So Far...

- Inter-domain routing:
  - hierarchy to scale routing and allow distributed management
  - at what price?

- Peering and Transit/Settlements
  - The commercial side to routing
  - Policies inside an AS and among AS’s

- For these reasons, routing is not the nice optimal routing we learned on flat topologies
  - We looked at Paxon’s paper to see the resulting overall effect

IP in IP (IP tunneling) – How

- The IP Packet is encapsulated by another IP header with the tunnel address
  - Protocol field is set to IP-in-IP protocol

- When tunnel destination receives the tunneled packet, it strips the header and sends the original packet on
  - For efficiency, an abbreviated IP header is defined for tunnelling
Virtual Private Networks

- Corporations would like private networks to connect their different campuses
- Companies cannot afford dedicated physical networks (or leased lines)
- Solution – use tunneling
  - Tunneling creates a virtual circuit
  - Encrypt payload if security is needed

Routing for Mobile Networks – Last Hop

- Mobile hosts can move to another network – Problem:
  - How to keep existing connections/IP address
- One solution – Mobile IP. Idea is, when a host moves:
  - Foreign Agent (FA) assigns a care-of-address
  - FA informs Home Agent (HA) of care-of-address
  - HA tunnels packets destined to the mobile to the care-of-address
  - FA strips tunnel header and delivers packet
- Any problems?
  - What about hosts on the home network?
  - How efficient? Can it be optimized?

Alternative Mobility Support

- Mobility: fundamental problem is
  - Need a globally unique name
  - Need to be able to route based on name in the presence of mobility
  - IP addresses: global name + routing locator
- Need – dissociate name from routing information
- Mobile IP – use tunneling, inefficient
- Snorenson and Balakrishnan (Mobicom 2000):
  - Use Fully Qualified Domain Name (FQDN) as a globally unique name
  - Dynamically obtain a new IP address when you move
  - Update DNS when you move so that the name is resolved to your new IP
  - Mechanism developed to migrate existing connections
- How does this compare with Mobile IP?
- Keep this in mind when you read the IPNL paper
Routing – Ad Hoc Networks

- Recall – ad hoc networks are networks made up completely of mobile devices
  - Scenario – handheld devices on infantry in enemy territory; search and rescue operations; conference; cars on a highway
- Problem: computer moves, network really dynamic
- Can we just use traditional routing algorithms such as RIP or OSPF?
  - RIP is too slow to converge
  - Change is an exception in wired protocol, it is the norm in ad hoc networks

Routing in Ad Hoc Networks

- Two flavors of routing algorithms
  - Table-driven, or proactive:
    - similar to traditional algorithms – construct a table of paths to all nodes in the network
    - High Overhead, but low startup time
  - On-demand, or reactive:
    - when you need a path to a node, flood the network to search for it
    - Lower overhead, higher startup time

Discussion

- Which of the two approaches is better?
- Are there security issues in either?

Generalizing the Idea: Overlay Networks

- Deploy an Overlay Network over multiple Internet Sites
  - Applications configure themselves and route traffic among participating overlay nodes
  - Many applications can use such a model to implement features not available in a vanilla IP network
  - End to End argument in action?
  - Do we have to use tunneling?
- E.g., Peer-to-Peer networks, Application Level Multicast, Secure Overlays, VPNs, Content Delivery systems, even your second class project
- Important and emerging research area. Many research issues. We will revisit later.
Example of Table Driven Protocol

- Best known example – Destination Sequenced Distance Vector (DSDV) – Perkins, 1994
  - Simple distance-vector based protocol
  - Sequence numbers used to avoid loops

- DSDV Operation
  - Nodes periodically broadcast their tables; increment sequence number of route to self each time
  - If a directly connected link is down
    - Set cost of all route-table entries using that link to infinity and add 1 to their sequence number
  - Triggered update – broadcast table immediately
  - When receiving a route table:
    - Accept entries containing sequence numbers at least as large as your own
    - Accept entries with equal sequence numbers but lower cost
  - Two flavors of update – full or incremental

Problems with DSDV/Table Driven

- Control message overhead grows with $O(n^2)$
  - A lot of overhead incurred proactively when a node moves in or out
  - What if no one is communicating with it?
- Each node must have complete table
- Alternative – Reactive routing
  - Compute route only when needed
  - Low routing overhead, but high latency
  - Big flood of the network when communication is needed

On-demand (or reactive) Routing

- When you need to find a path to a destination, a route discovery process is initiated
  - Flood the network with a route request packet
  - When a node receives a route request looking for it, it reverses the path and sends a route reply back
  - Intermediate nodes can set their routing table entries
- Once a reply is received, the packet can be sent
- Examples, Dynamic Source Routing (DSR, Johnson et al), Ad Hoc On Demand Distance Vector (AODV, Perkins et al)
- Any concerns?
  - Should we use a flood on every packet? Too expensive
  - What if the route disappears while in use?
  - It is too expensive to detect a route breaking (several MAC layer timeouts)

Example – DSR

- Source Routing protocol; intermediate nodes do not have to keep information
- Route request
  - Each node receives the route request, adds itself as a hop and forwards it along
  - When destination receives a route request, it reverses the path and sends a reply
- Source uses path with smallest number of hops
- When a link breaks, the node along the link that failed sends a route error packet to source
  - Source looks for a new route
Path Inflation – Background

- How to study this kind of stuff?
- Internet Mapping
  - Don’t have access to internal state of the network; measurement has to be end to end
  - Tools that use mechanisms such as traceroute/ping to map network
    * Augment with DNS to try and figure out AS and/or physical location
  - If you’re lucky, BGP router tables at some routers may be available
- Graph analysis is applied to discover connectivity properties

Characterizing Path Inflation

- SIGCOMM 2003 paper by Spring, Mahajan and Anderson
- They do a trace-driven study of 65 ISPs to
  - Find the causes of path inflation
    * Topology and routing policy choices within an ISP, between pairs of ISPs and across the global Internet
- Use techniques to infer intra-domain and peering policies from measurements (theirs and those of others)
- Chose mainly large ISPs that have interesting topologies and some smaller ISPs for diversity

Path Inflation

- In a well-provisioned well-operated network, shortest paths are preferred
- On the Internet, really long paths are prevalent (path inflation) – why?
  - Tangmunarunkit et al, (INFOCOM 2001) found that over 20% of paths were 5 hops or more longer than the shortest available path at the router level
- Possible reasons? (BGP’s number of ASs metric?)

Optimizations

- Route flood too expensive
  - Flood returns multiple paths – retain in a cache
  - Localize the flood
    * Based on location (Location Aided Routing, Vaidya et al)
    * Based on previous path (Query Localization, Das et al)
  - Reply from cache rather than continuing flood?
    * Cache inconsistency
- MAC level broadcast unreliable
Causes of Inflation

- Interaction of topology and policy at three levels
  - Selection of paths within an AS (intradomain topology/policy)
  - Peering to reach neighboring ASs (peering topology/policy)
    - Hot potato routing and MEDs
  - Paths across more than 2 ASs (inter-domain topology/policy)

Overall Results

Figure 1: Path inflation observed in our dataset as a cumulative distribution. The x-axis represents the extra distance traveled through the network beyond a hypothetical direct link between end points. ISP paths stay within one ISP's network; 1 & 2 ISP paths traverse a maximum of two ISPs; overall is the complete dataset.
Intra-Domain Results

- Intradomain topology and policy do not cause a lot of inflation
  - Well connected topologies
  - Effective/delay sensitive internal traffic engineering

Intra-Domain Topology

Figure 3: Path inflation due to intra-domain topological constraints. The left graph shows the CDF of additive inflation, and the right one shows the median and 95th percentile inflation as a function of latency of a hypothetical direct link.

Peering Effects

Figure 5: Path inflation due to intra-domain routing policy, when compared to shortest latency paths. The left graph shows the CDF of additive inflation, and the right one shows additive inflation as function of the latency of the shortest path.
Summary

<table>
<thead>
<tr>
<th></th>
<th>Median</th>
<th>Mean</th>
<th>95th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra-domain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topology</td>
<td>1.0 ms</td>
<td>2.4 ms</td>
<td>8.4 ms</td>
</tr>
<tr>
<td>Policy</td>
<td>1.4 ms</td>
<td>3.2 ms</td>
<td>11.5 ms</td>
</tr>
<tr>
<td>Peering</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topology</td>
<td>2.0 ms</td>
<td>5.0 ms</td>
<td>17.7 ms</td>
</tr>
<tr>
<td>Policy</td>
<td>3.0 ms</td>
<td>6.5 ms</td>
<td>24.5 ms</td>
</tr>
<tr>
<td>Inter-domain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topology</td>
<td>3.0 ms</td>
<td>7.8 ms</td>
<td>34.1 ms</td>
</tr>
<tr>
<td>Policy</td>
<td>6.9 ms</td>
<td>13.9 ms</td>
<td>60.3 ms</td>
</tr>
</tbody>
</table>

Figure 17: Cumulative path inflation caused by each of the six factors, computed with reference to a hypothetical direct link.

Peering Policies

- Tier 1 ISPs:
  - Late exit, often (15)
  - Late exit, sometimes (10)
  - Early-exit (19)
  - Single peering point (42)
  - Engineered but not late (13)

- Most ISPs used early exit most of the time; they study early exit inflation
- Top 5% of the paths suffer inflation of more than 12msec

Discussion

- Most inflation caused by BGP shortest AS path routing and inefficient peering
- Find that many ISPs use early exit, but also many that are not so greedy
- Not clear if policies contribute to path inflation
- Propose an informed BGP that carries location of exit links
- Other alternatives?
- Thoughts on Internet scale routing?

Interdomain Summary

- They used Gao's heuristics (No-valley, NV + preferred customer, shortest AS path)
- Study inflation vs. that; similar to the papers we've seen before
Medium-Term Roadmap

- Internet Structure
  - Router Design
  - Router Algorithmics
  - Internet Topology paper

- The future of IP/Internetworking
  - Redesign or retro fit? IPNL paper
  - If we have time, next generation BGP? HLP paper

- Moving on to Multicast