Misc.

• Project 2
  – Project on class webpage – please get started early – you may be already late!
  – Next time: help getting started – please try to read the document, even if you don’t understand everything

• Today
  – Link State Routing
  – Address assignment
  – Hierarchical Routing/BGP
Last Time

- Link state routing
- Link cost metrics
- Scalable Routing – reduce the overhead and computation required for routing
  1. Create Hierarchy – Autonomous Systems routed independently
     - Interior protocol (intra-domain protocol)
     - Exterior protocol (inter-domain protocol)
     - OSPF and areas to add additional hierarchy within a domain
  2. Address aggregation
     - Subnetting and supernetting – Classless routing
Link State Routing

• Idea: Send updates to *all nodes in the network* with information about *only your immediate neighbors*
  
  – The update consists of a Link State Packet (LSP) which holds
    * Source node id
    * Cost of the link to each directly connected node
    * A sequence number (SEQNO)
    * A time to live (TTL)
  
  – How to send the packet to all nodes in the network? Reliable Flooding

• Each node has the full topology of the network, and can carry out shortest path computation locally to find optimal paths
Forward Search Algorithm

Confirmed: D,0,− C,2,C B,5,C
Tentative: A,10,C B,11,B C,2,C A,12,C
Determining Link Cost

- Queue length – link properties ignored
- Queue delay + transmission delay + propagation delay?
- Third metric
  - Solution: base cost on utilization, compress the range of the cost to avoid circuitous routes
OSPF Example

- Example from OSPF RFC

- Note that the costs a connection can be different in each direction
Each router constructs its shortest path tree

Example shows the tree for router 6
Neither approach scales gracefully

Must reduce size of stored and exchanged information

– Use Autonomous Systems
  * Each autonomous system routed independently using an interior routing protocol (Intra-domain). Examples: RIP, OSPF
  * Routing across autonomous systems using an exterior routing protocol (inter-domain). Examples: EGP, BGP
Large Autonomous Systems and OSPF

- OSPF provides internal hierarchy to deal with a large autonomous area
  - Connections across different areas pass through the backbone
- Routers are classified depending on their location. They take part in routing among their peers
- Three types of routes
  - Intra-area routes (within area)
  - Inter-area route (area1 – backbone – area2)
  - Inter-AS routes
• Note R11, it is connected to two areas, but not to backbone
Break/Discussion

- Before we discuss Inter AS routing/the full routing picture on the internet
  - Need to understand the use of addresses
    * To ensure high utilization of the address space – it is severely limited
    * To provide hierarchy for the routing
  - Need to have a better view/model of what the internet looks like
IP Addresses

- Problem: severe underutilization of address space if we follow the address classes strictly
  - Example, using a class C network for a LAN with 3 machines, 252 addresses are wasted
  - Example, if you have a network of 400 machines, need a class B address (65536)
- Need a bigger address space (IPv6)
- Until then, need to develop more efficient ways to use the available space
  - Use subnetting to fit multiple sub networks within a single network
  - Use supernetting to assign multiple networks a single id
Subnetting

- Allow multiple networks to share the same “physical network” space

- Example, Computer science department has 100 machines. Electrical Engineering department has 100 machines. We need two networks that can each be managed and routed independently
  - Assign 2 class C networks?
  - Use subnetting to share one class C network
    * 128.226.123.xx is the network number
    * CS gets 128.226.123.0 to 128.226.123.127
    * EE gets 128.226.123.128 to 128.226.123.255
  - How?
    * Address mask is used to figure out what network you are on
      - Address Mask is 255.255.255.128
      - Any address in CS ANDED with the mask produces 128.226.123.0
      - Any address in EE ANDED with the mask produces 128.226.123.128
Another Subnetting Example

<table>
<thead>
<tr>
<th>Network number</th>
<th>Host number</th>
</tr>
</thead>
<tbody>
<tr>
<td>111111111111111111111111111111</td>
<td>00000000</td>
</tr>
</tbody>
</table>

Subnet mask (255.255.255.0)

<table>
<thead>
<tr>
<th>Network number</th>
<th>Subnet ID</th>
<th>Host ID</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Subnetted address

Subnet mask: 255.255.255.128
Subnet number: 128.96.34.0

128.96.34.15

H1

128.96.34.1

R1

128.96.34.130

128.96.34.129

R2

128.96.33.1

H3

128.96.33.14

Subnet mask: 255.255.255.128
Subnet number: 128.96.34.128

128.96.33.1

H2

128.96.33.14

Subnet mask: 255.255.255.0
Subnet number: 128.96.33.0
Supernetting

- Instead of assigning a corporation that has 400 hosts a class B address, assign it two class C addresses.
- If the networks are consecutive, they can be abstracted away as a single network.
- The combination of subnetting and supernetting gives us classless routing (Classless Internet Domain Routing/CIDR).
- The network portion of the address can be of any length.
- Use masks to sort things out.
- Use addresses/masks to help with routing.
- As you get closer to the destination, more details become available.
CIDR Routing

- Multi-regional 192.0.0.0 - 193.255.255.255
- Europe 194.0.0.0 - 195.255.255.255
- Others 196.0.0.0 - 197.255.255.255
- North America 198.0.0.0 - 199.255.255.255
- Central/South America 200.0.0.0 - 201.255.255.255
- Pacific Rim 202.0.0.0 - 203.255.255.255
- Others 204.0.0.0 - 207.255.255.255
- How can we make use of this information?
Assigning Addresses

- Don't have to worry about hardcoded network class boundaries

- Company kludge.com wants 1024 addresses
  - Assign it 198.15.0.0 to 198.15.3.255 with mask 255.255.252.0 (22 bit mask)

- Company bob.com wants 4096 addresses
  - Assign it 198.20.16.0 to 198.20.33.255 with mask 255.255.240.0 (20 bit mask), must start at 4096 boundary

- Company oops.com wants 2048 addresses, assign it what?
Decoding Addresses

- The mask indicates the prefix that all addresses in a range have in common
  - if B is the first address in the range and MB is its mask, then for any address A in the range, A AND MB == B
  - if C and D are the base addresses of two ranges such that MC is a prefix of MD, then D AND MC == D AND MD

- To decode an address A:
  - for each mask M and associated base address B, determine if A AND M == B
  - for all M/B for which this is so, A lies in the address range whose mask has the most bits

- A message to A can be forwarded to the router associated with the base of its address range
CIDR and Routers

- Original RIP does not understand CIDR
  - New version RIP-2 does

- OSPF understands CIDR and includes address mask in each record

- Authentication is a big concern for routing algorithms
  - Router messages should be authenticated
  - Not done in RIP
    * Simple plain-text password authentication in RIP-2
  - OSPF supports password and cryptographic based authentication
Internet Past and Present

- Past – tree structure

- Present – multi backbone

- How to route it?
Border Gateway Protocol (BGP)

- RFCs 1771–1774
- A distance-vector protocol
- Nodes of the routing tables are the ASs
- Avoids count to infinity by using path vectors
- ASs are classified as:
  - Stub – only connected to one other AS; carries local traffic only
  - Multihomed – connected to multiple ASs, but refuses to carry transit traffic
  - Transit
- Routers must enforce policy constraints (i.e., not route through multihomed ASs)
- BGP-4 (current version) can deal with arbitrarily connected ASs (not just the simple tree structure of old which was the limitation of EGP)
BGP – Very Briefly

• Challenges:
  – Scale – still a very large scale (10s of thousands of ASs)
  – How to specify cost metrics? Each AS is independent
  – Trusting advertisements
  – Specifying and enforcing flexible routing policies
  – Integration with Intradomain routing

• BGP – How it works
  – Each AS has a BGP speaker
  – Speaker exchanges reachability information with other speakers (no costs)
  – Distance vector but advertises full path to avoid count to infinity loops (no count since only reachability info is exchanged)
  – AS numbers centrally assigned and unique
  – Speaker can cancel previously advertised paths
Example and Discussion

- AS 2 advertises it has a path to 128.96 (AS2, AS4); AS 1 advertises it has a path to 128.96 (AS1, AS2, AS4)

Discussion

- How optimal are the paths?
  * Always trying to maintain a balance between scalability and optimality
- How scalable?
  * Number of ASs much smaller than number of networks
  * Updates of networks in ASs can be summarized as in OSPF
• BGP table size at a sample Backbone routers (courtesy Oregon Route Views Project http://www.antc.uoregon.edu/route-views/)

SUNY-BINGHAMTON – CS428 SPRING ’05 LEC. #12
BGP

- Allows ISPs/administrators to specify policy; but policy is not part of the protocol
- Example policy – hot potato – get the traffic off my network as fast as possible
- Policy shaped by commercial relationships between ISPs (peering and transit relationships)
- Can lead to strange behavior
  - e.g., asymmetric routes
  - e.g., traffic between UAE and Saudi Arabia was getting routed through Washington DC
Discussion/Summary

- Routing scalability:
  - Get routing table sizes down – how?
  - Get routing overhead down – how?

- Role of hierarchy and address aggregation

- Commercial factors and relationships, and not performance/technical factors, often play the most important role in how routing works
Policies

Routing internally and externally strongly influenced by policies. Example:

- Prefer shortest path
- Hot Potato Routing (get it off the AS as soon as possible)

No longer a problem of finding the best path
Peering vs. Transit

- **Peering**: Two ISPs provide connectivity for each other’s customers
  - Usually for free
  - Not a transitive relationship

- **Transit**: One ISP provides another with connectivity to every place it knows
  - $$$
Peering Example

- West Net
- USA Net
- East Net

Customers

Routing Tables
Transit

West Net

USA Net

East Net

Customers

Routing Tables

Internet as known to USA Net

peering

Transit

SUNY-BINGHAMTON – CS428 SPRING ’05 LEC. #12