Introduction to Routing in the Internet

ISP/IXP Workshops
Network Topology and Definitions

- Definitions and icons
- Network topologies
- PoP topologies
- Interconnections and IXPs
Some Icons

- Router (layer 3, IP datagram forwarding)
- ATM or Frame Relay switch (layer 2, frame or cell forwarding)
- Ethernet switch (layer 2, packet forwarding)
- Network Cloud
Definitions

- **PoP - Point of Presence**
  physical location of ISP’s equipment

- **vPoP - virtual PoP**
  apparent ISP location
  in reality a back hauled access point
  used mainly for dial access networks

- **Hub - large central PoP**
  links to many PoPs
Routed backbone

- Routers are the infrastructure
- HDLC or PPP links between routers
- Easier routing configuration and debugging
Switched backbone

- frame relay or ATM switches in the core surrounded by routers
- more complex routing and debugging
- traffic management
PoP Topologies

- **Core** routers - high speed trunk connections
- **Distribution** routers and **Access** routers - high port density
- **Border** routers - connections to other providers
- **Service** routers - hosting and servers
- Some functions might be handled by a single router
Pure routed PoPs

- Other PoPs
- Core Routers
- Border Routers
- Distribution Routers
- Access Routers
- Customer Premises Routers
- To other provider or interconnects

Other PoPs to other provider or interconnects other PoPs.
Definitions

• **Transit** - carrying traffic across a network, usually for a fee

• **Peering** - exchanging routing information and traffic

• **Default** - where to send traffic when there is no explicit match is in the routing table
A and B can peer, but need transit arrangements with D to get packets to/from C
Public Interconnect Points

- **IXP** - Internet eXchange Point
- **NAP** - Network Access Point
- **local IXPs**
  - peering point for a group of local/regional providers
- **transit IXPs**
  - connects local providers to backbone (transit) providers
- **hybrid IXPs**
  - combines the function of local and transit
Public Interconnect Point

- Centralised (in one facility)
- Distributed (connected via WAN links)
- Shared, switched or routed interconnect
  - Router, FDDI, Ethernet, ATM, Frame relay, SMDS, etc.
- Each provider establishes relationship with other provider at IXP
  - ISP border router peers with all other provider border routers
Public interconnect

Network 1
Network 2
Network 3

Network 4
Network 5
Network 6

each of these represents a border router in a different autonomous system
Route Server

• Advantages:
  reduces resource burden on border routers (CPU, memory, configuration complexity)
  reduces administrative burden on providers

• Disadvantages:
  must rely on a third party (for management, configuration, software updates, maintenance, etc)
The default free zone is made up of Internet routers which have explicit routing information about the rest of the Internet, and therefore do not need to use a default route.
High Level View of the Global Internet

Default Free Zone

Backbone Provider 1

Access Providers 1

Local NAP or IXP

Backbone Provider 2

Access Providers 2

Customer Networks
Categorizing ISPs

Tier 1 NSP → Tier 1 NSP
Tier 1 NSP → Tier 2 ISP
Tier 2 ISP → Tier 2 ISP
Tier 2 ISP → Tier 2 ISP
Tier 2 ISP → Tier 3 ISP
Tier 3 ISP → Tier 3 ISP
Tier 3 ISP → Tier 3 ISP
Tier 3 ISP → Tier 3 ISP
Tier 3 ISP → IXP
IXP → Tier 2 ISP
IXP → Tier 2 ISP
IXP → Tier 3 ISP
IXP → Tier 3 ISP
Internet Topology and Architecture

- Rapidly increasing complexity
  more providers and locations
  increased meshing

- Emergence of global providers
  capital, regulatory, and technical
  reasons limit the scope of coverage of a
  single provider

- Many new interconnect points (IXPs)
Routing Basics
Routing Concepts

- Routing
- Forwarding
- Some definitions
- Policy options
- Addressing
- Routing Protocols
What does a router do?
A day in a life of a router

find path

forward packet, forward packet, forward packet, forward packet...

find alternate path

forward packet, forward packet, forward packet, forward packet...

repeat until powered off
Routing versus Forwarding

- Routing = building maps and giving directions
- Forwarding = moving packets between interfaces according to the “directions”
IP Routing - finding the path

• Path derived from information received from a routing protocol

• Several alternative paths may exist
  
  best next hop stored in forwarding table

• Decisions are updated periodically or as topology changes (event driven)

• Decisions are based on:
  
  topology, policies and metrics (hop count, filtering, delay, bandwidth, etc.)
IP route lookup

- Based on destination IP packet
- “longest match” routing
  
  more specific prefix preferred over less specific prefix

example: packet with destination of 10.1.1.1/32 is sent to the router announcing 10.1/16 rather than the router announcing 10/8.
IP route lookup

- Based on destination IP packet

Packet: Destination IP address: 10.1.1.1

R2’s IP routing table:
- 10/8 -> R3
- 10.1/16 -> R4
- 20/8 -> R5
- 30/8 -> R6
- 10.1/16

R2’s routing table:
- All 10/8 except 10.1/16
**IP route lookup: Longest match routing**

- Based on destination IP packet

Packet: Destination IP address: 10.1.1.1

R1 -> R2 -> R3: All 10/8 except 10.1/16

R2’s IP routing table:

- 10/8 -> R3
- 10.1/16 -> R4
- 20/8 -> R5
- 30/8 -> R6

**Match!**

Packet: Destination IP address: 10.1.1.1

R3 vs.

10.1.1.1 && FF.0.0.0 vs.

10.0.0.0 && FF.0.0.0

10.1/16
IP route lookup: Longest match routing

- Based on destination IP packet

Packet: Destination IP address: 10.1.1.1

R2’s IP routing table:

- 10/8 -> R3
- 10.1/16 -> R4
- 20/8 -> R5
- 30/8 -> R6

R3:
All 10/8 except 10.1/16

R4:
10.1/16

Packet: Destination IP address: 10.1.1.1

10.1.1.1 && FF.FF.0.0 vs. 10.1.0.0 && FF.FF.0.0

Match as well!
IP route lookup: Longest match routing

- Based on destination IP packet

Packet: Destination IP address: 10.1.1.1

10/8 -> R3
10.1/16 -> R4
20/8 -> R5
30/8 -> R6

R2’s IP routing table

10.1.1.1 && FF.0.0.0 vs. 20.0.0.0 && FF.0.0.0

Does not match!

R3
All 10/8 except 10.1/16

R4
10.1/16
IP route lookup: Longest match routing

- Based on destination IP packet

Packet: Destination IP address: 10.1.1.1

R2's IP routing table

10/8 -> R3
10.1/16 -> R4
20/8 -> R5
30/8 -> R6
......

10.1.1.1 && FF.0.0.0 vs. 30.0.0.0 && FF.0.0.0

Does not match!

All 10/8 except 10.1/16

10.1/16
IP route lookup: Longest match routing

- Based on destination IP packet

Packet: Destination IP address: 10.1.1.1

R2's IP routing table

10/8 -> R3
10.1/16 -> R4
20/8 -> R5
30/8 -> R6
.....

Longest match, 16 bit netmask

All 10/8 except 10.1/16
IP Forwarding

- Router makes decision on which interface a packet is sent to
- Forwarding table populated by routing process
- Forwarding decisions:
  - destination address
  - class of service (fair queuing, precedence, others)
  - local requirements (packet filtering)
- Can be aided by special hardware
Explicit versus Default routing

- **Default:**
  - simple, cheap (cycles, memory, bandwidth)
  - low granularity (metric games)

- **Explicit (default free zone):**
  - high overhead, complex, high cost, high granularity

- **Hybrid**
  - minimise overhead
  - provide useful granularity
  - requires some filtering knowledge
Egress Traffic

- How packets leave your network
- Egress traffic depends on:
  - route availability (what others send you)
  - route acceptance (what you accept from others)
  - policy and tuning (what you do with routes from others)
- Peering and transit agreements
Ingress Traffic

- How packets get to your network and your customers’ networks

- Ingress traffic depends on:
  what information you send and to whom
  based on your addressing and AS’s
  based on others’ policy (what they accept from you and what they do with it)
Autonomous System (AS)

- Collection of networks with same routing policy
- Single routing protocol
- Usually under single ownership, trust and administrative control
Demarcation Zone (DMZ)

- Network shared between AS’s
Definition of terms

• Neighbours - AS’s which directly exchange routing information

• Announce - send routing information to a neighbour

• Accept - receive and use routing information sent by a neighbour

• Originate - insert routing information into external announcements (usually as a result of the IGP)

• Peers - routers in neighbouring AS’s or within one AS which exchange routing and policy information
Routing flow and packet flow

For networks in AS1 and AS2 to communicate:

- AS1 must announce to AS2
- AS2 must accept from AS1
- AS2 must announce to AS1
- AS1 must accept from AS2
• Traffic flow is always in the opposite direction of the flow of routing information

  filtering outgoing routing information inhibits traffic flowing in

  filtering incoming routing information inhibits traffic flowing out
AS99 uses red link for traffic going to the red AS and green link for traffic going to the green AS

To implement this policy for AS99:

• accept routes originating in the red AS on the red link
• accept all other routes on the green link
For packets flowing *toward* AS 99:

Unless AS 22 and all other intermediate AS’s co-operate in pushing *green* traffic to the *green* link then some reasonable policies can not be implemented.
Routing policy with multiple AS’s

For net N1 in AS1 to send traffic to net N16 in AS16:

• AS16 must originate and announce N16 to AS8.
• AS8 must accept N16 from AS16.
• AS8 must announce N16 to AS1 or AS34.
• AS1 must accept N16 from AS8 or AS34.

For two-way packet flow, similar policies must exist for N1.
Routing policy with multiple AS’s

As multiple paths between sites are implemented it is easy to see how policies can become quite complex.
Granularity of routing policy

- What to announce/accept
- Preferences between multiple accepts
  - single route
  - routes originated by single AS
  - routes originated by a group of AS’s
  - routes traversing specific path
  - routes traversing specific AS
  - routes belonging to other groupings (including combinations)
• 77000+ prefixes (not realistic to set policy on all of them individually)
• 7000+ origin AS’s (too many)
• routes tied to a specific AS or path may be unstable regardless of connectivity
• groups of AS’s are a natural abstraction for filtering purposes
Routing Policy Issues

• Destination based limitations
• Global topology not known (and constantly changing)
• Route groupings are not known
  AS membership or AS groups
• Set of all routes in the Internet is not known
IP Addressing

- Internet is **classless**
- Concept of Class A, class B or class C is **no more**

  engineers talk in terms of prefix length, for example the class B 158.43 is now called 158.43/16.

- **All routers must be CIDR capable**

  **Classless InterDomain Routing**
IP Addressing

- IP Address space is a resource shared amongst all Internet users

Regional Internet Registries delegated allocation responsibility by the IANA

RIRs allocate address space to ISPs and Local Internet Registries

ISPs/LIRs assign address space to end customers or other ISPs

- 61% of available address space allocated
• Geographical addressing
  ARIN - APNIC - RIPE NCC (the 3 IRs)
  APNIC serves the Asia Pacific region

• Provider-based addressing
  Addresses assigned by upstream provider
  Local Internet Registries
Geographical addressing

• Advantages:
  Not tied to local backbone provider
  Part of the regional registry process

• Disadvantages:
  Increases size of global routing table
  More difficult to get started
Provider based addressing

- **Advantages:**
  - Easy to get started
  - No increase in size of global routing table

- **Disadvantages:**
  - Must renumber when changing providers
  - May fragment provider address block when multihoming
  - No part in the regional registry process
What Is an IGP?

• **Interior Gateway Protocol**
• **Within an Autonomous System**
• **Carries information about internal infrastructure prefixes**
• **Examples - OSPF, ISIS, EIGRP…**
Why Do We Need an IGP?

- ISP backbone scaling
  
  Hierarchy
  
  Modular infrastructure construction
  
  Limiting scope of failure
  
  Healing of infrastructure faults using dynamic routing with fast convergence
What Is an EGP?

- **Exterior Gateway Protocol**
- Used to convey routing information between Autonomous Systems
- De-coupled from the IGP
- Current EGP is BGP
Why Do We Need an EGP?

• Scaling to large network
  Hierarchy
  Limit scope of failure

• Policy
  Control reachability to prefixes
  Merge separate organizations
  Connect multiple IGPs
Interior versus Exterior Routing Protocols

- **Interior**
  - automatic neighbour discovery
  - generally trust your IGP routers
  - prefixes go to all IGP routers
  - binds routers in one AS together

- **Exterior**
  - specifically configured peers
  - connecting with outside networks
  - set administrative boundaries
  - binds AS’s together
Interior versus Exterior Routing Protocols

- **Interior**
  
  Carries ISP infrastructure addresses only
  
  ISPs aim to keep the IGP small for efficiency and scalability

- **Exterior**
  
  Carries customer prefixes
  
  Carries Internet prefixes
  
  EGPs are independent of ISP network topology
# Default Administrative Distances

<table>
<thead>
<tr>
<th>Route Source</th>
<th>Default Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connected Interface</td>
<td>0</td>
</tr>
<tr>
<td>Static Route</td>
<td>1</td>
</tr>
<tr>
<td>Enhanced IGRP Summary Route</td>
<td>5</td>
</tr>
<tr>
<td>External BGP</td>
<td>20</td>
</tr>
<tr>
<td>Internal Enhanced IGRP</td>
<td>90</td>
</tr>
<tr>
<td>IGRP</td>
<td>100</td>
</tr>
<tr>
<td>OSPF</td>
<td>110</td>
</tr>
<tr>
<td>IS-IS</td>
<td>115</td>
</tr>
<tr>
<td>RIP</td>
<td>120</td>
</tr>
<tr>
<td>EGP</td>
<td>140</td>
</tr>
<tr>
<td>External Enhanced IGRP</td>
<td>170</td>
</tr>
<tr>
<td>Internal BGP</td>
<td>200</td>
</tr>
<tr>
<td>Unknown</td>
<td>255</td>
</tr>
</tbody>
</table>