CS551
Hierarchical Routing
[Tsuchiya88a]
Bill Cheng

http://merlot.usc.edu/cs551-f12
Context

- Fairly early in the Internet life
  - before BGP-3
  - before CIDR

- Example of SIGCOMM "wild idea" paper
Hierarchies

What?
- logical structure overlaid on collections of nodes

Why?
- together with information abstraction, the only known solution to scaling issues
Routing Hierarchies

- Flat routing doesn’t scale
  - each node cannot be expected to have routes to every destination (or destination network)

- Key observation
  - need less information with increasing distance to destination

- Two radically different approaches for routing
  - the area hierarchy
  - the landmark hierarchy
The Area Hierarchy
Areas

- Technique for hierarchically addressing nodes in a network
- Divide network into *areas*
  - areas can overlap
  - areas can have nested sub-areas
  - constraint:
    - there must exist at least one path between each pair of subareas in an area that does not exit the area
    - other areas can have one entry for entire area
Addressing

Address areas hierarchically

- sequentially number top-level areas
- sub-areas of area are labeled relative to that area
- nodes are numbered relative to the smallest containing area
  - nodes can have multiple addresses (when?)
Routing

- **Within area**
  - each node has routes to every other node

- **Outside area**
  - each node has routes for *other top-level areas only*
  - inter-area packets are routed to nearest border router

- Can result in sub-optimal paths
Path Sub-optimality

3 hop red path
v.s.
2 hop green path
Landmark Hierarchy

- Details about things nearby and less information about things far away
- Not defined by arbitrary boundaries
  - thus, not well suited to the real world that does have administrative boundaries
Key Idea

- Self-configuring hierarchy for routing with many routers
  - compare to the number of engineers needed to keep the Internet running
  - appropriate for 1000 node, unattended sensor networks?
Router 1 can be seen by routers 2, 3, 4, 5, and 6.

Router 1 is a landmark of radius 2.

Router 1 can be seen by routers 2, 3, 4, 5, and 6.
Landmark Overview

- Landmark routers have "height" which determines how far away they can be seen (visibility)

- Routers within the *radius* of landmark router d (this radius is denoted by r[d]) can *see* (landmark) router d (a.k.a LM[d])

- *See* means that those routers have LM[d]'s address in their routing tables and know next hop to reach it
  - Router x has an entry for router y if x is within the radius of y

- Distance vector style routing with simple metric

- Routing table: Landmark, Level, Next hop

Ex:

<table>
<thead>
<tr>
<th>Landmark</th>
<th>Level</th>
<th>Next hop</th>
</tr>
</thead>
<tbody>
<tr>
<td>LM2[d]</td>
<td>2</td>
<td>f</td>
</tr>
</tbody>
</table>
LM Hierarchy Definition

- Each LMi[d] associated with level i and radius (ri[d])
- Every node is an LM0 landmark
- Recursion: some LMi are also LM(i+1)
  - Every LMi router is seen by at least one LM(i+1) router, i.e., "there is at least one LM(i+1)[d] within ri[d] hops of each LMi[d]" (so you can route a message downward)
  - To route a message upward, use visibility

- Terminating state when all level H LMs is seen by the entire network, i.e., "rH[d] ≥ D, where D is the diameter of the network"
  - These routers at level H are called global landmarks
LM Addresses

- LM2.LM1.LM0 (e.g., x.a.b and y.a.b)

- LM level maps to radius (part of configuration), e.g.:
  - LM level 0: radius 2
  - LM level 1: radius 4
  - LM level 2: radius 8

- If destination is more than two hops away, will not have complete routing information, refer to LM1 portion of address, if not then refer to LM2...
  (c would forward based on y in y.a.b)
LM Routing

- LM does not imply hierarchical forwarding
- It is *not* a source route
- En route to LM1 may encounter router that is within LM0 radius of destination address (like longest match)
- Paths may be asymmetric
LM Self-configuration

- Bottom-up hierarchy construction algorithm
  - goal to bound number of children
- Every router is LM0 landmark
- All routers advertise themselves over a distance
- All LMi landmarks run election to self-promote one or more LM(i+1) landmarks
  - How is this done exactly?
    - HW2
    - see [Estrin99a] for some hints
- Dynamic algorithm to adapt to topology changes - Efficient hierarchy
Landmark Routing: Basic Idea

- Not shortest path
- Packet does not necessarily follow specified landmarks

Source wants to reach LM0[a], whose address is c.b.a:
- Source can see LM2[c], so sends packet towards c
- Entering LM1[b] area, first router diverts packet towards b
- Entering LM0[a] area, packet delivered to a
Recall: every LMi router is seen by at least one LM(i+1) router
Routing Table for Router "g"

<table>
<thead>
<tr>
<th>Landmark</th>
<th>Level</th>
<th>Next hop</th>
</tr>
</thead>
<tbody>
<tr>
<td>LM2[d]</td>
<td>2</td>
<td>f</td>
</tr>
<tr>
<td>LM1[i]</td>
<td>1</td>
<td>k</td>
</tr>
<tr>
<td>LM0[e]</td>
<td>0</td>
<td>f</td>
</tr>
<tr>
<td>LM0[k]</td>
<td>0</td>
<td>k</td>
</tr>
<tr>
<td>LM0[f]</td>
<td>0</td>
<td>f</td>
</tr>
</tbody>
</table>

r0 = 2, r1 = 4, r2 = 8 hops

How to go from d.i.g to d.n.t?

How does path length compare to shortest path?
Recap

- Strongest point: self configuration
- No administrative bounds, thus not suitable for Internet
- No policy routing
- Variable (and unstable) addresses
- Not really used at this point