

Canicula: An Improved Hybrid Overlay Networks

Yang CHEN

Tsinghua University



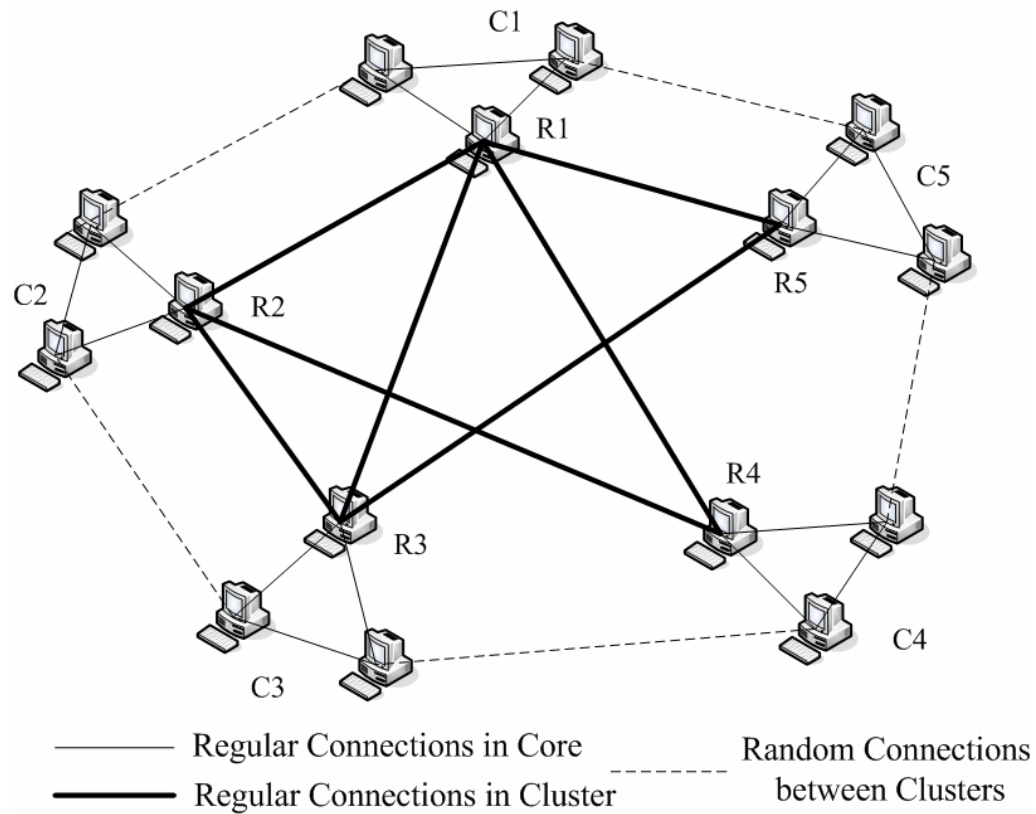
Outline

- Introduction
- Related Work
- Design of Canicula
- Simulation
- Experiment on PlanetLab
- Conclusion



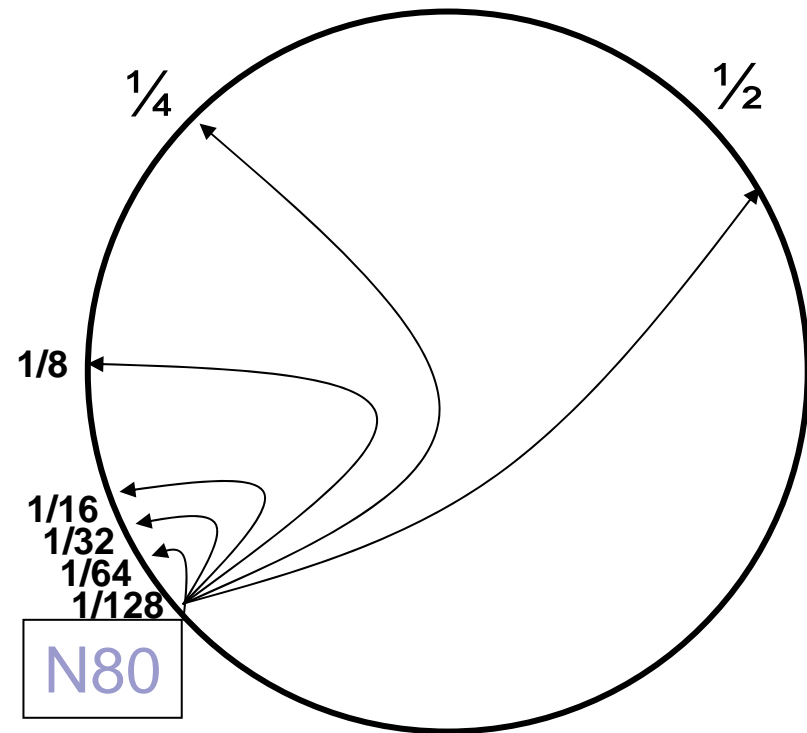
Related Work

Hybrid Overlay Network



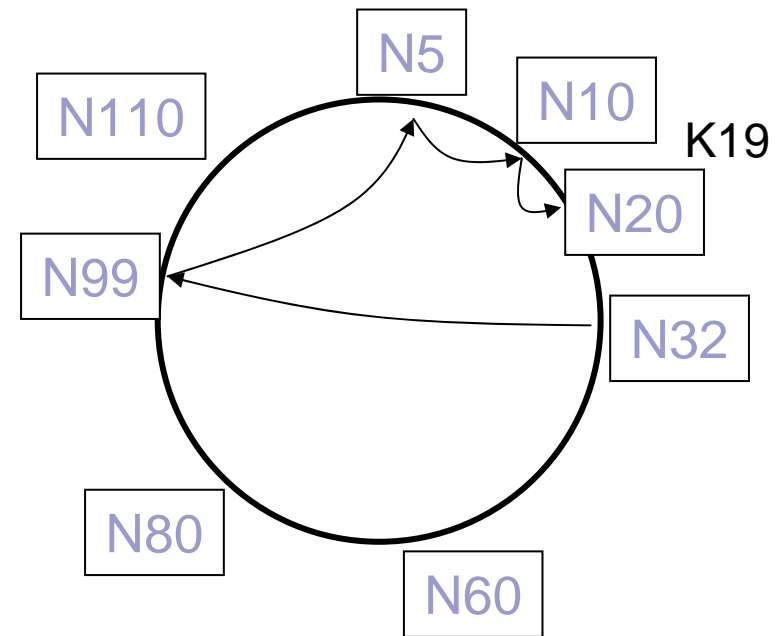
Chord

- Topology for Routing in P2P Networks
 - Nodes -- Machines
 - Edges -- TCP connections between Machines
- Routing table lists nodes:
 - $\frac{1}{2}$ way around circle
 - $\frac{1}{4}$ way around circle
 - $\frac{1}{8}$ way around circle
 - ...
 - next around circle
- $\log N$ entries in table



Chord

- Each step goes at least halfway to destination
- $\log N$ steps, like binary search



N32 does lookup for K19



Design of Canicula



Design of Canicula

- Network Distance Prediction
- Node Clustering
- Overlay Construction under Limited End-to-End Reachability
- Basic Overlay
- Message Routing
- Workflow



Network Distance Prediction

■ Triangulated Heuristic

- select N Landmark nodes
- node H is assigned coordinates which are simply given by the N-tuple of distances between H and the N base nodes
- The distance between H1 & H2 is bounded by

$$L = \max_{i \in \{1, 2, \dots, N\}} \left(\left| d_{H_1 B_i} - d_{H_2 B_i} \right| \right)$$

$$U = \min_{i \in \{1, 2, \dots, N\}} \left(\left| d_{H_1 B_i} + d_{H_2 B_i} \right| \right)$$



Network Distance Prediction

■ Distance Function


- Various weighted averages of L and U can then be used as distance functions to estimate the distance between $H1$ and $H2$.
- In our work, we use U as the distance function



Node Clustering

■ Node Join Step

- Get own network coordinates
- Find the closest cluster root
- Compare the distance with T
 - Distance $> T$:
 - Join the network as a new cluster root
 - P.S. : root nodes must not be guarded
 - Distance $< T$:
 - Join the cluster led by the closest cluster root



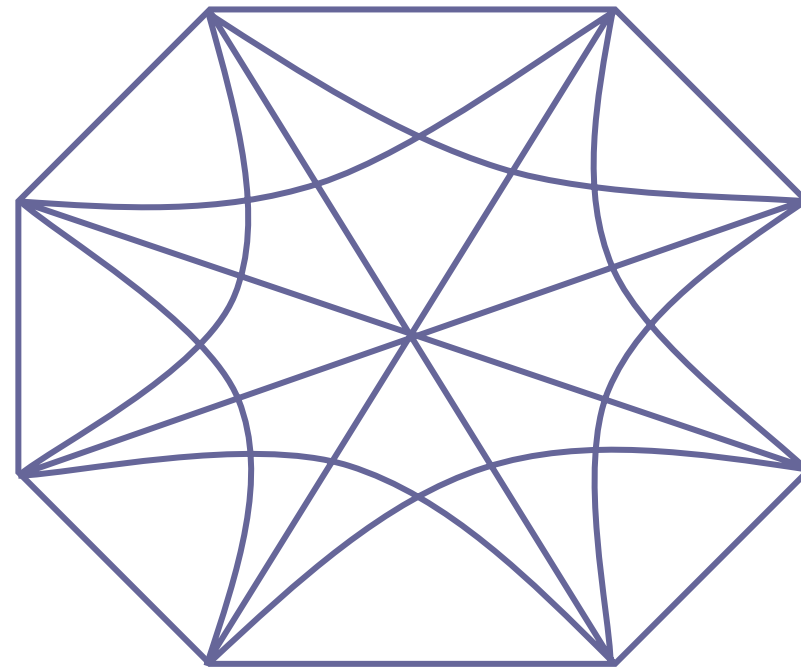
Overlay Construction under limited End-to-End Reachability

- Guarded Host Detection
- Root Election Protocol

$RRV = \langle open, lifetime, cluster \ dist \rangle$

Basic Overlay

- Bidirectional Chord
 - TCP connections are bi-directional.
 - Average Path Length
 - Uni-Chord: $b / 2$
 - Bi-Chord: $b / 3 + O(1)$



CHORD with $b=3$ (8 Nodes)



Message Routing

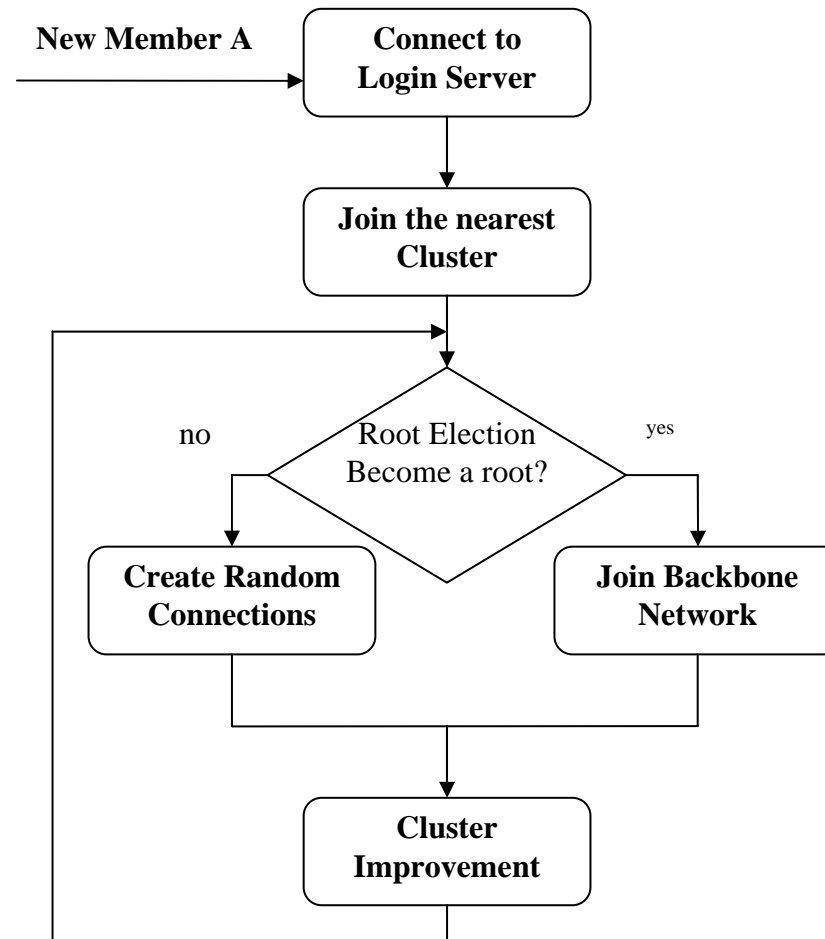
■ Hierarchical Routing

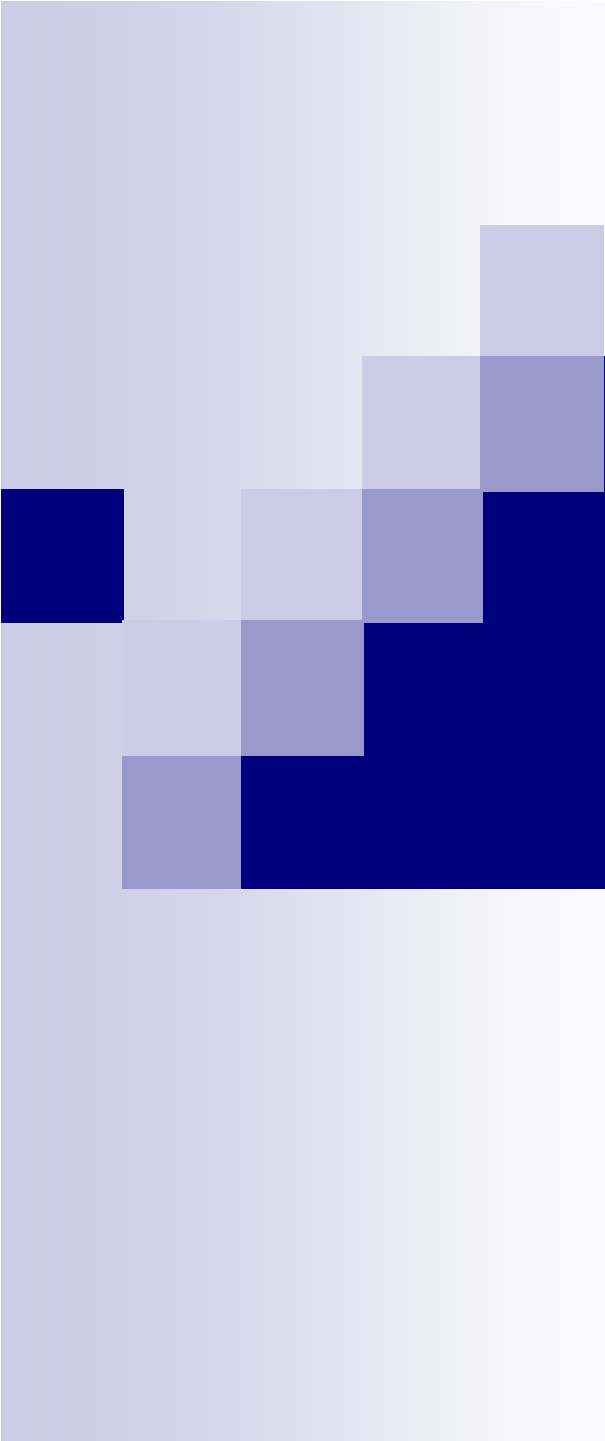
- Message are delivered from one cluster to another through the backbone network

■ Fast Routing

- Fast Routing utilizes the random connections between clusters as inter-cluster routing shortcuts

Workflow





Performance Evaluation



Main Performance Metric

- Average Relative Delay Penalty (ARDP)
 - Relative Delay Penalty (RDP) is the ratio of the overlay latency between nodes i and j to their unicast latency.

$$ARDP = \frac{1}{N} \sum_{i,j(i \neq j)}^N \frac{D'_{i,j}}{D_{i,j}}$$



Simulation

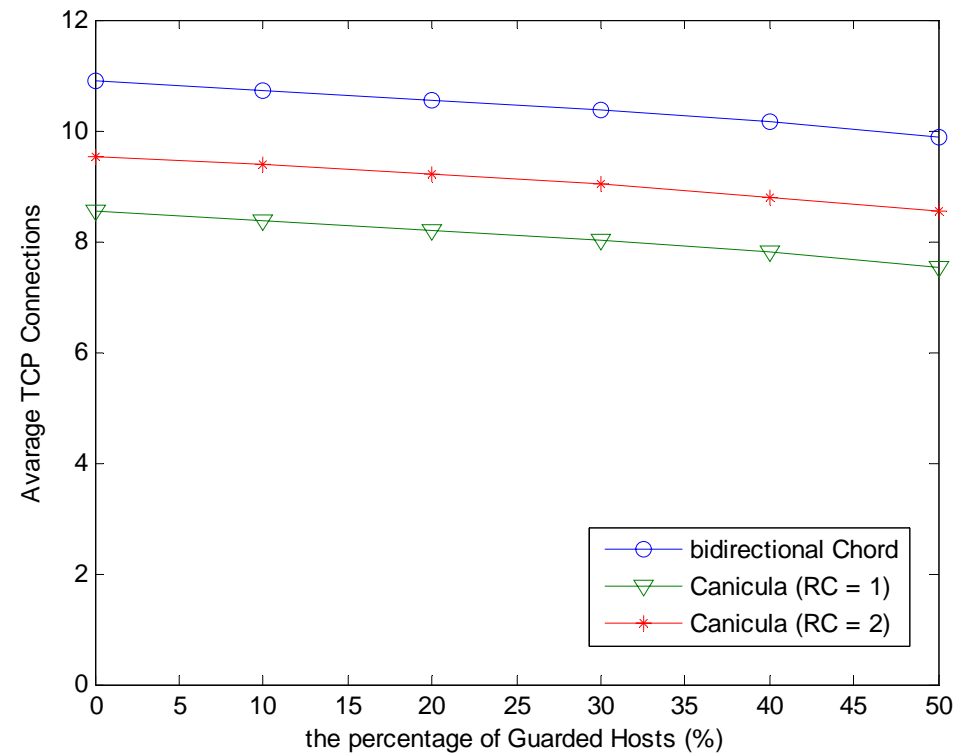


Simulation

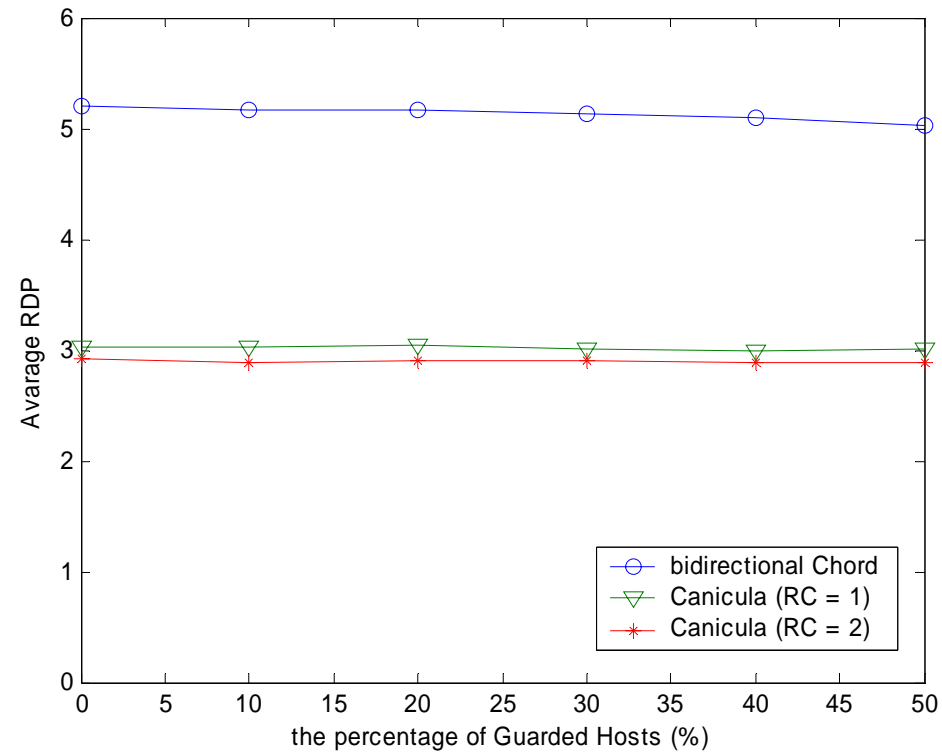
■ Simulation Setup

- transit-stub network topology
 - 287 transit nodes
 - 3000 stub nodes
- different distances to the edges in the topology
 - Intra-stub edge : 1
 - The edges between transit node and stub node : a random integer in [5; 15];
 - distance between transit nodes : using PlanetLab distance

Average TCP Connections



Average RDP





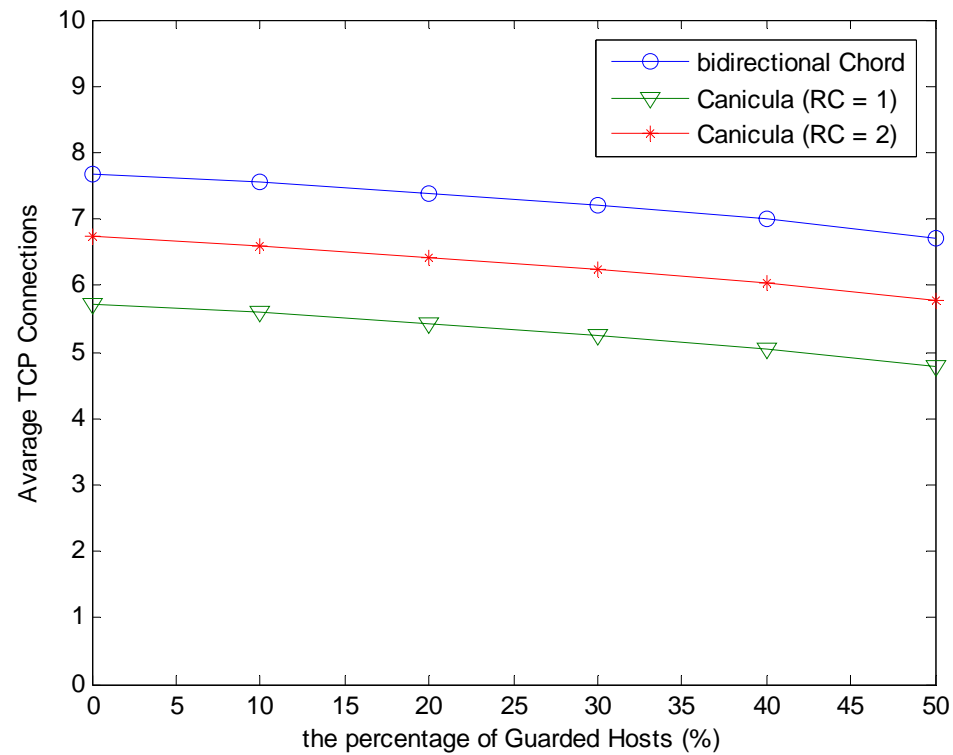
Experiment on PlanetLab



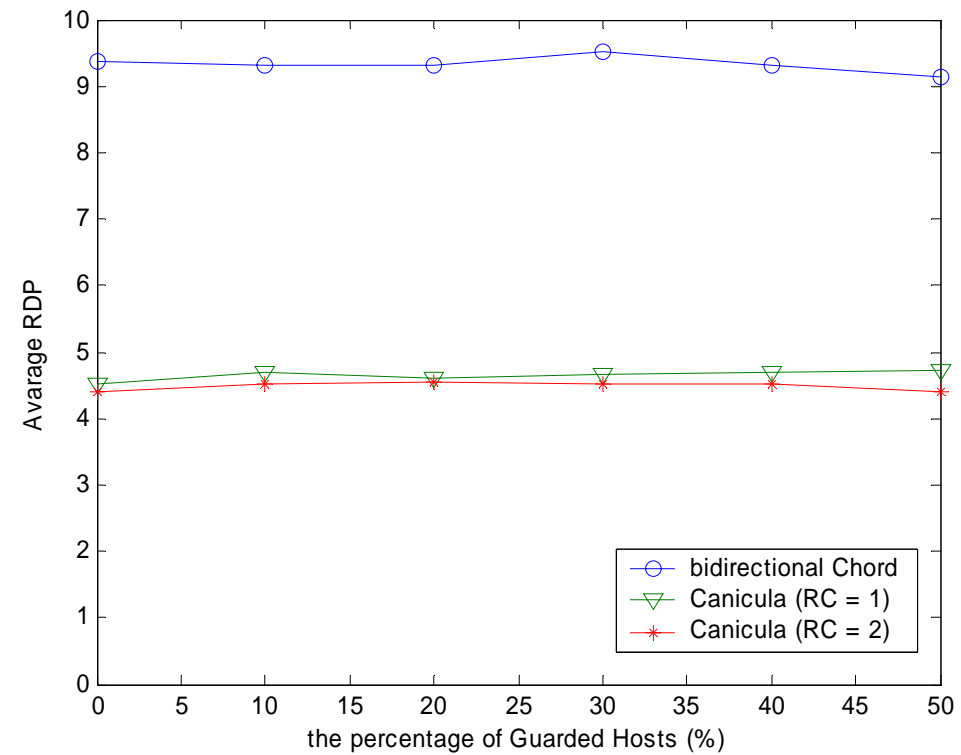
Experiment on PlanetLab

- 450 PlanetLab nodes
- Login Server : 128.31.1.14
- 15 Landmark nodes
- T is set as 40ms

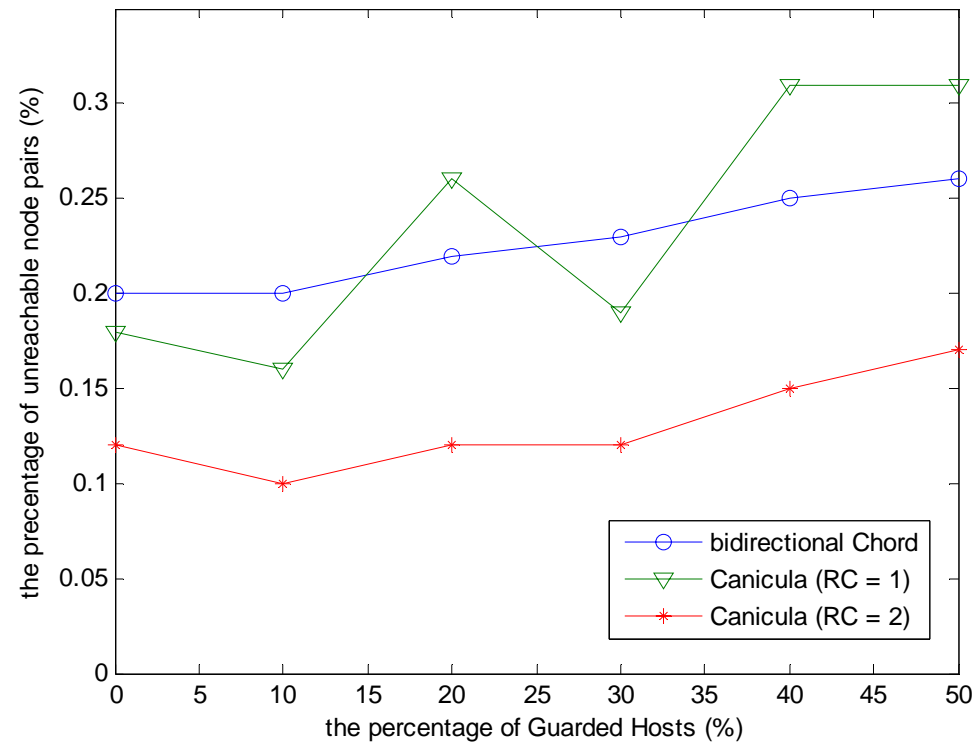
Average TCP Connections



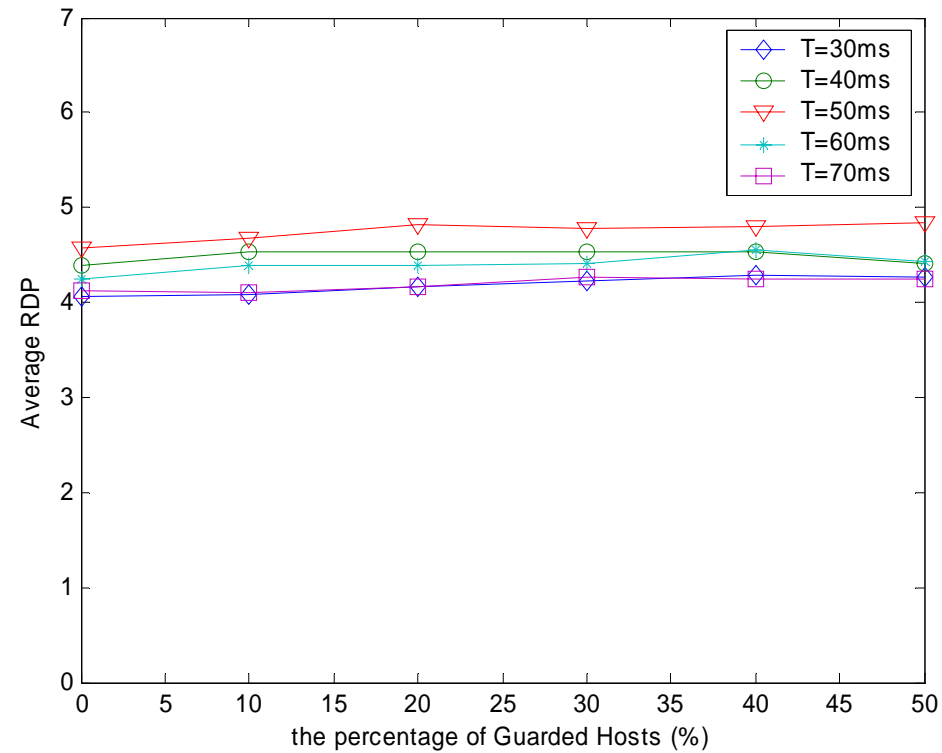
Average RDP



Unreachable node pairs



ARDP of different T





Conclusion