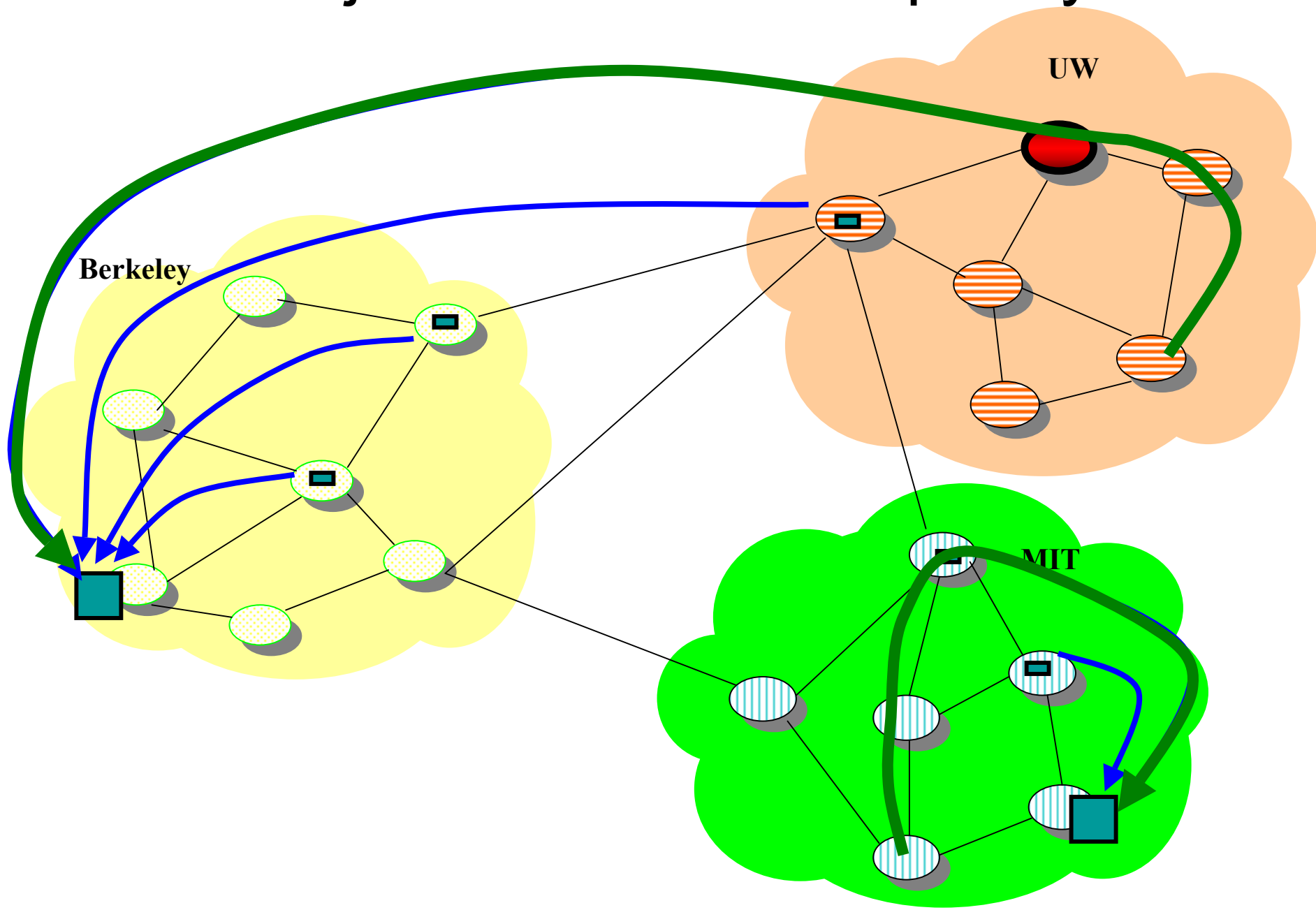
Locality Optimizations in Tapestry

Jeremy Stribling

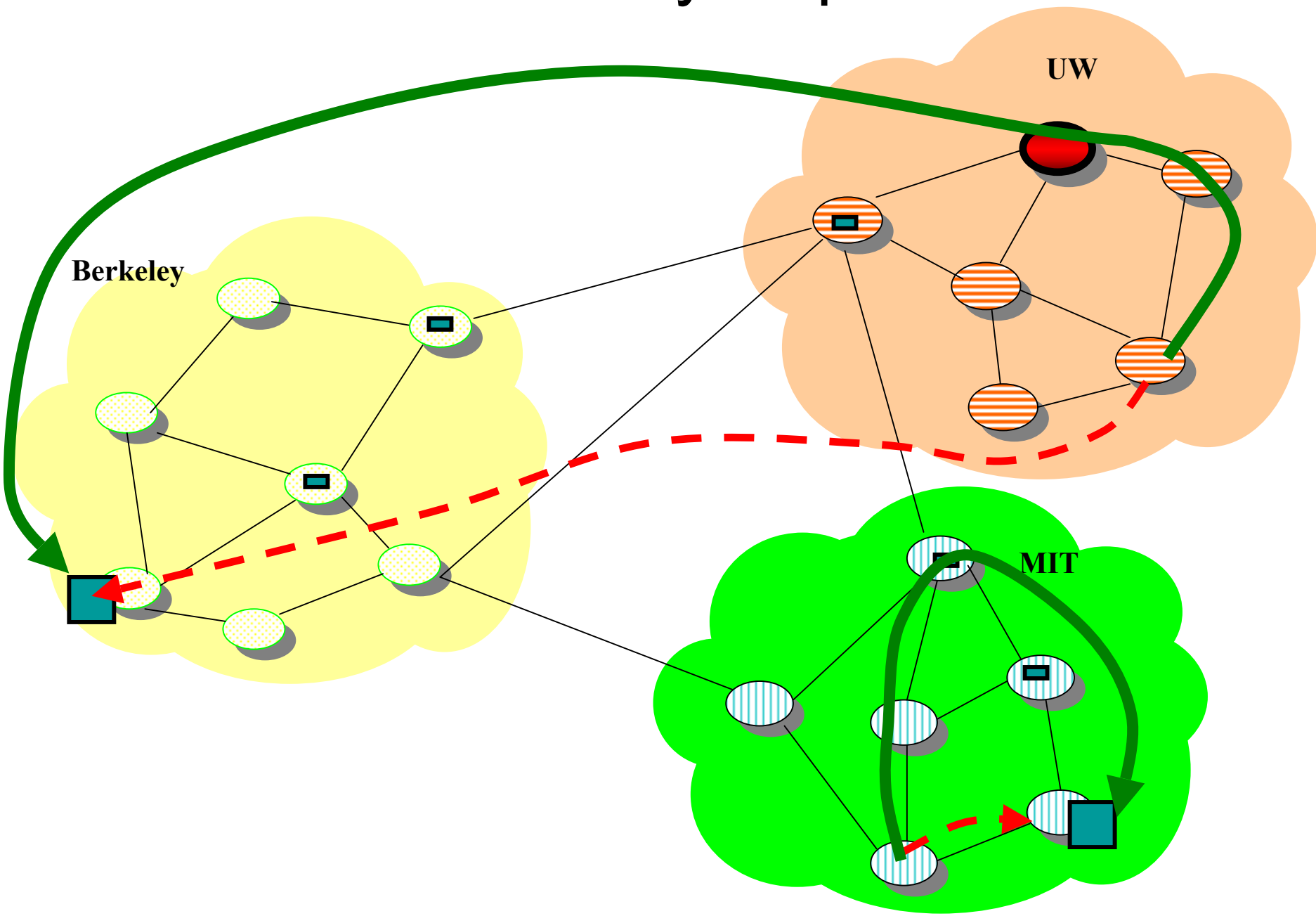
Joint work with:
Kris Hildrum
Ben Y. Zhao
Anthony D. Joseph
John D. Kubiatowicz

Sahara/OceanStore Winter Retreat
January 14, 2003

Object Location in Tapestry



Is This Always Optimal?

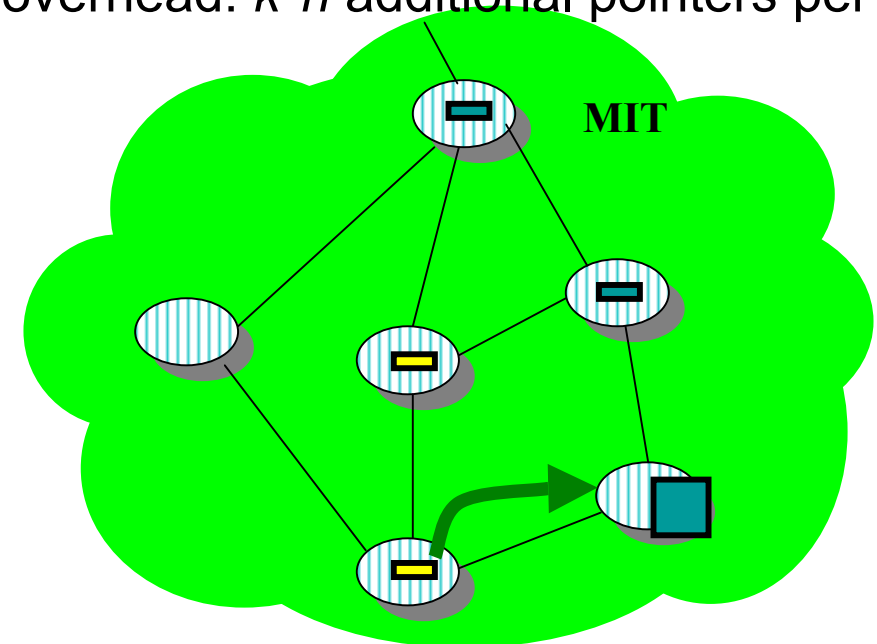


Why Is This a Problem?

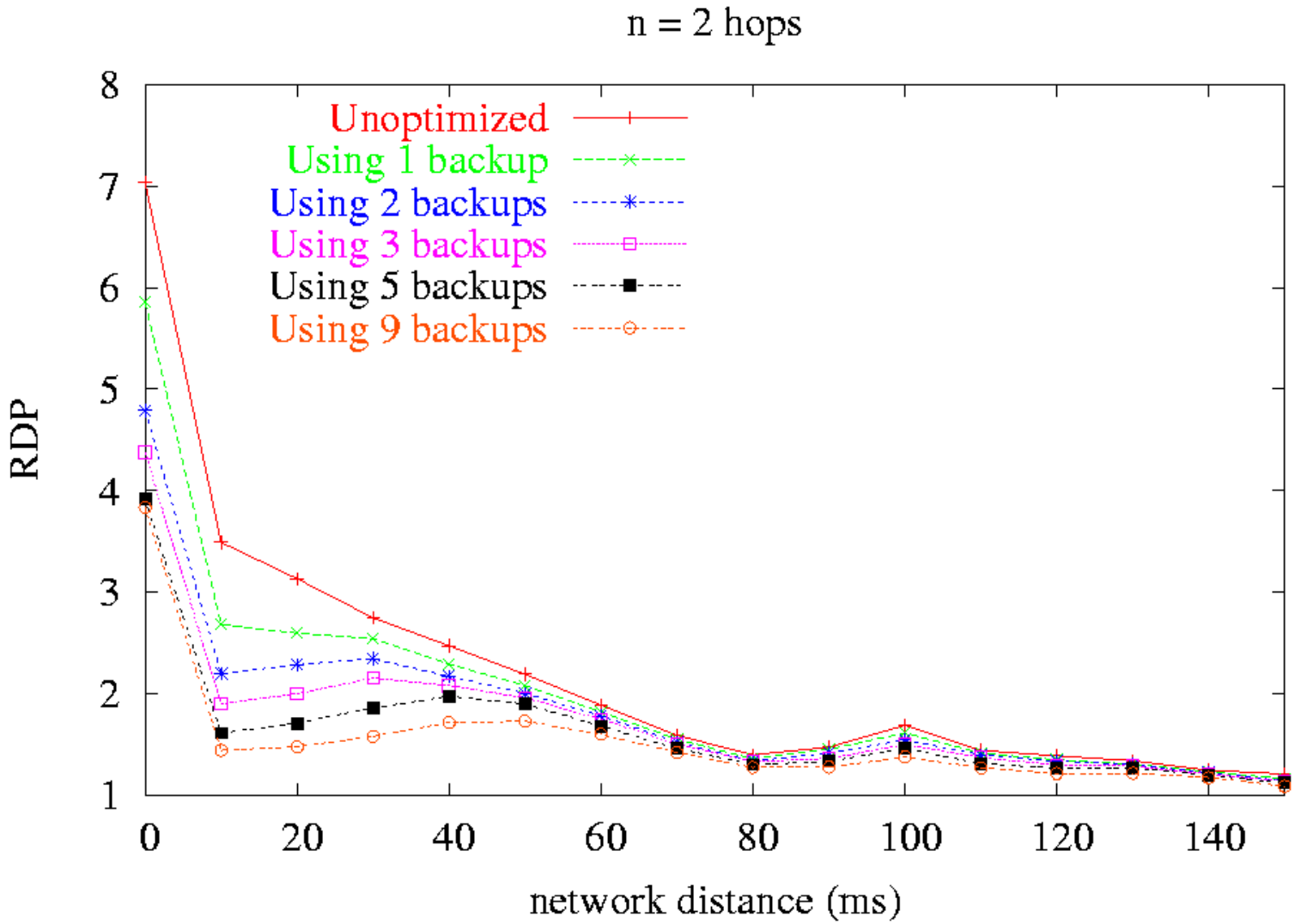
- **Example Application**: OceanStore web caching
  - If a nearby replica exists, we **must** find it quickly
- **Measure of Locality**: Relative Delay Penalty (RDP)
  - The ratio of the distance of an object in Tapestry to the minimum possible distance (i.e. over IP)
- **Problem**: finding nearby objects incurs a high RDP
  - Two extra hops have a huge relative impact if object is close
  - An issue for all similar systems, not just Tapestry
- **Solution**: trade storage overhead for low RDP

Optimization 1: Publish to Backups

- **Redundancy:** Routing table entries store up to *c* nodes
  - Closest node is the *primary neighbor*, *c*–1 nodes are *backups*
- **A simple optimization**: publish to *k* backups
  - Limit to the first *n* hops of the publish path
- **Result**
  - Nodes near the object more likely to encounter pointers while routing to the root
  - Storage overhead: *k*n* additional pointers per object

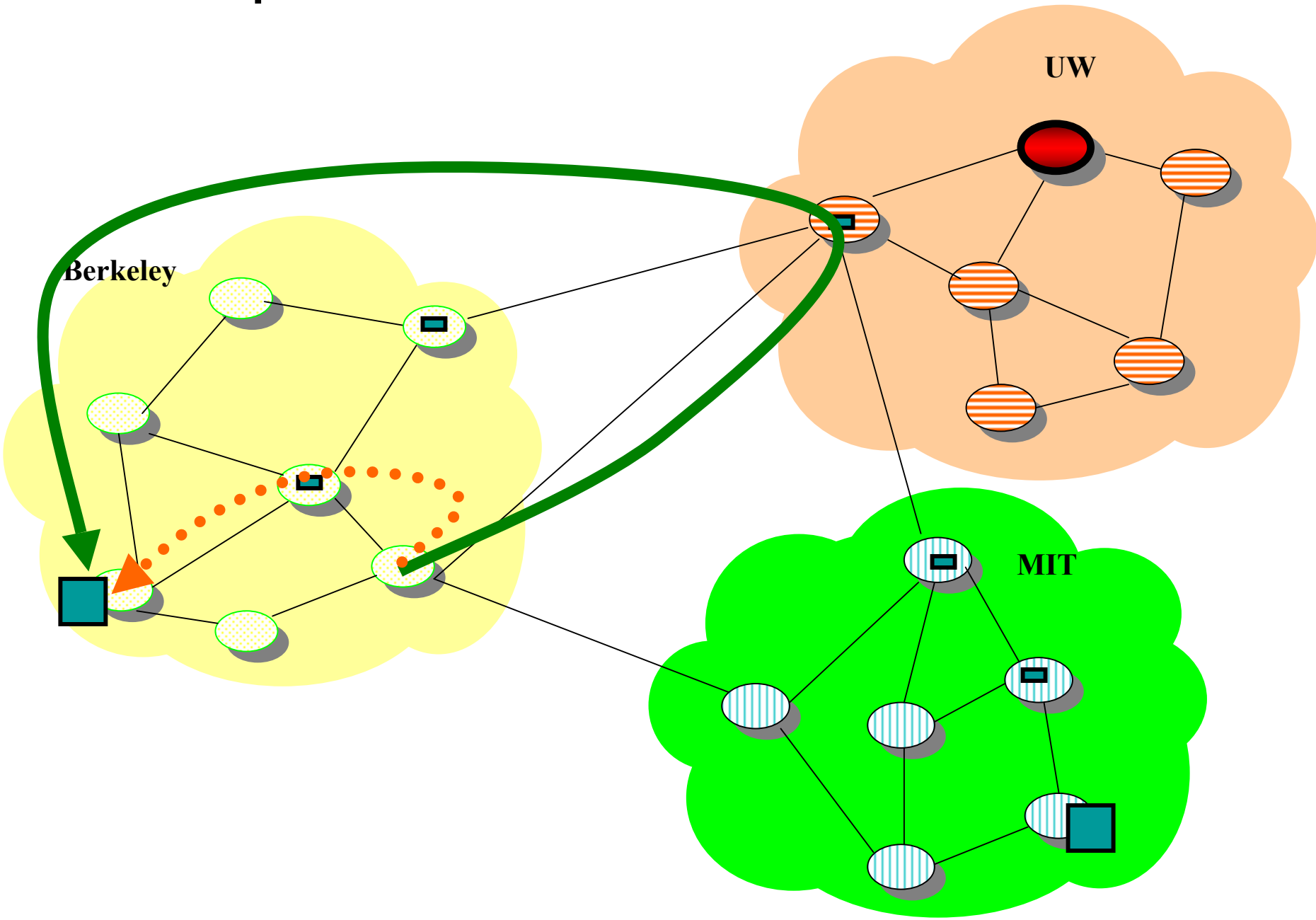


Optimization 1: Publish to Backups



Experiments run in simulation on a PlanetLab-based topology

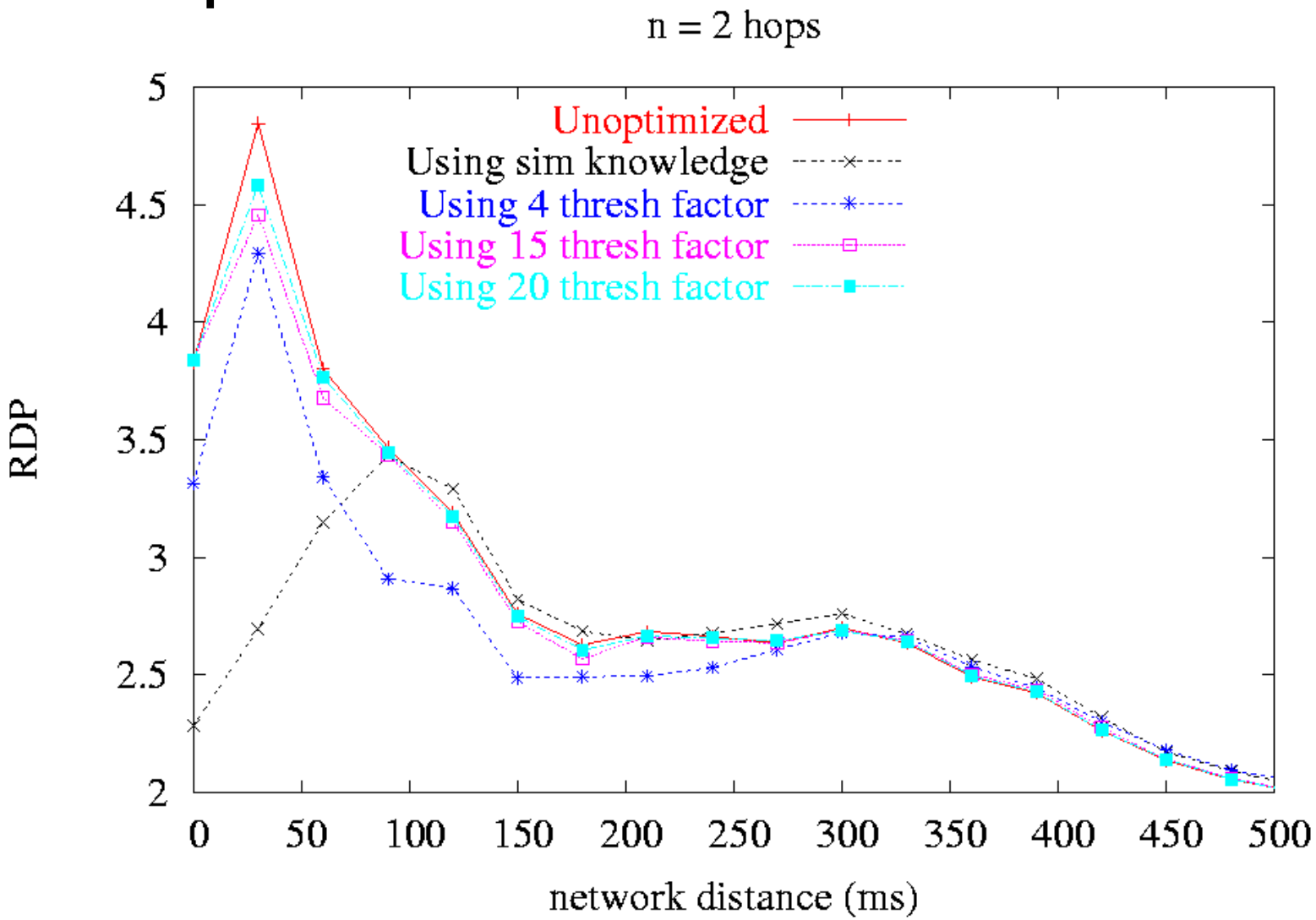
Optimization 2: Local Misroute



Optimization 2: Local Misroute

- **Solution**: Before taking “long” hop, misroute to closer node
  - Look a little harder in the local area before leaving
  - When publishing, place a pointer on *local surrogate*
- **Issue**: What determines a “long” hop?
  - One metric: if next hop is more than *m* times longer than last hop, consider it “long”
  - Call *m* the *threshold factor*

Optimization 2: Local Misroute



Experiments run in simulation on a transit-stub topology
“Using sim knowledge” indicates direct use of the topology file

Future Work

- Analyze more locality optimizations and different parameter configurations
- Take measurements on PlanetLab
- Test optimizations with real workloads (i.e. web caching)
- Complete cost analysis of storage overhead vs. RDP benefit across all optimizations