Masking Failures Using Anti Entropy and Redundant Independent Paths

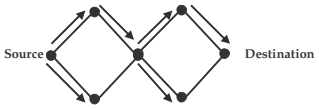
Rebecca Braynard and Amin Vahdat
Internet Systems and Storage Group
Duke University
http://issg.cs.duke.edu

Building Reliable Overlays

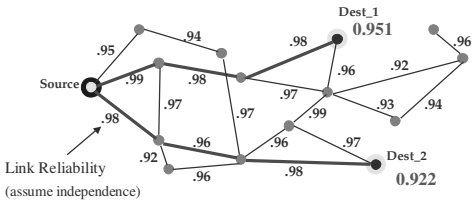
- Hide failures in the network from applications using the overlay
  - Scalable, reliable publish subscribe systems
- Goal:
  - Input: overlay with neighbor connection information and reliability target
  - Probe neighbors for reliability information
  - Construct end-to-end paths that match the target level of reliability
- Benefit: Proactively handle failures
  - Avoid convergence time after network failures through redundancy

Restricted Flooding

- All nodes in the overlay receive the data
- Messages do not travel along every link in the overlay
- Prevent nodes from receiving multiple copies
  - Idea is similar to Anti Entropy: used to maintain consistency between replicas in distributed computing

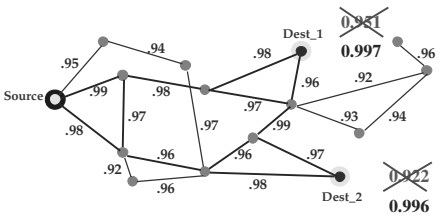


First Step: Maximum Reliability Tree



Highest reliability paths for source to destinations

Second Step: Multiple Disjoint Routes



Basic Flood: 14 Restricted Flooding: 8

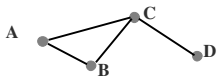
Tune connectivity to application specifications

Questions?

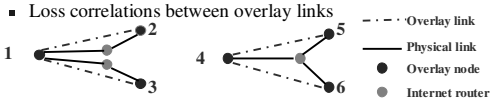
- Webpage:
  - http://issg.cs.duke.edu/malachi
- E-Mail:
  - Rebecca Braynard: rebecca@cs.duke.edu
  - Amin Vahdat: vahdat@cs.duke.edu

Research Problems

- Disjoint paths may not be available



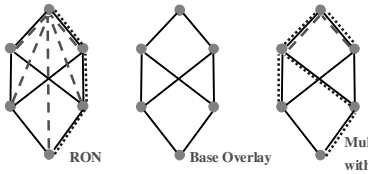
- Loss correlations between overlay links



- Nodes only have local network state information to make decisions
  - Only probe direct neighbors in the overlay

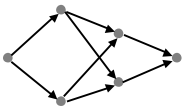
Differences: RON

- RON: $O(N^2)$ probing to obtain network state information
  - Each node probes every other node in the overlay
  - Global network state to make routing decisions
  - Limits scalability

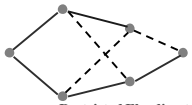


Differences: Mesh-Based Content Routing

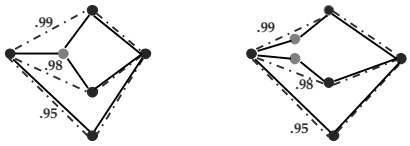
- MBCR: intermediate nodes receive messages multiple times
  - Single receiver
  - We address the problem of finding disjoint routes



Intermediate nodes receive the same message multiple times



Restricted Flooding to reduce network resource consumption

Link Loss Correlation

- Loss rates correlation between links
  - share a link in the underlying topology