



# BSCN Study Notes

S.J. Iveson

# Preface

**This document is only relevant:**

- **Up to version 12.0 of the Cisco IOS.**
- **To the CCNP v2.0 track.**
- **To the material taught on the BSCN course.**

You should not try to use this document as a cheat to be memorised or learnt by rote in order to pass the BSCN exam without attending a course or having any practical experience. This is generally not possible with Cisco exams anyway, due to their structure and depth.

This guide is intended for reference and last minute revision by candidates who are due to take the exam after having taken a course and/or having had a fair amount of relevant practical experience.

It will also be good in helping you to prepare prior to the course helping you to make the most of it once there.

These notes were created by the author while studying for the exam on a Global Knowledge/Geotrain course. The author passed first time with a score of 903.

If you have any comments regarding this document please e-mail: [sjiveson@routerzone.com](mailto:sjiveson@routerzone.com) .

For more information and resources for the BSCN course and exam visit my website at [www.routerzone.com](http://www.routerzone.com).

Steven Iveson 2001-2002

ALL INFORMATION PROVIDED IN THIS DOCUMENT IS PROVIDED "AS IS" WITH ALL FAULTS WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED. STEVEN IVESON DISCLAIMS ALL WARRANTIES, EXPRESSED OR IMPLIED INCLUDING, WITHOUT LIMITATION, THOSE OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT OR ARISING FROM A COURSE OF DEALING, USAGE, OR TRADE PRACTICE.

STEVEN IVESON SHALL NOT BE LIABLE FOR ANY INDIRECT, SPECIAL, CONSEQUENTIAL, OR INCIDENTAL DAMAGES INCLUDING, WITHOUT LIMITATION, LOST PROFITS OR REVENUES, COSTS OF REPLACEMENT GOODS, LOSS OR DAMAGE TO DATA ARISING OUT OF THE USE OR INABILITY TO USE THIS DOCUMENT OR ANY PRODUCT FEATURED, DAMAGES RESULTING FROM USE OF OR RELIANCE ON THE INFORMATION PRESENT, EVEN IF STEVEN IVESON HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

**This document is not sponsored by, endorsed by or affiliated with Cisco Systems, Inc.**

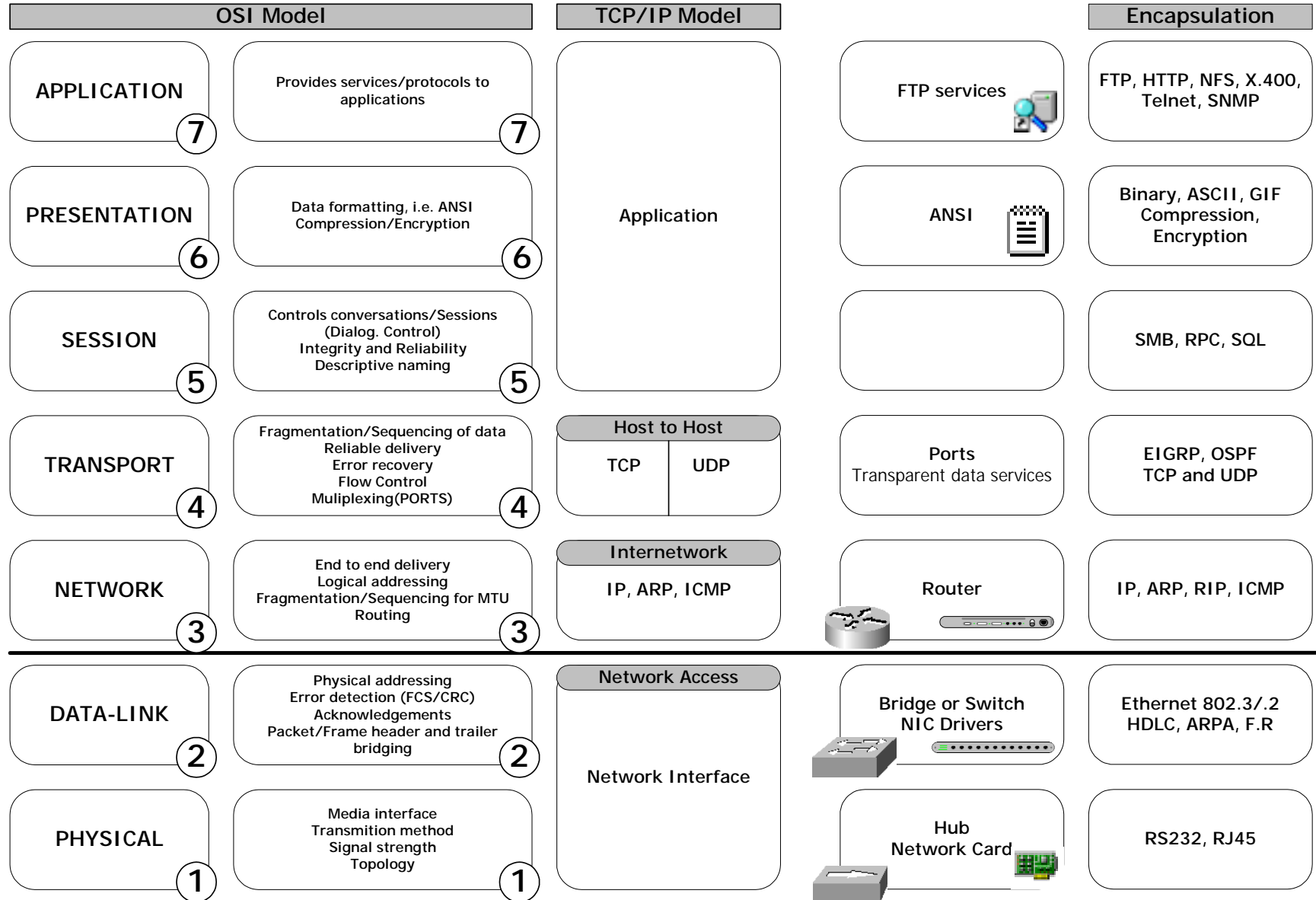
# CONTENTS

<b>CONTENTS</b> .....	<b>3</b>
<b>SECTION ONE: TCP/IP and Routing Protocol Review</b> .....	<b>6</b>
TCP/IP and the OSI Model .....	7
Subnetting.....	7
Subnetting.....	8
Interior Routing Protocols.....	9
Exterior Routing Protocols .....	9
What is Routing?.....	10
Routing Table Entries .....	10
Administrative Distance .....	11
Routing Protocol Types: .....	11
Classful Routing .....	12
Classless Routing .....	12
VLSM .....	13
Route Summarisation .....	13
When Not to Use Route Summarisation .....	14
CIDR (Classless InterDomain Routing).....	16
IP Unnumbered .....	16
IP Helper Address .....	17
<b>SECTION TWO: OSPF</b> .....	<b>18</b>
OSPF Facts .....	19
OSPF Terminology.....	20
Supported OSPF Topologies.....	21
Neighbor Relationships.....	21
Operation over Point to Point Links .....	25
Operation over Non-Broadcast Multi-Access (NBMA) Networks.....	25
Configuring OSPF - Internal Routers.....	27
OSPF Optional Commands .....	28
Configuring OSPF over NBMA.....	29
Verifying OSPF Operation .....	30
<b>SECTION THREE: Multiple OSPF Areas</b> .....	<b>32</b>
The Need for Multiple Areas.....	33
Router, LSA and Area Types .....	34
Virtual Links.....	36
Configuring OSPF for Multiple Areas .....	37
Stub & Totally Stubby Areas .....	37
Configuring Stub & Totally Stubby Areas .....	38
Summarization and VLSM .....	38
Costs for Summary & External Routes .....	39
Verifying OSPF Operation .....	39
<b>SECTION FOUR: EIGRP</b> .....	<b>40</b>
EIGRP Facts.....	41
EIGRP Terminology .....	42
Packets .....	42
Neighbor Relationships.....	43
Neighbor Tables.....	44
Reliability.....	44
Route Discovery.....	46
Route Selection.....	46
DUAL .....	47
Configuring EIGRP .....	48
Summarization.....	49
Load Balancing .....	50

WAN Links .....	50
Scalability .....	51
Queries .....	52
Verifying EIGRP Operation .....	53
<b>SECTION FIVE: BGP.....</b>	<b>54</b>
BGP Facts .....	55
BGP Terminology .....	55
Using Static Routes Instead .....	56
Characteristics .....	57
Policy Based Routing .....	58
BGP Update Message Attributes .....	58
The Synchronisation Rule .....	61
BGP Messages .....	62
Route Selection Decision Process .....	63
CIDR and Aggregate Addresses .....	63
Configuring BGP.....	64
Verifying BGP Operation .....	65
<b>SECTION SIX: BGP Scalability.....</b>	<b>66</b>
The Split Horizon Rule .....	67
Route Reflectors .....	67
Route Reflector Design & Configuration .....	69
Policy Control.....	70
Prefix Lists .....	70
Configuring Prefix Lists.....	71
Prefix List Example .....	72
Verifying Prefix Lists .....	72
Multihoming.....	73
Configuring Weight & Local Preference.....	74
Redistribution with IGP .....	74
Redistributing BGP into an IGP.....	75
<b>SECTION SEVEN: Optimising Routing Update Operation .....</b>	<b>76</b>
Redistribution .....	77
Selecting the Best Route.....	77
The Seed Metric.....	78
Configuring Redistribution .....	78
Configuring Static Routes .....	80
Configuring default-network .....	81
<b>Acronyms and Terms.....</b>	<b>82</b>

# SECTION ONE: TCP/IP and Routing Protocol Review

# TCP/IP and the OSI Model



# Subnetting

Binary									
→	128	192	224	240	248	252	254	255	N/W Mask
←	128	64	32	16	8	4	2	1	Decimal
←	2 <sup>7</sup>	2 <sup>6</sup>	2 <sup>5</sup>	2 <sup>4</sup>	2 <sup>3</sup>	2 <sup>2</sup>	2 <sup>1</sup>	2 <sup>0</sup>	
	128	64	32	8	2	1			= 235.
	1	1	1	0	1	0	1	1	

## Address Classes

Class	Range	Networks	Hosts	High Order Bits
A	001-126	126	16,777,214	0
B	128-191	16,384	65,534	10
C	192-223	2,097,152	254	110
D	224-239	Multicast		1110

### To Subnet By Number of Networks:

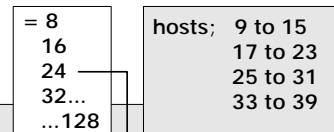
- Take no of networks req. and calculate in binary:  =  If all 1's, add another byte
- The amount of binary bits used is the size of the mask:  =
- Convert to decimal using table above:
- No of networks = 2 to the power of no. of bits used:  =
- No of hosts = 2 to the power of no. of bits remaining -2  =
- To calculate subnets 256 - mask, and then keep adding:

### By No. of Hosts

- =
- netmask =
- =
- =
- =

If class C, then ignore the last host address for each network. (This is the broadcast address.)

→ To calculate valid hosts question do just this bit →



\*The **ip subnet-zero** command allows the use of all available networks; For CCNA you are not expected to know this, calculate  $n = 2^5 - 2 = 30$ . For CCNP exams you are and do not subtract the 2, thus  $n = 2^5 = 32$ .

\*Allowing **ip directed broadcasts** will cause problems with 0 subnets. This feature should be disabled if you use them.

**ip subnet-zero is enabled and ip directed broadcasts is disabled** by default with v12 of the IOS

Subnet a single subnet further for WAN links between these networks

## Interior Routing Protocols

Used to connect networks within a single Autonomous System under the control of a single technical administration

Protocol	Standard?	Link State or Distance Vector	Updates: Intermittent   Periodic	Metric Used	Convergence Speed	Classless   Classful (Carry Subnet Information?)	VLSM Support?	Protocol Support
<b>RIP</b>	RFC what?	Distance Vector	P – 30s	Hops	SLOW	v1 – CLASSFUL v2 - CLASSLESS	v1 – No v2 – Yes	IP
<b>OSPF v2</b>	RFC 2328	Link State	I – 10s hello time 40s dead interval 30m database refresh	Cost: Bandwidth = $10^8/\text{bandwidth}$	<b>QUICK</b>	CLASSLESS	YES	IP
<b>IGRP</b>	Cisco	Distance Vector	P – 90s	Cost: Bandwidth & Delay(fixed)	SLOW	CLASSFUL	NO	IP
<b>EIGRP</b>	Cisco	Adv. Distance Vector (Both)	I – 5s hello time x3s dead interval	Cost: Bandwidth & Delay(fixed)	<b>QUICK</b>	CLASSLESS	YES	IP IPX Appletalk
<b>IPX RIP</b>	Novell	Distance Vector	P – 60s	Tics/Hops	SLOW	CLASSFUL	N/A	IPX
<b>RTMP</b>	Apple	Distance Vector	P – 10s	Hops	SLOW	CLASSFUL	N/A	Appletalk

## Exterior Routing Protocols

Used to connect between Autonomous Systems generally controlled by separate technical administrations

Protocol	Standard?	Link State or Distance Vector	Updates: Intermittent   Periodic	Metric Used	Convergence Speed	Classless   Classful (Carry Subnet Information?)	VLSM Support?	Protocol Support
<b>BGP v4</b>	RFC	Adv. Distance Vector (Both)		Path Vector	N/A	CLASSLESS	YES	IP

# What is Routing?

Cisco now classifies routers tasks into two areas:

- Learning and maintaining awareness of the network topology is considered **routing**
- The movement of data is now considered to be **switching**
  - Process switching involves looking up a packets destination in a routing table; this is the slowest form of switching. (Using access lists generates a lot of process switching)
    - Is the protocol suite enabled on the router?
    - Is there an entry for the destination network in the routing table?
    - Is the route available?
    - Which interface should be used?
  - Fast Switching involves simply switching the packets in a stream to an interface after the initial routing decision has been made based on the first packet; this is the fastest form of switching)

## Routing Table Entries

D 168.74.59.0 [100/118654] via 168.74.58.12, 00:02:36 Serial0

<b>D</b>	The protocol used to learn the route (D = EIGRP)
<b>168.74.59.0</b>	The destination logical network or subnet
<b>[100/118654]</b>	Administrative distance (trustworthiness)/Metric value (reachability)
<b>via 168.74.58.12</b>	Next hop logical address (next router)
<b>00:02:36</b>	Age of entry (hh:mm:ss)
<b>Serial0</b>	Interface through which the route was learnt and through which the packet will leave

## Administrative Distance

Administrative distance is a selection method for IP routing protocols

- The lower the administrative distance the more trusted the learning mechanism
- Static routes are preferred to dynamic routes
- Routing protocols with sophisticated metrics are preferred to those with simple metrics

Route Source	Default Administrative Distance
Connected interface	0
Static route out of an interface	0
Static route to a next hop	1
EIGRP summary route	5
External BGP	20
Internal EIGRP	90
IGRP	100
OSPF	110
IS-IS	115
RIP v1 or v2	120
EGP	140
External EIGRP	170
Internal BGP	200
Unknown	255

## Routing Protocol Types:

### Link State – OSPF:

- Uses a bandwidth and delay metric
- Avoids loops by transmitting and maintaining detailed topology information
- Fast convergence
- Suitable for larger networks

### Distance Vector – RIP:

- Uses a hop count metric, pure counting
- Slow network convergence, flash updates help
- Suitable for small networks only – has a maximum hop count of 15

## Classful Routing

RIP v1 and IGRP are classful routing protocols

Network/subnet masks are not carried within routing updates.

A router applies masks in one of two ways

- If the receiving device shares the same subnet mask as the advertising device and advertised route the mask is maintained
- If the mask does not match then the router summarises the received route into the appropriate class boundary, A B or C.

## Classless Routing

RIP v2, OSPF, EIGRP and BGP are classless routing protocols

Network/subnet masks ARE carried within routing updates

This enables the use of

- VLSM
- Automatic and manual route summarization

## VLSM

VLSM, Variable Length Subnet Masking allows you to:

- Have multiple subnet masks within a network
- Subnet an already subnetted network address

Benefits include:

- A more efficient use of IP addresses, particularly with small networks (i.e. WAN Links)
- Enhanced ability to use route summarisation, (more hierarchy thus smaller routing tables)

## Route Summarisation

Route summarisation is a useful way of minimizing routing table entries:

- A single route can be used to represent multiple networks
- This reduces router memory and CPU usages/requirements

In the following example a class B network address has been subnetted with a further 8 bits.

Without summarization these networks would each be listed individually in the routing table of a router.

By finding the maximum amount of matching bits shared by each network address and then reducing the subnet mask accordingly a single entry can be created in the routing table, representing all 8 networks:

Subnet Add.	Octet 1	Octet 2	Octet 3	Octet 4	
140.88.120.0/24	10001100	01011000	01111	000	00000000
140.88.121.0/24	140	88	01111	001	0
140.88.122.0/24	140	88	01111	010	0
140.88.123.0/24	140	88	01111	011	0
140.88.124.0/24	140	88	01111	100	0
140.88.125.0/24	140	88	01111	101	0
140.88.126.0/24	140	88	01111	110	0
140.88.127.0/24	140	88	01111	111	0
21 Matching bits – Summary Address = <b>140.88.120.0/21</b>				11 dissimilar bits	

To calculate the number of networks that are summarized by a summary address, use the following formula:  $2^n$  where n is the number of bits of the original network or subnet mask no longer used.

For example in the summary address above 3 bits have been removed from the original mask.  $2^3 = 8$  networks have been/can be summarized.

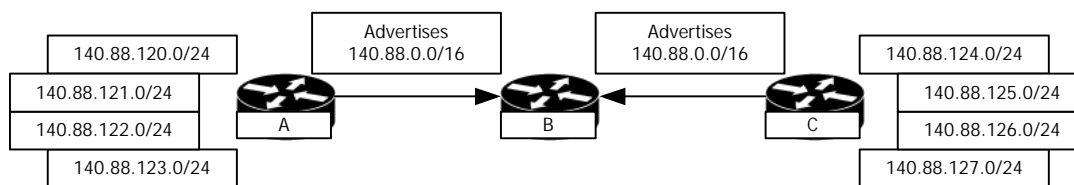
## When Not to Use Route Summarisation

Route summarisation should not be used where an internetwork has discontinuous subnets and particularly where a router advertising a particular summary route cannot actually reach all routes within that summary.

As RIPv1 and IGRP are classful they do not support discontinuous networks. Classless routing protocols do but the issue shown below can arise without careful thought.

**Note:** Automatic route summarisation is enabled by default when using EIGRP and RIPv2.

In the example below routers A and C both auto-summarise their 140.88.x.0/24 subnets using a 16-bit mask, thus router B receives advertisements for the same networks via two separate interfaces:

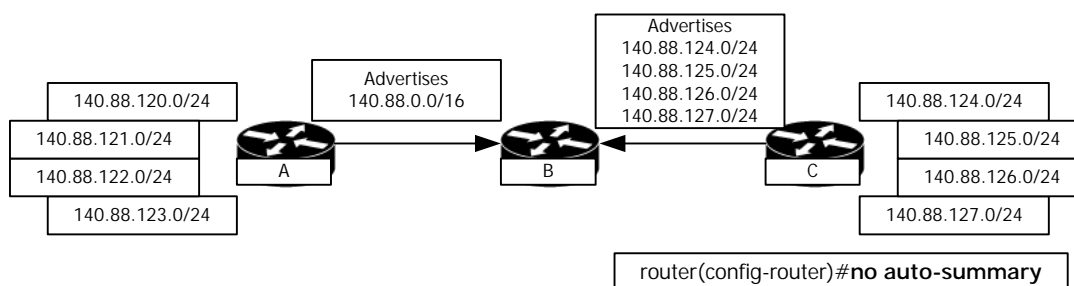


This issue can be solved in two ways; the least efficient is to disable auto summarisation on either router A, router C or both, as shown below:

- This example results in 5 routing table entries. As routers use the most specific/longest match found in routing tables, router A's summary route does not cause a problem.
- To disable route summarisation for a RIPv2 or EIGRP use the command:

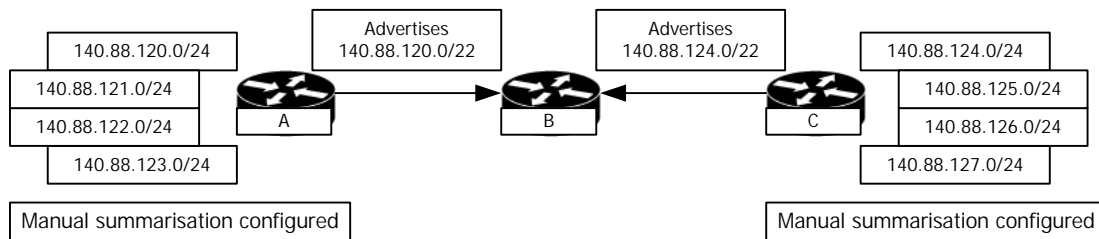
```
router(config-router)#no auto-summary
```

- Disable automatic summarization



The second solution is to manually summarise the subnets to be advertised by both router A and router C, as shown below:

- See the table for information on how the manual summarisation has been calculated



Subnet Add.	Octet 1	Octet 2	Octet 3		Octet 4
140.88.120.0/24	10001100	01011000	011110	00	00000000
140.88.121.0/24	140	88	011110	01	0
140.88.122.0/24	140	88	011110	10	0
140.88.123.0/24	140	88	011110	11	0
22 Matching bits – Summary Address = <b>140.88.120.0/22</b>					10 dissimilar bits
140.88.124.0/24	140	88	011111	00	0
140.88.125.0/24	140	88	011111	01	0
140.88.126.0/24	140	88	011111	10	0
140.88.127.0/24	140	88	011111	11	0
22 Matching bits – Summary Address = <b>140.88.124.0/22</b>					10 dissimilar bits

To manually configure EIGRP to summarize the above summary addresses for router A and C respectively, enter the following commands on the relevant interfaces:

```
router(config)#router eigrp 64666
router(config-router)#no auto-summary
router(config-if)#ip summary-address eigrp 64666
140.88.120.0 255.255.192.0
```

```
router(config)#router eigrp 64666
router(config-router)#no auto-summary
router(config-if)#ip summary-address eigrp 64666
140.88.124.0 255.255.192.0
```

- Router A
- Router C

To manually configure OSPF to summarize the above summary addresses for router A and C respectively, enter the following commands:

```
router(config)#router ospf 100
router(config-router)#summary-address 140.88.120.0
255.255.192.0
```

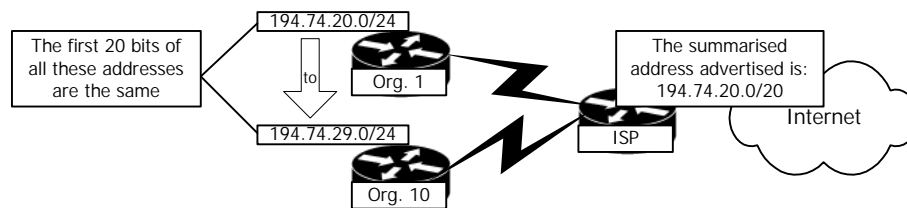
```
router(config)#router ospf 100
router(config-router)#summary-address 140.88.124.0
255.255.192.0
```

- Router A
- Router C

## CIDR (Classless InterDomain Routing)

Developed to alleviate the exhaustion of IP addresses and to reduce Internet routing tables

- Blocks of class C addresses are allocated to ISP's
- These blocks of addresses can be combined (supernetted) to create a larger pool of host addresses.
- The ISP assigns addresses as required
- These addresses are then summarised by the ISP on its internet connected routers routing tables, resulting in far fewer entries and advertisements than if all the ISP's clients were directly connected on an individual basis.



## IP Unnumbered

Allows IP processing on a serial interface only, without assigning an explicit address to the interface

Use the command:

```
router(config-if)#ip unnumbered int int-no
```

• i.e. Eth 0

When the unnumbered interface generates a packet it uses the address of the specified interface, (i.e. Eth 0,) as the source address of the packet.

### Benefits

- Another solution allowing you to save on network addresses

### Drawbacks

- You are unable to ping the interface, SNMP must be used for monitoring
- The specified interface, (i.e. Eth 0,) must be enabled and 'up' (Use loopback for reliability)
- The interface will not show in routing tables
- Unavailable with X.25 or SMDS interfaces

## IP Helper Address

By default routers do not forward broadcasts, this prevents broadcast storms and unnecessary traffic crossing WAN links.

In some situations, client broadcasts need to pass through a router to fully facilitate network applications or access. For example a client attempting to gain an IP address from a DHCP server on a network separated from its own by a router.

Helper addresses solve this problem by forwarding these broadcasts to another network or directly to a server. In doing so the router changes the broadcast to either a unicast address or directed broadcast address, (i.e. broadcasting to a specific network only.)

**Note:** All relevant broadcasts are sent to all helper addresses.

**Note:** When using **ip subnet 0** you must disable directed broadcasts with the command **no ip directed broadcast**, this is the default with v12, this means helper addresses must be specific, not a directed broadcast address such as 172.16.32.255.

- If more than one server on the same network needs to be accessed by clients it is recommended you use a directed broadcast to enable access to all servers.

```
router(config-if)#ip helper-address 172.16.32.255
```

- Use a specific address if only one server needs to be accessed, if servers are in separate networks or if using subnet 0.

```
router(config-if)#ip helper-address 172.16.32.66
```

The following UDP ports are enabled automatically: Time (37), TACACS (49), DNS (53), BOOTP server (67), TFTP (69), NetBIOS name service (137) and NetBIOS datagram service (138).

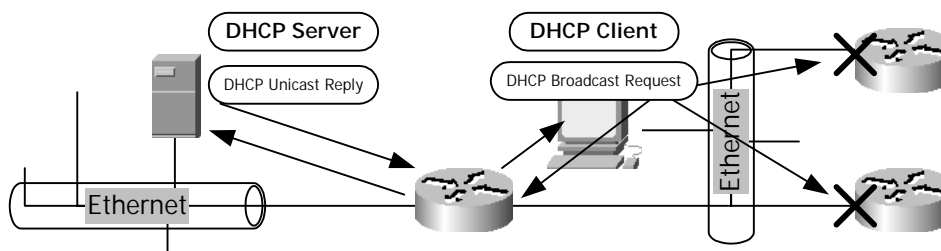
To add or remove ports from this list use the `ip forward-protocol udp port` command, for example:

To disable the NetBIOS name service enter:

```
router(config)#no ip forward-protocol udp 137
```

To enable the Timeserver service enter:

```
router(config)#ip forward-protocol udp 525
```



# SECTION TWO: OSPF

# OSPF Facts

## General

- Open Shortest Path First
- Interior Gateway Protocol (IGP)
- RFC 2328 (OSPF v2)
- Uses the SPF or Dijkstra algorithm
- Link state
- Protocol no. 89
- Administrative Distance: 110

Frame Header	Frame Payload		CRC
	IP Header	Protocol No.	
		89 6 17	OSPF TCP UDP

## Benefits

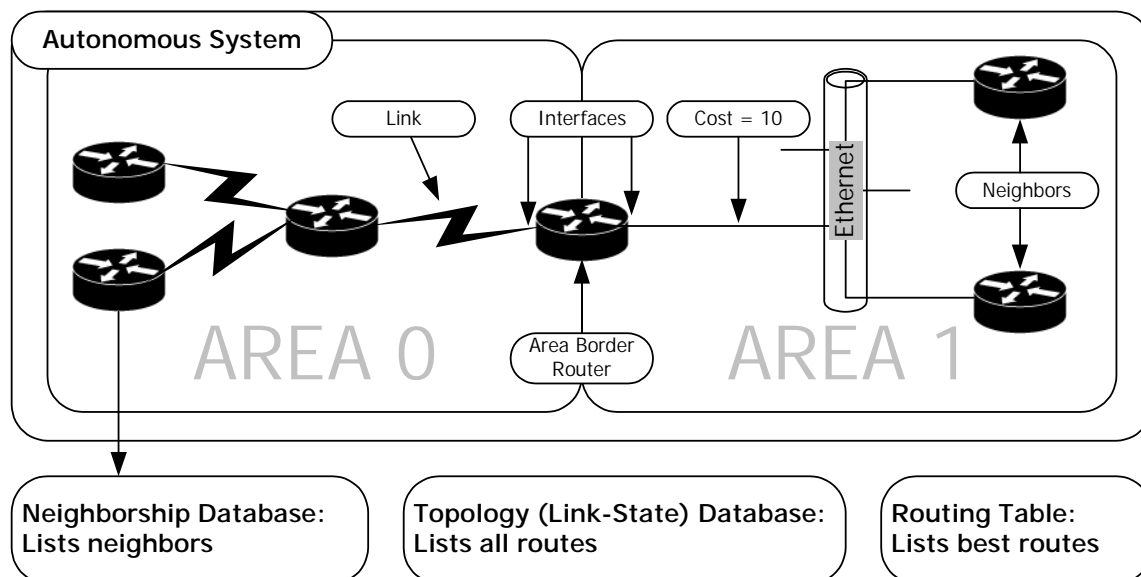
- Fast convergence due to routing changes being flooded immediately and computed in parallel
- Supports VLSM and subnet masking
- No reachability (distance) limitations
- Link state updates are multicast and sent only when there is a change in the network, this reduces bandwidth usage, particularly over WAN links
- Path selection is based on speed of connections not pure hop count
- Support for equal cost multipath

## Addresses Used

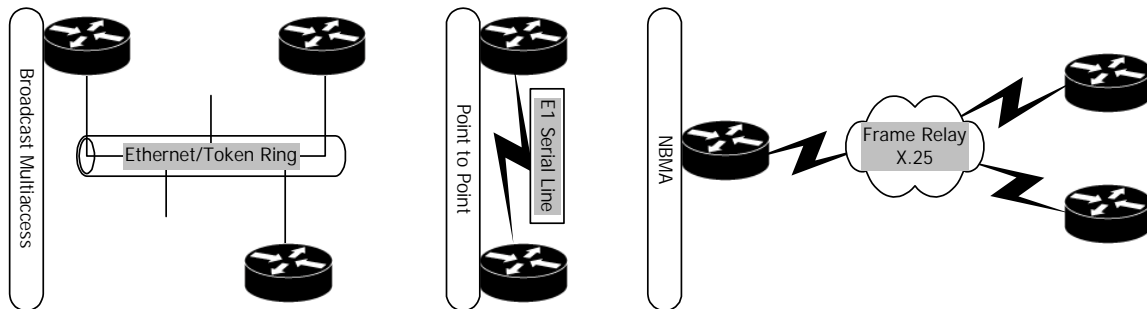
- 224.0.0.5      All OSPF routers multicast address
- 224.0.0.6      All OSPF DRs and BDRs multicast address

## OSPF Terminology

Interface	Connection between a router and attached network a.k.a link
Link State	Status of a link between two routers, advertised to other routers in Link State Advertisements
Cost	Value assigned to a link, based on link speed, referred to as Interface Output Cost
Autonomous System - AS	Group of routers exchanging routing information using a common routing protocol
Area	Routers and networks sharing the same area identification. All routers in an area have the same link state information. Routers within an area are known as internal routers
Neighbors	Two routers connected by a common network. The relationship is maintained using the Hello protocol
Hello	Protocol used to create and maintain neighbor relationships
Neighborship Database	A list of all neighbors a router has established relationships with
Link State Database	(Topological database) List of link state entries of all routers in the network. Shows the network topology. All routers have the same link state database. Created using LSAs.
Routing Table	(Forwarding database) Generated when the SPF algorithm is run on the link state database.



## Supported OSPF Topologies



## Neighbor Relationships

OSPF is dependent on the status of links between routers; therefore neighboring routers must recognize each other on a network before sharing information. This is done using the Hello protocol. Hello packets are sent periodically out of all interfaces running OSPF using IP multicast address 224.0.0.5.

Fields contained in a Hello packet are:

Router ID	A 32-bit number uniquely identifying the router with an AS. Highest IP address on all active interfaces.
Hello and dead intervals (Must be the same on both neighbors)	The hello interval is the frequency in seconds that a router sends hellos, (10secs on BMA networks.) The dead interval is the time in seconds that a router waits for a Hello packet before declaring the neighboring router down. (4 times the hello interval by default.)
Neighbors	The neighbors with which a relationship has been established. A relationship is indicated when the router sees itself listed in the neighbors hello packet.
Area ID (Must be the same on both neighbors)	Two routers must be in the same area, the same network segment and have the same subnet and mask in order to become neighbors.
Router Priority	An 8-bit number indicating the priority of this router. Used when selecting a designated DR and BDR.
DR & BDR IP Addresses	The IP addresses of the DR and BDR for the relevant network.
Authentication Password (Must be the same on both neighbors)	If enabled two routers must exchange the same password. If authentication is used all peer routers must have the same password.
Stub Area Flag (Must be the same on both neighbors)	Two routers must agree on the stub area flag.

## DR and BDR

Routers on a segment must elect a DR, (designated router,) and a BDR, (backup designated router,) to represent a network, using Hellos.

DR's and BDR's reduce routing update traffic by acting as a central point of contact for link state information exchange on a given network. Each router must therefore establish an adjacency with the DR and BDR. Rather than each router exchanging link state information with every other router on the network, each router sends link state information to the DR and BDR instead. The DR then sends each routers link state information to all other routers on the network. This flooding process reduces router related traffic on the network segment.

The DR and BDR ensure that the other routers in the network have the same link state information, reducing routing errors.

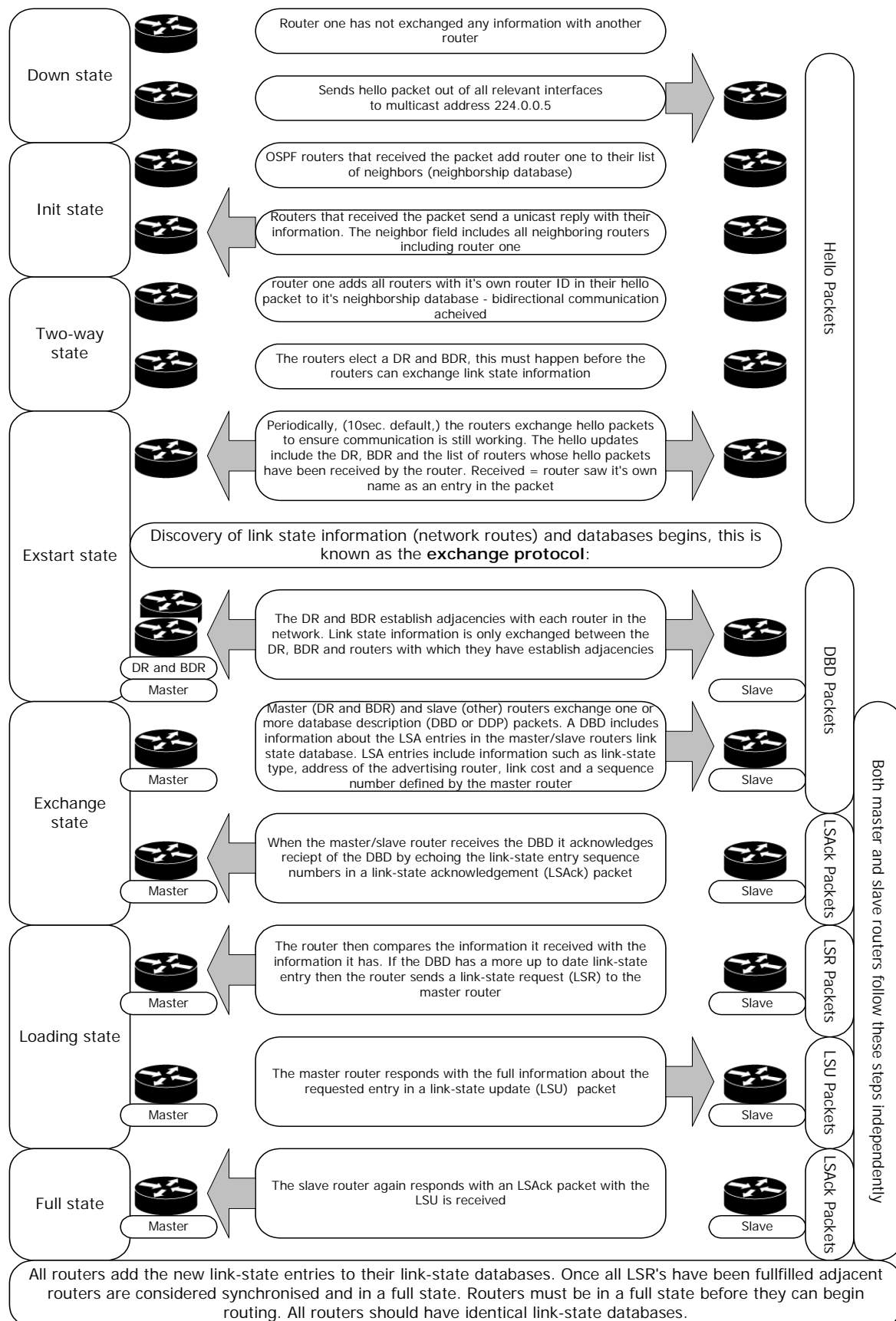
## DR and BDR Election

When electing a DR and BDR routers use each other's priority value during the hello packet exchange and use the following criteria to determine which is elected:

- The router with the highest priority value is the DR
- The router with the second highest priority value is the BDR
  - Default priority is 1
  - In a tie the router with the highest Router ID is DR and so on
  - Use a loopback interface with a high IP address (the router ID) to ensure DR or BDR election
  - A router with a priority of 0 will never become DR or BDR
- If a router with a higher priority is added to the network the DR and BDR do not change. Changes occur only if a DR or BDR goes down
- The BDR uses a timer to ensure the DR is up. If the BDR does not receive LSAs from the DR before the timer expires the BDR assumes the DR is down.

Each network segment has it's own DR and BDR. A router connected to multiple segments can have multiple roles, DR in one segment, normal router in another and so on.

## Exchange Process



## Cost & Route Selection

Once a router has a complete link-state database it can create a routing table.

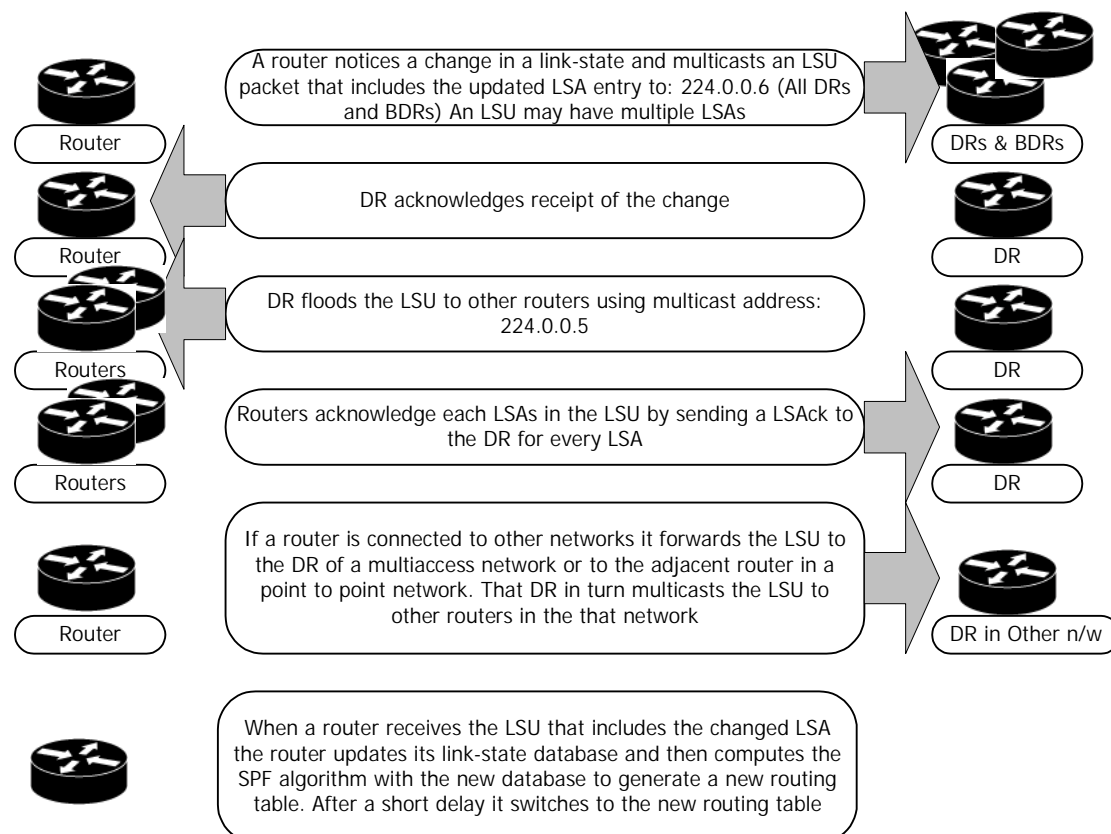
OSPF selects best path based on lowest cost, default cost metrics are based on media bandwidth and therefore 10Mbps Ethernet has a lower cost than a 1Mbps serial line. Cost is calculated as:  $100,000,000/\text{bandwidth in bps}$  so the cost of a 10Mbps Ethernet would = a cost of 10.

OSPF uses the Dijkstra algorithm to calculate the lowest cost to a destination. Using its link-state database as input, a router runs the Dijkstra algorithm to build its routing table. The algorithm adds up the total costs between itself, (the root,) and each destination network and selects the lowest cost path. Up to 6 equal cost paths can be kept in the routing table for load balancing.

To avoid 'flapping' causing constant LSU's and possibly preventing convergence routers wait 5 seconds by default before recalculating its routing table when an LSU is received.

### Route Updates/Changes

When there is a change in a link-state routers use a flooding process to notify other routers in the network of the change. LSUs are used for flooding LSAs.



In Cisco routers if a route already exists the routing table is used at the same time the SPF algorithm is calculating. However if the SPF is calculating a new route the use of the routing table occurs after the SPF calculation is complete.

LSA entries have an ageing timer with a default value of 30 minutes.

When an LSA entry ages the originating router sends an LSU to the network to verify the link is still active. (Distance Vector protocols send the whole routing table.)

When each router receives the LSU:

If the entry doesn't exist it adds the entry to its link-state database

If the entry exists but includes new information it adds the entry to its link-state database

- AND Sends an LSack to the DR, floods the info. to other routers and updates its router table

If the entry exists and the LSU has the same information the router ignores the LSA entry

If the entry exists but the LSU contains old information it sends an LSU to the sending router with the newer info.

## Operation over Point to Point Links

Hello and dead intervals are the same as for Broadcast Multi-access Networks, 10 and 40 seconds.

Routers dynamically detect neighbors using Hello packets, neighboring routers become adjacent when they can communicate directly.

No election, DR or BDR.

Source address in the OSPF packet is usually the address of the outgoing interface but may be that of another interface if **ip unnumbered** is used.

## Operation over Non-Broadcast Multi-Access (NBMA) Networks

Frame relay, ATM and X.25 are examples of NonBroadcast MultiAccess networks, which support many hosts but have no broadcast ability.

Hello and dead intervals are 30 and 120 seconds.

Due to the different possible Frame Relay topologies, full mesh, partial mesh and star, election of a DR and BDR becomes a problem as the DR and BDR need to have full physical connectivity with all routers in the network.

OSPF can run in any of 5 modes of operation depending on the NBMA topology

Attribute	NBMA (Default)	Point to Multipoint (Problematic)	Point to Multipoint NonBroadcast	Broadcast	Point to Point
Standard	RFC 2328	RFC 2328	Cisco	Cisco	Cisco
Command	ip ospf network non-broadcast	ip ospf network point-to-multipoint	ip ospf network point-to-multipoint non-broadcast	ip ospf network broadcast	ip ospf network point-to-point
Topology	Full Mesh (PVC's/Map)	Partial Mesh/Star	Partial Mesh/Star	Full Mesh (PVC's/Map)	Partial Mesh/Star (using subints.)
DR/BDR	YES	NO	NO	YES	NO
Neighbor Statements	YES	NO	YES	NO	NO
Subnet	Same	Same	Same	Same	Different
Hello/Dead	30/120	30/120	30/120	10/40	10/40
Notes	Same as OSPF in a broadcast network. Extra config. Is required to work properly. <b>CPU and b/w intensive</b>	Treats the nonbroadcast network as a collection of point-to-point links. Neighbor relationships only.			

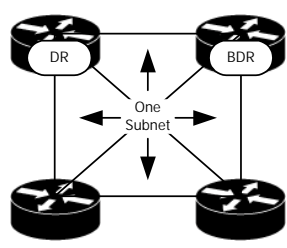
Generally when configuring routers in a NBMA network subinterfaces are used to combat split horizon and problems with distance vector routing protocols.

Subinterfaces can be either point-to-point or multipoint.

- The default OSPF mode on a point-to-point subinterface is point-to-point mode (Cisco)
- The default OSPF mode on a point-to-multipoint subinterface is NBMA (RFC)

### NBMA Mode Neighborhood – RFC 2328

**NBMA Mode (RFC)**



Full or Partially Meshed - If partial, DR and BDR must have full connectivity to other routers

DR and BDR elected (LSA generated for network)

Neighbors are statically configured (administrative overhead/misconfiguration)

LSA's are replicated to be sent to all of an interfaces neighbors

Cannot be used in Frame Relay partial mesh using physical interfaces, not subinterfaces

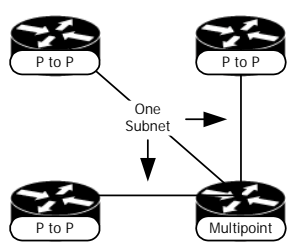
If a PVC on a subinterface goes down the interface is still up, therefore a down link will not be reported

Link state changes are flooded across the network and require the forwarding database to be recalculated, generating traffic and using CPU resources

Link state changes are flooded to the whole area, (other segments,) as per normal, generating traffic, (even though other routes are available.)

**Point-to Multipoint Mode (RFC)**



Full or Partially Meshed - all router to router connections treated as point-to-point links

**NO** DR and BDR elected (**NO** LSA generated for network)

Neighbors are dynamically discovered using OSPF multicast hello packets (as with BMA)

Works by exchanging additional LSUs containing information elements describing connectivity to neighboring routers

LSA's are replicated to be sent to all of an interfaces neighbors

## NBMA Mode Neighborship – Cisco Extensions

### Point-to-Multipoint Non-broadcast Mode

Extension of the RFC allowing static configuration of neighbors when nonbroadcast media is being used, (such as classic IP over ATM.)

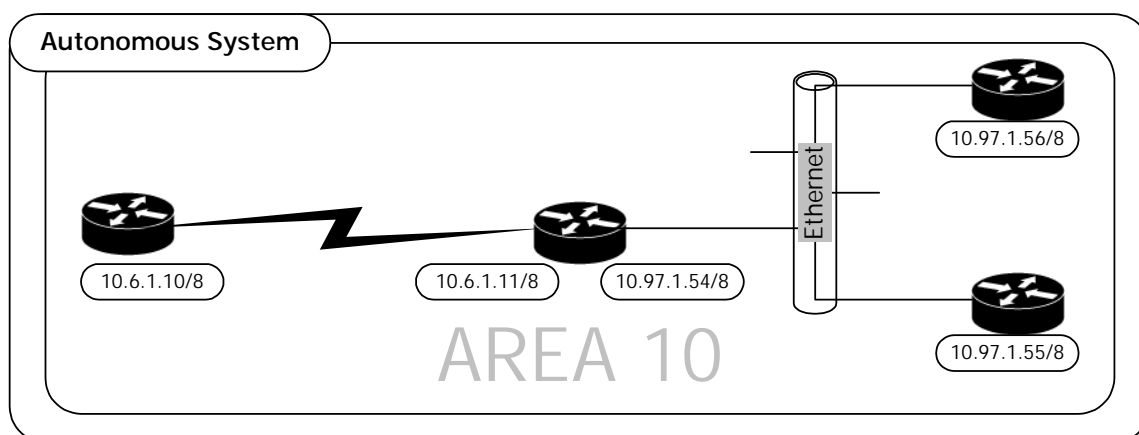
### Broadcast Mode

This is a workaround to statically listing all neighbors. Interfaces are logically set to broadcast and will behave as if connected to a LAN. A DR and BDR are still elected so care should be taken to ensure either a full mesh topology or static selection of the DR based on the interface priority, (or router ID.)  
SIMPLE and FLEXIBLE

### Point-to-Point Mode

Used on point to point subinterfaces, each point-to-point connection has an IP subnet and forms an adjacency. No DR or BDR election

## Configuring OSPF - Internal Routers



```
router(config)# router ospf process-id
```

- Process I.D is a number used internally to identify if multiple OSPF processes are running.
- The I.D does not need to match on other routers, although I would recommend this.
- It is not recommended that you run multiple instances of OSPF.

```
router(config-router)# network 10.0.0.0 0.255.255.255  
area 10
```

- Network to advertise, link to advertise and link to listen on.
- This command specifies the network address.

```
router(config-router)# network 10.6.1.11 0.0.0.0 area 10  
router(config-router)# network 10.9.1.54 0.0.0.0 area 10  
router(config-router)# network 10.9.1.55 0.0.0.0 area 10
```

- Network to advertise, link to advertise and link to listen on.
- This command specifies the interface address.

More specific network commands should be entered first to increase speed, i.e. 0.0.255.255 should be entered before 0.255.255.255.

## OSPF Optional Commands

These commands will only take effect on the election/re-election of the DR and BDR.

### To adjust the router I.D or increase/decrease chances of election as a BDR :

```
router(config)# interface loopback number
router(config)# ip address ip-address network-mask
```

- The router I.D is it's highest IP address. To override this configure a loopback interface with a suitable address.
- This interface is always active and never goes down.
- Recommended for key routers, needs to be a different subnet for every router used on.

An unadvertised, (i.e. no network command entered,) loopback address:

- Does not appear in the OSPF routing table
- Saves on real address space
- Cannot be PINGed

An advertised, (i.e. network command entered,) loopback address:

- Appears in the OSPF routing table
- Uses address space
- Can be PINGed

```
router# show ip ospf interface
```

To display the router I.D

### To control election as a DR or BDR :

```
router(config-if)# ip ospf priority number
```

- Number from 0 to 255, default of 1
- A value of 0 disallows an interface from ever being DR or BDR

### To adjust the OSPF cost of a link to interoperate with non-Cisco equipment:

```
router(config-if)# ip ospf cost cost
```

Number from 1 to 65535

All interfaces connected to the same link must have the same cost.

Cisco routers calculate the cost as follows:

$$10^8/\text{Bandwidth, so the cost of 10Mbps Ethernet} = 10^8/10000000 = 10$$

( $10^8$  is 100M)

To adjust the reference bandwidth used to calculate cost:

```
router(config-router)# auto-cost reference-bandwidth ref-bw
```

Bandwith in Mbps from 1 to 4294967, default is 100Mbps or  $10^8$

## Configuring OSPF over NBMA

### RFC Compliant Modes:

router(config-if)# ip ospf network non-broadcast      Default mode, NBMA mode

router(config-if)# ip ospf network point-to-multipoint      Point to multipoint mode

### Cisco Extension Modes:

router(config-if)# ip ospf network point-to-multipoint non-broadcast      Point to multipoint non-broadcast mode

router(config-if)# ip ospf network broadcast      Broadcast mode

router(config-if)# ip ospf network point-to-point      Point to point mode, default for point-to-point subinterfaces

### Configuring NBMA Mode:

```
router(config)#int serial0
router(config-if)#ip address 192.168.10.23...
router(config-if)#encapsulation frame-relay
router(config-if)#ip ospf network non-broadcast
router(config)#router ospf 50
router(config-router)#network 192.168.10.0 0.0.0.255
area 10
router(config-router)#neighbor 192.168.10.22
router(config-router)#neighbor 192.168.10.21
router(config-router)#neighbor 192.168.10.20
```

Assign interface address

Frame Relay interface  
No need to enter as this is the default mode  
Process I.D 50  
Listen on/advertise this network, area 10  
Static neighbor statement  
"  
"

DR and BDR are elected.

### Configuring Point-to-Multipoint Mode:

```
router(config)#int serial0
router(config-if)#ip address 192.168.10.23...
router(config-if)#encapsulation frame-relay
router(config-if)#ip ospf network point-to-multipoint
router(config)#router ospf 50
router(config-router)#network 192.168.10.0 0.0.0.255
area 10
```

Assign interface address

Frame Relay interface  
Set mode  
Process I.D 50  
Listen on/advertise this network, area 10

No DR and BDR election.

Neighbor statements not required. Single subnet.  
Extra LSUs exchanged. Star topology.

If `ip ospf network point-to-multipoint non-broadcast` is used instead then neighbor statements are required.

### Configuring Broadcast Mode:

```
router(config)#int serial0
router(config-if)#ip address 192.168.10.23...
router(config-if)#encapsulation frame-relay
router(config-if)#ip ospf network broadcast
router(config)#router ospf 50
router(config-router)#network 192.168.10.0 0.0.0.255
area 10
```

```
Assign interface address
Frame Relay interface
Set mode
Process I.D 50
Listen on/advertise this network, area 10
```

DR and BDR are elected.

Neighbor statements are not required.

A full mesh topology or the static selection of the DR based on priority is required.

### Configuring Point-to-Point Mode:

```
router(config)#int serial0
router(config-if)#no ip address
router(config-if)#encapsulation frame-relay
router(config)#interface serial0.1 point-to-point
router(config-subif)#ip address 192.168.12.10...
router(config-subif)#frame-relay interface-dlci 28
router(config)#interface serial0.2 point-to-point
router(config-subif)#ip address 192.168.11.10...
router(config-subif)#frame-relay interface-dlci 27
router(config)#router ospf 50
router(config-router)#network 192.168.0.0 0.0.255.255
area 10
```

```
Frame Relay interface
Create subinterface 0.1
Assign interface address
Assign Frame Relay DLCI
Create subinterface 0.2
Assign interface address
Assign Frame Relay DLCI
Process I.D 50
Listen on/advertise this network, area 10
```

Automatic adjacency.

Default mode for point-to-point subinterfaces.

## Verifying OSPF Operation

```
router#show ip protocols
```

```
Verifies OSPF is configured
```

```
router#show ip route
```

```
Displays routes learnt by the router
Verifies connectivity to the rest of the network
```

```
router#show ip ospf interface (eth 0)
```

```
Displays area ID and adjacency information
Verifies interfaces are in the correct area
Router ID and timer intervals
```

```
router#show ip ospf
```

```
Displays OSPF timers and statistics
Link state update interval
How many times the SPF interval has been run
```

router#show ip ospf neighbor (detail)

Displays information about DR, BDR and neighbors and their state, i.e. init, exstart or full

### Output on a broadcast multi-access network:

Neighbor ID	Pri	State	Dead Time	Address	Interface
10.1.20.10	1	FULL/DR	00:00:27	10.1.20.10	Ethernet0
10.1.20.8	1	2WAY/DROTHER	00:00:31	10.1.20.8	Ethernet0
10.1.20.9	1	FULL/BDR	00:00:21	10.1.20.9	Ethernet0
10.1.20.7	1	2WAY/DROTHER	00:00:10	10.1.20.7	Ethernet0

2WAY/DROTHER means the router has reached adjacency with the DR and BDR and the DR is another router.

### Output on a point-to-point network:

Neighbor ID	Pri	State	Dead Time	Address	Interface
10.1.20.15	1	FULL/ -	00:00:37	10.1.20.15	Serial2

### Output on an NBMA network, NBMA mode:

Neighbor ID	Pri	State	Dead Time	Address	Interface
10.1.20.10	1	FULL/DROTHER	00:00:57	10.1.20.10	Serial2
10.1.20.8	0	FULL/DROTHER	00:00:41	10.1.20.8	Serial2
10.1.20.9	1	FULL/BDR	00:00:38	10.1.20.9	Serial2

Neighbor statements have been used to allow adjacencies to be established, this router is the DR!

### Output on an NBMA network, Broadcast mode:

Neighbor ID	Pri	State	Dead Time	Address	Interface
193.10.78.10	1	FULL/DR	00:00:57	10.1.20.10	Serial2
193.10.78.8	1	FULL/DROTHER	00:00:41	10.1.20.8	Serial2
193.10.78.9	1	FULL/DROTHER	00:00:38	10.1.20.9	Serial2

This router is the BDR, this is a fully meshed network.

router#show ip ospf database(-summary)

Displays the link-state database, router ID and OSPF process ID (summary)  
Additional keywords show different databases

Router Link States are LSA1 point-to-point links  
Net Link States are LSA2 ethernet/token ring links  
Some interfaces generate 2 link counts, for example point-to-point interfaces

router#debug ip ospf adj

Displays adjacency debugging information

# SECTION THREE: Multiple OSPF Areas

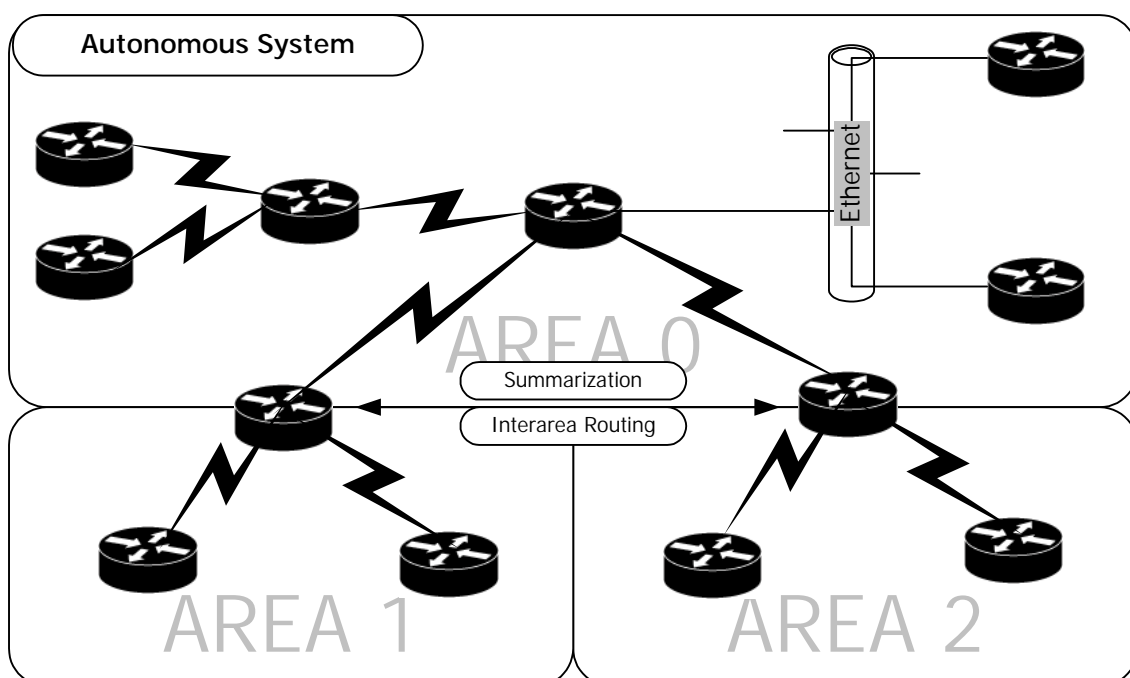
## The Need for Multiple Areas

When an OSPF area becomes too large the following problems may be encountered:

- Repeated SPF algorithm calculations. In a large network changes occur regularly causing routers to recalculate the routing table frequently, using up CPU cycles. A router may be too busy running the SPF algorithm to route.
- All routers will have a large routing table, using up memory and resources.
- All routers will have a large link-state table, the complete topology of the network, again using up memory and resources.

OSPF allows large areas to be split into smaller, more manageable areas, which exchange routing information. This is known as hierarchical routing and has the following benefits:

- Routing still occurs between areas but operations such as recalculating routing tables, are kept within an area.
- Problems in one area are isolated to that area.
- Route summarization between areas reduces routing tables, reducing LSAs.
- Route summarization between areas reduces link-state tables.



## Router, LSA and Area Types

### Multi-area Router Types:

- Internal
    - All interfaces are in a single area
    - Routers within the same area share the same link-state database
  - ABR (Area Border Router)
    - A router with interfaces attached to multiple areas
    - Maintains separate link-state databases for each area and routes between them
    - Summarizes the link-state database for each area and distributes the information into the backbone area
  - ASBR (Autonomous System Boundary Router)
    - A router with an interface in an external network including non OSPF networks
    - Can redistribute into and out of the network to other routing protocols etc.
  - Backbone
    - Placed on the perimeter of the backbone area, at least one interface is connected to area 0
    - Similar to internal routers
- Routers can have more than one role

### Multi-area LSA Types:

LSA Type	Entry Type	Generated By	Shown in Routing Table as
Type 1	Router Link Entry	Router	O
Type 2	Network Link Entry	DR	O
Type 3 & 4	Summary Link Entry	ABR	IA-OSPF Inter Area
Type 5	AS External Link Entry	ASBR	E1-OSPF external type 1 E2-OSPF external type 2

### Area Types:

Area Type	LSAs Accepted	Details
Backbone	Accepts ALL	Interconnects all other areas, always area 0. <b>Prevents routing loops</b>
Standard	Accepts ALL	
Stub	NO external	Uses default route to route outside the AS
Totally Stubby	NO external or summary	Uses default route to route outside the AS and for inter-area routing

### Multi-Area LSUs

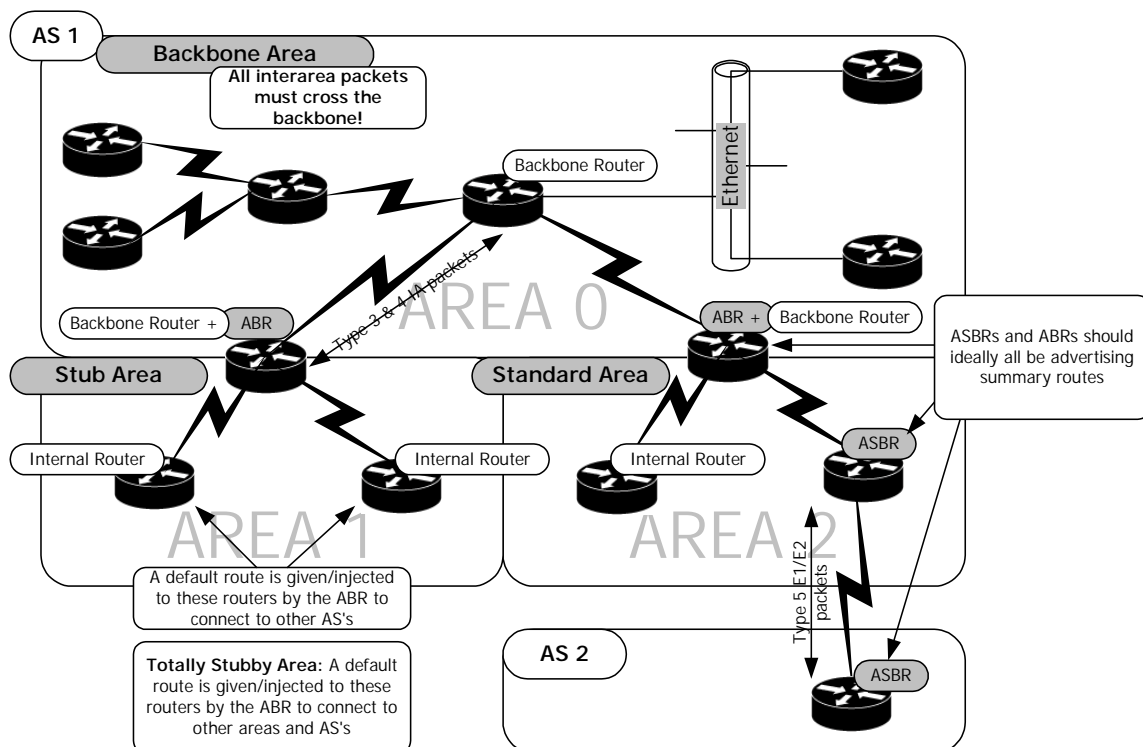
ABRs are responsible for generating routing information about each area to which they are connected and flooding this information through the backbone area to other areas

- The intra-area, (internal) routing process takes place. Before the ABR can begin sending summary LSAs the area must be synchronised.
- The ABR checks the resulting link-state database and generates summary LSAs.

- By default the ABR sends summary LSAs for every network it is aware of, route summarization should be used to reduce the number of entries
- The summary LSAs (type 3 and 4) are placed in an LSU and distributed through all ABR interfaces, except for:
  - If the interface is connected to a neighboring router that is in a state lower than the exchange state
  - If the interface is connected to a totally stubby area
  - If the interface is connected to a stub area and the LSA includes a external (type 5) route
- When an ABR or ASBR receives the summary LSAs it adds them to it's link-state database and floods them to the local area. The internal routers then assimilate the information into their own link-state databases. (Use a stub area to reduce the size of internal routers routing tables.)

Now all routers have received the routing updates and have added them to their link-state databases they must recalculate their routing tables, paths are calculated in the following order:

1. Paths to networks within the area (type 1 and 2) are calculated and added to the routing table.
2. Paths to networks in other areas (type 3 and 4) are calculated. (If a router has inter and intra area paths to the same network it will use the intra-area path, i.e. same area.)
3. Paths to external networks (type 5 – E1/E2) are then calculated.



A packet from a network router in Area 1 destined for a network in Area 2 would follow this route:

- The internal router would send the packet to an ABR for it's area
- The ABR sends the packet through the backbone to the ABR of the destination area
- The destination area ABR then forward the packet to the destination network/router

## Virtual Links

With multi-area OSPF all areas must be physically connected to the backbone area 0.

Sometimes, due to acquisition or organisational changes it may not be possible to connect an area directly to the backbone area. In these cases a virtual link can be used to connect to the backbone.

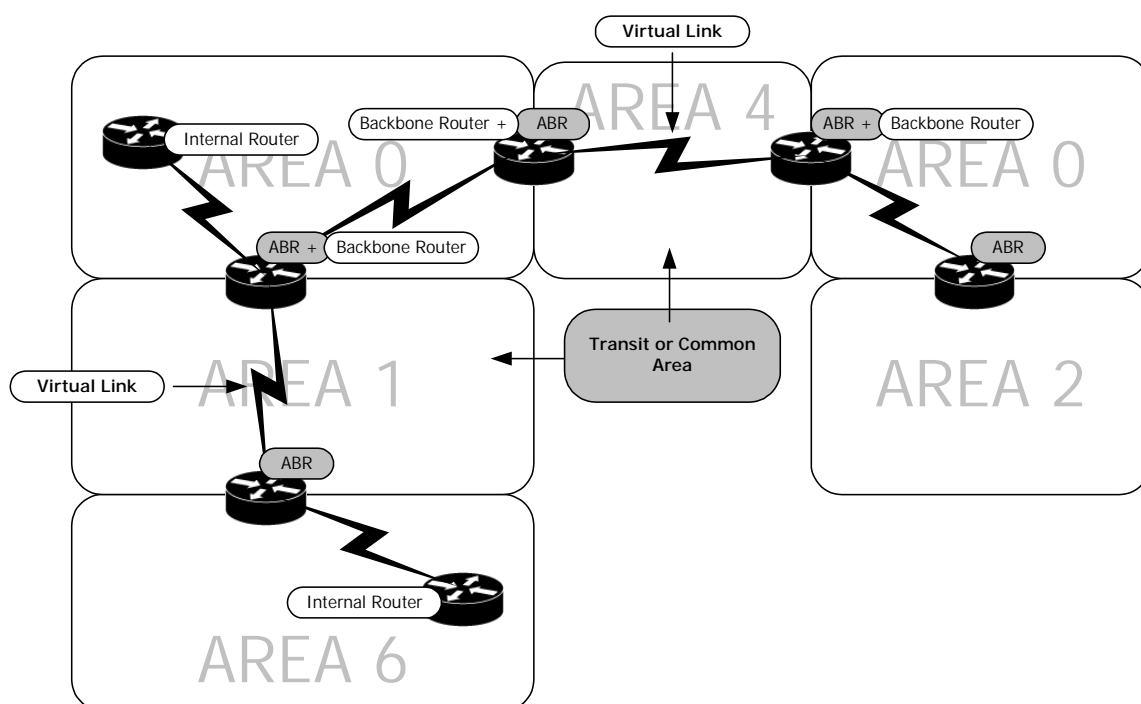
There are only two requirements for creating a virtual link:

- It must be established between two routers that share a common area
- One of the two routers must be connected to the backbone

When used, virtual links require extra processing during the SPF calculation so that the 'real' next-hop router can be determined and the true cost to the destination across the backbone can be calculated.

- Virtual links can also be used to patch discontinuity in area 0
- Provide redundancy if the failure of one router splits area 0 into two

For adjacency purposes two routers connected by a virtual link are treated as if they were connected by an unnumbered point-to-point backbone network.



### Configuring a Virtual Link

On both ABRs in the link:

```
router(config)#router ospf 100
router(config-router)#area 1 virtual-link router-id
```

```
Enable/configure OSPF, process ID 100
Configure a virtual link across area 1 using the router
id of the virtual link neighbor
The area specified should be the transit area!!
```

This is a good reason to use loopback interfaces as this ensures that the router ID does not change.

## Configuring OSPF for Multiple Areas

### ABRs and ASBRs

No special commands are required to configure a router as an ABR or ASBR. The router assumes this role depending on the areas to which it is connected.

For a router to become an ABR two or more network statements are configured with different areas

```
router(config)#router ospf 100
router(config-router)#network 10.20.5.0 0.0.0.255 area 10
router(config-router)#network 192.168.20.0 0.0.255.255
are 30
```

- Enable/configure OSPF, process ID 100
- Route/advertise etc. this network, area 10
- Route/advertise etc. this network, area 30

For a router to become an ASBR an interface must be configure/connect to a non-OSPF network

## Stub & Totally Stubby Areas

### Stub Area (RFC)

- Reduces internal router memory requirements, link-state database etc.
- Denies external networks, type 5 LSAs
- ABR sends a 0.0.0.0 LSA which all internal routers use to route to any network not in their routing tables
- Typically used in hub and spoke topology with the stub being a spoke

### Totally Stubby Area (Cisco)

- Further reduces internal router memory requirements, link-state database etc.
- Denies external networks, type 5 LSAs AND summary inter-area routes, type 3 and 4 LSAs
- ABRs sends a 0.0.0.0 LSA which all internal routers use to route to any network not with the area
- Better than stub areas unless non-Cisco equipment used

### Stub Area Restrictions

- Ensure there is only a single exit point from the area, if there are more than one ABRs all will inject a default route and a packet may take a less than optimal path to another area or AS as it may use the default route to an ABR furthest from the destination.
- All internal and ABR routers must be configured as stub routers.
- No ASBR is internal to the stub area.
- The area is not the backbone area 0.

**Stub or Totally Stubby Areas are particularly useful in reducing traffic over NBMA networks.**

## Configuring Stub & Totally Stubby Areas

### Stub Area

Configure on all routers:

```
router(config)#router ospf 100
router(config-router)#area 10 stub
router(config-router)#area 10 default-cost x
```

- Configure OSPF, process ID 100
- Make area 10 a stub area
- Define cost of default route injected into area
- Optional for ABRs, default is 1

### Totally Stubby Area

Configure as stub on all routers and this on ABRs only:

```
router(config)#router ospf 100
router(config-router)#area 10 stub no-summary
router(config-router)#area 10 default-cost x
```

- Configure OSPF, process ID 100
- Make area 10 a totally stubby area
- Define cost of default route injected into area
- Optional for ABRs, default is 1

## Summarization and VLSM

Summarization has the following benefits in a multi-area OSPF topology

- It minimizes the number of routing table entries
- It localizes the impact of topology changes
- It reduces LSAs and saves on CPU and memory usage

OSPF supports VLSM because

- It carries subnet mask information (classless)
- It allows for the use of hierarchical addressing schemes

### Configuring Route Summarization

- Summarization is not enabled by default.
- If more than one ABR is present between two areas you may not wish to summarize as a sub-optimal path may be selected. Ensure you don't summarize discontinuous networks.

On an ABR:

```
router(config)#router ospf 100
router(config-router)#area 10 range address mask
```

```
Configure OSPF, process ID 100
Address range and mask to summarise
```

On an ASBR:

```
router(config)#router ospf 100
router(config-router)#summary-address address mask
```

```
Configure OSPF, process ID 100
Address range and mask to summarise
```

To a summary for a group of networks where you don't own only a few subnets use a very specific network statement pointing to Null0.

## Costs for Summary & External Routes

### Summary Routes:

- The **smallest** cost of a given inter-area route that appears in the summary
- **Plus** the cost of the ABR link to the backbone

### External Routes:

- Cost depends on the external type packet configured on the ASBR Router
  - Type 1 (E1) The sum of the cost of the internal and external links the packet crosses. Used when multiple ASBRs are advertising a route to the same AS.
  - Type 2 (E2) (Default) The cost of the external links only. Used if only one ASBR router advertises a route to the AS. Preferred over type 1 routes unless two same-cost routes exist to the same destination.

## Verifying OSPF Operation

router#show ip ospf border-routers

Lists ABRs in the AS

router#show ip ospf virtual-links

Displays info. on the current status of virtual links

router#show ip ospf *process-id*

Displays info. on each area that the router is connected to. Indicates if the router is an ABR or ASBR

# SECTION FOUR: EIGRP

# EIGRP Facts

## General

- Enhanced Interior Gateway Routing Protocol
- Interior Gateway Protocol (IGP)
- RFC ??
- Uses the DUAL, Diffusing Update Algorithm
- Advanced distance vector (really a hybrid, both link state and distance vector)
- Protocol no. 88
- Administrative Distance: 90 (IGRP is 100)

Frame Header	Frame Payload			CRC
	IP Header	Protocol No.		Packet Payload
		88	EIGRP	
		6	TCP	
		17	UDP	

## Benefits

- Rapid convergence due to the storing of backup routes and neighbor querying
- Reduced bandwidth usage as there are no periodic updates, EIGRP sends partial updates only when paths or metrics for a route change
- Updates send only the changed information and only to routers that require it
- EIGRP supports Appletalk, IP and Novell Netware (Multiple network layers)

## Features

- 100% Loop free (claimed)
- Easy configuration (no areas)
- Fewer design constraints compared to OSPF
- Classless, supporting summarization and VLSM
- Compatible with existing IGRP networks, (if the same AS)

## Advantages

- Uses multicast instead of broadcast reducing bandwidth usage
- Cost utilizes link bandwidth and delay (32bit as opposed to 24bit IGRP)
- Unequal cost path load balancing (unavailable with OSPF and IGRP)
- Summary routes can be used anywhere, not just at major network boundaries

## Addresses

- 224.0.0.10      Hello packet multicast address

## EIGRP Terminology

Interface	Connection between a router and attached network a.k.a link
Neighbor Table	Lists adjacent routers. Ensures bi-directional communication between each directly connected neighbor. One table for each protocol supported.
Topology Table	Route entries for all the destinations that the router has learnt. Could be multiple routes to the same destination. One table for each protocol supported.
Routing Table	The best (successor) routes to a destination network taken from the topology table. One table for each protocol supported.
Successor	The primary route selected to be used to reach a destination. These are the entries kept in the routing table.
Feasible Successor	A backup secondary route selected at the same time as the successor, but kept in the topology table. There can be many feasible successors.
Hello	Protocol used to create and maintain neighbor relationships

## Packets

### Hello

- Used for neighbor discovery
- Sent as multicasts and with a 0 acknowledgement number

### Update

- Communicates routes that a particular router has converged on
- Sent as multicasts when a new route is discovered or when convergence has completed
- Sent as unicasts when neighbors start up to synchronize topology tables (as updates are not sent periodically)

### Queries

- Sent when a router does not have a feasible successor
- Multicast packet sent to neighbors asking if they have a feasible successor for the destination

### Replies

- Sent in response to a query packet
- Unicast, sent only to the originator of a query

### ACK

- Used for acknowledging other types of packets
- ACKs are hello packets that are sent as unicasts with a non-zero acknowledgement number

**Update, Query and Reply packets are all sent reliably and require an acknowledgement**

## Neighbor Relationships

Routers send periodic hello packets out of all EIGRP configured interfaces.

When a router receives another routers hello packet, with the same AS number, it establishes a neighbor relationship – adjacency.

Neighbor tables are used to record neighbors, which interface the neighbor was learnt through, information learnt from the neighbor and a hold time.

The hold time, specified by the neighbor in its hello packets, is the amount of time a router treats the neighbor as reachable, this is reset every time a hello packet is received. The hold time is normally three times the hello interval

Hello interval of 5 seconds/Hold time of 15 seconds on:

- Ethernet, Token Ring, FDDI,
- Point-to-Point Serial Links (PPP, HDLC)
- Point-to-Point FR and ATM Subinterfaces
- Multipoint Circuits with a bandwidth greater than T1 such as ISDN PRI, SMDS, Frame Relay

Hello interval of 60 seconds/Hold time of 180 seconds on:

- Multipoint Circuits with a bandwidth LESS than T1 such as ISDN BRI, Frame Relay, SMDS

When a hello packet is not heard before the hold time expires

- A topology change is detected
- The neighbor adjacency is deleted
- All topology table entries learnt from the neighbor are deleted

To adjust the hello interval (per interface):

```
router(config-if)#ip eigrp hello-interval xx
```

Change hello interval to xx seconds

To adjust the hold time (per interface):

```
router(config-if)#ip eigrp hold-time xx
```

Change hold time to xx seconds

Neighbors will form relationships even if the hello interval and hold time do not match

**Neighbors will not form relationships in the following circumstances:**

- Over secondary addresses, the primary address is always used
- If routers reside in different autonomous systems
- If routers' metric calculation mechanism (K value) differs for a particular link

## Neighbor Tables

Neighbor tables include:

- H Value – The order in which neighbors were learnt
- Neighbor Address – Network Layer address of the neighbor
- Queue – The number of packets waiting to be sent to this neighbor. A non-zero value indicates congestion
- Smooth Round Trip Timer (SRTT) – The average time it takes to send and receive packets to/from a neighbor. Used to determine the retransmit interval, (RTO.)
- Hold Time – The amount of time a router treats the neighbor as reachable, this is reset every time a hello packet OR any EIGRP packet is received from the neighbor

To display the neighbor table:

```
router#show ip eigrp neighbors
```

Displays the neighbor tables

H	Address	Interface	Hold (sec)	Uptime	SRTT (ms)	RTO	Q Cnt	Seq Num
1	10.1.20.6	To1	6	08:51:22	9	160	0	7
2	10.1.18.9	Eth0	9	04:32:48	26	160	0	6
0	10.1.18.5	Eth0	12	04:30:12	21	160	0	5

## Reliability

Reliable Transport Protocol is responsible for the guaranteed ordered delivery of EIGRP packets to neighbors.

Only Update, Query and Reply packets are transmitted reliably.

On multi-access networks with multicast abilities hello packets are sent via a multicast that does not require acknowledgement

RTP can send multicast packets quickly when unacknowledged packets are pending, which ensures low convergence times in the presence of varying speed links.

In general reliability is achieved through sequence numbers and acknowledgements to routing information packets.

RTP keeps a neighbor list and a retransmission list for every neighbor. Packets that have not been acknowledged by a neighbor are retransmitted up to 16 times, this ensures delivery of critical routing information to neighboring routers and a loop free topology.

The neighbor relationship is reset when the retry limit of 16 is reached.

Potential delays can occur on multi-access media with multiple neighbors

Each reliable packet sent as a multicast to the neighbors must be acknowledged by all neighbors before the next packet can be sent (window size of 1)

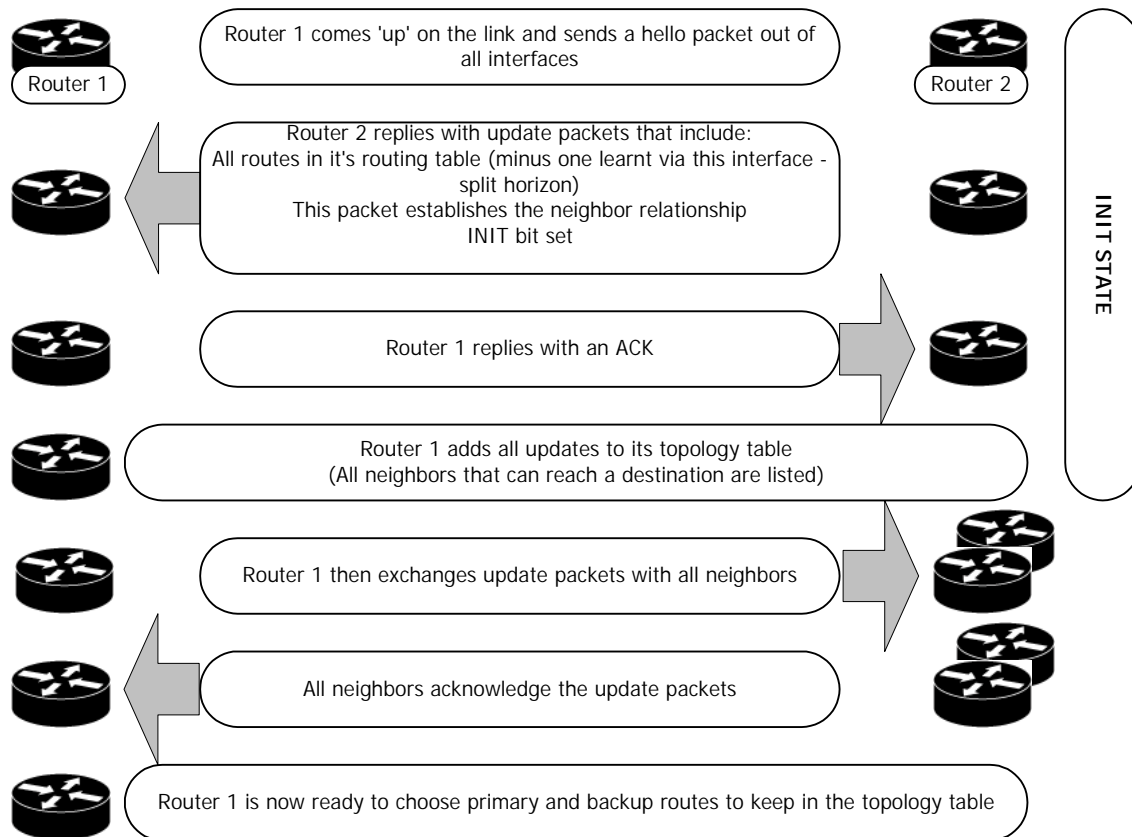
If a peer is slow to respond it delays the next transmission and affects the faster peers

To avoid this RTP

Retransmits the packet as a unicast to the slow neighbor(s)

It can then continue sending the next multicast packet(s) to all neighbors

## Route Discovery



## Route Selection

EIGRP route selection distinguishes it from other routing protocols.

- EIGRP selects primary and backup routes that are kept in the topology table (up to 6 per destination)
- Primary routes are then moved to the routing table

The metric used to determine the best path is based on five criteria:

1. Bandwidth – smallest between source and destination
2. Delay – Cumulative interface delay long the path

The following are not recommended for use as they may cause frequent recalculation of the topology table:

3. Reliability – Worst between source and destination based on keepalives
4. Loading – Worst between source and destination based on bits per second
5. MTU – Smallest MTU in path

The DUAL algorithm is used to calculate the best route to a destination, based on the composite metric and ensures a loop routing table.

# DUAL

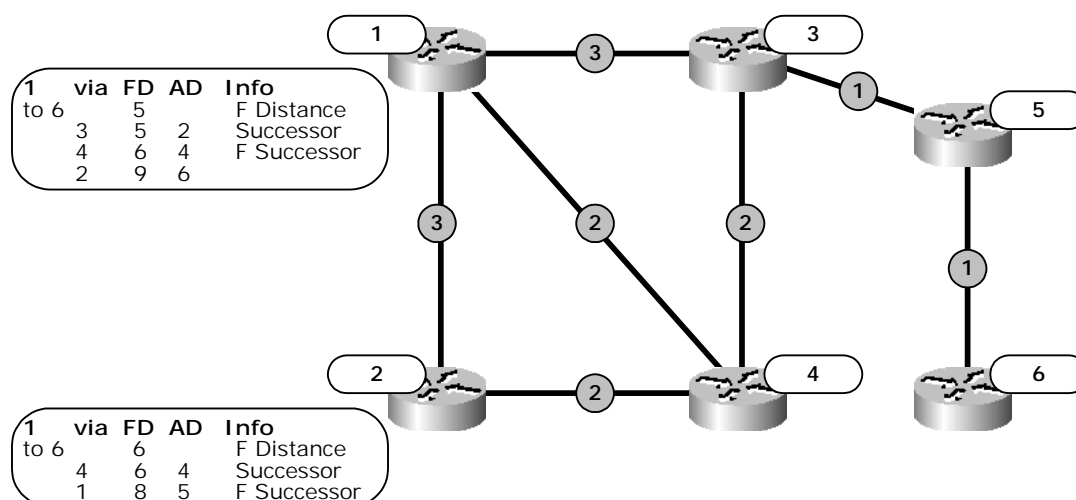
DUAL, Diffusing Update Algorithm:

- Tracks all routes advertised by neighbors
- Distance information is used to select efficient loop free paths
- Lowest cost routes are calculated as follows
  - The cost between the next hop router and the destination (advertised distance – AD)
  - PLUS the cost of between this router and the next hop router
  - The total is the feasible distance – FD
- The successor is the neighboring router that has the least cost path to a destination (and not part of a routing loop)
- Multiple successors can exist if
  - They have the same feasible distance
  - They use different next hop routers

All successors are added to the routing table

Next hop routers for the backup path are feasible successors; they must have an advertised distance less than the feasible distance of the current successor route.

**This prevents loops; an AD greater than the FD may be using this very router to reach the destination. A router with an equal or greater AD is as far or further away from the destination than this router!**



An active route is one where a change has occurred and there is no successor/feasible successor, i.e. queries are being sent or waiting to be received

A passive route is one where a successor has been found.

If a route goes active the path that changed/failed is shown as unused.

## Configuring EIGRP

router#**router eigrp *as-number***

- Enable EIGRP using this AS number
- The AS number must match on all routers in the network

router(config-router)#**network 10.0.0.0**

- Network to advertise
- Also determines which interfaces will participate in EIGRP

On serial links such as Frame Relay or SMDS ensure the bandwidth is correctly specified otherwise it will be assumed to be of T1 speed and the router may not be able to converge or routing updates may be lost:

router(config-if)#**bandwidth 64**

- Set the bandwidth to 64Kbs

# Summarization

## Automatic

Automatic summarization is enabled by default and takes place at major class or network boundaries.

Subnets are summarized into a single **classful** network.

To disable automatic summarization:

```
router(config-router)#no auto-summary
```

## Manual

Can be configured on an interface on any router within the network

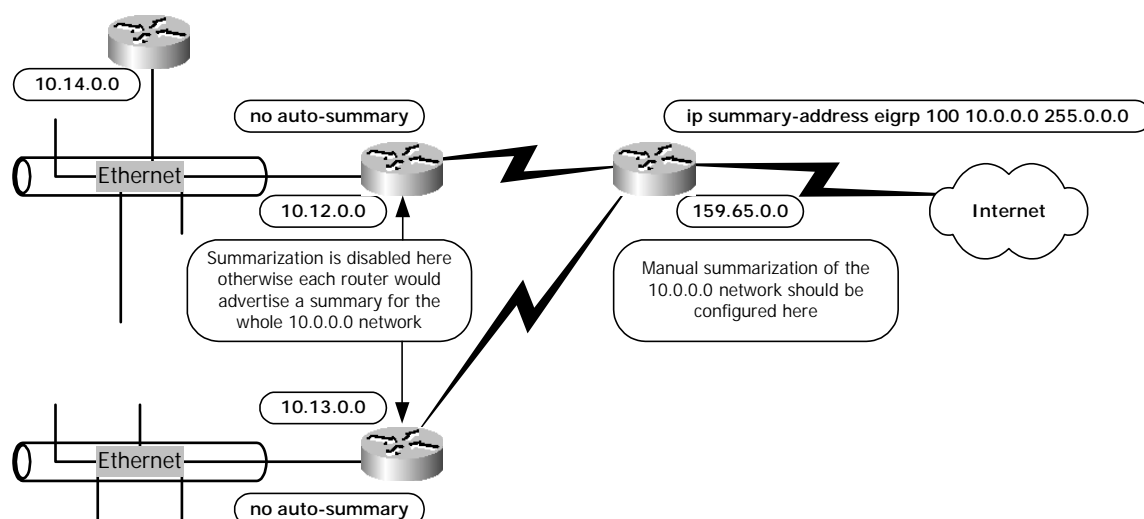
When a summary route is created it is added to the routing table with a reference to Null0. This prevents the router from trying to forward traffic to other routers in search of a longer match contained within the aggregate route and prevents traffic from looping within the network.

The minimum metric of the specified routes is used as the metric of the summary route.

```
router(config-if)#ip summary-address eigrp as address  
mask
```

- Create a summary route on this interface for eigrp.
- as is the autonomous system number
- Summary address and mask

A manual summary is only advertised if an entry represented in the summary is present in the routing table.



## Load Balancing

### Equal-Cost Load Balancing

Routes with a metric equal to the **minimum** metric will be added to the routing table.

There can be up to 6 entries to the same destination in the routing table but the default is 4.

### Unequal-Cost Load Balancing

EIGRP can be configured to route across routes that have different metrics using the variance command.

The variance multiplier can be a number from 1 to 128 and has a default of 1 (equal cost)

The variance multiplier defines the range of metrics that can be used for unequal-cost load balancing

If a successor has a feasible distance of 12 a variance of 2 would allow routes with a feasible distance of up to 24 to become successors.

```
router(config-router)#variance 2
```

Allow load balancing of routes under twice the lowest feasible distance

## WAN Links

EIGRP support the following WAN link types:

- Point-to-Point
- NBMA
  - Multipoint
  - Point-to-point

Configuration of these links must address:

### Bandwidth

EIGRP utilizes up to 50% of the bandwidth of a link by default, to change this use the following command on the interface:

```
router(config-if)#ip bandwidth-percent eigrp as nnn
```

- EIGRP can only use nnn% of the bandwidth available on this link.
- This number can be over 100% if you have set the bandwidth artificially low for any reason.

When using Frame Relay point-to-point subinterfaces the IOS assumes the bandwidth is that of a T1 link. As this is normally not the case you should configure the bandwidth as the CIR of the PVC.

When using multipoint Frame Relay, ATM, SMDS or ISDN PRI EIGRP uses 50% of the bandwidth of the main interface divided by the number of neighbors. For example with ISDN PRI using 4 64k lines set the bandwidth to 256Kbs so EIGRP will use  $256 / 4 / 2 = 32k$  of the bandwidth.

If using multipoint interfaces you should configure the bandwidth to be the minimum CIR times the number of circuits so the circuits with low CIRs are not congested. (Note the higher CIR circuits may not be fully utilized.)

Alternatively configure the low CIR circuits as point-to-point subinterfaces so the bandwidth can be properly set. Hybrid Multipoint – this is the preferred solution.

Or use subinterfaces for each DLCI.

## Scalability

**These factors affect how scalable an EIGRP network is:**

- The amount of information neighbors exchange, if more information than is necessary is passed in order for routing to function correctly EIGRP has to work harder at neighbor start-up and when reacting to network changes
- When a change occurs, the resources used by EIGRP are directly related to the number of routers that must be involved in the change
- Depth of topology is an issue if information must be propagated through many hops for convergence. An example would be a very large network with no summarization
- Too many alternative paths and therefore complexity can create problems with EIGRP convergence

**Solution Overview:**

- Fine tune EIGRP in large networks
- Limit the query range
- Summarize as much as possible and use summaries and default routes at regional or remote sites

## Queries

**When searching for an alternative for a lost route a router puts the route in active mode and sends queries to all neighboring routers, these are propagated throughout as far as all the routers that had knowledge of the route.**

The router must get replies (acknowledgments) from all of the routers that have been queried before recalculating successor information.

If a router fails to reply to a query in 3 minutes the route is STUCK IN ACTIVE and the router resets the neighbor in question. This is a very serious problem!

To overcome this problem you should limit the scope of query propagation through the network limiting the chances of an isolated failure in another part of the network causing convergence problems on the local router.

Queries do cross AS boundaries, however the boundary router replies to the original query and then sends a new query to the other AS transferring the problem to the other AS.

### **Limiting the query range using summarization**

Manual or automatic summarization is the best way to limit queries

When a query reaches an interface with a summary the router in question replies with infinity and the query stops there.

### **Essentially good design overcomes this problem**

Intelligent addressing schemes aid summarization; regions should have a contiguous address space

A tiered network design should be used

Summarization limits the routes/subnets a router is aware of (this removes remote routers from the convergence process)

The use of default routes limits the routes/subnets a router is aware of (this removes remote routers from the convergence process)

Minimize the redundancy deployed in the topology otherwise a single router may receive many queries about the same route/subnet

Packet filtering can also help

Ensure routers at convergence points on the network have the necessary memory for the task

WAN links should have the required capacity for the EIGRP overhead and other traffic, lost packets will cause convergence problems

Ensure WAN links are configured with the correct bandwidth using the **bandwidth xx** command

## Verifying EIGRP Operation

router#show ip eigrp neighbors	Display neighbors discovered by EIGRP
router#show ip eigrp topology	Display the topology table, status of routes, successors and feasible distances
router#show ip route eigrp	Display EIGRP entries in the routing table
router#show ip protocols	<ul style="list-style-type: none"><li>• Displays parameters and the current state of active routing processes.</li><li>• Displays the AS number</li><li>• Displays neighbors and distance information</li></ul>
router#show ip eigrp traffic	<ul style="list-style-type: none"><li>• Displays EIGRP packet statistics including hellos, updates, queries, replies and ACKs</li></ul>
router#show ip eigrp traffic	<ul style="list-style-type: none"><li>• Displays EIGRP packet statistics including hellos, updates, queries, replies and ACKs</li></ul>
router#debug ip eigrp packet	Displays all types of packets sent and received
router#debug ip eigrp neighbor	Displays neighbor interaction
router#debug ip eigrp route	Displays advertisements and routing table changes
router#debug ip eigrp summary	Displays brief report of routing activity
router#debug ip eigrp events	Displays the types of packets sent and received Displays statistics on route compilation

# SECTION FIVE: BGP

# BGP Facts

## General

- Border Gateway Protocol
- Exterior Gateway Protocol (EGP)
- RFC 1771
- Uses the ? Algorithm
- Advanced distance vector
- PORT NO. 179 (NO protocol number – carried within TCP packets)
- Administrative Distance: 20
- AS numbers 64512 to 65530 are for private use

Frame Header	Frame Payload			CRC
	IP Header	Protocol No.	Packet Payload	
		6 17	TCP UDP	Port Number: 179

## Benefits

- Does not require a hierarchical topology
- The effects of BGP are well understood

## Features

- Provides an interdomain routing system that guarantees the loop free exchange of routing information between AS's.
- Routers exchange information on paths to destination networks
- v4 Used extensively in the Internet

# BGP Terminology

Path Vector or attribute	BGP's metric; network reachability information that includes a list of the full path (as AS no's) that a route should take in order to reach a destination network.
--------------------------	---

### A BGP Internet router has

- A routing table larger than 30Mb
- Over 90,000 routes
- Around 5000 AS numbers
- These figures are growing constantly

Telnet to: [route-server.cerf.net](http://route-server.cerf.net) for the latest statistics from a live Internet router

### Appropriate When

- An AS allows packets to traverse it to reach another AS (ISP)
- An AS has multiple connections to other AS's

### Not Appropriate For

- Single connections to the internet or another AS
- Routing policy and route selection are not a concern
- Underpowered routers without the necessary memory or CPU resources
- If you don't understand route filtering and the BGP path selection process
- When there is low bandwidth between AS's
- **USE STATIC ROUTES INSTEAD**

## Using Static Routes Instead

```
router(config)#ip route 0.0.0.0 0.0.0.0 address / interface
(distance)
```

- Prefix and mask for the destination network
- The address of the next hop router of the interface to be used
- Administrative distance

### Floating Static Routes

The default administrative distance of a static route using the *address* parameter is 1; using the *interface* parameter it is 0.

To create a floating static route specify a *distance* that is larger than the dynamic routing protocols administrative distance. The route will then only be used when there is no other path available.

### Default Routes with RIP

If there is no matching route in the routing table then the default route will be used. RIP propagates the default route automatically.

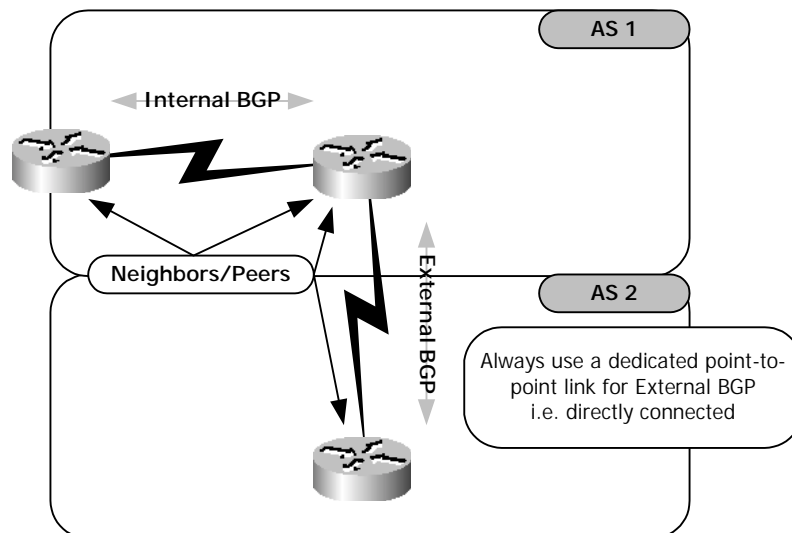
### Default Routes with OSPF

Use the command: **default-information originate always** to propagate the route into the OSPF routing domain. The **always** keyword ensures the static default route is always advertised whether or not the router already has a default route (from an ABR for example) and even if the next hop router or interface go down.

## Characteristics

- Reliable updates – runs on top of TCP, port 179 (Doesn't have its own retransmission or error recovery mechanism)
  - Two BGP peers or neighbors form a TCP connection with each other and exchange messages to open and confirm the connection parameters
  - Full routing tables are then exchanged
- Incremental, triggered updates only
  - Only changes are sent after the full routing tables have been exchanged
- Keepalives to verify TCP connectivity
  - Similar to OSPF or EIGRP hello packets
- Rich metrics (path vectors/attributes)
  - Routers exchange network reachability information, called path vectors that include a list of the full path (of BGP AS numbers) that a route should take to a destination network
  - This information is used to create a graph of AS's
  - Routing restrictions can then be applied, enforcing restrictions on routing behaviour
  - BGP is loop free as it will not accept a routing update that includes it's own AS in the path list
- Design to scale massively
- Has it's own routing table, separate from that of the IGP although information can be exchanged between the two

BGP routers can be internal or external to an AS



Internal BGP (IBGP) routers do not need to be directly connected as long as they can reach each other (via an IGP running in the AS)

## Policy Based Routing

A BGP router can advertise to its peers in neighboring AS's only those routes that it itself uses; this reflects the hop-by-hop routing paradigm used throughout the Internet.

Some policies cannot be supported by the hop-by-hop routing paradigm and require techniques such as source routing to enforce. You cannot influence how a neighbor AS will route your traffic but you can influence how your traffic gets to a neighbor AS.

## BGP Update Message Attributes

### Attributes Can Be:

#### Well Known Attributes

- Must be recognised by all compliant BGP implementations
- Are propagated to other neighbors
  
- Well Known Mandatory
  - Must be present in all update messages (route descriptions)
- Well Known Discretionary
  - Could be present in update messages (route descriptions)

#### Optional Attributes

- Recognised by some implementations (could be private)
- No expected to be recognised by everyone
- Recognised optional attributes are propagated to other neighbors based on their meaning
  
- Optional Transitive (Partial)
  - If not recognized they are marked as partial and propagated to other neighbors
- Optional Nontransitive
  - Discarded if not recognised

### Attributes Are:

AS-path – (Well Known Mandatory)  
Next Hop – (Well Known Mandatory)  
Origin - (Well Known Mandatory)

Local Preference – (Well Known Discretionary)  
Atomic Aggregate – (Well Known Discretionary)

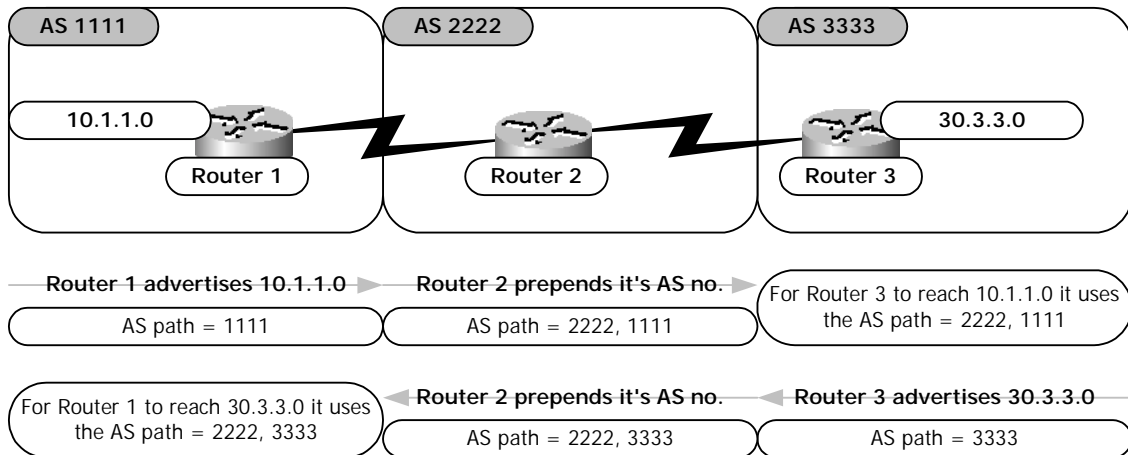
Aggregator – (Optional Transitive)  
Community – (Optional Transitive)

Multi-Exit-Discriminator (MED) – (Optional Nontransitive)

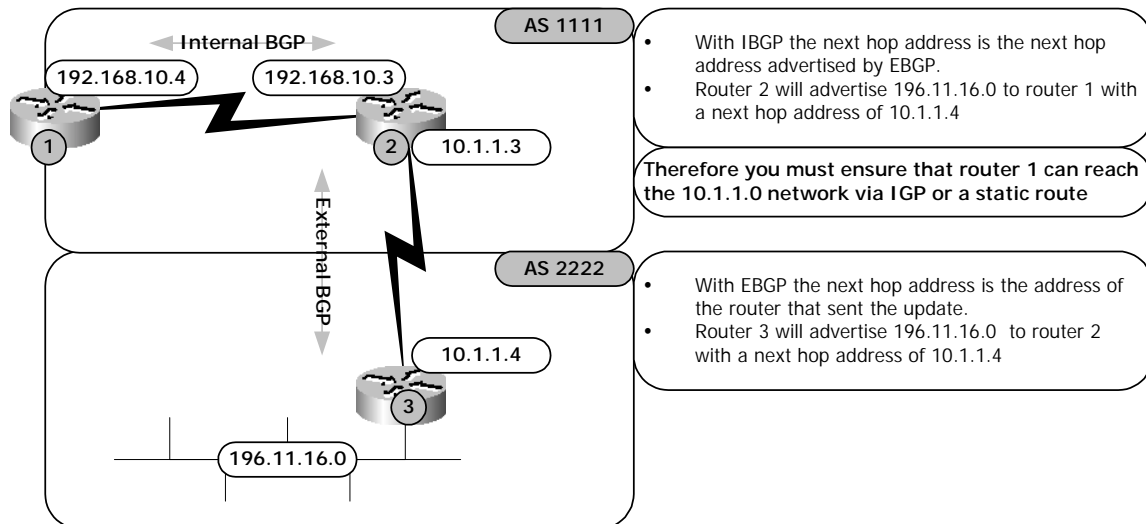
Cisco has also defined a weight attribute for BGP

### AS-Path Attribute – Well Know Mandatory

A list of the AS's a route update has passed through in order to reach a destination. As the update passes through an AS the AS number is prepended to the front of the list.

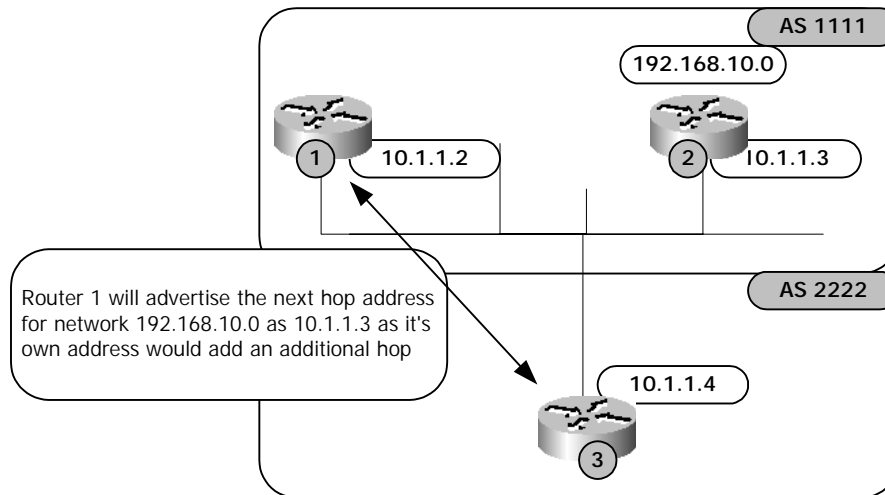


### Next Hop Attribute – Well Known Mandatory



### Next Hop Attribute on a Multi-access Network

When running BGP over a multi-access network such as Ethernet a BGP router will advertise a network with the most appropriate next hop address, avoiding placing additional hops into the network.



### Next Hop Attribute on an NBMA Network (Frame Relay)

Same as above, however if router 3 does not have a connection to router 2, (not fully meshed,) a problem will arise. In this case router 1 will have to be configured to advertise itself.

### Local Preference Attribute – Well Known Discretionary

**IBGP only.** Provides an indication to routers in an AS which is the preferred path used to exit the AS. A higher local preference is preferred. Default of 100 on Cisco routers. Passed to other routers.

### MED Attribute – Optional Nontransitive

Also known as the metric. An indication to EXTERNAL neighbors about the preferred path INTO an AS, a dynamic way of influencing another AS as to which path it should choose to reach a certain route if there are multiple entry points into an AS. (Not normally used between ISPs)

A lower value is preferred. The MED is exchanged between AS's but not propagated.

A router, by default, only compares MED's for paths from neighbors in the same AS.

### Origin Attribute – Well Known Mandatory

Defines the origin of path information, can be one of 3 values

- i – IGP, the route origin is interior to the AS. Normally occurs when the network command is used in BGP
- e – EGP, the route origin is learnt via Exterior Gateway Protocol
- ? – Incomplete, the route origin is unknown or learnt via some other means. Normally due to route redistribution

### Community Attribute – Optional Transitive

Used to filter incoming or outgoing routes. Allows routers to tag routes with a community indicator and allow other routers to make decisions/route selections based upon the tag.

By default, communities are stripped in outgoing BGP updates.

### Weight Attribute – Cisco

A Cisco defined attribute used for path selection. Configured locally on a router and not propagated to other routers.

Value from 0 to 65535.

By default paths that a router originates have a weight of 32768 and other paths have a weight of 0.

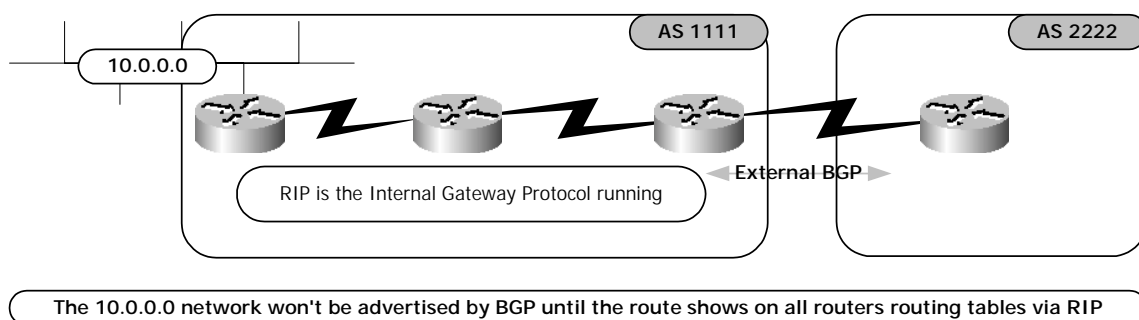
Higher weight is preferred when multiple routes exist to the same destination.

## The Synchronisation Rule

The rule: Do not use or advertise to an external neighbor a route learnt by IBGP until a matching route has been learnt from an IGP or unless the route is local

- This ensures consistency of information throughout the AS, all routers in the AS should know about the route
- Avoids black holes within the AS
- Can be turned off when all routers in the AS are running BGP

On by default



# BGP Messages

## Open Message

Once a TCP connection is established the first message sent by each peer is an open message.

If the open message is acceptable a keepalive message confirming the open is sent back. The BGP connection is now established and other messages may now be exchanged.

An open message contains:

- Hold time – maximum time in seconds that may elapse before receipt of a keepalive or update message from the sender. Routers use the smallest hold time; it's own or one sent in an open message.
- BGP router identifier – 32bit field router ID, same selection as with OSPF.

## Keepalive Message

Exchanged between peers often enough to keep the hold timer from expiring. If set to 0 keepalive messages are not sent. The message contains only a message header.

## Update Message

Information on one path only, multiple paths require multiple messages. May include the follow fields

- Withdrawn routes – List of IP address prefixes for routes that are being withdrawn from service
- Path Attributes – As discussed, includes attribute type, length and value
- Network layer reachability information – List of IP address prefixes that can be reached by this path

## Notification Message

Send when an error is detected. The BGP connection is closed once the message is sent. Includes and error code, subcode and data related to the error.

## Route Selection Decision Process

Once updates regarding destinations are received from different AS's BGP decides which path to chose to reach a destination, only a single path is chosen.

This is how a Cisco router decides on the best route when there are multiple routes to the same destination:

1. If the path is internal, synchronisation is on and the route is not synchronised do not consider it
2. If the next hop address of the route is not reachable do not consider it
3. Prefer the route with the highest weight (local setting)
4. If multiple routes with the same weight prefer route with highest local preference (internal AS setting)
5. If multiple routes with same local preference prefer route that was originated by the local router
6. If multiple routes with same local preference and none originated by the local router prefer the route with the shortest AS path
7. If AS path is the same prefer the lowest origin code (IGP<EGP<Incomplete)
8. If origin codes are the same, prefer the path with the lowest MED (External AS setting)
9. If the routes have the same MED prefer external paths to internal paths
10. If synchronisation is disabled and only internal paths remain, prefer the path through the closest IGP neighbor. (Use shortest internal path to reach the destination)
11. For EGBP paths select the oldest route, this minimises the effects of routes flapping
12. Prefer the route with the lowest neighbor router ID

The path is then put in the routing table and propagated to the routers BGP neighbors

## CIDR and Aggregate Addresses

BGP v4 supports CIDR

- BGP update messages include both the prefix and prefix length
- Addresses can be aggregated when advertised by a BGP RouterZone
- The AS path attributes can include a combined list of all AS's that all of the aggregated routes have passed through and should be considered to ensure the route is loop free.

Two attributes relate to aggregate addressing

- Atomic Aggregate – A well know discretionary attribute that informs the neighbor AS the originating router has aggregated routes.
- Aggregator – An optional transitive attribute that specifies the BGP router ID and AS no. of the router that performed the route aggregation

## Configuring BGP

<pre>router(config)#router bgp <i>autonomous-system-no</i></pre>	<ul style="list-style-type: none"> <li>• Activate and configure BGP for this AS no.</li> </ul>
<pre>router(config-router)#neighbor (<i>ip-address/peer-group-name</i>) remote-as <i>autonomous-system-no</i></pre>	<ul style="list-style-type: none"> <li>• Add neighbor with this IP address or add to peer group (Activates session)</li> <li>• Neighbor is in this AS</li> <li>• If same AS IBGP used</li> <li>• If different AS EBGP used</li> </ul>
<pre>router(config-router)#neighbor (<i>ip-address/peer-group-name</i>) next-hop-self</pre>	<ul style="list-style-type: none"> <li>• All updates from this router are advertised with this router as the next hop.</li> <li>• Use for NBMA networks if required</li> </ul>
<pre>router(config-router)#network <i>network-no. mask network-mask</i></pre>	<ul style="list-style-type: none"> <li>• Advertise this network</li> <li>• Should list ALL networks in the AS not just those locally connected</li> <li>• Will only advertise if this route is already in the IGP table</li> <li>• Use the mask if subnetting</li> </ul>
<pre>router(config-router)#no synchronization</pre>	<ul style="list-style-type: none"> <li>• If synchronisation is not required</li> <li>• Router will advertise routes in BGP before learning them in IGP</li> <li>• <b>Only use if all routers in your AS run BGP:</b></li> </ul>
<pre>router(config-router)#aggregate-address <i>ip-address mask (summary-only) (as-set)</i></pre>	<ul style="list-style-type: none"> <li>• Creates an aggregate (summary) entry in the BGP table</li> <li>• Use summary only to only advertise the summary and not the specific routes</li> <li>• The as-set option includes a list of all the AS numbers that the more specific routes have passed through</li> <li>• The atomic aggregate attribute is set unless the as-set command is used</li> </ul>
<p><b>This can seriously disrupt routing:</b>  Router#clear ip bgp (*   address) (soft (in out))</p>	<ul style="list-style-type: none"> <li>• Removes entries from the BGP table</li> <li>• Resets BGP sessions</li> <li>• Use after every configuration change to ensure the change is activated and peer routers are informed</li> </ul>

## Verifying BGP Operation

router#**show ip bgp (summary| neighbors)**

- Displays entries in the BGP tables
- Use a network number for more specific info. on a network
- Summary displays the status of all BGP connections
- Neighbors displays information about the TCP and BGP connections to neighbors

router#**show ip bgp**

Add example here

router#**debug ip bgp events**

- Displays BGP events

router#**debug ip bgp keepalives**

- Displays BGP keepalives

router#**debug ip bgp updates**

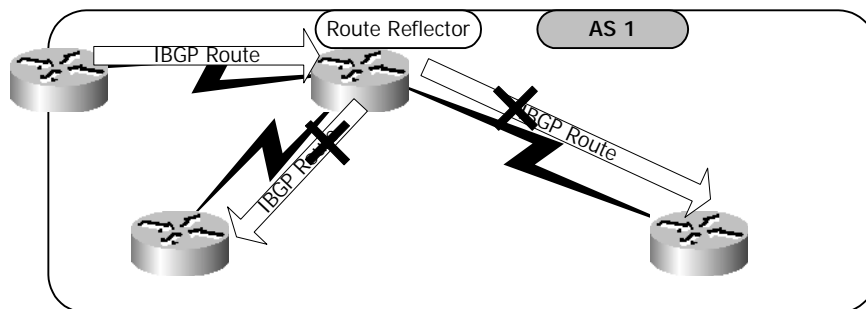
- Displays BGP updates

# SECTION SIX: BGP Scalability

## The Split Horizon Rule

BGP rules specify that routes learnt via IBGP are never propagated to other IBGP peers.

- This ensures there are no routing loops.
- Consequently IBGP peers must be fully meshed within an AS



Unfortunately fully meshed IBGP is not scalable, as with Frame Relay. For 10 routers to be fully meshed  $10(10-1)/2$  IBGP session would be required, that's 45!

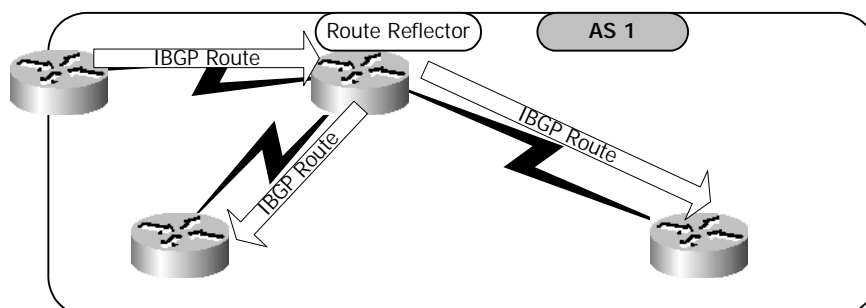
Traffic may also be replicated on some links as it travels to each IBGP peer, this is a particular problem with WAN links.

## Route Reflectors

Route reflectors overcome the problems with BGP split horizon.

Route reflectors adjust the split horizon rule by allowing a route reflector router to propagate routes learned by IBGP to other IBGP peers.

Used mainly by ISPs when the number of internal neighbor statements becomes excessive.

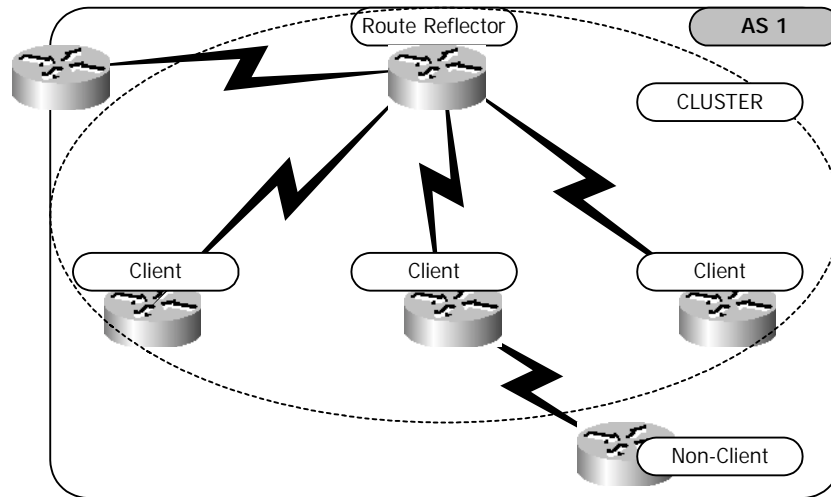


### Benefits

- Reduces the number of BGP neighbor relationships using key routers to replicate updates.
- Only affects routing information distribution not actual routes.
- There can be multiple route reflectors for redundancy and to further reduce the number of IBGP sessions.

- Involves minimal configuration and can be done in stages as non route-reflector routers can co-exist with route reflectors within an AS.

### Terminology

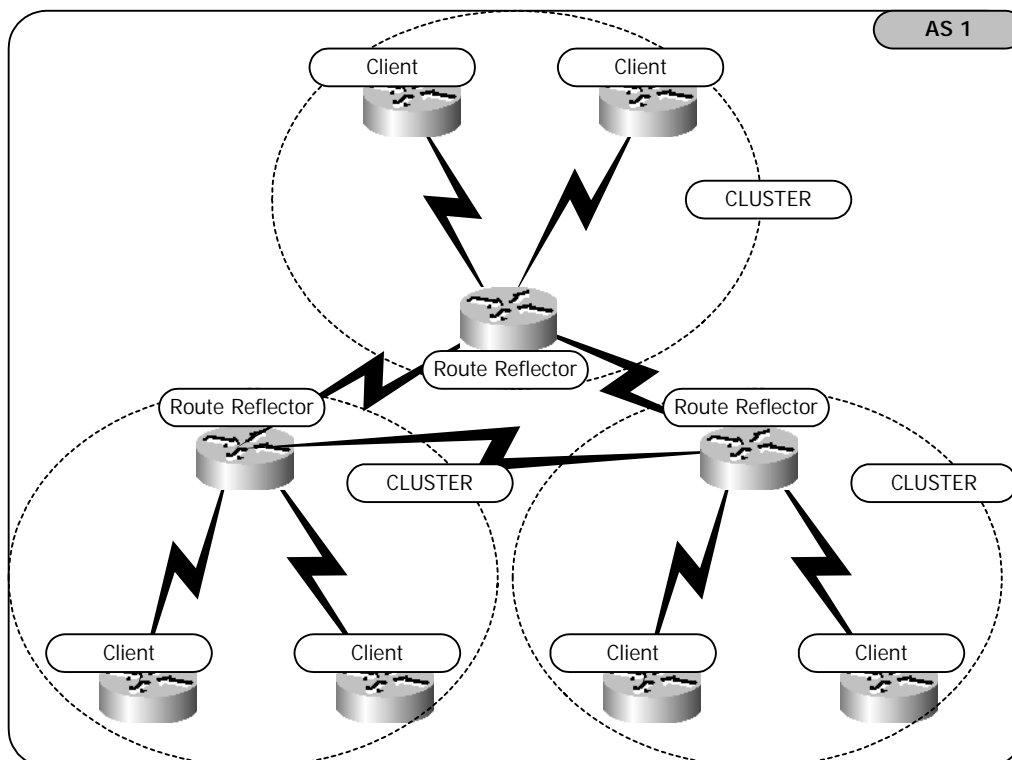


- Originator ID, an optional, non-transitive attribute created by the route reflector. Carries the router ID of the originator of the route in the local AS. If the update returns to the originator it is ignored. Prevents loops.
- Cluster ID, used when there are more than one route reflectors in a cluster. Manually configured on all the route reflectors within the cluster. This allows the route reflectors to recognise updates from other route reflectors in the same cluster. The router ID when there is only a single route reflector in a cluster.

# Route Reflector Design & Configuration

## Design Rules

- The AS should be divided into multiple clusters
- Each cluster should have only a few clients
- Route reflectors need to be fully meshed with IBGP to ensure all routes learnt are propagated throughout the AS
- An IGP is still being used to carry local routes and next-hop addresses



## Operation

If a route reflector receives an update is from:

- A client peer, it sends it to all non-client peers (other route reflectors) and to all client peers except the originator
- A non-client peer (other route reflector), it sends it to all clients in the cluster
- An EBGP peer, it sends it to all non-client peers and to all client peers

## Migration Tips

- Follow physical topology when selecting reflectors and clients. This ensures forwarding paths are not affected. Not doing so may result in routing loops.
- Configure one reflector at a time and then remove the redundant IBGP sessions between clients.

## Configuration

```
router(config-router)#neighbor ip-address route-reflector-client
```

- Configures this router as the route reflector
- Configures the specified neighbor as it's client

## Verifying Route Reflector Operation

```
router(config-router)#show ip bgp neighbor
```

- Displays information on whether this router is a client or route reflector. If a client the BGP neighbor is the route reflector

# Policy Control

Policy control is the process of filtering BGP routing updates to and from particular neighbors to restrict the routing information that is learnt or advertised. This can be done using two features:

- Distribute lists, which use access lists, these are now obsolete for use with BGP. (Access lists are really designed for packet filtering)
- Prefix lists

## Prefix Lists

Prefix lists:

- Are new in IOS v12.0
- Are an alternative to using access lists in many BGP route filtering commands
- Cannot be used with other routing protocols but this is likely to change in the future

The advantages of using prefix lists, over access lists, are:

- Significant performance improvements when loading and looking up routes in large lists
- Support for incremental changes
- User friendly CLI
- Greater flexibility

Filtering with prefix lists involves matching the prefixes of routes with those listed in the prefix list.

A prefix is permitted or denied based upon these rules:

- An empty prefix list permits all prefixes
- A router will search for a match from the top down, the top is the lowest statement sequence number
- Once a match is found the rest of the list is ignored
- Most frequent matches should be at the top of the list to improve efficiency
- An implicit deny is assumed at the end of a prefix list (unless empty)

## Configuring Prefix Lists

### Creation:

```
router(config)#ip prefix-list name (seq seq-value)
deny|permit network/len (ge ge-value) (le le-value)
```

- Creates a prefix list with the name you specify
- Seq = Sequence number to be used when adding a statement to this prefix list
- Deny|Permit = The action taken on a match
- Network/Len = The prefix to be matched and its length. Network is 32-bit address. Len is a decimal number
- ge and le = optional, can be used to specify a prefix range

An exact match is assumed when ge and le are not specified.

GE and LE can be used to specify the range of the prefix length to be matched. The value range is:  
Len <ge-value <le-value <=32

Entries can be added or removed individually

```
router(config)#no ip prefix-list name
```

- Deletes the named prefix list

```
router(config)#ip prefix-list name description text
```

- Adds a text description to a prefix list

### Application:

```
router(config-router)#neighbor ip-address/peer-group-name prefix-list name in|out
```

- Distribute BGP neighbor information as specified in this list
- Ip-address|peer-group-name = IP address of neighbor or peer group for which routes will be filtered
- In = Prefix list applied to incoming advertisements from the neighbor
- Out = Prefix list applied to outgoing advertisements to the neighbor

### Sequence Numbers:

Sequence numbers are generated automatically using values of 5, 10...

You can use sequence numbers to insert entries in a very specific order

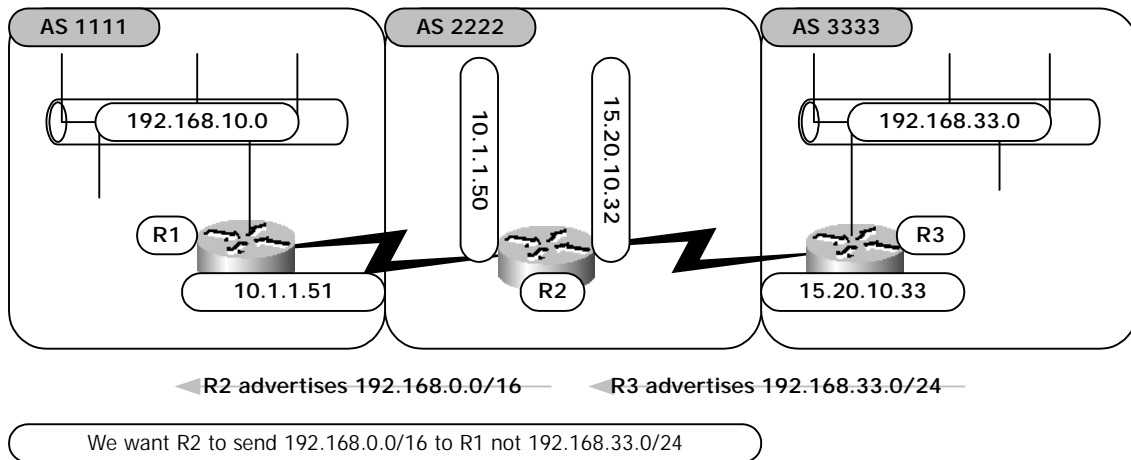
```
router(config)#no ip prefix-list sequence number
```

- Disables generation of automatic sequence numbers

If you choose not to use automatically generated numbers you must always specify a number when using the **ip prefix-list** command.

Sequence numbers do not need to be used when removing entries!

## Prefix List Example



### The configuration:

```
R2(config)#ip prefix-list shorty permit 192.168.0.0/16
R2(config)#router bgp 2222
R2(config-router)#neighbor 10.1.1.51 remote-as 1111
R2(config-router)#neighbor 15.20.10.33 remote-as 3333
R2(config-router)#aggregate-address 192.168.0.0
255.255.0.0
R2(config-router)#neighbor 10.1.1.51 prefix-list shorty out
R2(config-router)#end
```

- Only allow advertisement of 192.168.0.0/16, not more specific networks
- Configure BGP, AS no. 2222
- Configure R1 as a peer/neighbor
- Configure R3 as a peer/neighbor
- Configure summarisation of this class C network
- Apply prefix list to outbound updates to this neighbor

No other routes will be sent to R1 as there is only one statement in the prefix-list and it has an implicit DENY ANY at the end.

## Verifying Prefix Lists

```
router(config)#show ip prefix-list (detail)
```

- Displays information on all prefix lists
- The detail keyword also shows descriptions and the hit count of prefix list entries

```
router(config)#show ip prefix-list (detail) list-name
```

- Displays entries in a specific list

```
router(config)#clear ip prefix-list
```

- Clears hit counts on prefix list entries

# Multihoming

Term used to describe when an AS is connected to more than one ISP

- Increases reliability
- Increases performance, a specific ISP may have better paths to common Internet destinations

## **Types of multihoming:**

Connections to multiple ISPs can be configured in the following ways:

- The ISPs pass only default routes to the AS
- The ISPs pass default routes to the AS and a few specific routes to the AS (specific routes may be from customers with whom the AS exchanges large amounts of traffic)
- The ISPs pass all routes to the AS

### **Default Routes from All Providers:**

This solution has low memory and CPU requirements

The providers send a BGP default route only

The AS sends all routes to the providers

The ISP that a router uses to reach another AS is decided by the IGP metric used to reach the default route within the AS

Inbound packets are routed appropriately by the ISPs routers

### **Default & Few Specific Routes from All Providers:**

This solution has medium memory and CPU requirements

The providers send a BGP default route and selected routes (from customers with whom the AS exchanges large amounts of traffic)

The AS sends all routes to the providers

The ISP that a router uses to reach a specific customer AS will be the shortest AS path although this can be overridden. The path to other destinations will be decided by the IGP metric used to reach the default route within the AS

Inbound packets are routed appropriately by the ISPs routers

## Full Routes from All Providers

This solution has high memory and CPU requirements

The ISP that a router uses to reach a specific AS will be the shortest AS path although this can be overridden

The AS sends all routes to the providers

Inbound packets are routed appropriately by the ISPs routers

## Configuring Weight & Local Preference

These commands are used to influence the path taken to external routes

Higher weight routes are preferred

```
router(config-router)#neighbor ip-address/peer-group-name weight weight
```

- Assign a weight to this neighbor connection
- Valid values are 0 to 65535
- Default value is 32768 for local routes

Higher local preference is preferred

```
router(config-router)#bgp default local-preference value
```

- Changes the default local preference value
- Valid values are 0 to 4294967295

## Redistribution with IGP

You can redistribute networks into BGP in three ways

- Using **network** commands, this allows BGP to advertise routes that are already in the IP table.
- Redistributing static routes
- Redistributing dynamic IGP routes (not recommended as it may cause instability)

### Redistributing Static Routes into BGP

Use the command **redistribute static**.

I can't work out the rest but the **aggregate-address** command is preferred to using **redistribute static** to avoid creating 'black holes'. Null0 figures somewhere.

## Redistributing Dynamic IGP Routes into BGP

This is not recommended as any change in the IGP routes may cause a BGP update. This could result in unstable BGP tables if routes go up and down or flap frequently.

Care must be taken to ensure only local routes are redistributed, routes learnt from other ASs by redistributing BGP into the IGP must not be redistributed back to BGP otherwise loops can occur. Filtering for this can be complicated.

# Redistributing BGP into an IGP

## ISP's – No redistribution from BGP into an IGP

As ISPs normally have all routers in it's AS running BGP redistribution is not necessary. This is a full mesh IBGP environment and IBGP would be used to carry EBGP routes across the AS.

The **no synchronisation** command would be used on all routers in the AS disabling synchronisation between IGP and BGP.

The IGP would only need to route information within the AS and the routes to the next hop addresses of the BGP routes.

Advantages:

- The IGP carries fewer routes
- BGP converges faster as it doesn't have to wait for the IGP to advertise the routes

## Redistribution from BGP into IGP

Non-ISPs would not normally be running BGP on all routers in the AS and would not have a full mesh IBGP environment.

If knowledge of external routes were required within the AS BGP would need to be redistributed into an IGP. Because of the huge no. of routes in the BGP table filtering would be required.

Alternatively default routes and selected external routes could be received from the ISP(s).

# SECTION SEVEN: Optimising Routing Update Operation

## Redistribution

### Why redistribute between/use multiple routing protocols?

- During a conversion/migration to another IGP until the conversion is complete
- When you wish to use a specific protocol but need to continue using the old one due to the requirements of legacy systems
- Political or department boundaries. Departments may have different views on the protocols to use.
- Proprietary protocol issues, if using a vendor specific protocol you may need to use a 'standards based' protocol on other vendors equipment

### What is redistribution?

The exchange and advertisement of routing information between networks using different routing protocols, Cisco defines internetworks running different routing protocols as Autonomous Systems, (not to be confused with BGP and EIGRP AS's.)

Routers connecting AS's are called boundary routers, this is where actual redistribution takes place.

### Considerations and Issues:

- Redistribution should be used only when necessary as it can be complex and cause routing confusion
- Routing feedback can occur, routers may send routes received from one AS back into the same AS
- Suboptimal paths may be selected particularly because most protocols use different metrics and these cannot be translated exactly between protocols. i.e. hops to cost
- Convergence times differ between routing protocols and therefore changes may be learnt about in one protocol before the other
- You can only redistribute between routing protocols which support the same protocols such as IP, IPX etc.
- Redistribution occurs automatically between IGRP and EIGRP if they have the same AS no.

## Selecting the Best Route

To understand how issues with redistribution can occur you need to understand how routers select the best route to a destination

### Administrative Distance (believability)

- If a router receives a route to the same destination network from two different routing protocols the administrative distance of the routing protocol is used to decide which route is placed in the routing table. Refer to Section One for a list of protocols and their administrative distances.
- You may need to modify a protocols administrative distance if you wish for it to be preferred.

### Metric

- If multiple routes to the destination network exist (within the same routing protocol) then the route with the lowest metric is used.

## The Seed Metric

If a router were advertising a link directly connected to one of its interfaces it would use an initial, seed metric derived from the characteristics of the interface. The metric would be incremented as it passed between other routers.

For instance, the seed metric for a RIP route would be 0 and incremented by 1 each time it passed through a router.

However a redistributed route is not physically connected to a router, it is learnt from another protocol. Boundary routers must translate the metrics used between redistributed routing protocols.

You must use the **default-metric** command to configure the seed metric to be used for redistributed routes.

This metric will be incremented normally within the AS, (except OSPF E2 routes)

When configuring the default metric it should be set to a value larger than the largest metric within the receiving AS to prevent routing loops.

## Configuring Redistribution

Before configuring redistribution determine:

- Which router is to be the boundary router where redistribution is to be configured
- Which routing protocol is the core/backbone protocol (normally OSPF or EIGRP)
- Which routing protocol is the edge protocol (migrating from this protocol maybe)
- Which protocol you will redistribute into, normally the backbone protocol

### Configuring Redistribution into the Backbone Protocol:

#### Into OSPF

```
router(config)# router ospf 100  
router(router-config)#redistribute eigrp 64666 metric  
metric-value (subnets)
```

- Configure this OSPF process ID 100
- Redistribute EIGRP AS no. 64666 into OSPF using the seed metric you specify (default is 20)
- The subnets keyword allows subnetted routes to be redistributed. If not used only routes that are not subnetted are redistributed.

## Into EIGRP

```
router(config)# router eigrp 64666
router(router-config)#redistribute ospf 100 (match
internal|external 1|external 2) metric metric-value
```

- Configure this EIGRP AS
- Redistribute OSPF process ID 100 into EIGRP using the seed metric you specify (default is 0 – may not be redistributed)
- The match keyword determines which OSPF routes are redistributed.
- Internal – redistribute OSPF internal routes
- External 1 – OSPF type 1 external routes
- External 2 – OSPF type 2 external routes

## Configuring the default-metric Parameter

Use this command when redistributing into IGRP and EIGRP

```
router(config)# router eigrp 64666
router(router-config)#default-metric bandwidth delay
reliability loading mtu
```

- Configure this EIGRP AS
- Bandwidth = max. bandwidth of the route in kbps
- Delay = route delay in tens of microseconds
- Reliability = likelihood of successful packet transmission, values from 0 to 255 where 255 = route 100% reliable
- Loading = Loading of the route, values from 0 to 255 where 255 = 100% loaded
- MTU = maximum packet size along the route, in bytes, must be 1 or more

Use this command when redistributing into OSPF, RIP, EGP and BGP

```
router(config)# whatever
router(router-config)#default-metric number
```

- Configure this routing protocol
- Value of the metric, no. hops for RIP etc.

## Configuring Redistribution into the Edge Protocol:

How you configure redistribution into the edge protocol is dependant on your network configuration and which techniques you wish to use to prevent routing loops. You can:

- Redistribute a default route about the core AS into the edge AS
- Redistribute multiple static routes about the core AS into the edge AS
- Redistribute all routes from the core AS into the edge AS using a distribution filter to filter out unwanted routes
- Redistribute all routes from the core AS into the edge AS and then modify the administrative distance of the received routes so they are not selected when multiple routes exist for the same destination

## The passive interface Command

This command prevents all routing updates for the routing protocol from being sent into a network from a specific interface, but does not prevent the reception of updates.

When using a link-state protocol, neighbor adjacency cannot be established on the link this interface is attached to.

For all routing protocols:

router(router-config)#**passive-interface** *type number* (i.e. E0)

- Configure specified interface as passive, no updates or advertisements will be sent

## Configuring Static Routes

Used most often to:

- Define specific routes to use when two ASs must exchange routing information, rather than exchanging entire routing tables
- Define routes to destinations over a WAN link to remove the need for a dynamic routing protocol and the associated bandwidth overhead

If you have a route to the defined address:

- Requires redistribution

router(config)#**ip route** *prefix mask address* (distance) (tag *tag*) (permanent)

- Add static route to the network defined by the prefix and mask, using the specified next hop address.
- Distance is optional and specifies the administrative distance to be used
- The permanent keyword ensures that the route is not removed when the interface associated with the route is down.

Use if you do not have a route to the next-hop address: (i.e. using ip unnumbered)

- Automatically redistributed in some cases
- Only use on point-to-point interfaces, i.e. wouldn't work on an Ethernet with multiple routers

router(config)#**ip route** *prefix mask interface* (distance) (tag *tag*) (permanent)

- Add static route to the network attached to this interface
- Distance is optional and specifies the administrative distance to be used
- The permanent keyword ensures that the route is not removed when the interface associated with the route is down.

When using RIP default static routes are redistributed automatically.

# Configuring default-network

# Acronyms and Terms

Acronym	Stands For:
AD	Advertised Distance (EIGRP-DUAL)
AS	Autonomous System (OSPF, BGP, Redistribution)
ABR	Area Border Router (OSPF)
ASBR	Autonomous System Boundary Router (OSPF)
BDR	Backup Designated Router (OSPF)
BGP	Border Gateway Protocol
CIDR	Classless InterDomain Routing
DBD (aka DDP)	Database Description Packet (OSPF Exchange Protocol)
DR	Designated Router (OSPF)
DLCI	Data Link Connection Identifier (Frame Relay)
DUAL	Diffusing Update Algorithm (EIGRP)
EBGP	External BGP
EIGRP	Enhanced Interior Gateway Routing Protocol
FD	Feasible Distance (EIGRP-DUAL)
FDDI	Fibre Distributed Data Interface
FS	Feasible Successor (EIGRP)
HDLC	High Level Data Link Control
IBGP	Internal BGP
IGRP	Interior Gateway Routing Protocol
IOC	Interface Output Cost (OSPF)
MPR	Multi-Protocol Routing
LSAck	Link-State Acknowledgement (OSPF)
LSAs	Link-State Advertisements (OSPF)
LSR	Link-State Request (OSPF)
LSU	Link-State Update (OSPF)
MED	Multi-Exit-Discriminator (BGP Attribute)
MTU	Maximum Transmission Unit
NBMA	Non-Broadcast Multi-Access
OSPF	Open Shortest Path First
OSI	Open Systems Interconnection
PDU	Protocol Data Unit
RIP	Routing Information Protocol
RTO	ReTransmit Interval (EIGRP)
RTP	Reliable Transport Protocol (EIGRP)
SDLC	Synchronous Data Link Control
SLIP	Serial Line Internet Protocol
SRTT	Smooth Round Trip Timer (EIGRP)
VLSM	Variable Length Subnet Masking
Term	Stands For:
Vector	Path
Metric	Numerical Value