

TOMORROW starts here.



Cisco *live!*

Troubleshooting Routing Protocols – BGP/OSPF/EIGRP

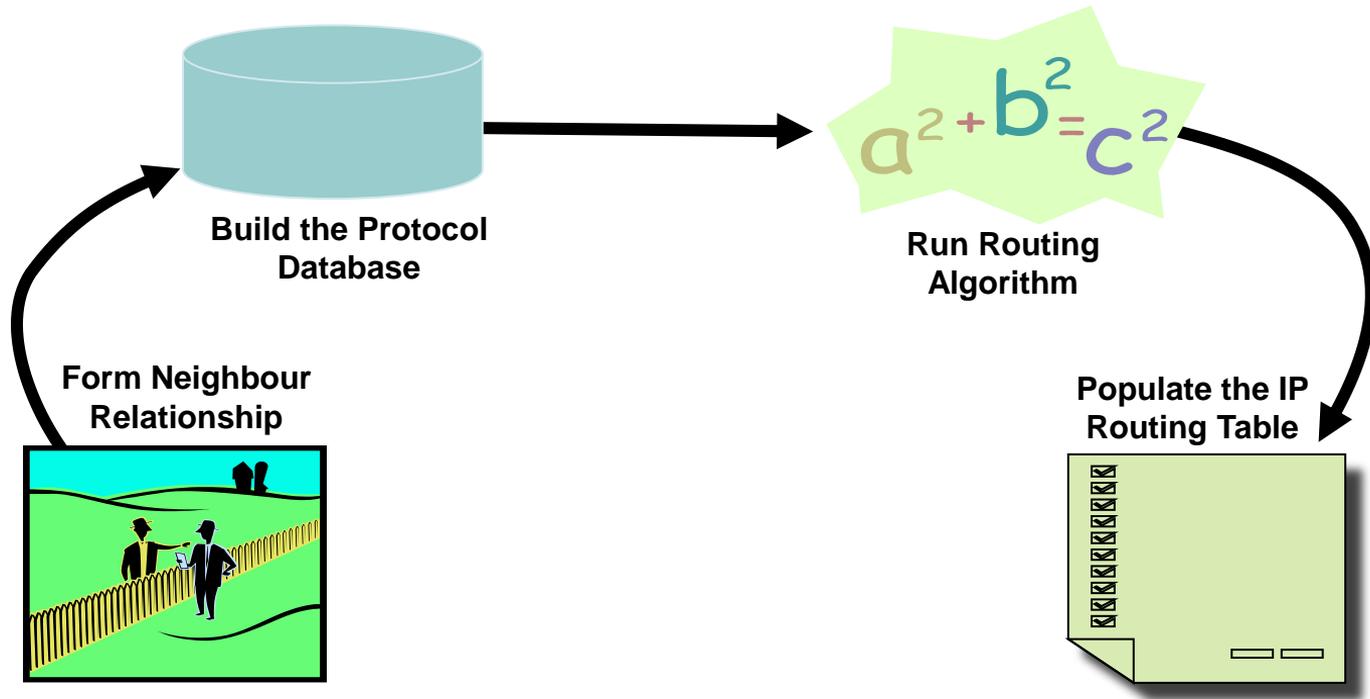
BRKRST-2619

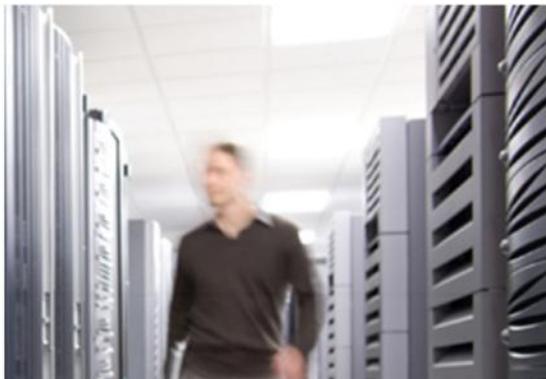
Ruchir Jain
Customer Support Engineer

Agenda

- View from 50,000 metres
- Generic Lifecycle approach to troubleshoot the following protocols:
 - EIGRP
 - OSPF
 - BGP
- Troubleshooting best practices

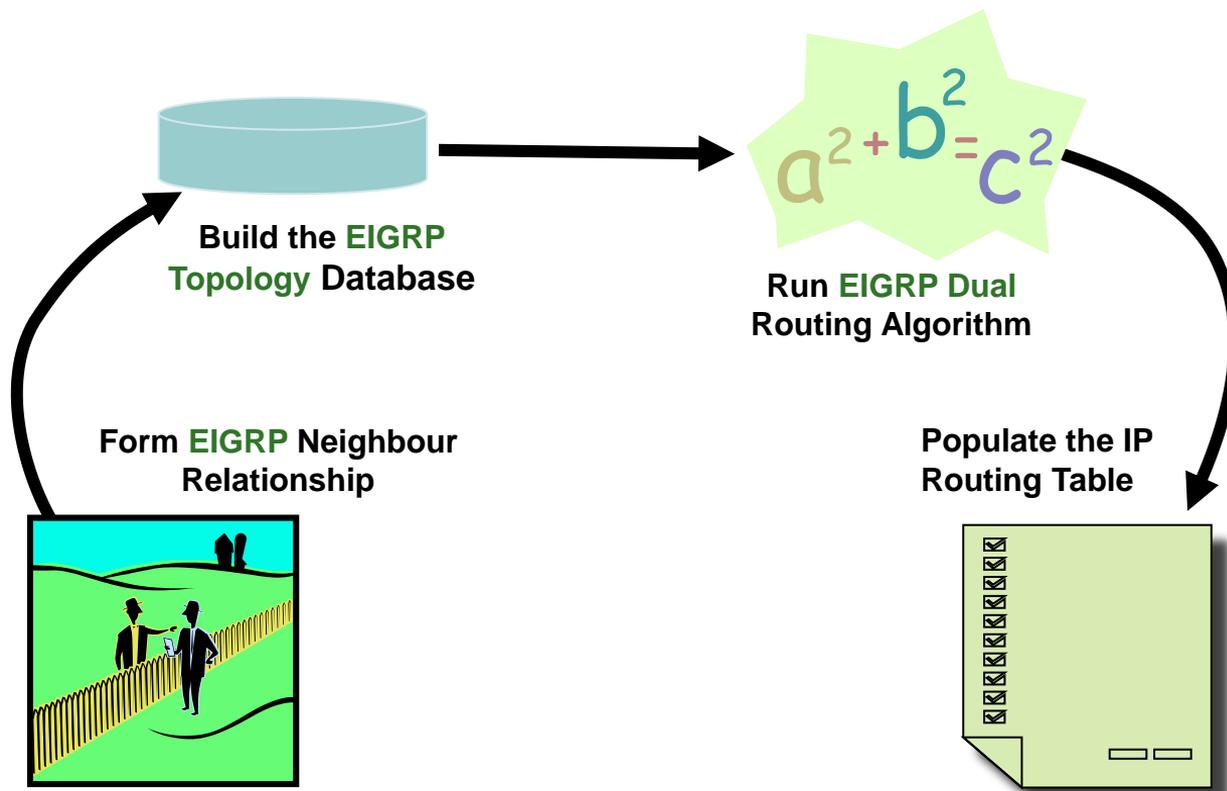
View From 50,000 Metres



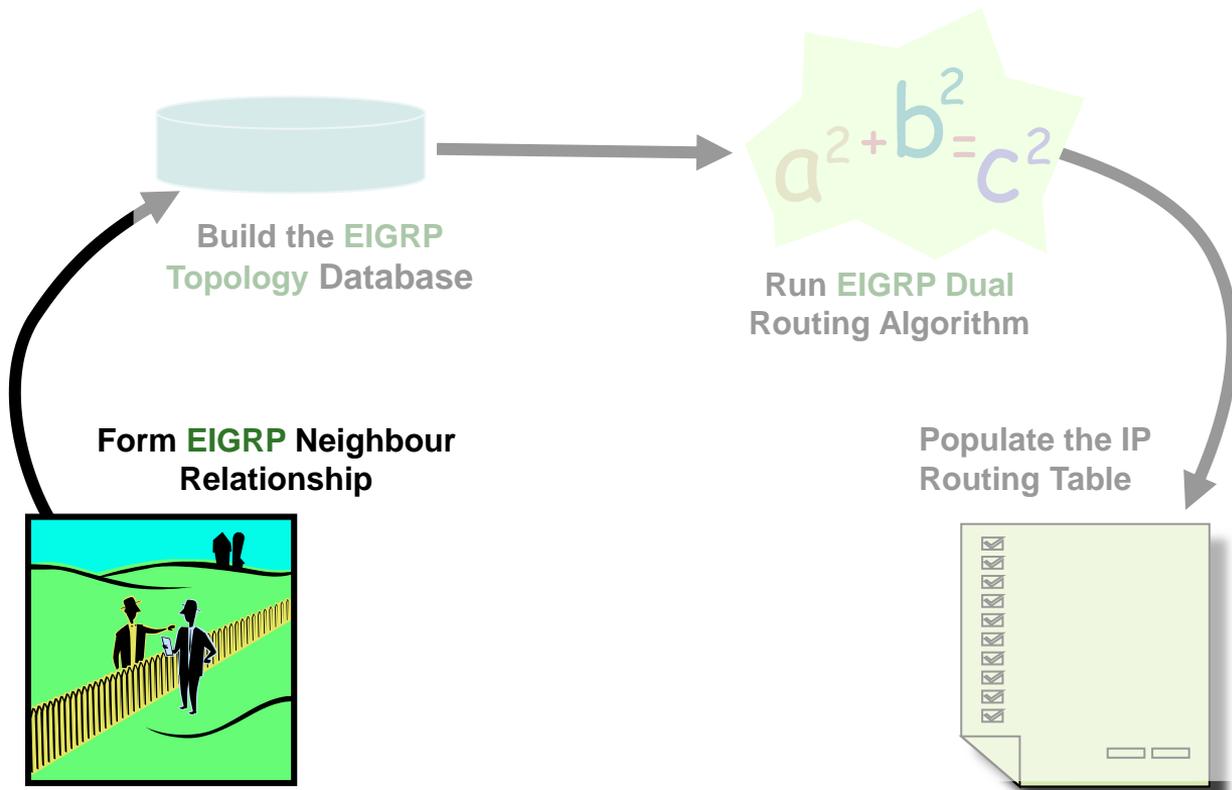


Troubleshooting EIGRP

Lifecycle View of the EIGRP Routing Process



Lifecycle View of the EIGRP Routing Process



EIGRP Neighbour Process

- Hello process used for neighbour discovery and maintenance

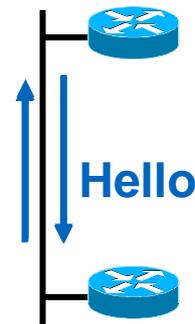
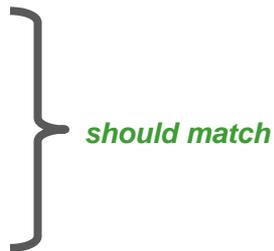
Hello Types

K-value

AS number

IP subnet

Primary/secondary IP address



- Multicast hellos (by default)
224.0.0.10 (0100.5e00.000a)
- Hello Interval
60 seconds for low-speed NBMA
5 seconds for all other interfaces

EIGRP Neighbour Process

```
RTRA#show ip eigrp neighbors
IP-EIGRP neighbors for process 1
H   Address      Interface Hold Uptime  SRTT  RTO  Q   Seq
                (sec)          (ms)                Cnt   Num
 2   10.1.1.2     Et0          12    6d16h  20   200  0   233
 1   10.1.4.3     Et1          13    2w2d  87   522  0   452
 0   10.1.4.2     Et1          10    2w2d  85   510  0    3
```

Seconds remaining before declaring neighbour down

How long since the last time neighbour was discovered

How long it takes for this neighbour to respond to reliable packets

How long we'll wait before retransmitting if no acknowledgement

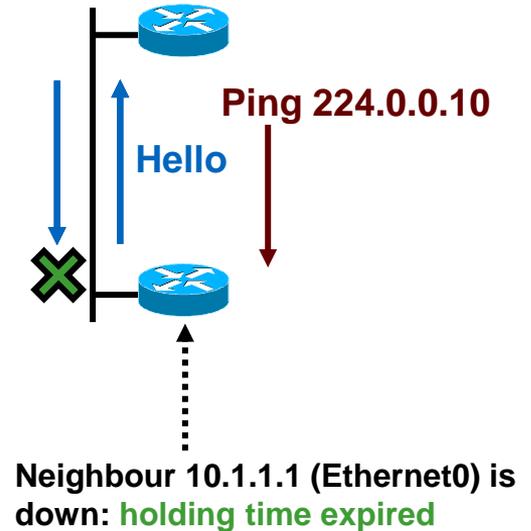
* Normal to have 5000 during the startup

Number of EIGRP packets unacknowledged by neighbour

EIGRP Neighbour Problems

Hold Time Expired

- The hold time expires when an **EIGRP packet** is not received during hold time
- Ping the multicast Address (224.0.0.10) from the Other Router
- Other checks:
 - Access-lists
 - Debug EIGRP packet hello
 - Ping neighbour with small & large packets
 - Interface errors
 - Sniffer Captures – external or internal
 - Configure neighbour statements



EIGRP Neighbour Problems

Manual Changes

- Manual configuration changes can also reset EIGRP neighbours:
 - Summary changes (manual and auto)
 - Route filter changes

- Graceful restart to resynchronise neighbour relationships:
 - `clear ip eigrp neighbor <address> soft`

EIGRP Neighbour Log Messages

eigrp log-neighbor-changes must be enabled

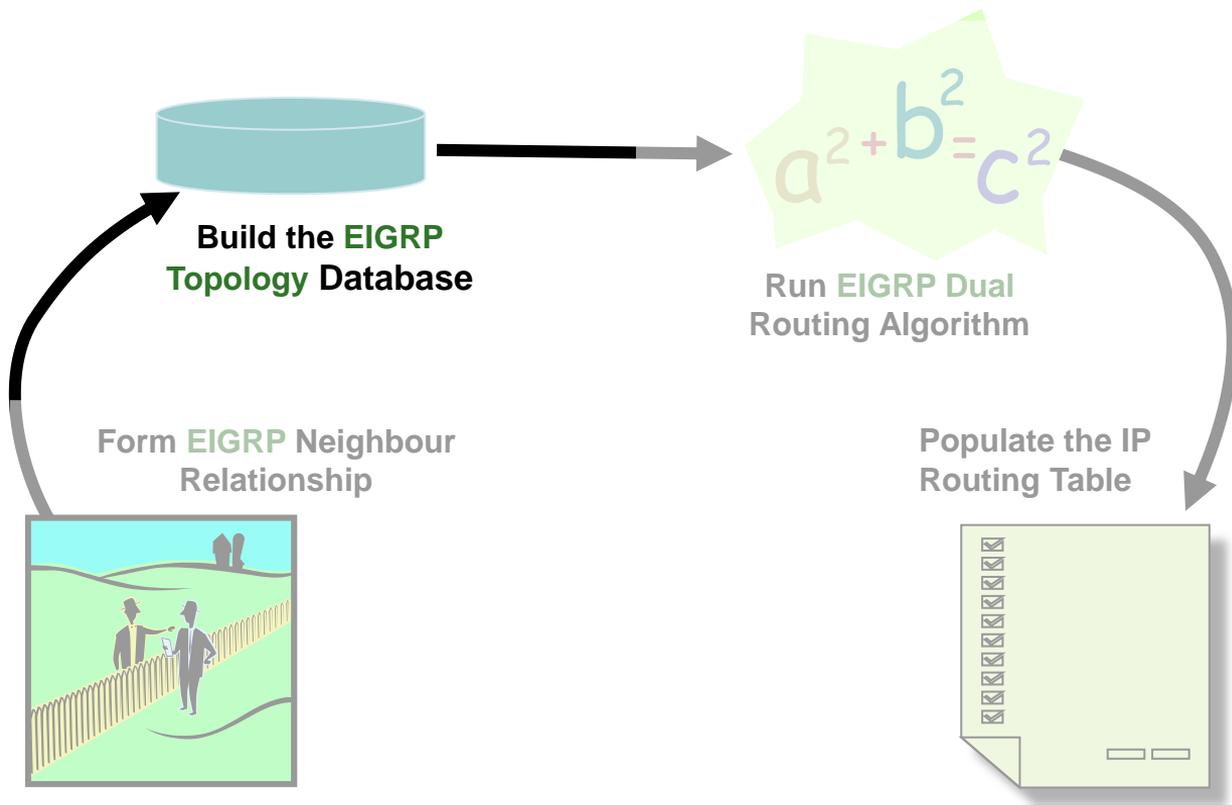
Example logs

Neighbor 10.1.1.2 (Ethernet0) is down: **peer restarted**
Neighbor 10.1.1.2 (Ethernet0) is down: **holding time expired**
Neighbor 10.1.1.2 (Ethernet0) is down: **retry limit exceeded**
Neighbor 10.1.1.2 (Ethernet0) is down: **route filter changed (old)**
Neighbor 10.1.1.2 (Ethernet0) is down: **K-value mismatch**
Neighbor 10.1.1.2 (Ethernet0) is down: **manually cleared**
Neighbor 10.1.1.2 (Ethernet0) is down: **Interface Goodbye received**

Neighbor 10.1.1.2 (Ethernet0) is resync: **route configuration changed**
Neighbor 10.1.1.2 (Ethernet0) is resync: **manually cleared**
Neighbor 10.1.1.2 (Ethernet0) is resync: **peer graceful-restart**

There are others, but not seen very often...

Lifecycle View of the EIGRP Routing Process



EIGRP Topology Exchange

```
RouterA#debug ip eigrp
```

```
IP-EIGRP Route Events debugging is on
```

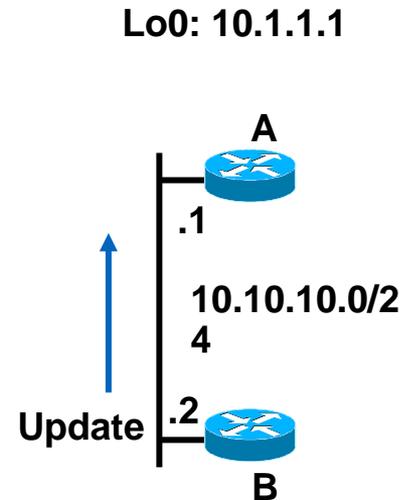
```
RouterA#debug ip eigrp 10 10.2.2.0 255.255.255.0
```

```
IP-EIGRP AS Target Events debugging is on
```

```
02:13:13.765: IP-EIGRP(Default-IP-Routing-Table:10): Processing incoming UPDATE packet
```

```
02:13:13.765: IP-EIGRP(Default-IP-Routing-Table:10): Int 10.2.2.0/24 M 409600 – 256000
```

```
153600 SM 128256 - 256 128000
```



Lo0: 10.2.2.2

EIGRP Topology Exchange

```
RouterA#show ip eigrp events
```

```
Event information for AS 10:
```

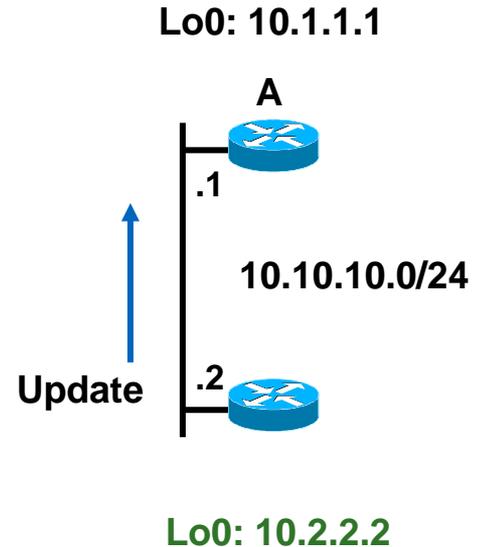
```
...
```

```
...
```

```
10 13:20:43.289 Rcv update met/succmet:  
      409600 128256
```

```
11 13:20:43.289 Rcv update dest/nh:  
      10.2.2.0/24 10.1.1.2
```

```
...
```



EIGRP Event Log

- Always running
- Separate event log for each AS
- Default 500 lines (configurable):

`eigrp event-log-size ##` (where ## are number of lines)

`size 0` will disable log

- Most recent events at top.
- Clearing the event log:

`clear ip eigrp event`

EIGRP Topology Exchange

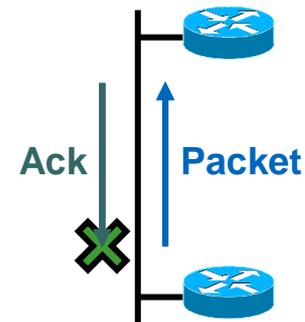
Retry Limit Exceeded

- Reliable packets: re-sent after Retransmit Time Out (Typically $RTO = 6 \times SRTT$)
Minimum 200 ms and maximum 5000 ms
- After 16 retransmissions: re-initialise the neighbour

```
RtrB#show ip eigrp neighbors
```

```
IP-EIGRP neighbors for process 1
```

H	Address	Interface	Hold (sec)	Uptime	SRTT (ms)	RTO	Q Cnt	Seq Num
1	10.1.102.2	Et0	14	00:00:15	0	5000	4	0

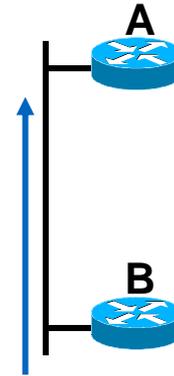


Neighbour 10.1.102.2 (Ethernet0) is down: **retry limit exceeded**

EIGRP Topology Exchange

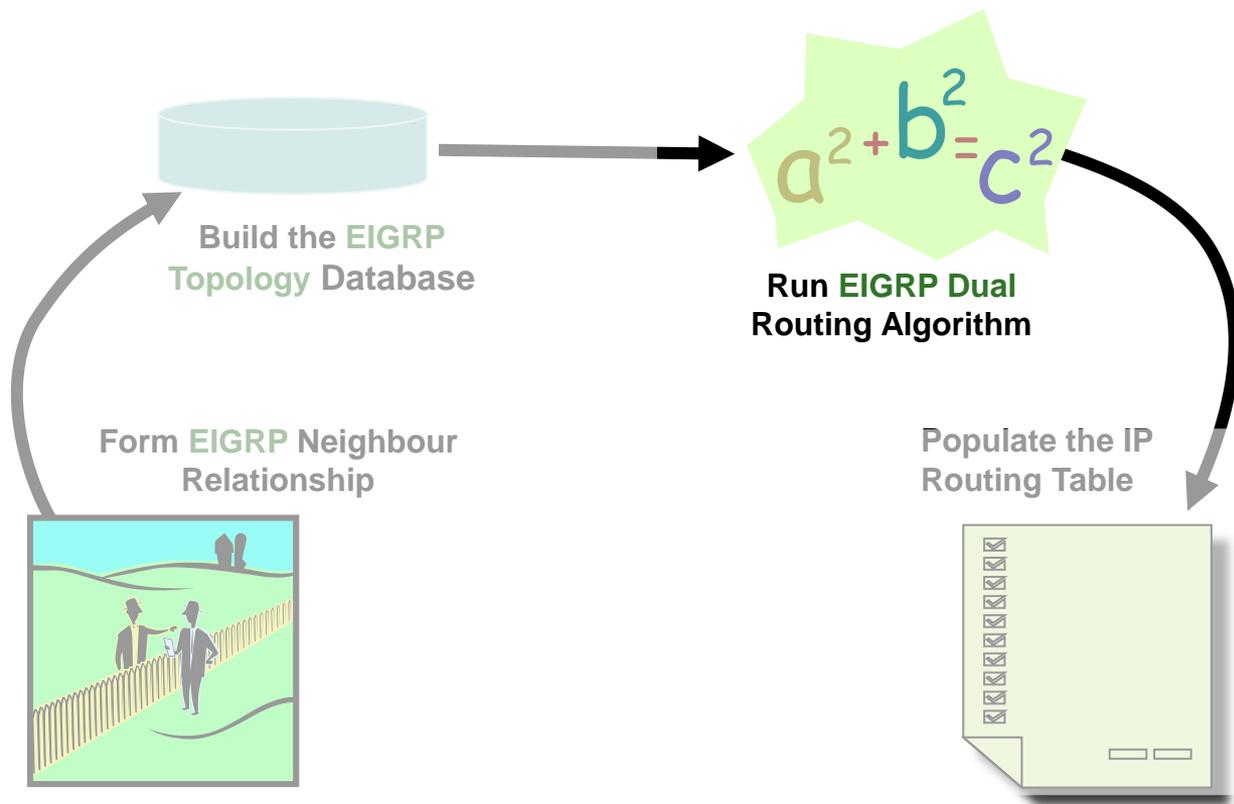
Retry Limit Exceeded - Troubleshoot and Causes

- Ping the neighbour's unicast address
 - Vary the packet size
 - Try large numbers of packets
- This Ping Can Be Issued from Either Neighbour
- Common causes
 - Mismatched MTU
 - Unidirectional link
 - Dirty link
 - Bad Hardware



```
RtB# ping
Protocol[ip]:
Target IP address: 10.1.1.1
Repeat count [5]: 100
Datagram Size: 1500
Timeout in seconds[2]:
Extended commands[n]: y
```

Lifecycle View of the EIGRP Routing Process



EIGRP DUAL Routing Algorithm Concepts

- Feasible Distance (FD): minimum distance (metric) to a destination network.
- Reported distance (RD): distance (metric) towards a destination advertised by upstream neighbour.
- A neighbour meets the feasibility condition (FC) if the RD is smaller than the FD of this router.

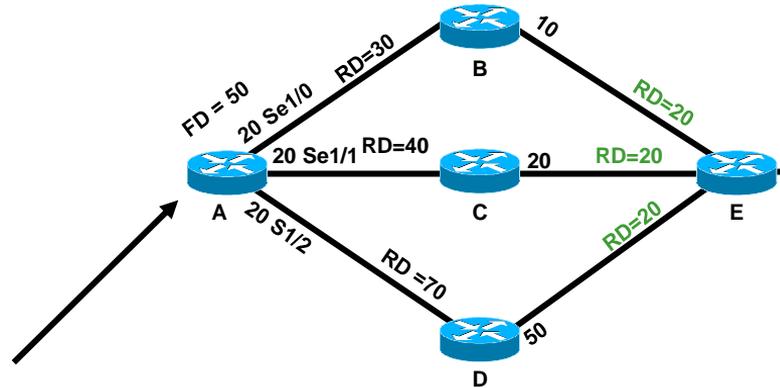
DUAL FD, RD, FS and Successor

A-B-E = 20+30 = 50

A-C-E = 20+40 = 60

A-D-E = 20+70 = 90

FD = 50



```
RtrA#show ip eigrp topology
IP-EIGRP Topology Table for AS(1)/ID(10.1.6.1)
..snip....
P 10.200.1.0/24, 1 successors, FD is 50
  via 10.1.1.2 (50 / 30), Serial1/0
  via 10.1.2.2 (60 / 40), Serial1/1
```

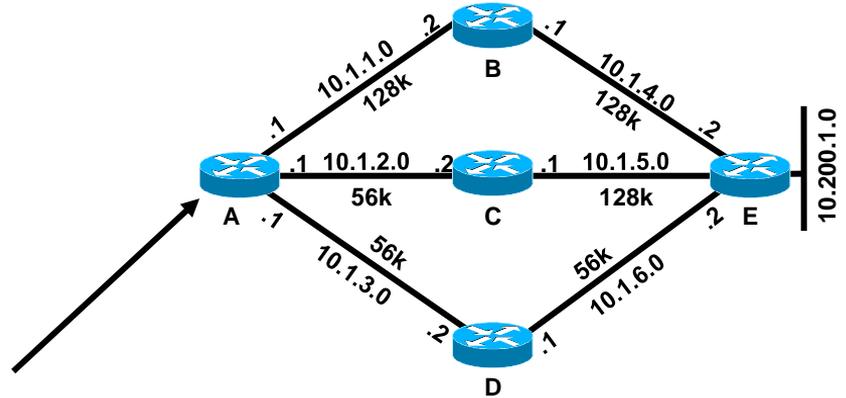
← Feasible Distance
← Successor
← Feasible Successor

↑
Computed
Distance

↑
Reported
Distance

Show IP EIGRP Topology All-Links

Show ip eigrp topology all-links displays a list of All neighbours



```
RtrA#show ip eigrp topology all-links
IP-EIGRP Topology Table for AS(1)/ID(10.1.6.1)
....snip....
P 10.200.1.0/24, 1 successors, FD is 21026560
  via 10.1.1.2 (21026560/20514560), Serial1/0
  via 10.1.2.2 (46740736/20514560), Serial1/1
  via 10.1.3.2 (46740736/46228736), Serial1/2
```

← Successor
← Feasible Successor
← Possible Successor

Reported Distance

Show IP EIGRP Topology Summary

Total number of routes in the local topology table

Number of queries this router is waiting on replies for

Internal data structures

```
RtrA#sh ip eigrp topology sum
IP-EIGRP Topology Table for AS(200)/ID(40.80.0.17)
Head serial 1, next serial 1526
589 routes, 0 pending replies, 0 dummies
IP-EIGRP(0) enabled on 12 interfaces, neighbors present on 4 interfaces
Quiescent interfaces: Po3 Po6 Po2 Gi8/5
```

Interfaces with no outstanding packets to be sent or acknowledged

EIGRP DUAL Routing Algorithm

- Finite-State-Machine

 - Track all routes

 - Select a Successor

 - Remember Feasible Successors

 - If Successor lost

 - Use Feasible Successor

 - If no Feasible Successor exists

 - Query Neighbours and Recompute new Successor

EIGRP DUAL Routing Algorithm

Active Process

- Normal state of a route is **passive**
- Going **active** is the normal during convergence
 - Successor is lost and no feasible successors
 - It means sending Queries to neighbours
 - SIA timer is 180 seconds

EIGRP DUAL Routing Algorithm

Stuck in Active

%DUAL-3-SIA: Route 10.64.5.0 255.255.255.192
stuck-in-active state in IP-EIGRP 100. Cleaning up

- Two (probably) unrelated causes of the problem **stuck** and **active**
 - Need to troubleshoot both parts
 - Cause of active often easier to find
 - Cause of stuck more important to find

EIGRP DUAL Routing Algorithm

Troubleshooting the Stuck Part of SIAs

- Show ip eigrp topology active

Useful only during problem

```
rtrA#show ip eigrp topology active
IP-EIGRP Topology Table for AS(1)/ID(20.1.1.1)
A 10.10.10.0/24, 1 successors, FD is Inaccessible
  1 replies, active 00:01:17, query-origin: Local origin
    via Connected (Infinity/Infinity), Ethernet1/0
  Remaining replies:
    via 10.1.1.2, r, Ethernet0/0
```

↑
rtrA is waiting on reply from
10.1.1.2

EIGRP DUAL Routing Algorithm

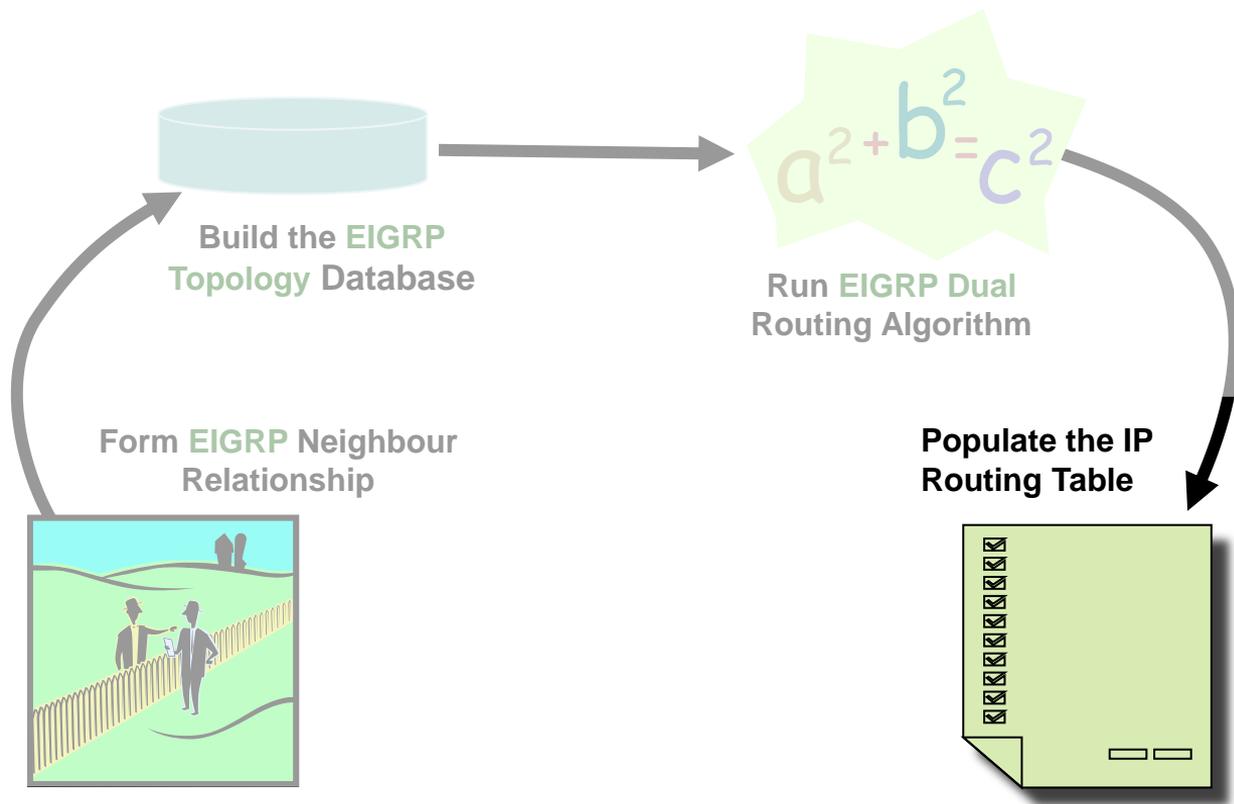
Likely Causes for Stuck-in-Active

- Bad or congested links
- Query range is too long (Reduce Scope or Summarise)
- Excessive redundancy
- Overloaded router (high CPU)
- Router memory shortage
- Dead-lock queries (rare)
- Software defects (seldom)

Minimising SIA Routes

- Decrease query scope
 - Summarisation (manual or auto)
 - Distribute-lists
 - Define remote routers as stubs
- New Cisco IOS handle this condition better

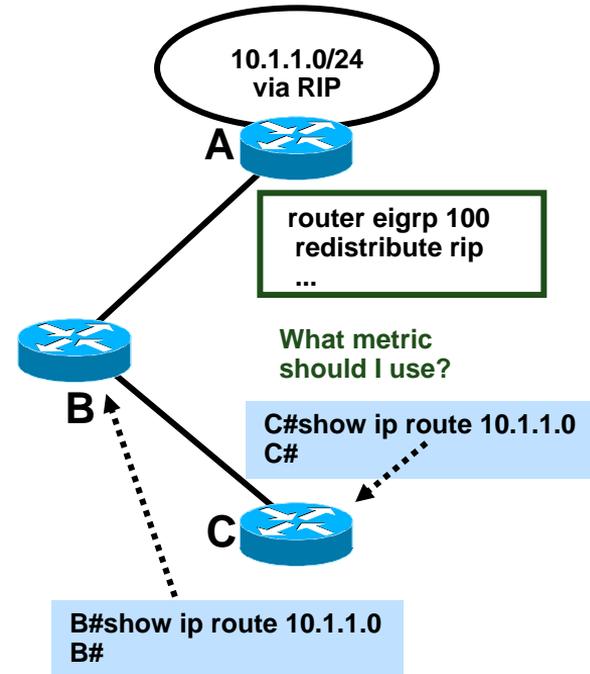
Lifecycle View of the EIGRP Routing Process



EIGRP Routing Table

Problem with External Routes

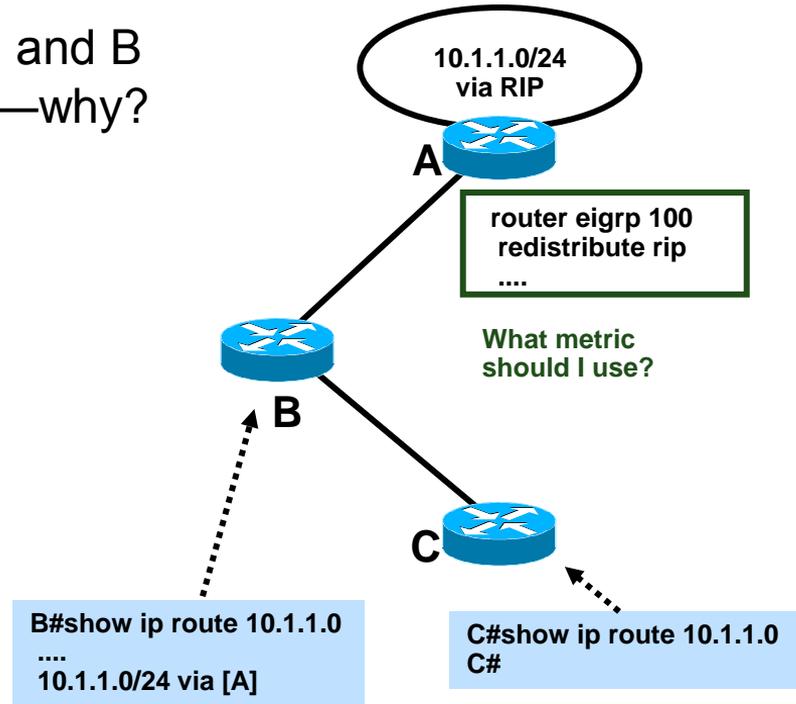
- EIGRP not installing external routes
- The first thing to check is:
 - default metric configured
 - metric tied to the redistribution statement
 - default-metric
 - redistribute <metric>



EIGRP Routing Table

Problem with External Routes

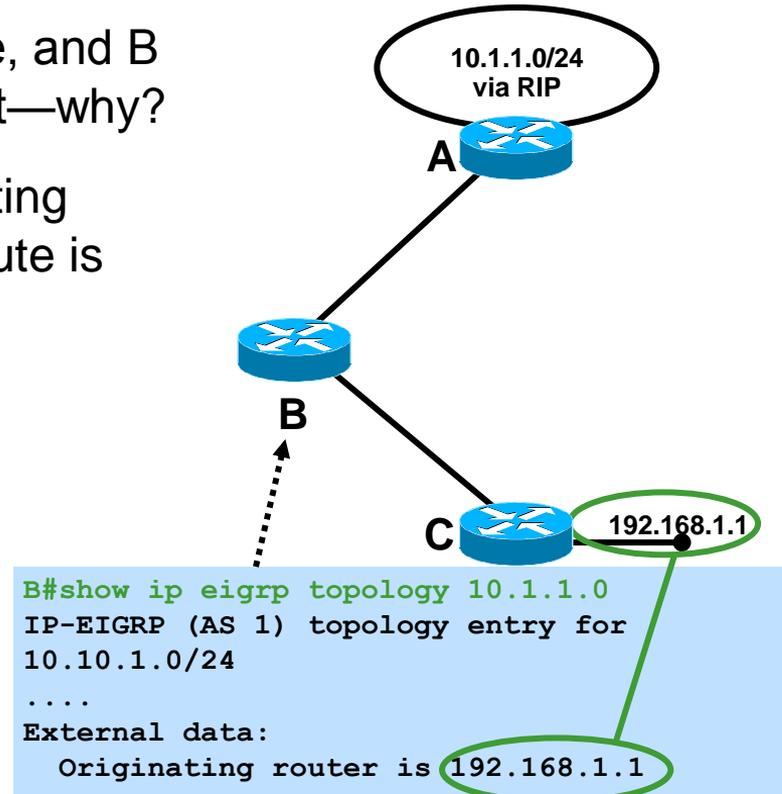
- We've resolved the metric issue, and B picked the route up, but C didn't—why?



EIGRP Routing Table

Problem with External Routes

- We've resolved the metric issue, and B picked the route up, but C didn't—why?
- B's topology table show originating router ID field in the external route is set to 192.168.1.1
- But, that's Router C's loopback address!



EIGRP Routing Table

Problem with External Routes

- Router ID is listed in the output of `show ip eigrp topology`:

```
router-1# show ip eigrp topology
IP-EIGRP Topology Table for AS(7)/ID(192.168.1.1)
....
```

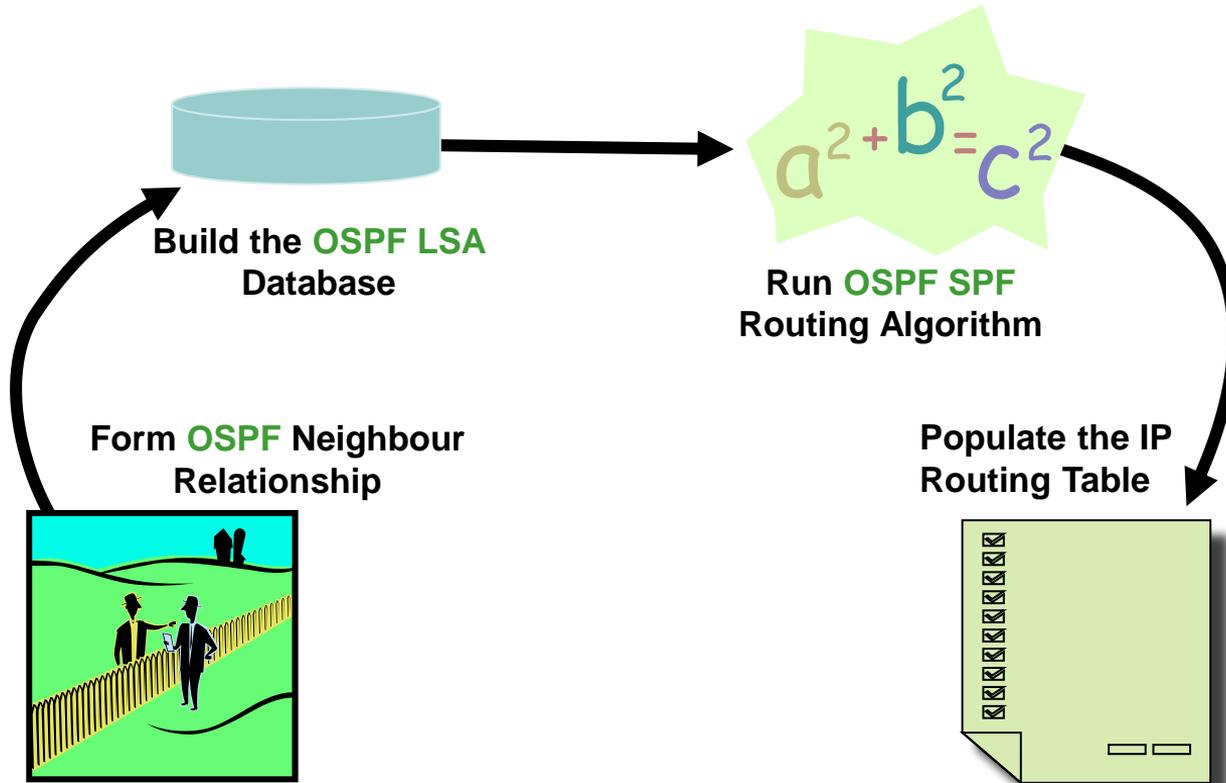
- Problem indicated in event log:

```
1 02:30:18.591 Ignored route, metric: 192.168.1.0 2297856
2 02:30:18.591 Ignored route, neighbor info: 10.1.1.0/24 Serial0/3
3 02:30:18.591 Ignored route, dup router: 192.168.1.1
```

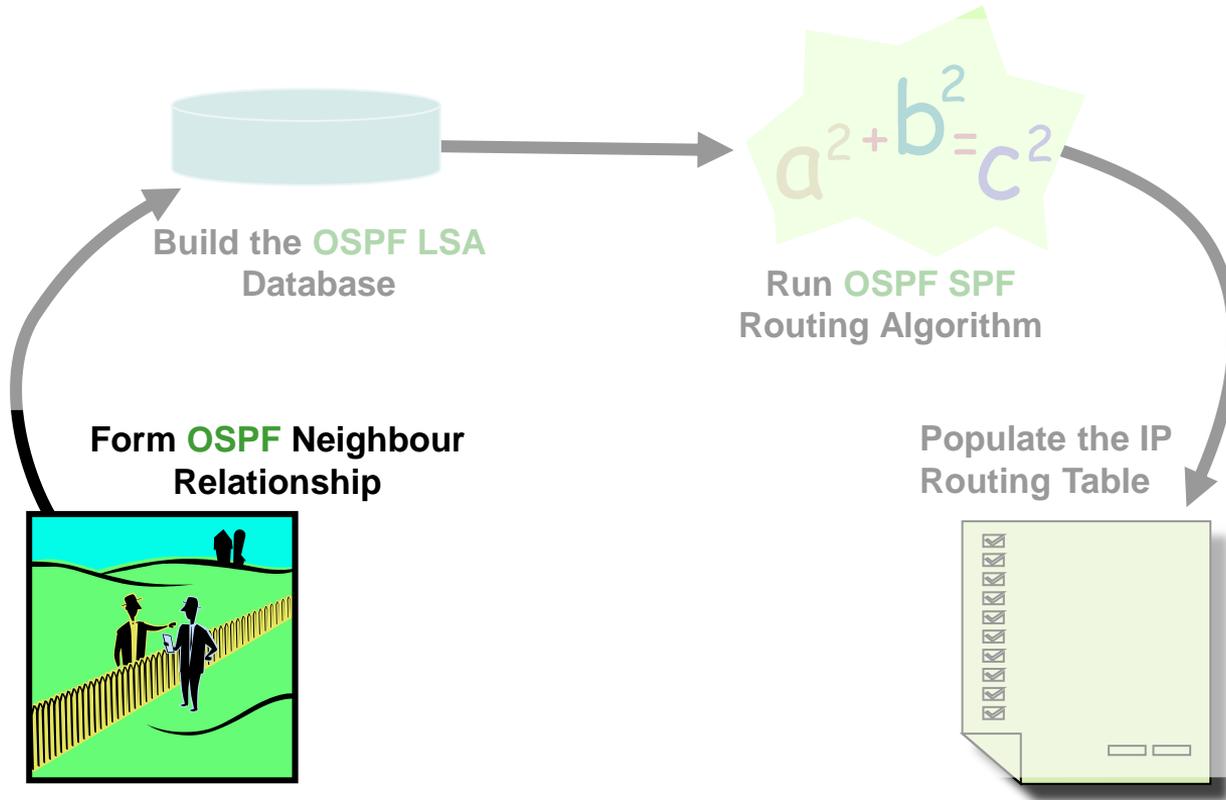


Lifecycle View of the OSPF Routing Process

Lifecycle View of the OSPF Routing Process

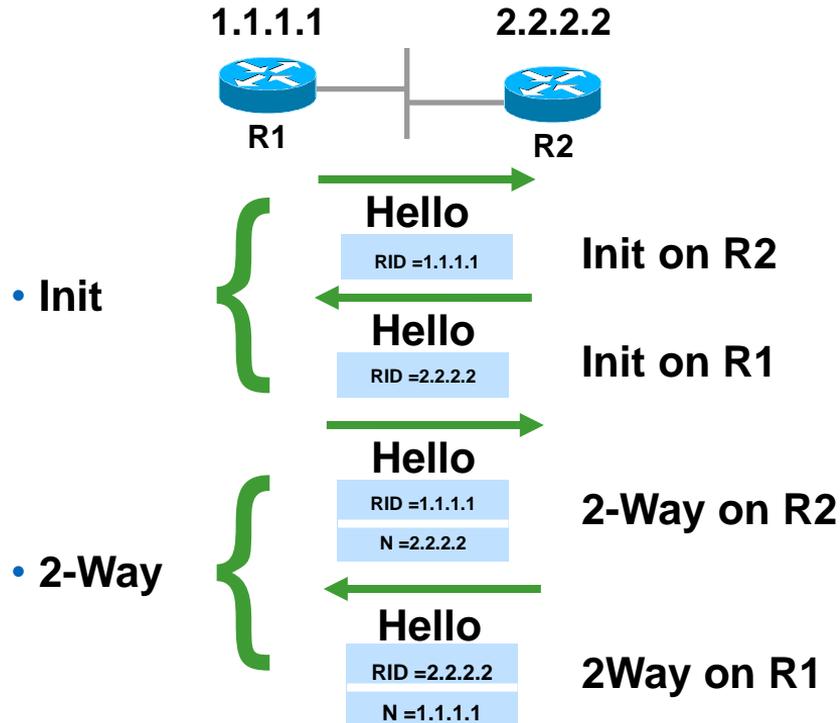


Lifecycle View of the OSPF Routing Process



OSPF Neighbour Adjacency Process

Init, 2-Way



OSPF Neighbour Adjacency Process

Init, 2-Way



```
R1#debug ip ospf adj
```

OSPF adjacency events debugging is on

```
20:23:45.539: OSPF: 2 Way Communication to 2.2.2.2 on Ethernet0/0, state 2WAY
```

```
20:23:45.539: OSPF: Neighbor change Event on interface Ethernet0/0
```

```
R1#debug ip ospf events
```

OSPF events debugging is on

```
13:51:14.146: OSPF: Rcv hello from 2.2.2.2 area 0 from Ethernet0/0 10.10.100.2
```

```
13:51:14.146: OSPF: Send immediate hello to nbr 2.2.2.2, src address 10.10.100.2, on Ethernet0/0
```

OSPF Neighbour Adjacency

Stuck in Init

Possible Reasons for Stuck in Init

- Access-list
- One side multicast capabilities
- OSPF authentication
- Dialer map or frame-relay map missing keyword 'broadcast'
- Link-Local Signalling capability not compatible

Troubleshooting Steps

- Debug ip ospf adj
- Debug ip ospf events
- Debug ip packet 101 detail – Use access list to filter
- Ping 224.0.0.5 from both sides

OSPF Neighbour Adjacency

Stuck in 2-WAY

Possible Reasons for Stuck in 2-WAY

- Normal in broadcast network type
- If all the routers are configured with priority equal to “0”

OSPF Neighbour Adjacency

Common Issues

- Mismatched Subnet Mask
- Mismatched Hello/Dead Interval
- Mismatched Authentication Key
- Mismatched Area Id
- Mismatched Transit/Stub/NSSA Option

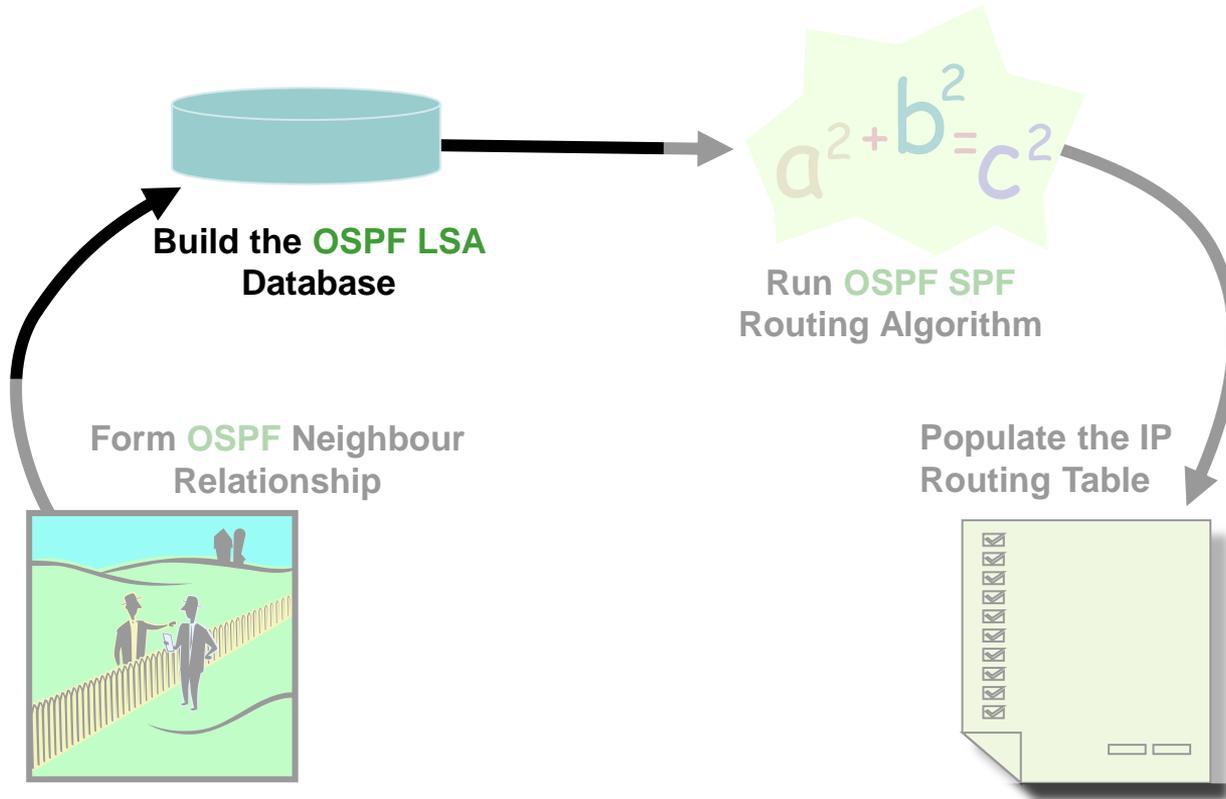
- OSPF Graceful Shutdown; Router or Interface mode
`shutdown / ip ospf shutdown`

- OSPF TTL Security Check
`ttl-security all-interfaces [hops hop-count]`
`ip ospf ttl-security [hops hop-count | disable]`

Interface Scoped Debugging

- limiting the OSPF debug to selected interfaces
- Example:
 - debug condition interface Ethernet 0/0
 - debug condition interface Ethernet 1/0
 - debug ip ospf hello
 - debug ip ospf adjacency

Lifecycle View of the OSPF Routing Process



OSPF LSA Database

EXSTART / EXCHANGE



- EXSTART

MTU = 1500
Flag = 0x7
Seq = 0x2499



- EXCHANGE

MTU = 1500
Flag = 0x3
Seq = 0x837

OSPF LSA Database

EXSTART



```
R1#debug ip ospf adj
```

```
23:42:08.259: OSPF: Send DBD to 2.2.2.2 on Ethernet0/0 seq u opt 0x52 flag 0x7 len 32
```

```
23:42:08.339: OSPF: Rcv DBD from 2.2.2.2 on Ethernet0/0 seq 0x836 opt 0x52 flag 0x7 len 32 mtu 1500 state EXSTART
```

```
23:42:08.339: OSPF: NBR Negotiation Done. We are the SLAVE
```

```
R2#debug ip ospf adj
```

```
23:42:08.423: OSPF: Send DBD to 1.1.1.1 on Ethernet0/0 seq 0x836 opt 0x52 flag 0x7 len 32
```

```
23:42:08.423: OSPF: First DBD and we are not SLAVE
```

```
23:42:08.511: OSPF: Rcv DBD from 1.1.1.1 on Ethernet0/0 seq 0x836 opt 0x52 flag 0x2 len 52 mtu 1500 state EXSTART
```

```
23:42:08.511: OSPF: NBR Negotiation Done. We are the MASTER
```

OSPF LSA Database

The Flag Field

OSPF: Send DBD to 1.1.1.1 on Ethernet0/0 seq 0x836 opt 0x52 flag 0x7 len 32

Flag 0x7 --> 111 means I(Initial) = 1, M = 1(More), MS = 1(Master)

Flag 0x6 --> 110 not possible

Flag 0x5 --> 101 not possible

Flag 0x4 --> 100 not possible

Flag 0x3 --> 011 means master has more data to send

Flag 0x2 --> 010 means slave has more data to send

Flag 0x1 --> 001 means master has no more data left to send

Flag 0x0 --> 000 means slave has no more data left to send

0	0	0	0	0	I	M	MS
---	---	---	---	---	---	---	----

OSPF Neighbour Adjacency

The Options Field

Normal area: OSPF: Send DBD to 141.108.97.1 on Serial0 seq 0xBC4 **opt 0x2** flag 0x3 len 492

E bit is 1, Allow externals, option: 0x2(HEX) = 0000010(Bin)

Stub area: OSPF: Send DBD to 141.108.97.1 on Serial0 seq 0x1866 **opt 0x0** flag 0x3 len 372

E bit is 0, no external allowed, options: 0x0 = 00000000

NSSA: OSPF: Send DBD to 141.108.97.1 on Serial0 seq 0x118 **opt 0x8** flag 0x3 len 372

N/P bit is on, options: 0x8 = 00001000

DC: OSPF: Send DBD to 141.108.97.1 on Serial0 seq 0x1A1E **opt 0x20** flag 0x3 len 392

DC bit is negotiated, options: 0x20 = 00100000



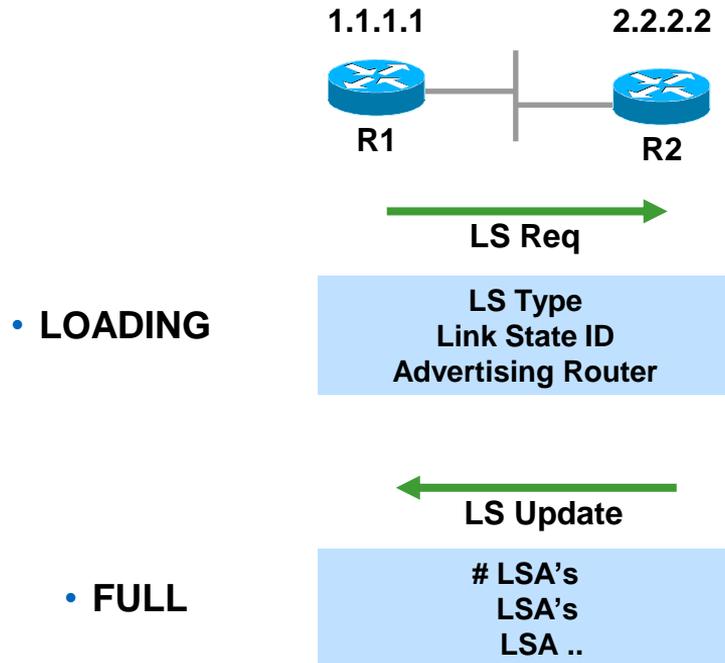
OSPF LSA Database

Stuck in EXSTART or EXCHANGE - Common Causes

- MTU mismatch— Stuck in EXCHANGE/EXSTART
- Neighbour RID is same as ours— Stuck in EXSTART
- Unicast is broken—Stuck in EXCHANGE
 - a. Wrong VC/DLCI mapping
 - b. MTU problem
 - c. Access-list blocking unicast

OSPF LSA Database

Loading and Full



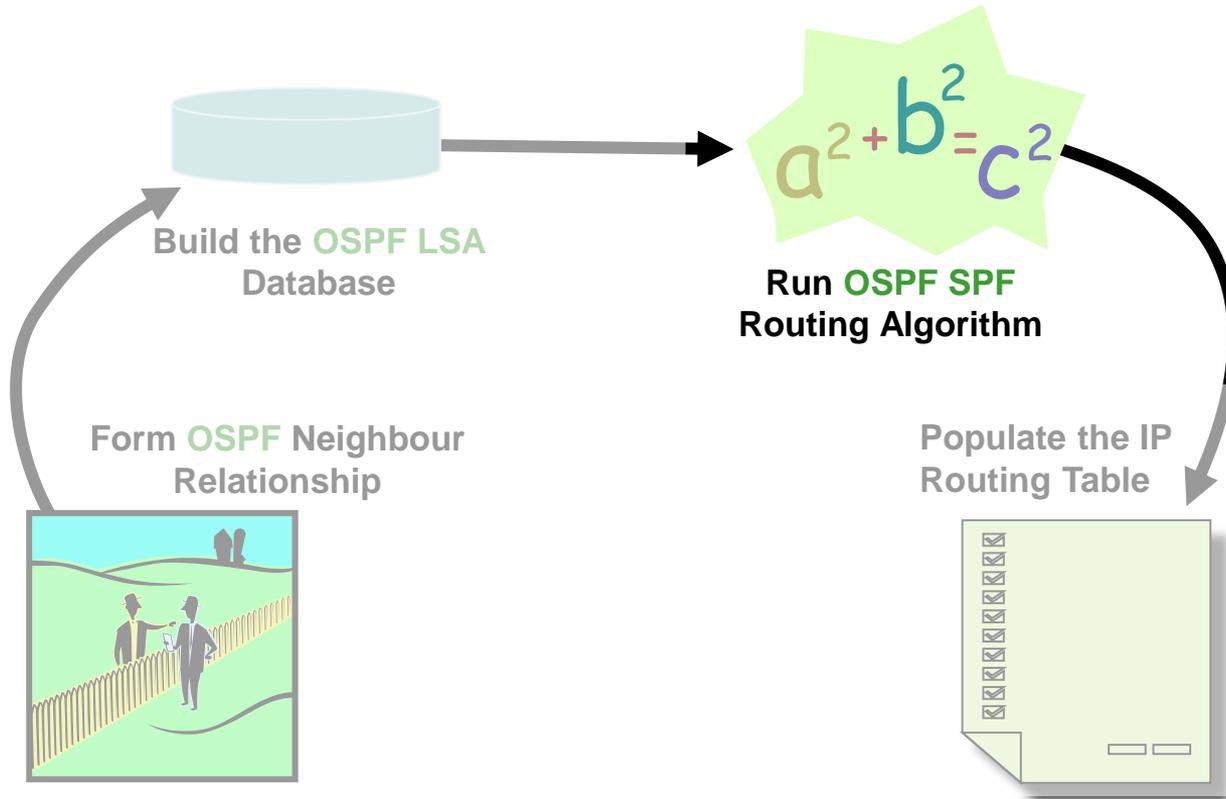
OSPF LSA Database

Stuck in LOADING

- **LS request is being made and neighbour is sending bad packet or memory corrupt**
 - a. Do `show ip ospf bad-checksum`
 - b. Show log will show OSPF-4-BADLSATYPE message
- **LS request is being made and neighbour is ignoring the request**
- **debug ip ospf flood**

Enhancement to allow an access-list filter and a “detail” option

Lifecycle View of the OSPF Routing Process



OSPF LSA Database - LSA Type Review

Type	LSA
1	Router
2	Network
3	Summary Network
4	Summary ASBR
5	External
6	Group Membership
7	NSSA
8	External Attributes
9-11	Opaque

Router LSA of R3 for Area 0

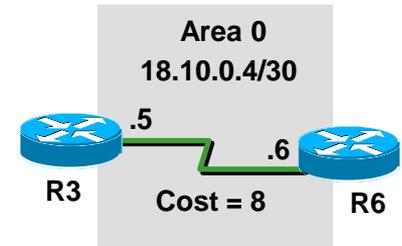
```
R3#show ip ospf database router 3.3.3.3
```

Router Link States (Area 0)

```
LS age = 0
Options = (No TOS-capability, DC)
LS type = Router Links           This is a Type 1 LSA
Link State ID = 3.3.3.3
Advertising Router = 3.3.3.3
It is an area border router bit B = 1

# links = 2
Link ID = 6.6.6.6                Router id of the neighbor
Link Data = 18.10.0.5            IP interface address of the router
Type = 1                         This is a point-to-point link
# TOS metrics = 0
metric = 8

Link ID = 18.10.0.4              IP subnet address
Link Data = 255.255.255.252      Subnet mask
Type = 3                         This is a stub link
# TOS metrics = 0
metric = 8
```



Router LSA of R3 for Area 1

Router Link States (Area 1)

LS age = 253

Options = (No TOS-capability, DC)

LS type = Router Links This is a Type 1 LSA

Link State ID = 3.3.3.3

Advertising Router = 3.3.3.3 Router ID of R3

It is an area border router bit B = 1

links = 2

Link ID = 192.1.1.4

IP address of the DR

Link Data = 192.1.1.3

Interface address of this router

Type = 2

This is a transit network

TOS metrics = 0

metric = 1

Cost to reach the interface

Link ID = 192.1.4.0

IP network number

Link Data = 255.255.255.0

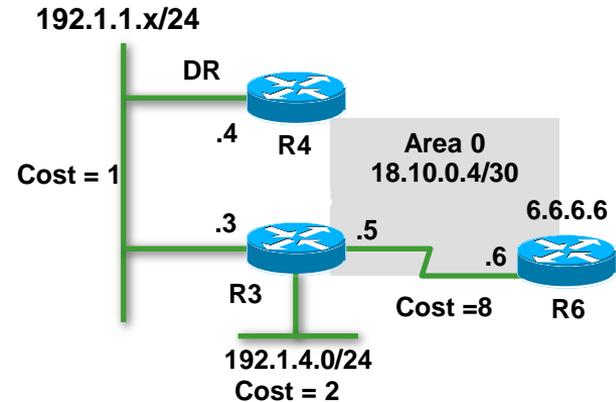
Subnet mask of the interface

Type = 3

Stub network

TOS metrics = 0

metric = 2



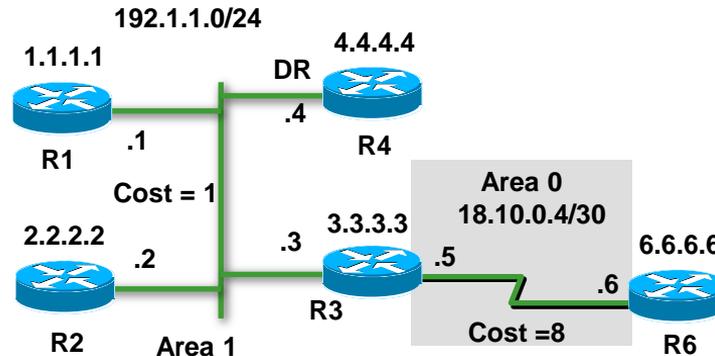
Link Details

Type	Description	Link ID	Link Data
1	Point-to-Point Numbered	Neighbors' RID	Interface IP Address
1	Point-to-Point Unnumbered	Neighbors' RID	MIB-II Ifindex Value
2	Transit	IP Address of the DR	Interface IP Address
3	Stub	IP Network Number	Subnet Mask
4	Virtual Link	Neighbors' RID	Interface IP Address

Network LSA for 192.1.1.0

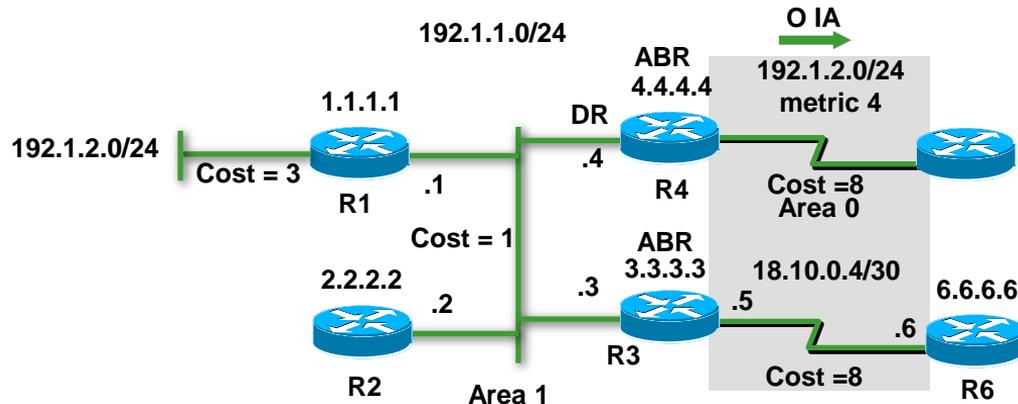
```
R3#show ip ospf database network 192.1.1.4
Network Link States (Area 1)

LS age = 0
Options = (No TOS-capability, DC)
LS type = Network Links                Type 2 LSA
Link State ID = 192.1.1.4              IP interface address of DR
Advertising Router = 4.4.4.4          RID of DR
Network Mask = 255.255.255.0
Attached Router = 4.4.4.4
Attached Router = 3.3.3.3RID of attached routers FULL with DR
Attached Router = 2.2.2.2
Attached Router = 1.1.1.1
```



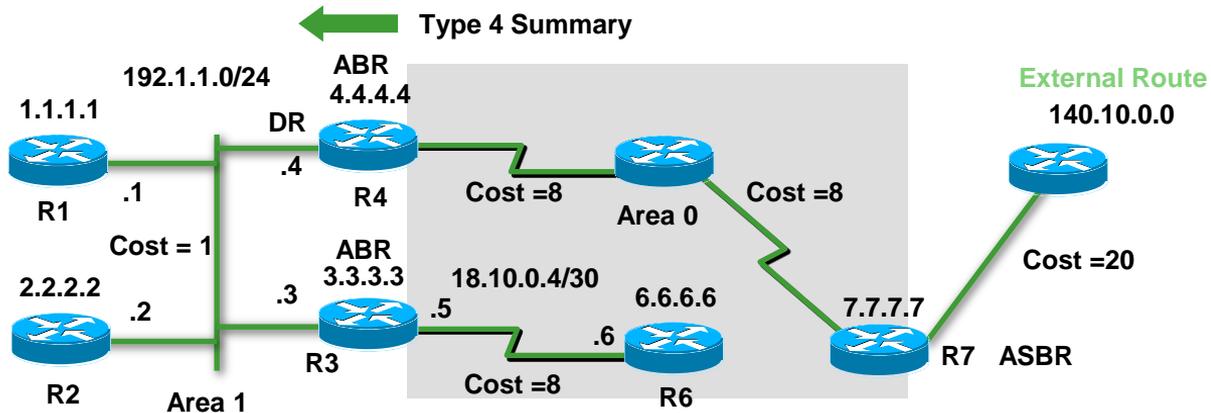
Type 3 Details

```
R4#show ip ospf database summary 192.1.2.0
      Summary Net Link States (Area 0)
LS age = 0
Options = (No TOS-capability, DC, Upward)
LS type = Network Links                Type 3 LSA
Link State ID = 192.1.2.0              Summary IP network number
Advertising Router = 4.4.4.4          RID of ABR
Network Mask = 255.255.255.0
metric = 4
```



Type 4 Details

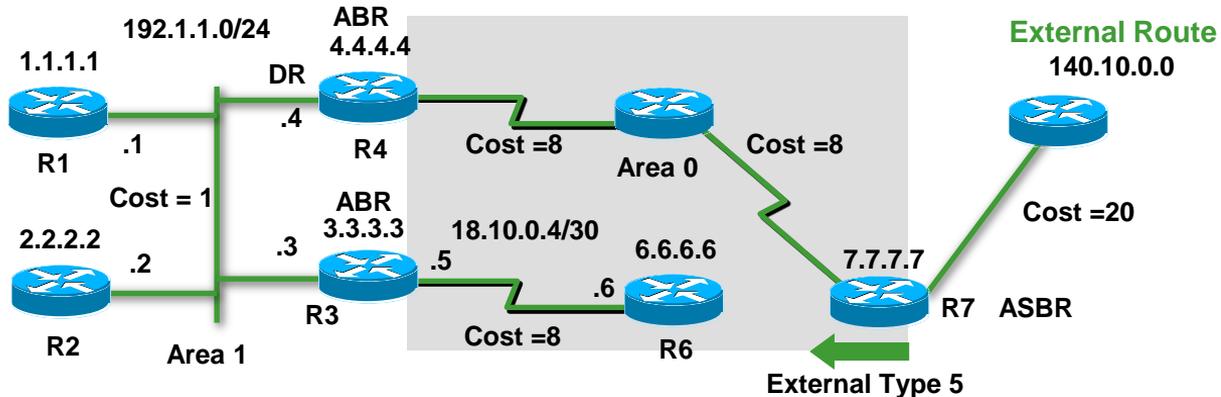
```
R4#show ip ospf database asbr-summary 7.7.7.7
      Summary ASB Link States (Area 1)
LS age = 0
Options = (No TOS-capability, DC, Upward )
LS type = Summary Links (ASBR)
Link State ID = 7.7.7.7           RID of ASBR
Advertising Router = 4.4.4.4     RID of ABR
Network Mask = 0.0.0.0
metric = 16
```



Type 5 Details

```
R4#show ip ospf database external 140.10.0.0
Routing Bit Set on this LSA
LS age = 0
Options = (No TOS-capability, DC)
LS type = AS External Link
Link State ID = 140.10.0.0
Advertising Router = 7.7.7.7
Network Mask = 255.255.0.0
Metric Type: 2
metric = 20
Forwarding address = 0.0.0.0
```

Type 5
IP network number
Router ID of R7
Bit E = 1 -> O E2 (Default)
Metric is 20 in all redistributed E2 routes
Traffic should be forwarded to the ASBR



Forwarding Address (Type 5)

- When will it set to non-zero?

OSPF is enabled on the ASBR's next hop interface

The ASBR's next hop interface is non-passive to OSPF

The ASBR's next hop interface network-type is NOT p2p or p2mp

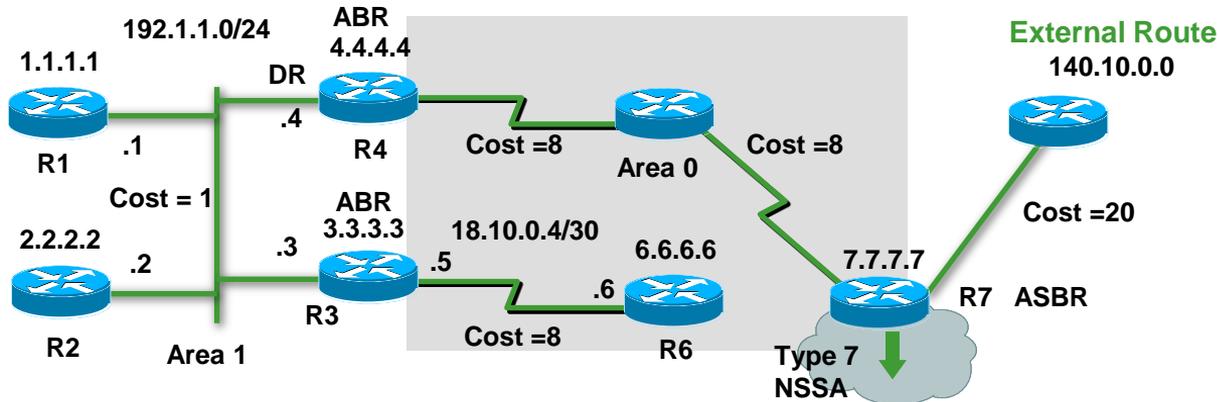
The ASBR's next hop interface address falls into OSPF network range

Type 7 Details

```
R7#show ip ospf database nssa-external 140.10.0.0
LS age = 0
Options = (No TOS-capability, No Type 7/5 translation, DC)
LS type = AS External Link           Type 7 LSA
Link State ID = 140.10.0.0           IP network number
Advertising Router = 7.7.7.7         Router ID of R7 (NSSA ASBR)
Network Mask = 255.255.0.0

P = 0 -> This router is an NSSA ASBR+ ABR
P = 1 -> This router is an NSSA ASBR

metric = 20
Forwarding address = 0.0.0.0         Traffic should be forwarded to the ASBR
```



OSPF SPF Algorithm

SPF Running Constantly

```
R3#show ip ospf database
```

```
OSPF Router with ID (3.3.3.3) (Process ID 1)
```

```
Router Link States (Area 0)
```

Link ID	ADV Router	Age	Seq#	Checksum	Link count
3.3.3.3	3.3.3.3	106	0x80000009	0xC3F1	3

```
Summary Net Link States (Area 0)
```

Link ID	ADV Router	Age	Seq#	Checksum
18.10.0.0	7.7.7.7	3 (DNA)	0x80000008	0x3DC2
18.10.0.0	8.8.8.8	1396	0x80000004	0x27D8

```
Router Link States (Area 1)
```

Link ID	ADV Router	Age	Seq#	Checksum	Link count
1.1.1.1	1.1.1.1	2	0x80000016	0xE6CD	2
. . .					

OSPF SPF Algorithm

SPF Running Constantly

```
R3#show ip ospf statistics <details>
```

```
Area 0: SPF algorithm executed 42 times
```

```
Area 1: SPF algorithm executed 38 times
```

```
SPF calculation time
```

Delta T	Intra	D-Intra	Summ	D-Summ	Ext	D-Ext	Total	Reason
00:22:00	0	0	0	0	0	0	0	R, N, SN,
00:21:44	0	0	4	0	0	0	4	R, SN, X
00:21:34	0	0	4	0	0	0	4	R, SN, X
00:21:24	0	0	0	4	0	0	4	R, SN, X
00:20:44	0	0	4	0	0	0	4	R, SN, X
00:20:34	0	0	0	0	0	0	0	X
00:00:17	4	0	0	0	0	0	4	R, N, SN, SA

```
. . .
```

```
R=Router LSA; N=NetworkLSA; SN=Summary Network LSA; SA=Summary ASBR LSA; X=External LSA
```

OSPF SPF Algorithm

SPF Running Constantly - Debugging Commands

```
R3#debug ip ospf monitor
```

```
OSPF: Schedule SPF in area 1
```

```
Change in LS ID 1.1.1.1, LSA type R,
```

```
OSPF: schedule SPF: spf_time 0ms wait_interval 861421816s
```

```
OSPF: Begin SPF at 0x33585480ms, process time 752ms
```

```
spf_time 0ms, wait_interval 861421816s
```

```
OSPF: End SPF at 0x33585488ms, Total elapsed time 8ms
```

```
R2#debug ip ospf spf ?
```

```
external OSPF spf external-route
```

```
inter OSPF spf inter-route
```

```
intra OSPF spf intra-route
```

```
statistic OSPF spf statistics
```

```
<cr>
```

```
R2#
```

```
R2#debug ip ospf spf external ?
```

```
<1-99> Access list
```

```
<1300-1999> Access list (expanded range)
```

```
<cr>
```

```
R2#
```

OSPF Traffic Statistics

show/clear ip ospf [process-id] traffic [interface]

Global section

Per-process sections

OSPF queues

Interface details

Per-process summary

```
router2#show ip ospf traffic
```

```
OSPF statistics:
```

```
Rcvd: 29 total, 0 checksum errors
```

```
  7 Hello, 8 database desc, 2 link state req
```

```
  8 link state updates, 4 link state acks
```

```
Sent: 29 total
```

```
  8 Hello, 6 database desc, 2 link state req
```

```
  8 link state updates, 5 link state acks
```

OSPF Traffic Statistics

Per Process Summary

Per process filter:

```
show ip ospf <process_id> traffic
```

Summary traffic statistics for process ID 1:

OSPF packets received/sent

Type	Packets	Bytes
RX Invalid	0	0
RX Hello	8	384
RX DB des	8	496
RX LS req	2	72
RX LS upd	8	740
RX LS ack	4	236
RX Total	30	1928
TX Failed	0	0
TX Hello	10	792
TX DB des	6	624
TX LS req	2	112
TX LS upd	8	708
TX LS ack	5	460
TX Total	31	2696

OSPF header errors

Length 0, Checksum 0, Version 0, Bad Source 0,
No Virtual Link 0, Area Mismatch 0, No Sham Link 0,
Self Originated 0, Duplicate ID 0, Hello 0,
MTU Mismatch 0, Nbr Ignored 0, LLS 0,
Authentication 0, TTL Check Fail 0,

OSPF LSA errors

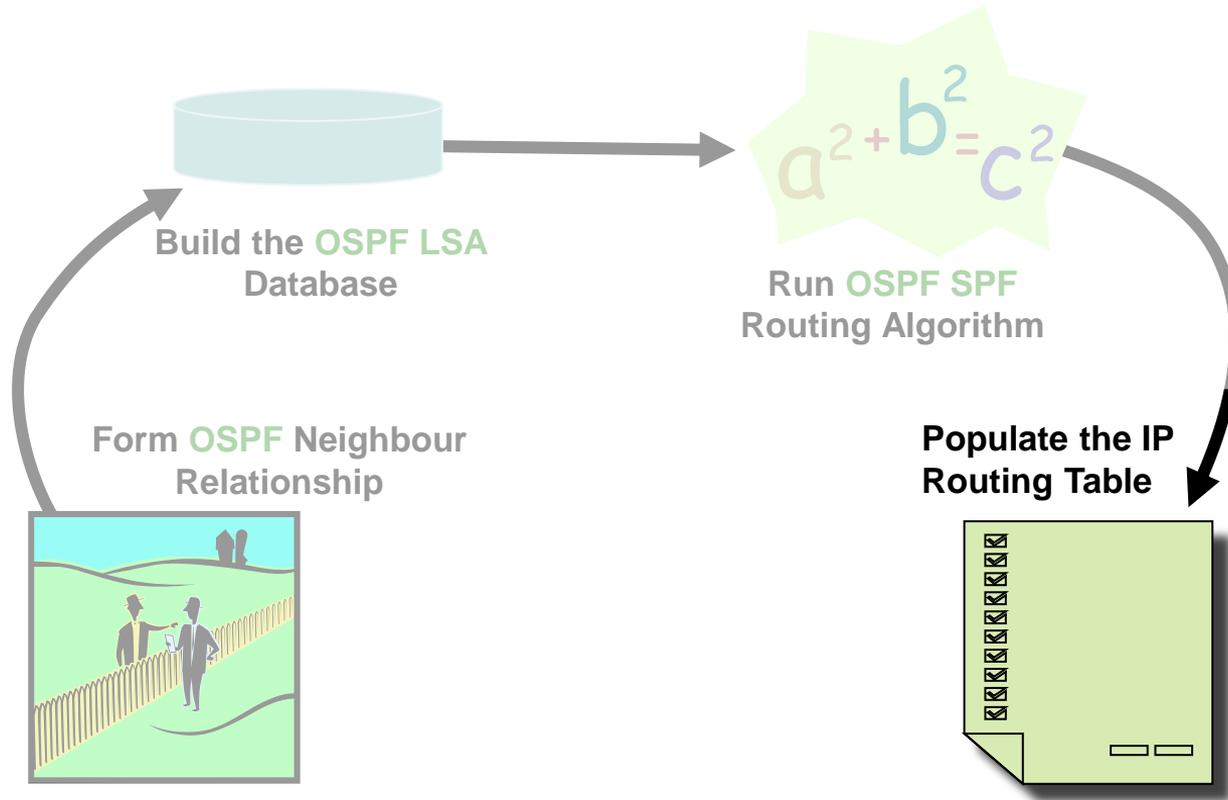
Type 0, Length 0, Data 0, Checksum 0,

OSPF Traffic Statistics

Queues

	InputQ	UpdateQ	OutputQ
Limit	0	200	0
Drops	0	8881	0
Max delay [msec]	1076	21188	28
Max size	3961	200	6
Invalid	0	0	0
Hello	3961	0	0
DB des	0	0	0
LS req	0	0	0
LS upd	0	200	0
LS ack	0	0	6
Current size	0	0	0
Invalid	0	0	0
Hello	0	0	0
DB des	0	0	0
LS req	0	0	0
LS upd	0	0	0
LS ack	0	0	0

Lifecycle View of the OSPF Routing Process



OSPF Routing Table

Prefix in Database but not in the Routing Table

```
R1#sh ip ospf nei
```

Neighbor ID	Pri	State	Dead Time	Address	Interface
201.1.1.1	0	FULL/ -	00:00:30	10.1.1.2	Ethernet0/0

```
R1#
```

```
R2#sh ip ospf nei
```

Neighbor ID	Pri	State	Dead Time	Address	Interface
101.1.1.1	1	FULL/BDR	00:00:30	20.1.1.1	Ethernet0/0

```
R2#
```

```
R1#sh ip route ospf
```

```
R1#
```



OSPF Routing Table

Prefix in Database but not in the Routing Table

```
R1#sh ip ospf data router 201.1.1.1
```

OSPF Router with ID (101.1.1.1) (Process ID 1)

Adv Router is not-reachable

LS age: 1254

Options: (No TOS-capability, DC)

LS Type: Router Links

Link State ID: 201.1.1.1

Advertising Router: 201.1.1.1

Link connected to: a Transit Network

(Link ID) Designated Router address: 10.1.1.2

(Link Data) Router Interface address: 10.1.1.2

Number of TOS metrics: 0

TOS 0 Metrics: 10

```
R1#sh ip ospf data router 101.1.1.1
```

OSPF Router with ID (201.1.1.1) (Process ID 1)

Adv Router is not-reachable

LS age: 1670

Options: (No TOS-capability, DC)

LS Type: Router Links

Link State ID: 101.1.1.1

Advertising Router: 101.1.1.1

Link connected to: a Stub Network

(Link ID) Network/subnet number: 10.10.10.0

(Link Data) Network Mask: 255.255.255.0

Number of TOS metrics: 0

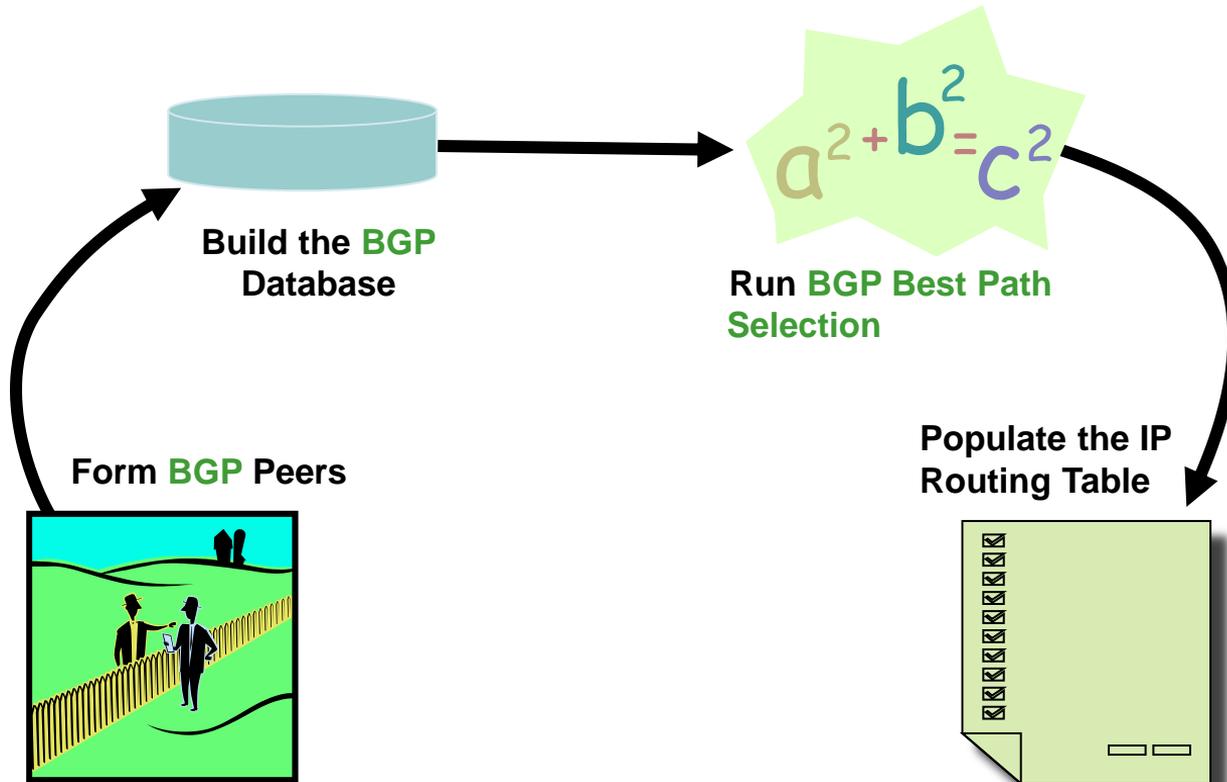
TOS 0 Metrics: 10



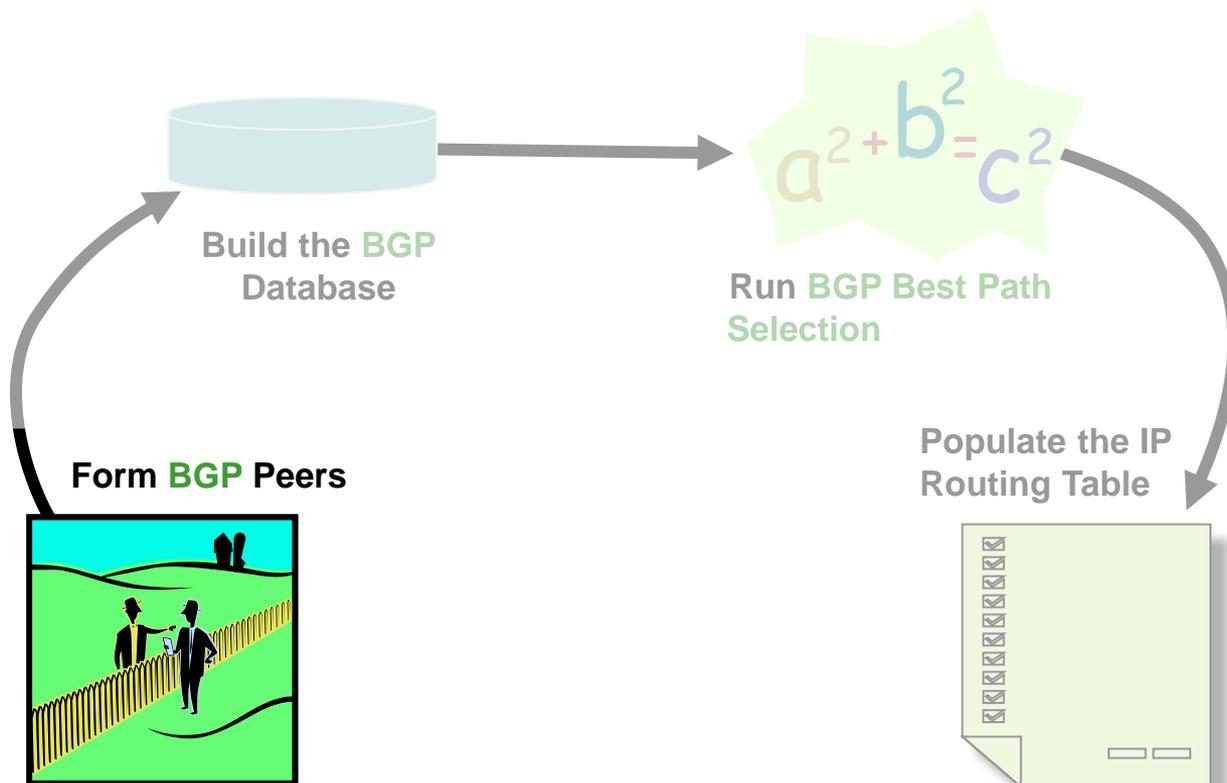


Lifecycle View of the BGP Routing Process

Lifecycle View of the BGP Routing Process



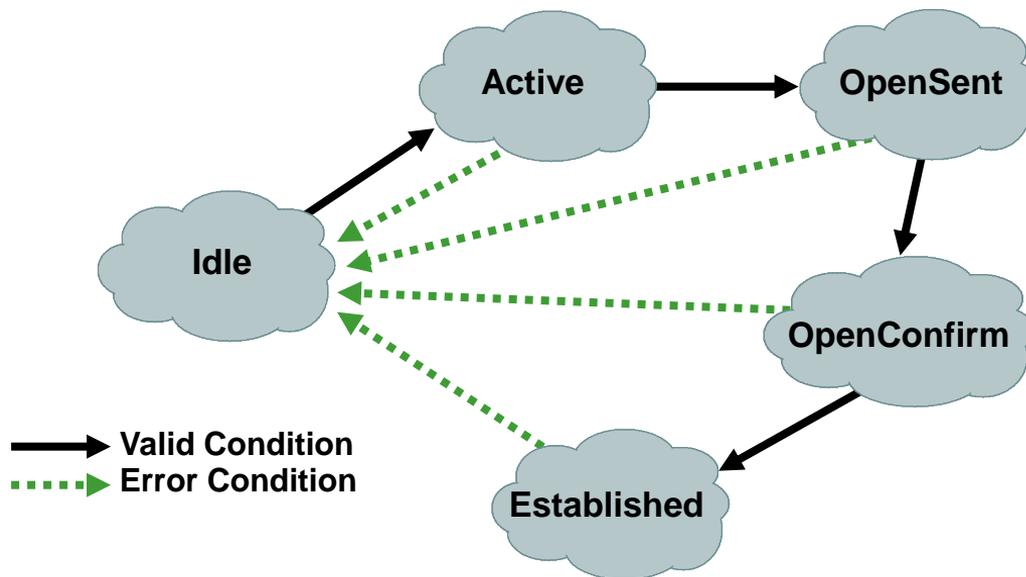
Lifecycle View of the BGP Routing Process



Lifecycle of BGP Peers

- BGP uses TCP port 179
- Peers exchange OPEN messages:
 - Router ID
 - AS #
 - Capabilities
 - Hold Time
- FSM (Finite State Machine) is used to negotiate to Established state
- Initial exchange of entire table
- Incremental updates
- Keepalive messages

BGP Peers - Peering Finite State Machine



- If everything is okay, proceed to the next state
- If not, reset back to Idle state

BGP Peers - Peering Finite State Machine

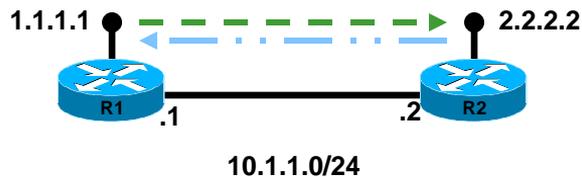
```
R1#debug ip bgp (shows the state transitions)
BGP: 2.2.2.2 went from Idle to Active

BGP: 2.2.2.2 active open failed - TCP session must be opened passively

BGP: 2.2.2.2 passive open to 192.168.1.1
BGP: 2.2.2.2 open active, local address 1.1.1.1
BGP: 2.2.2.2 read request no-op
BGP: 2.2.2.2 went from Active to OpenSent
BGP: 2.2.2.2 sending OPEN, version 4, my as: 100, holdtime 180 se
BGP: 2.2.2.2 rcv OPEN, version 4, holdtime 180 seconds
BGP: 2.2.2.2 rcv OPEN w/ OPTION parameter len: 16
BGP: 2.2.2.2 rcvd OPEN w/ optional parameter type 2 (Capability) len 6
BGP: 2.2.2.2 OPEN has CAPABILITY code: 1, length 4
BGP: 2.2.2.2 OPEN has ROUTE-REFRESH capability(new) for all address-families
BGP: 2.2.2.2 rcvd OPEN w/ remote AS 200
BGP: 2.2.2.2 went from OpenSent to OpenConfirm
BGP: 2.2.2.2 went from OpenConfirm to Established
%BGP-5-ADJCHANGE: neighbor 2.2.2.2 Up
R1#
```

BGP Peers – The TCP Connection

- R1 to R2 TCP connection 
 - neighbor 2.2.2.2 remote-as 100
 - neighbor 2.2.2.2 update-source loopback 0
- R2 to R1 TCP connection 
 - neighbor 1.1.1.1 remote-as 100
 - neighbor 1.1.1.1 update-source loopback 0
- neighbor x.x.x.x transport connection-mode <active|passive>
- neighbor x.x.x.x transport path-mtu-discovery disable
- neighbor N.N.N.N ttl-security hops <hops away>



Keepalive & Holdtime

- Holdtime is negotiated via OPEN messages
- Lowest holdtime wins
 - 0 seconds : infinite holdtime
 - 3 seconds : lowest non-zero setting
 - 180 seconds is default

```
R1(config-router)# neighbor x.x.x.x timers X Y Z  
  
X (0-65535) is keepalive  
Y (0-65535) is holdtime  
Z (0-65535) minimum acceptable holdtime
```

BGP Peers

Show ip bgp neighbours - AFI independent BGP info

```
R1#show ip bgp neighbors 2.2.2.2
BGP neighbor is 2.2.2.2, remote AS 200, external link
  BGP version 4, remote router ID 2.2.2.2
  BGP state = Established, up for 00:02:07
  Last read 00:00:06, last write 00:00:13, hold time is 180, keepalive
    interval is 60 seconds
  Neighbor capabilities:
    Route refresh: advertised and received(new)
    Address family IPv4 Unicast: advertised and received
  Message statistics:
    InQ depth is 0
    OutQ depth is 0

                Sent          Rcvd
  Opens:                6           6
  Notifications:        0           0
  Updates:               4           0
  Keepalives:           175          177
  Route Refresh:         0           0
  Total:                 185          183
  Default minimum time between advertisement runs is 30 seconds
```

BGP Peers

Show ip bgp neighbours - AFI specific BGP info

```
R1#show ip bgp neighbors
<snip>
For address family: IPv4 Unicast
  BGP table version 2, neighbor version 2/0
  Output queue size : 0
  Index 1, Offset 0, Mask 0x2
  1 update-group member

Prefix activity:
  Sent      Rcvd
  ----      ----
  Prefixes Current:      1      0
  Prefixes Total:        1      0
  Implicit Withdraw:     0      0
  Explicit Withdraw:     0      0
  Used as bestpath:      n/a     0
  Used as multipath:     n/a     0

Local Policy Denied Prefixes:
  Outbound  Inbound
  -----  -----
  Total:                0      0
Number of NLRIs in the update sent: max 1, min 1
```

BGP Peers

Show ip bgp neighbours (cont) BGP specific TCP info

```
<snip>
```

```
Connections established 6; dropped 5
```

```
Last reset 00:02:09, due to User reset
```

```
External BGP neighbor may be up to 255 hops away.
```

```
Connection state is ESTAB, I/O status: 1, unread input bytes: 0
```

```
Local host: 1.1.1.1, Local port: 12348
```

```
Foreign host: 2.2.2.2, Foreign port: 179
```

BGP Peers

Show ip bgp neighbours (cont) Generic TCP info

```
Enqueued packets for retransmit: 0, input: 0  mis-ordered: 0 (0 bytes)
```

```
Event Timers (current time is 0x5817B38):
```

Timer	Starts	Wakeups	Next
Retrans	5	0	0x0
TimeWait	0	0	0x0
AckHold	4	3	0x0
SendWnd	0	0	0x0
KeepAlive	0	0	0x0
GiveUp	0	0	0x0
PmtuAger	0	0	0x0
DeadWait	0	0	0x0

```
iss: 3541899715  snduna: 3541899871  sndnxt: 3541899871      sndwnd:  
16229  
irs: 2288128196  rcvnxt: 2288128318  rcvwnd:      16263  delrcvwnd:  
121
```

```
SRTT: 146 ms, RTTO: 1283 ms, RTV: 1137 ms, KRTT: 0 ms
```

```
minRTT: 0 ms, maxRTT: 300 ms, ACK hold: 200 ms
```

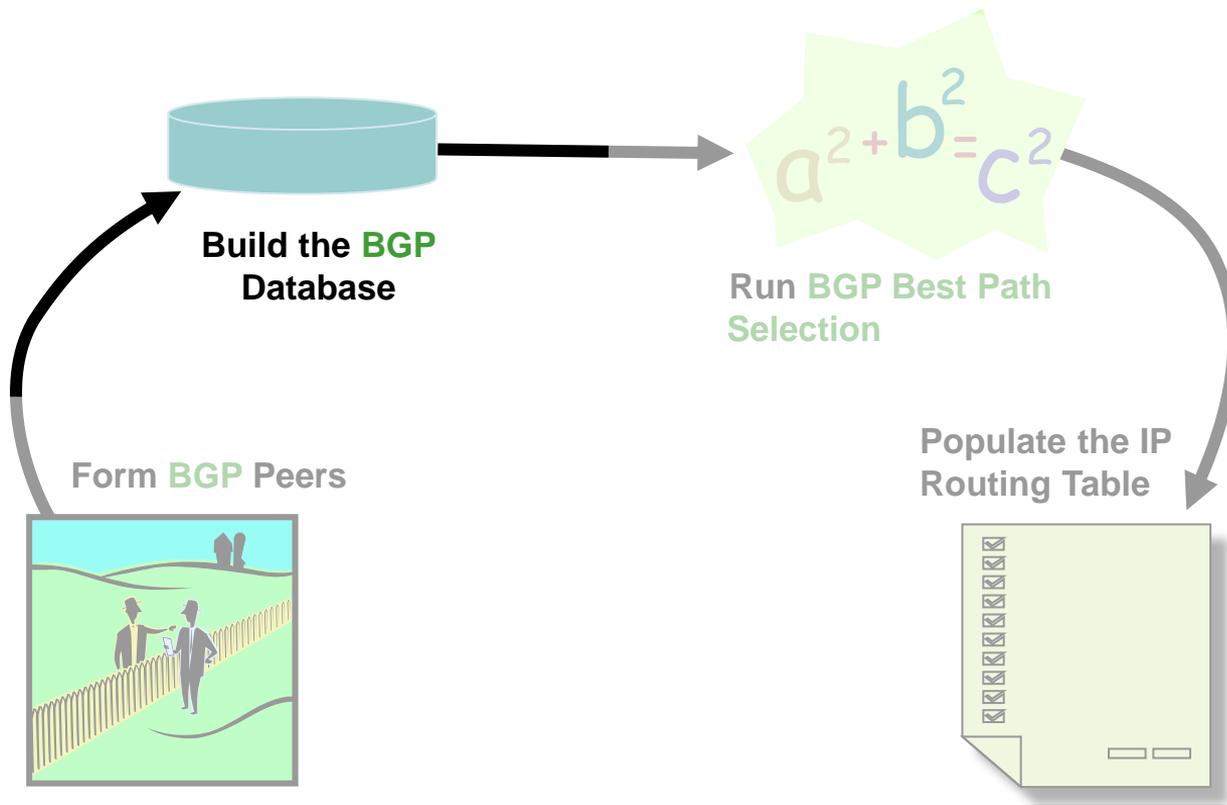
```
Flags: higher precedence, retransmission timeout, nagle, path mtu capable
```

```
Datagrams (max data segment is 1460 bytes):
```

```
Rcvd: 7 (out of order: 0), with data: 4, total data bytes: 121
```

```
Sent: 10 (retransmit: 0), with data: 5, total data bytes: 155
```

Lifecycle View of the BGP Routing Process



BGP Database Table

BGP Peers exchange UPDATE packets.

All locally known routes

Only the bestpath is advertised

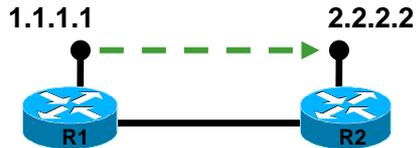
What Information is populated in the BGP Database Table?

- Locally generated prefixes via the network command
- Prefixes learnt from BGP Peers that have passed through neighbour filters

BGP Database Table

```
R1#  
router bgp 2  
no synchronization  
bgp log-neighbor-changes  
network 100.1.1.0 mask 255.255.255.0  
network 101.1.1.0 mask 255.255.255.0  
neighbor 2.2.2.2 remote-as 2  
neighbor 2.2.2.2 update-source Loopback0
```

```
R2#  
router bgp 2  
no synchronization  
bgp log-neighbor-changes  
network 200.1.1.0 mask 255.255.255.0  
network 201.1.1.0 mask 255.255.255.0  
neighbor 1.1.1.1 remote-as 2  
neighbor 1.1.1.1 update-source Loopback0  
neighbor 1.1.1.1 prefix-list From_R1 in  
neighbor 1.1.1.1 prefix-list To_R1 out  
no auto-summary  
  
ip prefix-list To_R1 seq 5 deny 201.1.1.0/24  
ip prefix-list To_R1 seq 7 permit 0.0.0.0/0 ge 1  
  
ip prefix-list From_R1 seq 5 deny 101.1.1.0/24  
ip prefix-list From_R1 seq 7 permit 0.0.0.0/0 ge 1
```

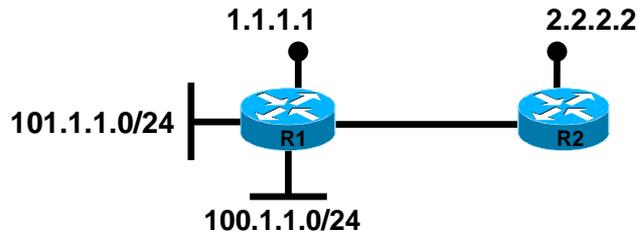


BGP Database Table

- `show ip bgp neighbor x.x.x.x received-routes`
- `soft-reconfiguration inbound` must be configured

```
R2#sh ip bgp neighbors 1.1.1.1 received-routes
BGP table version is 6, local router ID is 2.2.2.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
```

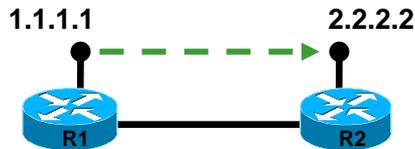
Network	Next Hop	Metric	LocPrf	Weight	Path
*>i100.1.1.0/24	1.1.1.1	0	100	0	i
* i101.1.1.0/24	1.1.1.1	0	100	0	i



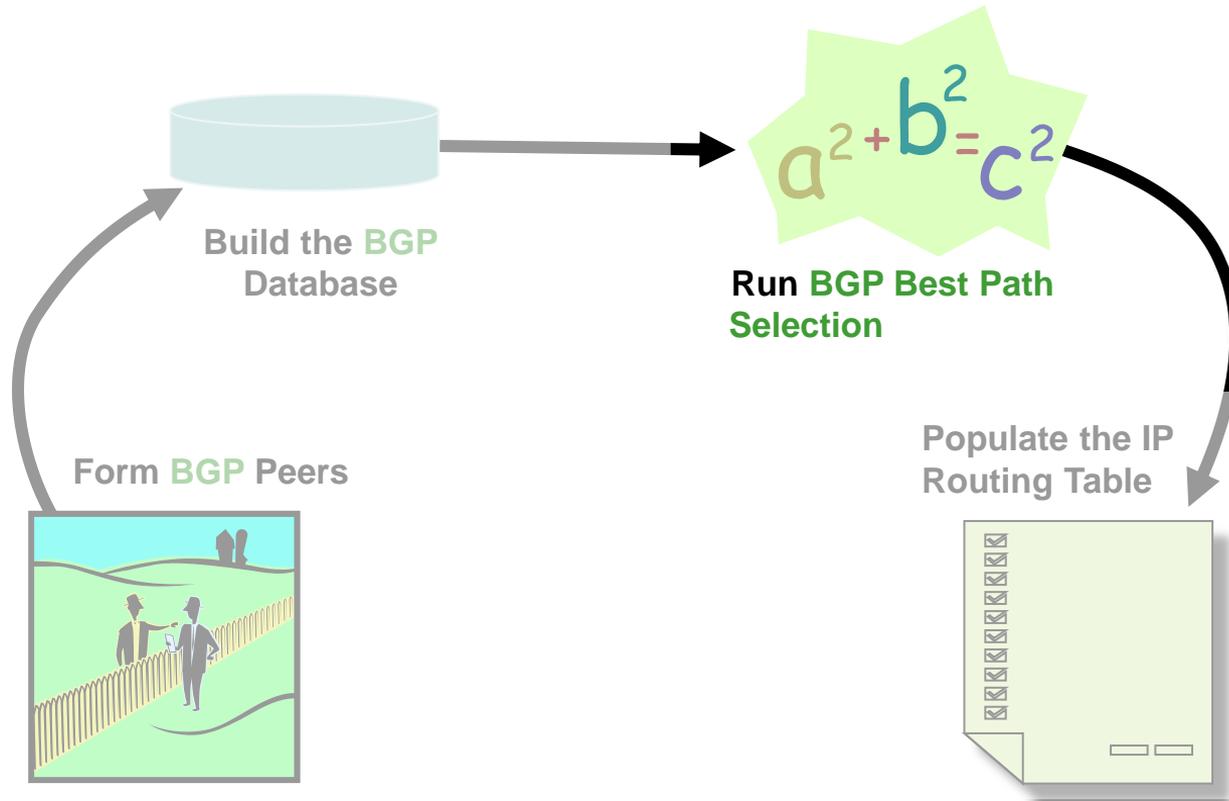
BGP Database Table

```
R2#debug ip bgp update
BGP updates debugging is on

13:49:45.878: BGP(0): 1.1.1.1 rcvd UPDATE w/ attr: nexthop 1.1.1.1, origin i, localpref 100, metric 0
13:49:45.878: BGP(0): 1.1.1.1 rcvd 101.1.1.0/24 -- DENIED due to: distribute/prefix-list;
13:49:45.878: BGP(0): 1.1.1.1 rcvd 100.1.1.0/24
13:49:46.146: BGP(0): Revise route installing 1 of 1 routes for 100.1.1.0/24 -> 1.1.1.1(main) to main IP table
13:49:46.146: BGP(0): nettable_walker 200.1.1.0/24 route sourced locally
13:49:46.146: BGP(0): nettable_walker 201.1.1.0/24 route sourced locally len is 4
13:49:46.146: BGP(0): 1.1.1.1 send UPDATE (format) 200.1.1.0/24, next 2.2.2.2, metric 0, path
R2#
```



Lifecycle View of the BGP Routing Process



BGP Best Path Algorithm – BGP Router

- BGP Router Process decides the best path

http://www.cisco.com/en/US/partner/tech/tk365/technologies_tech_note09186a0080094431.shtml

- Keep the RIB up to date
- Lots of things must happen when bestpaths change
 - RIB must be notified
 - Peers must be informed

BGP Path Selection Algorithm Simplified

1. Next-hop has to be accessible (in the routing table)
2. Route must be synchronised (better turn synchronisation off)
3. Largest weight (Admin Preference, local to the router)
4. Largest local preference (Admin Preference. Spread within AS)
5. Router originated (Metric= "0 ASes" - Better if we originated it)
6. Shortest as-path (Metric in AS's)
7. Lowest origin (igp < egp < incomplete)
8. Lowest MED (metric information from the next AS)
9. External over internal (Metric better if we are the border router)
10. Closest next-hop (IGP metric - the next-hop must be close)
11. Lowest router-id of Originator (tie-breaker)
12. Shortest Cluster-list (tie-breaker)
13. Lowest IP address of Neighbour (tie-breaker)

BGP Best Path Algorithm

How to Identify constant bestpath changes

- Prefix Table Version

A prefix's table version is **bumped** up for every bestpath change

```
R1#show ip bgp 10.0.0.0
BGP routing table entry for 10.0.0.0/8, version 31
Paths: (1 available, best #1, table Default-IP-Routing-Table)
Flag: 0x820
    Not advertised to any peer
    200
        2.2.2.2 from 2.2.2.2 (2.2.2.2)
            Origin IGP, metric 0, localpref 100, valid, external, best
R1#
```

BGP Best Path Algorithm

How to Identify constant bestpath changes

- RIB and Peer Table versions

Used to keep track of which bestpath changes have been propagated to whom

```
R2#show ip bgp summ
```

```
BGP router identifier 2.2.2.2, local AS number 200
```

```
BGP table version is 13, main routing table version 13
```

```
3 network entries using 351 bytes of memory
```

```
3 path entries using 156 bytes of memory
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
1.1.1.1	4	100	4386	4388	13	0	0	01:20:24	1

```
R2#
```

BGP Best Path Algorithm – BGP Scanner

- The Scanner performs the following Housekeeping Tasks
 - Validate nexthop reachability
 - Validate bestpath selection
 - Route redistribution and network statements
 - Conditional advertisement
 - Route dampening
 - BGP Database cleanup
- Full Scanner Run happens every 60 Seconds
 - bgp scan-time X
- CPU spike is normal when scanner runs

BGP Best Path Algorithm – BGP Scanner

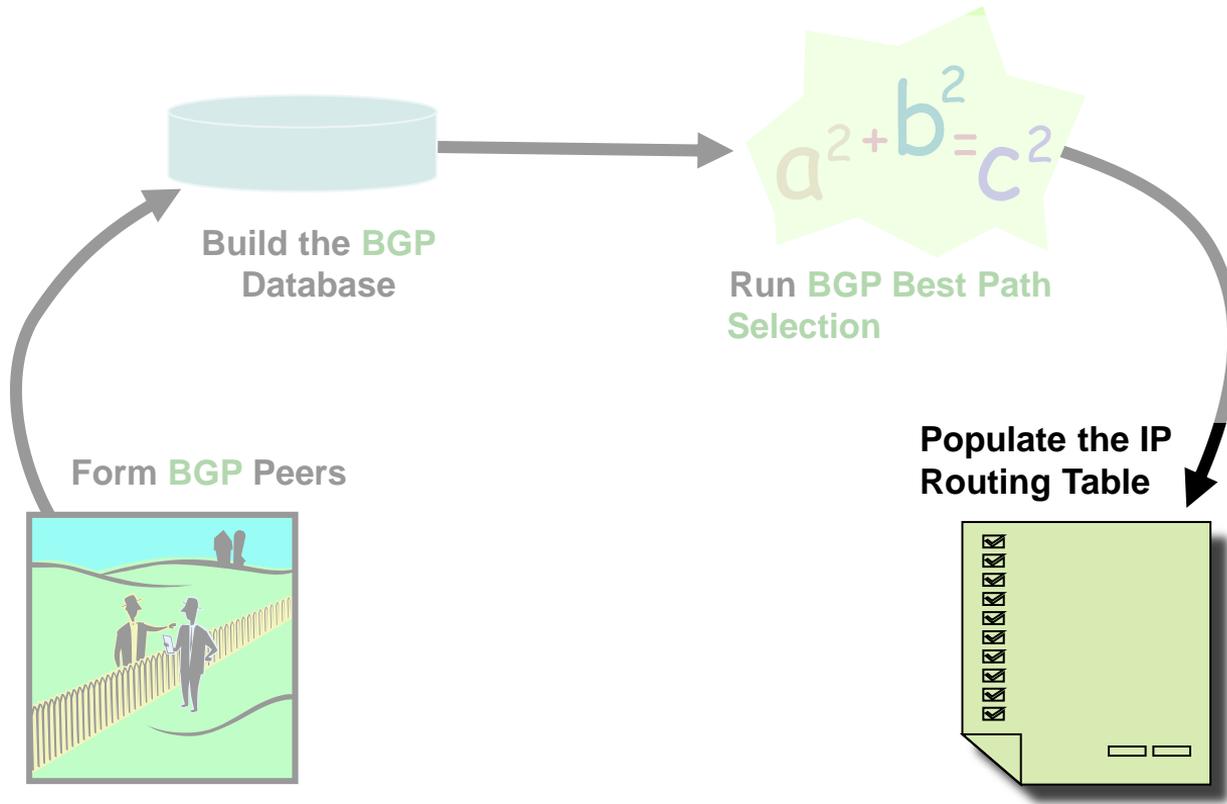
- `"debug ip bgp events"` will show you when scanner ran for each address-family
 - BGP: Performing BGP general scanning
 - BGP(0): scanning IPv4 Unicast routing tables
 - BGP(IPv4 Unicast): Performing BGP Nexthop scanning for general scan
 - BGP(0): Future scanner version: 7, current scanner version: 6
 - BGP(1): scanning IPv6 Unicast routing tables
 - BGP(IPv6 Unicast): Performing BGP Nexthop scanning for general scan
 - BGP(1): Future scanner version: 13, current scanner version: 12
 - BGP(2): scanning VPNv4 Unicast routing tables

- Improvements have been made to reduce CPU Impact

Route Redistribution and Network Statements are now event driven

Next-Hop AddressTracking

Lifecycle View of the BGP Routing Process



The BGP Routing Table

- The Best path selected is put in Routing Table (RIB)

```
R2#show ip bgp nei 1.1.1.1 routes
```

```
<..snip..>
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*>i100.1.1.0/24	1.1.1.1	0	100	0	i

```
Total number of prefixes 1
```

```
R2#show ip bgp nei 1.1.1.1 advertised-routes
```

```
<..snip..>
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 200.1.1.0	0.0.0.0	0		32768	i

```
Total number of prefixes 1
```

The BGP Routing Table

Why is the prefix not in the Routing Table?

```
R2#show ip bgp 101.1.1.0
```

```
BGP routing table entry for 101.1.1.0/24, version 7
```

```
Paths: (1 available, best #1, RIB-failure(17))
```

```
Not advertised to any peer
```

```
Local
```

```
1.1.1.1 (metric 11) from 1.1.1.1 (101.1.1.1)
```

```
Origin IGP, metric 0, localpref 100, valid, internal, best
```

```
R2#sh ip bgp rib-failure
```

Network	Next Hop	RIB-failure	RIB-NH Matches
101.1.1.0/24	1.1.1.1	Higher admin distance	n/a

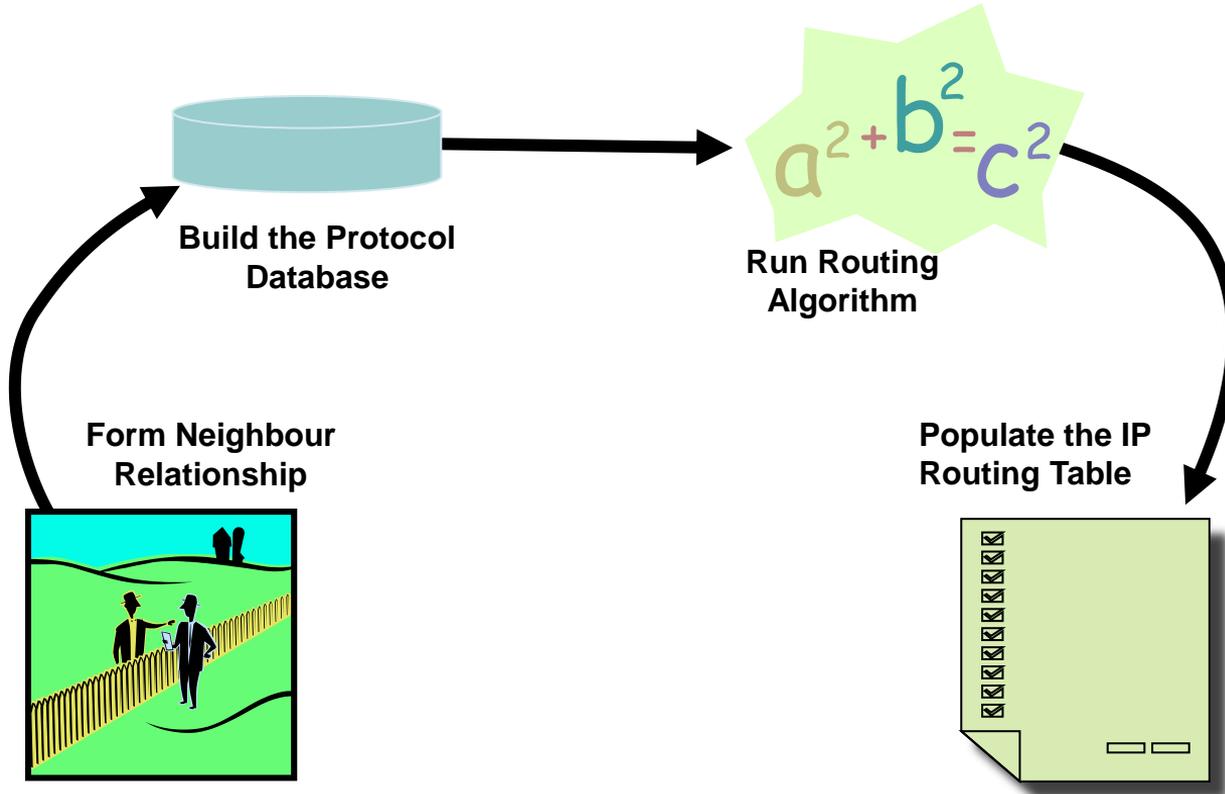
```
R2#
```

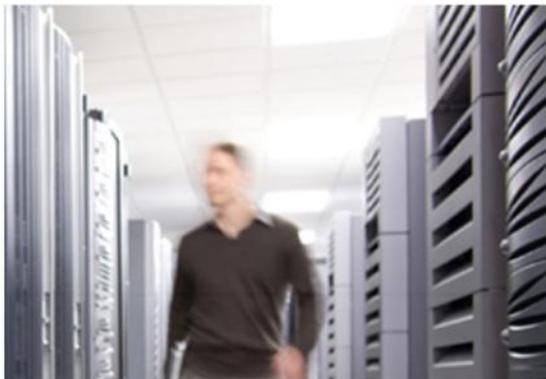
Troubleshooting Best Practices

- Large logging buffer
- Correct logging level
- Synchronise router clock
- Time Stamps
- Configuration Backup
- Familiarity with H/W specific troubleshooting tools

<https://supportforums.cisco.com/community/netpro/service-providers/blog/2013/05/29/packet-capture-capabilities-of-cisco-routers-and-switches>

Summary





Q & A

Complete Your Online Session Evaluation

Give us your feedback and receive a Cisco Live 2014 Polo Shirt!

Complete your Overall Event Survey and 5 Session Evaluations.

- Directly from your mobile device on the Cisco Live Mobile App
- By visiting the Cisco Live Mobile Site www.ciscoliveaustralia.com/mobile
- Visit any Cisco Live Internet Station located throughout the venue

Polo Shirts can be collected in the World of Solutions on Friday 21 March 12:00pm - 2:00pm



Learn online with Cisco Live!

Visit us online after the conference for full access to session videos and presentations.

www.CiscoLiveAPAC.com



CISCO TM