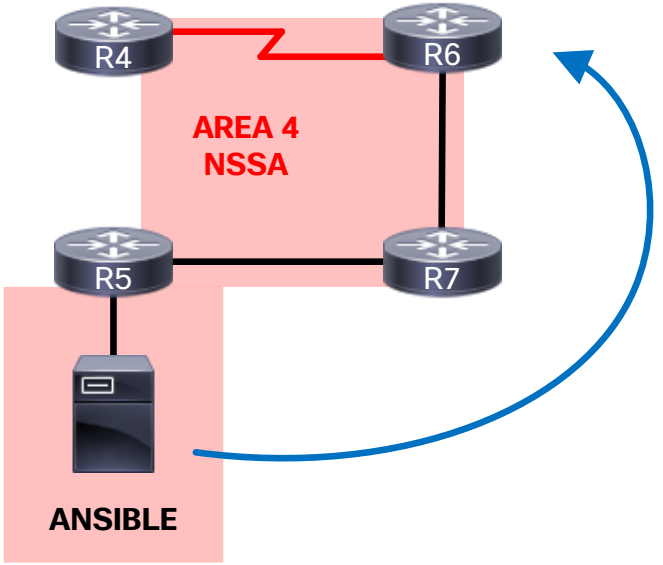
- Low risk of outage
- Saves time
- Consistent results
- **YOU** define success

Solution in One Slide



General Operation

- Ansible needs IP reachability
- Requires SSH access
- Check proper OSPF state
- Fail on error with details



Nick’s OSPF TroubleShooter (nots)

Common Performance Problems

You make the power of data **possible**

Duplicate RIDs

- How can you tell?
  - LS sequence number increasing fast
  - Low LS age due to constant resets
  - Very frequent SPF runs
  - High CPU usage from OSPF process
  - Obvious syslog messages

- But I didn't see any of that?
  - Use the Force (aka the LSDB)

Most basic method first

```
R12#show ip ospf 1 3 database database-summary
[snip]
Area 3 database summary
  LSA Type      Count    Delete   Maxage
  Router        3        0        0
[snip]

R12#show ip ospf 1 3 database
[snip]
Link ID         ADV Router      Age         Seq#       Checksum Link count
10.0.0.1        10.0.0.1        305         0x80000006 0x0053AD 5
10.0.0.13       10.0.0.13       381         0x80000006 0x00E9E8 5
10.0.0.14       10.0.0.14       305         0x80000005 0x00C1E4 5

R12#show ip ospf | include ID
 Routing Process "ospf 1" with ID 10.0.0.13
```

- Just count the number of nodes
  - We should have 4 router LSAs, not 3
  - We don't see 10.0.0.12, which likely means R12 has a misconfig
- The syslog below was not observed but would have been nice

```
%OSPF-4-DUP_RTRID1: Detected router with duplicate router ID 10.0.0.13 in area 3
```

Something trickier

```
R2#show ip ospf database external 10.0.0.17
[no output]
```

```
R3#show ip ospf database external 10.0.0.17
[no output]
```

```
R15#show ip ospf database external 10.0.0.17
[snip]
  Link State ID: 10.0.0.17 (External Network Number )
  Advertising Router: 10.0.0.2
  [snip]
  Network Mask: /32
        Metric Type: 2 (Larger than any link state path)
        MTID: 0
        Metric: 1
        Forward Address: 10.0.15.17
        External Route Tag: 42518
```

- Suppose there is another duplicate RID
  - No angry syslogs, which makes it harder
  - What happened?

Let's observe R15 and R2

```
R15#debug ip ospf lsa-generation
OSPF LSA generation debugging is on

R15#clear ip ospf process
Reset ALL OSPF processes? [no]: y

*Feb 26 01:14:48.642: OSPF-1 LSGEN: Build external LSA 10.0.0.17, mask 255.255.255.255, type 5, age 0, seq 0x80000001
*Feb 26 01:14:48.642: OSPF-1 LSGEN: MTID      Metric      Metric-type      FA              Tag         Topology Name
*Feb 26 01:14:48.642: OSPF-1 LSGEN: 0         1           2                10.0.15.17      42518       Base
```

```
R2#show ip ospf 1 0 database asbr-summary 10.0.0.2
[snip]
  Advertising Router: 10.0.0.3
  [snip]

R2#show ip ospf border-routers | include _10.0.0.2_
[no output]
```

- Four key points
  - R15 created and retained the LSA5 (note the seq number)
  - R3 created the LSA4 to describe R15 to area 0
  - R2 received that LSA4, but appears to reject it
  - Both R2 and R3 rejected the LSA5

Signaling the problem

```
R2#debug ip ospf lsa-generation
OSPF LSA generation debugging is on

*Feb 26 01:30:58.692: OSPF-1 LSGEN: Premature external LSA 5/10.0.0.17/10.0.0.2
*Feb 26 01:30:58.692: OSPF-1 LSGEN: Build external LSA 10.0.0.17, mask 255.255.255.255, type 5, age 3600, seq 0x80000002
*Feb 26 01:30:58.692: OSPF-1 LSGEN: MTID      Metric      Metric-type      FA              Tag         Topology Name
*Feb 26 01:30:58.692: OSPF-1 LSGEN: 0         16777215    2                0.0.0.0         0           Base
```

- How does R2 cry foul?
  - It originates an LSA5 itself, basically saying “I didn’t make this!”
  - Because the seq number is greater, it is newer, and thus more trusted
  - Other routers will purge this LSA from their LSDBs

- The syslog below was not observed but would have been nice

```
%OSPF-4-DUP_RTRID2: Detected router with duplicate router ID 10.0.0.2 in Type-4 LSA advertised by 10.0.0.3
```

Consider R3's perspective

```
R3#debug ip ospf spf external
OSPF SPF external debugging is on

*Feb 26 01:10:36.442: OSPF-1 EXTER: Start processing AS External LSA 5/10.0.0.17/10.0.0.2, mask 255.255.255.255
*Feb 26 01:10:36.442: OSPF-1 EXTER:  age 4, seq 0x80000001, metric 1, metric-type 2, fw-addr 10.0.15.17
*Feb 26 01:10:36.442: OSPF-1 EXTER:    Did not find route to ASBR 10.0.0.2

*Feb 26 01:10:37.217: OSPF-1 SPF  : Detect MAXAGE in LSA type 5, LS ID 10.0.0.17, from 10.0.0.2
*Feb 26 01:10:37.217: OSPF-1 SPF  : Detect generic change in LSA type 5, LSID 10.0.0.17, from 10.0.0.2 area 0
*Feb 26 01:10:37.217: OSPF-1 SPF  : Do not schedule partial SPF type 5, LSID 10.0.0.17,
                                    adv_rtr 10.0.0.2, area dummy area: INTRA/INTER spf scheduled

R3#show ip ospf database external 10.0.0.17
[no output]
```

- Now it makes sense
  - It first heard about the LSA5 from R15 in area 2
  - It then got a MAXAGE copy from R2 in area 0
  - When R15 re-generates it, process repeats

A Call to Action

- The graph is the core technology of OSPF
- Make time to draw it
- Empirical evidence is truth

Want more?

- Free stuff on GitHub
  - Session configurations
  - https://github.com/nickrusso42518/ospf_brkrst3310
  - OSPF troubleshooter (Ansible)
  - https://github.com/nickrusso42518/nots
- Reference content
  - Troubleshooting adjacencies
  - Troubleshooting performance
- Twitter @nickrusso42518

Cisco Webex Teams

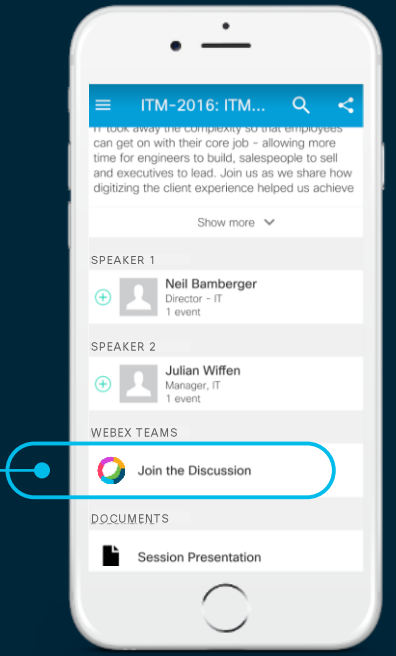
Questions?

Use Cisco Webex Teams to chat with the speaker after the session

How

1. Find this session in the Cisco Live Mobile App
2. Click “Join the Discussion”
3. Install Webex Teams or go directly to the team space
4. Enter messages/questions in the team space

Webex Teams will be moderated by the speaker until June 16, 2019.



cs.co/ciscolivebot#BRKRST-3310

Complete your online session evaluation

- Please complete your session survey after each session. Your feedback is very important.
- Complete a minimum of 4 session surveys and the Overall Conference survey (starting on Thursday) to receive your Cisco Live water bottle.
- All surveys can be taken in the Cisco Live Mobile App or by logging in to the Session Catalog on ciscolive.cisco.com/us.

Cisco Live sessions will be available for viewing on demand after the event at ciscolive.cisco.com.

Continue your education

Demos in the Cisco campus

Walk-in labs

Meet the engineer 1:1 meetings

Related sessions

Thank you

You make **possible**