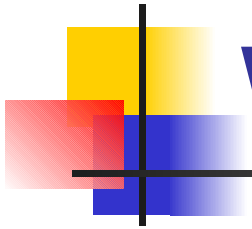
Advanced IP Networking Part 2: BGP - Service Provider, and Internet Architecture

Rik Wade

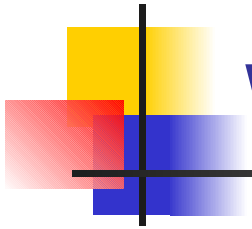
rik@rikwade.com

WYLUG 12/July/2004



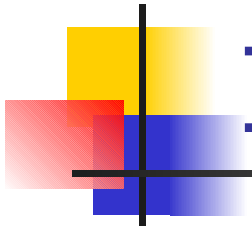
What We'll Cover

- A quick review
- Introduction to BGP
- BGP in action
- Peering and Transit
- BGP in Zebra / Quagga



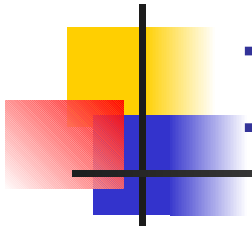
What We've Covered

- Ethernet
  - ARP (IP to MAC resolution)
- Ethernet switching
- Static IP routing
- Dynamic IP routing
  - RIP (distance vector protocol)
  - OSPF (link state protocol)



IP Routing Recap

- Medium and large networks use dynamic routing protocols
  - Automatic distribution of routes
  - (Usually) Simple administration
  - Scalable
  - Self-healing

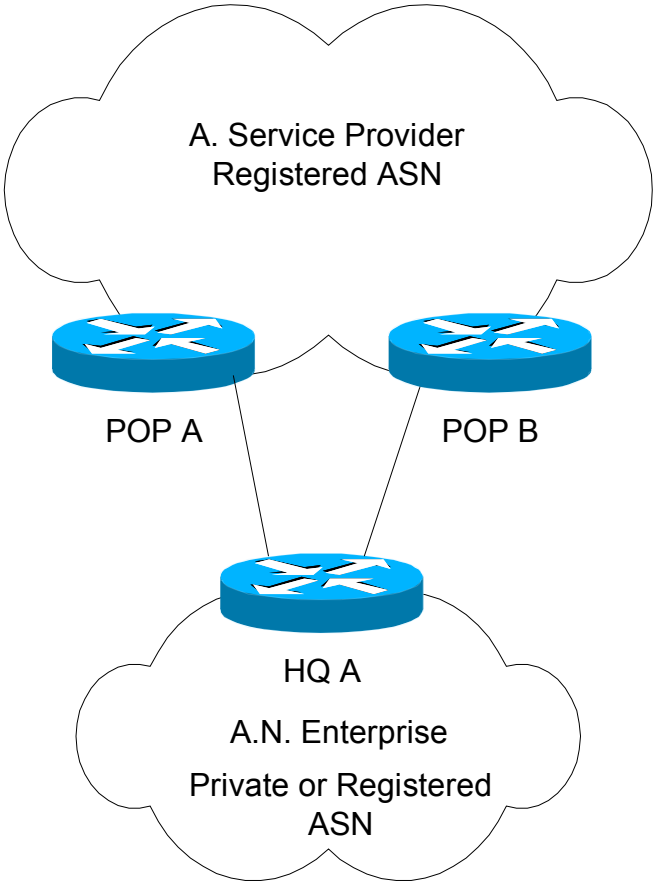


Introduction to BGP

- Border Gateway Protocol (BGP)
- RIP, OSPF, ISIS etc. are Interior Gateway Protocols
- BGP is an Exterior Gateway Protocol
- IGP routes are Summarised by the EGP
- Summarised routes can then be communicated to other BGP processes
- A BGP process represents an Autonomous System – see RFC 1930

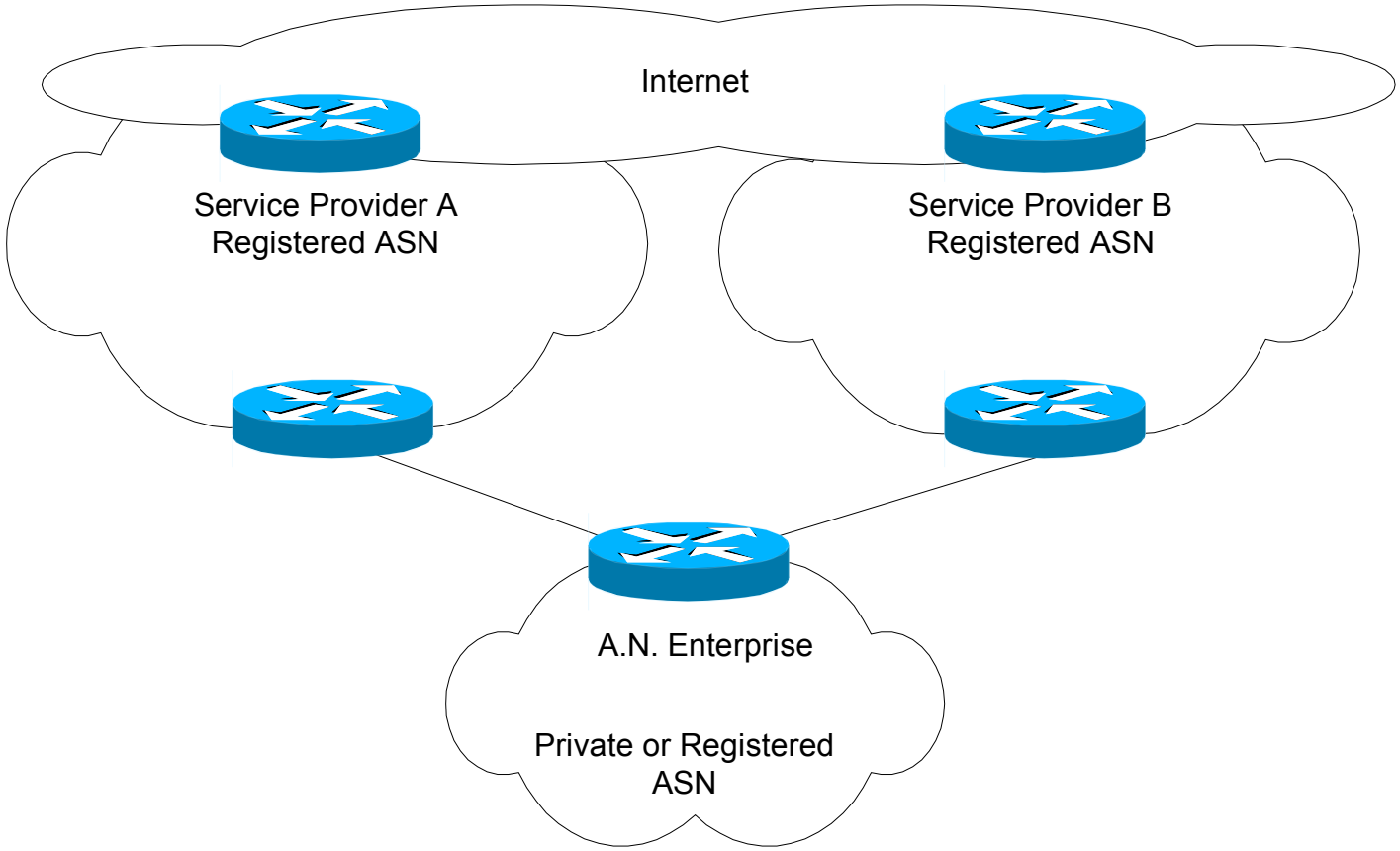
When to use BGP

- Enterprise with multiple links to a service provider



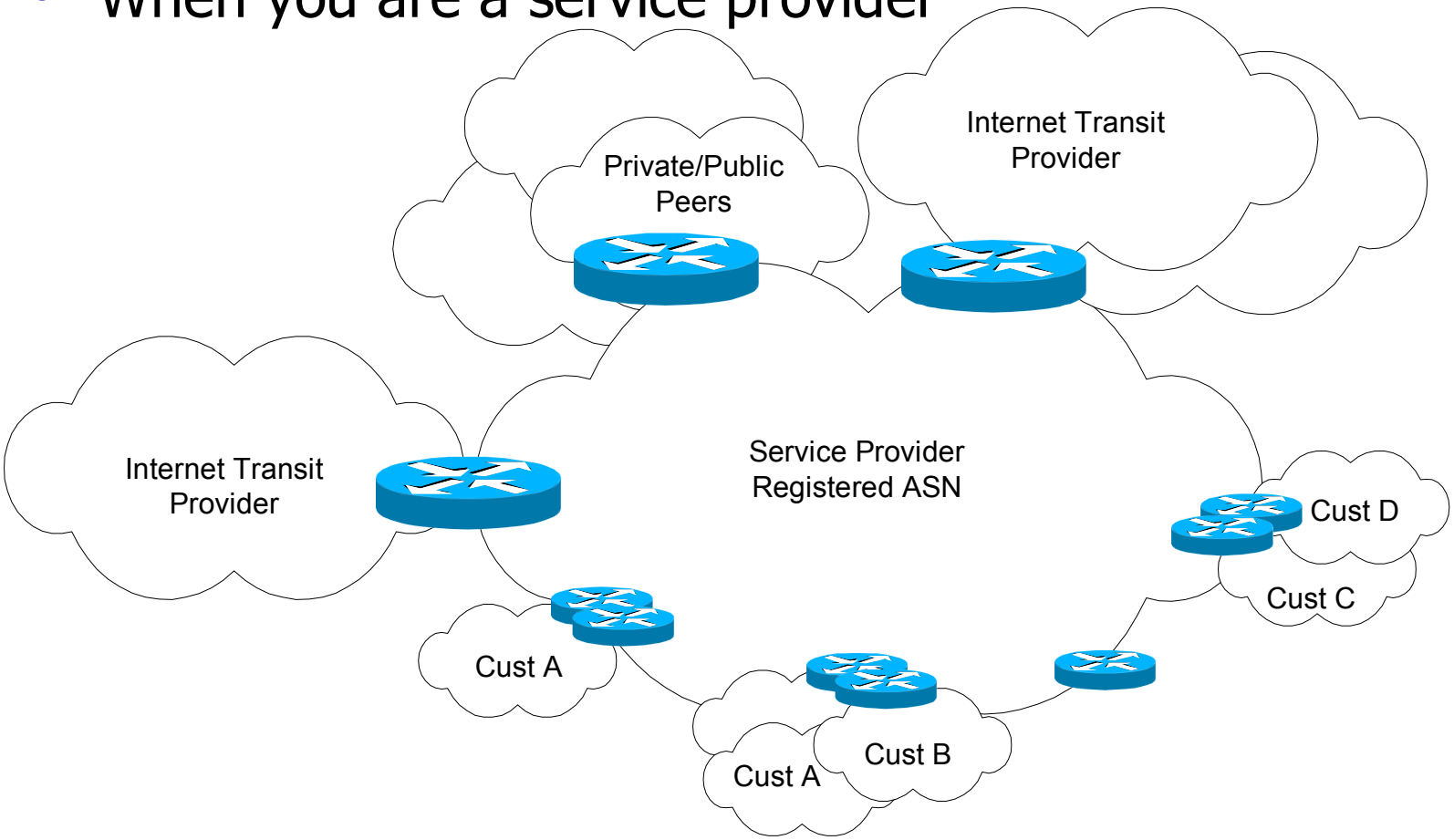
When to use BGP

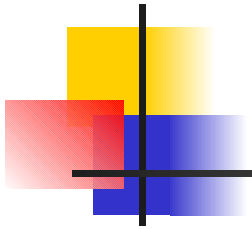
- Enterprise with multiple service providers



When to use BGP

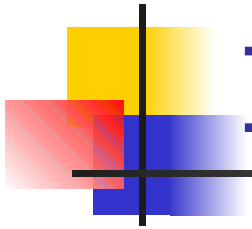
- When you are a service provider





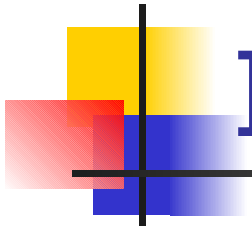
Introduction to BGP

- BGP communicates using standard TCP sockets
- Port TCP/179
- Currently on BGP version 4
- Defined in RFC 1771
- Distance Vector protocol
- Defined route selection rules
- RFC 1772 descibes its application



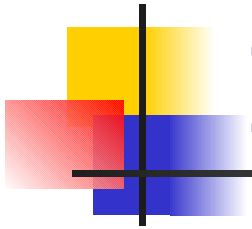
Introduction to BGP

- BGP is an inter-Autonomous System protocol
- Good support for
  - Very large, unstable, routing tables
  - Detection of routing loops
  - Routing policy enforcement
    - Route filtering
    - Load balancing
    - Basic logic decisions (policies)



Introduction to BGP

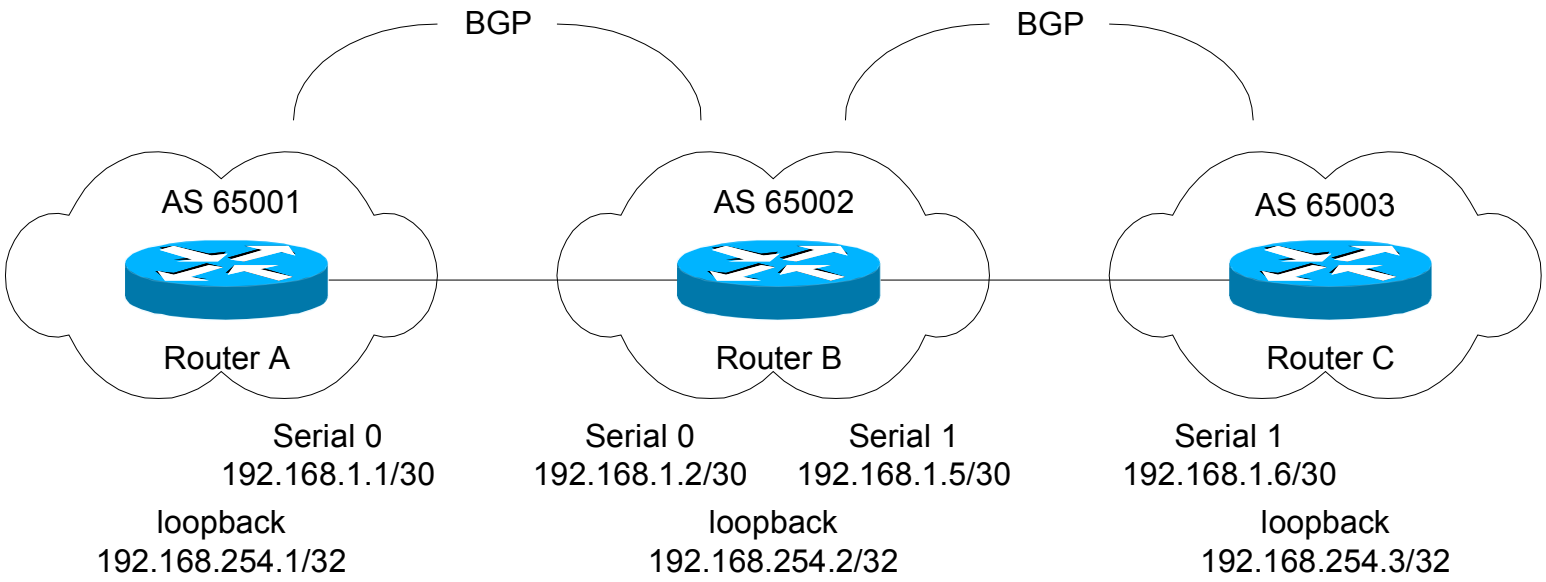
- Uses partial table updates
  - Triggered when the table changes
  - Convergence is slow
- Scalable to hundreds of thousands of routes
  - Some trade-off in speed of convergence
  - Flexibility, stability, speed – choose two
- Did I say that convergence is slow?



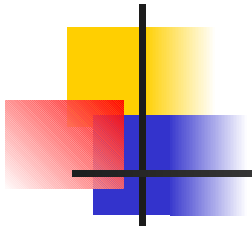
Introduction to BGP

- Announcing routes via BGP:
- Redistribution
  - Redistribute routes learnt from IGP in to EGP
  - Apply some rules and configuration!
- Explicit configuration
  - Announcement requires a component of the configured route to appear in the IP routing table (Cisco, not Quagga/Zebra)

BGP in Action



- Three routers connected using serial interfaces
- Each link has two usable host addresses (a 255.255.255.252 netmask or /30)
- Each router has a loopback IP interface (a local, logical, IP interface)
- Router B peers with routers A and C using BGP
- Router C has a static route to network 192.168.253.0/24



BGP in Action

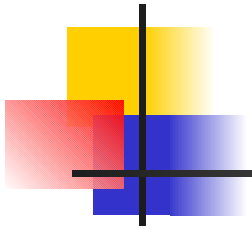
```
Router-B# sh run | begin router bgp
router bgp 65002
 bgp log-neighbor-changes
 redistribute connected
 neighbor 192.168.1.1 remote-as 65001
 neighbor 192.168.1.6 remote-as 65003
 no auto-summary
```

BGP in Action – Router A

```
Router-A#sh ip route
Codes: C - connected, B - BGP

Gateway of last resort is not set

     192.168.254.0/32 is subnetted, 3 subnets
B       192.168.254.2 [20/0] via 192.168.1.2, 01:02:02
B       192.168.254.3 [20/0] via 192.168.1.2, 01:02:02
C       192.168.254.1 is directly connected, Loopback0
     192.168.1.0/30 is subnetted, 2 subnets
C       192.168.1.0 is directly connected, Serial0
B       192.168.1.4 [20/0] via 192.168.1.2, 01:02:02
B    192.168.253.0/24 [20/0] via 192.168.1.2, 00:34:27
```

BGP in Action – Router B

```
Router-B#sh ip route
Codes: C - connected, B - BGP

Gateway of last resort is not set

     192.168.254.0/32 is subnetted, 3 subnets
C       192.168.254.2 is directly connected, Loopback0
B       192.168.254.3 [20/0] via 192.168.1.6, 01:05:18
B       192.168.254.1 [20/0] via 192.168.1.1, 01:05:18
     192.168.1.0/30 is subnetted, 2 subnets
C       192.168.1.0 is directly connected, Serial0
C       192.168.1.4 is directly connected, Serial1
B    192.168.253.0/24 [20/0] via 192.168.1.6, 00:37:42
```

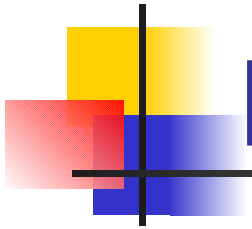


BGP in Action – Router C

```
Router-C#sh ip route
Codes: C - connected, S - static, B - BGP

Gateway of last resort is not set

     192.168.254.0/32 is subnetted, 3 subnets
B       192.168.254.2 [20/0] via 192.168.1.5, 01:07:56
C       192.168.254.3 is directly connected, Loopback0
B       192.168.254.1 [20/0] via 192.168.1.5, 01:07:56
     192.168.1.0/30 is subnetted, 2 subnets
B       192.168.1.0 [20/0] via 192.168.1.5, 01:07:56
C       192.168.1.4 is directly connected, Serial1
S    192.168.253.0/24 is directly connected, Null0
```

BGP in Action – AS PATH

```
Router-A#sh ip bgp
BGP table version is 21, local router ID is 192.168.254.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
   internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*  192.168.1.0/30   192.168.1.2             0             0 65002 ?
*>                  0.0.0.0                  0         32768 ?
*> 192.168.1.4/30   192.168.1.2             0             0 65002 ?
*> 192.168.253.0    192.168.1.2                           0 65002 65003 i
*> 192.168.254.1/32 0.0.0.0                  0         32768 ?
*> 192.168.254.2/32 192.168.1.2             0             0 65002 ?
*> 192.168.254.3/32 192.168.1.2                           0 65002 65003 ?
```



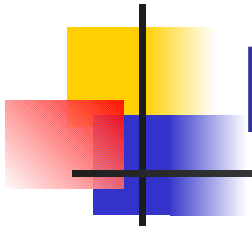
BGP in Action – Update

Debug on Router B

```
02:24:52: BGP(0): 192.168.1.6 rcv UPDATE about 192.168.253.0/24 -- withdrawn
02:24:52: BGP(0): no valid path for 192.168.253.0/24
02:24:52: BGP(0): nettable_walker 192.168.253.0/24 no best path
02:24:52: BGP(0): 192.168.1.1 computing updates, afi 0, neighbor version 11,
   table version 12, starting at 0.0.0.0
02:24:52: BGP(0): 192.168.1.1 send unreachable 192.168.253.0/24
02:24:52: BGP(0): 192.168.1.1 send UPDATE 192.168.253.0/24 -- unreachable
02:24:52: BGP(0): 192.168.1.1 1 updates enqueued (average=27, maximum=27)
02:24:52: BGP(0): 192.168.1.1 update run completed, afi 0, ran for 8ms,
   neighbor version 11, start version 12, throttled to 12
02:24:52: BGP(0): 192.168.1.6 computing updates, afi 0, neighbor version 11,
   table version 12, starting at 0.0.0.0
02:24:52: BGP(0): 192.168.1.6 update run completed, afi 0, ran for 0ms,
   neighbor version 11, start version 12, throttled to 12
```

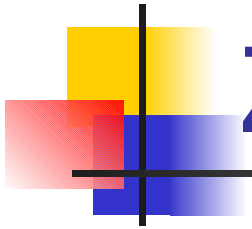
BGP in Action - Update

```
Router-A#sh ip bgp
BGP table version is 22, local router ID is 192.168.254.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
   internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*  192.168.1.0/30   192.168.1.2             0             0 65002 ?
*>                  0.0.0.0                  0         32768 ?
*> 192.168.1.4/30   192.168.1.2             0             0 65002 ?
*> 192.168.254.1/32 0.0.0.0                  0         32768 ?
*> 192.168.254.2/32 192.168.1.2             0             0 65002 ?
*> 192.168.254.3/32 192.168.1.2                           0 65002 65003 ?
```



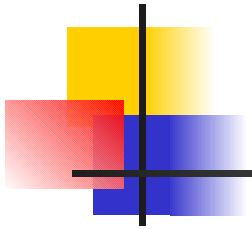
Peering and Transit

- Peering
  - The exchange of BGP routes between Autonomous Systems (Public / Private)
- Transit
  - The exchange of BGP routes between Autonomous Systems (Private)
  - The re-announcement of those routes to upstream BGP peers
  - The transport of traffic for the announced routes



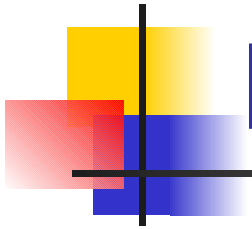
Zebra and Quagga

- Zebra (http://www.zebra.org)
  - A GNU form of a commercial package
  - Certain protocols and features are only available in the commercial package
  - The community forked this project
- Quagga (http://www.quagga.net)
  - Same architecture as Zebra
  - Many more features
  - Community-driven



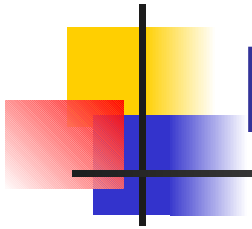
BGP in Zebra & Quagga

- Excellent BGP support
  - http://www.quagga.net/docs/docs-multi/BGP.html
- Supports
  - Redistribution from all its IGPs
  - Prefix filtering
  - Route maps (policy enforcement)
  - Route reflector and server capability
  - Multi-protocol BGP (for MPLS VPNs – RFC 2547bis)
- Highly scalable and efficient BGP router



MPLS VPN

- Multi-Protocol Label Switching
  - A fast way of switching data across an IP network using MPLS-tagged packets
  - Enables the addition of complex services to a plain routed network
- Virtual Private Network
  - A private IP network across a shared IP core
  - Achieved using virtual BGP routing tables



FIN

- BGP (O'Reilly)
  - Iljitsch van Beijnum, ISBN 0-596-00254-8
- Internet Routing Architectures (Cisco Press)
  - Sam Halabi, ISBN 1-57870-233-X
- Routing TCP/IP Volume 2 (Cisco Press)
  - Jeff Doyle, ISBN 1-57870-089-2
- Quagga
  - http://www.quagga.net
- Cisco References
  - http://www.cisco.com/en/US/tech/tk365/tech_protocol_families.html
  - Excellent reference for all protocols