


CS644

Advanced Networks

Lecture 9

Intra Domain Routing
Andreas Terzis


Spring 2004 1



Outline

- So far we have talked about E2E mechanisms
- Routing is the other big component of the network
 - Largest distributed system ever
- Two-level routing system
 - Inter-domain
 - Intra-domain
- Intra-Domain
 - Dynamic Routing
 - Revised ARPANET metric [AZ89]
 - Routing Architectures
 - Landmark [Tsu88]


Spring 2004 2



Dynamic Routing in the Internet

- What is routing?
 - Distributed algorithm to calculate *shortest* path to all network destinations
- Benefits of routing
 - Dynamically circumvent failures
 - Remember Baran's paper
- How do we define shortest?
 - Each link has an associated metric
 - Can be static or dynamic

Spring 2004 3



Original ARPANET routing algorithm

- Distributed version of Bellman-Ford (DV)
- Exchange table every 2/3 sec
- Link metric was equal to instantaneous queue length plus constant
- Results:
 - Persistent routing loops
 - Routing oscillations

Spring 2004 4

SPF Algorithm

■ Algorithm Overview

- Routers periodically advertise adjacencies to *everyone*
- Each router calculates shortest path using Dijkstra's algorithm

■ Metric calculation

- Average link delay over last 10 seconds
- If new average crosses a threshold then update is generated
- Each link has lower bound (bias)
 - Bias is function of link speed, prevents idle link from reporting zero delay

Spring 2004

5

Limitations of Delay Metric

$$\text{Delay} = Q_{\text{delay}} + T_{\text{delay}} + \text{Prop}_{\text{delay}}$$

■ Low, medium load

- Queuing delay is low → Reported metric is good prediction of delay after traffic reroute

■ High load

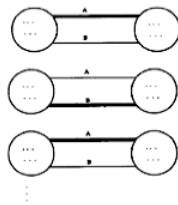
- Queuing delay is high → Metric does not correspond to delay after traffic is rerouted

Spring 2004

6

Routing Oscillation

- No range on queuing delay
- All nodes adjust routes in response to delay metric updates
- Negative Effects
 - Oscillations
 - Unused capacity
 - Congestion
 - Increased routing traffic, CPU utilization



Spring 2004

7

Revised Link Metric

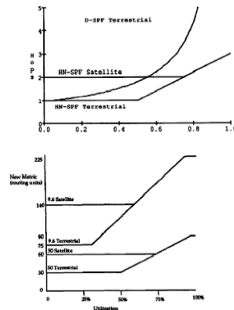
- Translate delay to link utilization (M/M/1)
- Take EWMA of calculated link utilization
- Normalize metric
- Limit the absolute change per update

Spring 2004

8

Metric calculation example

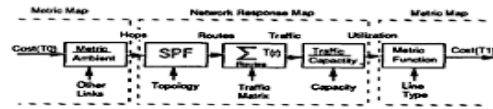
- Transformations are parameterized based on link type
- Lower bound depends on link type
 - Satellite links have higher lower bound
 - Discourages use under low-link load
 - Metric is normalized on link bandwidth



Spring 2004

9

Details on Metric Calculation



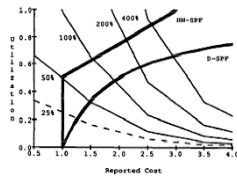
- Link Cost is input
- SPF calculates routes
- Traffic translates to link utilization
- Link utilization translates to new Link Costs
 - Equilibrium: New Cost equal to Old Cost

Spring 2004

10

Stability

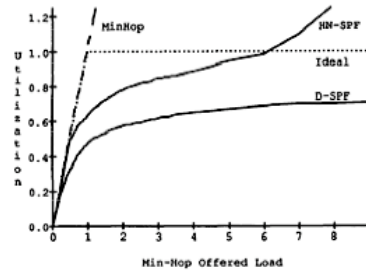
- HN-SPF is stable over larger area compared to D-SPF



Spring 2004

11

Utilization

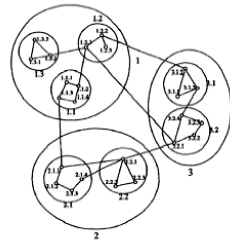


Spring 2004

12

Area Hierarchy

- Group routers into *areas*
 - Each individual router is an area
 - Create a hierarchy of areas
- Pros and cons
 - + Reduced size of routing table
 - - Increased path
- Question:
 - How do you create the hierarchy?

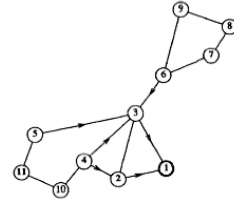


Spring 2004

13

Landmark

- What is a landmark?
 - A router whose neighbor routers within a certain number of hops "know" about
 - Landmark "radius"

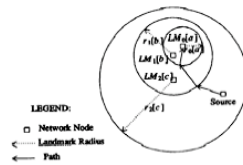


Spring 2004

14

Hierarchy of Landmarks

- LM_i : Landmark at level i
- Every router is LM_0 with r_0
- Some routers are LM_1
 - $r_1 > r_0$
 - Reachable by LM_0 routers
- Stop at level H where $r_H > D$



Spring 2004

15

Routing Table

- Each router keeps
 - Next hop on the shortest path to each of the Landmarks it knows
 - Routers have full
- Result
 - Routers have full knowledge of immediate vicinity
 - Increasingly less information about *distant* locations

Spring 2004

16

Addressing

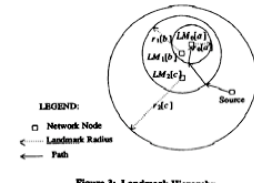
- What is the address of a router?
 - Series of Landmark Ids: $LM_H, LM_{H+1}, \dots, LM_0$
 - Analogous to other hierarchical systems
 - Landmark at level i must be within radius of landmark at level $i-1$
 - Router might have multiple addresses

Spring 2004

17

Routing

- Source routes first towards highest possible landmark
- Once packet enters radius of lower landmark it will follow more specific path
- Observations
 - Not shortest path
 - Path does not necessarily go through landmarks



Spring 2004

18

Hierarchy Management Algorithm

- Bottom-Up Approach
 - Each router advertises itself at distance r_0
 - If it hears from a Landmark at level 1 it grafts to the hierarchy
 - Otherwise do leader-election among peers at level 0
- Also required:
 - Algorithm to find path to Landmark
 - Use a variant of Distance Vector routing

Spring 2004

19

Performance

- Compare against area hierarchy
- Metrics
 - Size of routing table
 - Ratio R_L/R_{shortest}
- Size of routing table
 - Area: $HN^{1/4}$
 - Landmark: $4N^{1/3}$ ($4N^{1/2}$)

Spring 2004

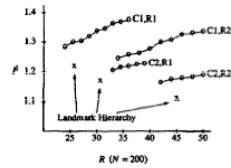
20

Routing Table Size

- Theoretical Results

- Area: $HN^{1/4}$
- Landmark: $4N^{1/3}$ ($4N^{1/2}$)

- Simulations



Spring 2004

21

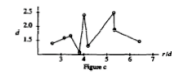
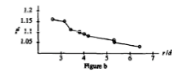
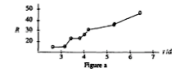
Landmark Routing Performance

- Effect of radius

- Large radius: larger table, shorter paths

- Effect of number of Landmarks

- More Landmarks at level i : shortest avg. distance to Landmark



Spring 2004

22

Additional Issues

- Mapping between well known IDs and Landmark addresses
- Support for Administrative Boundaries

Spring 2004

23