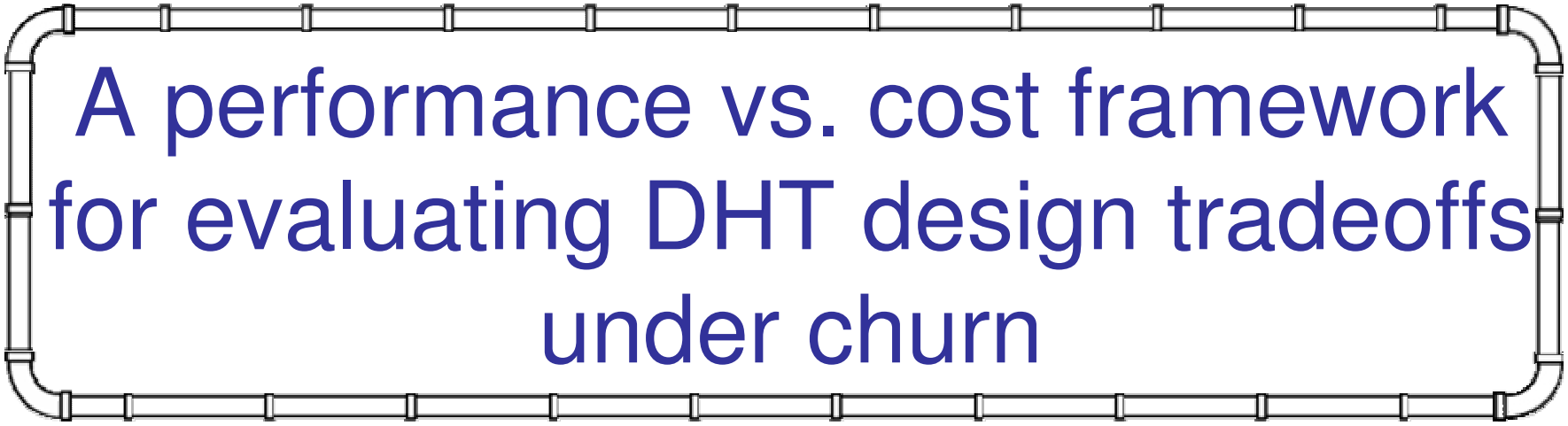
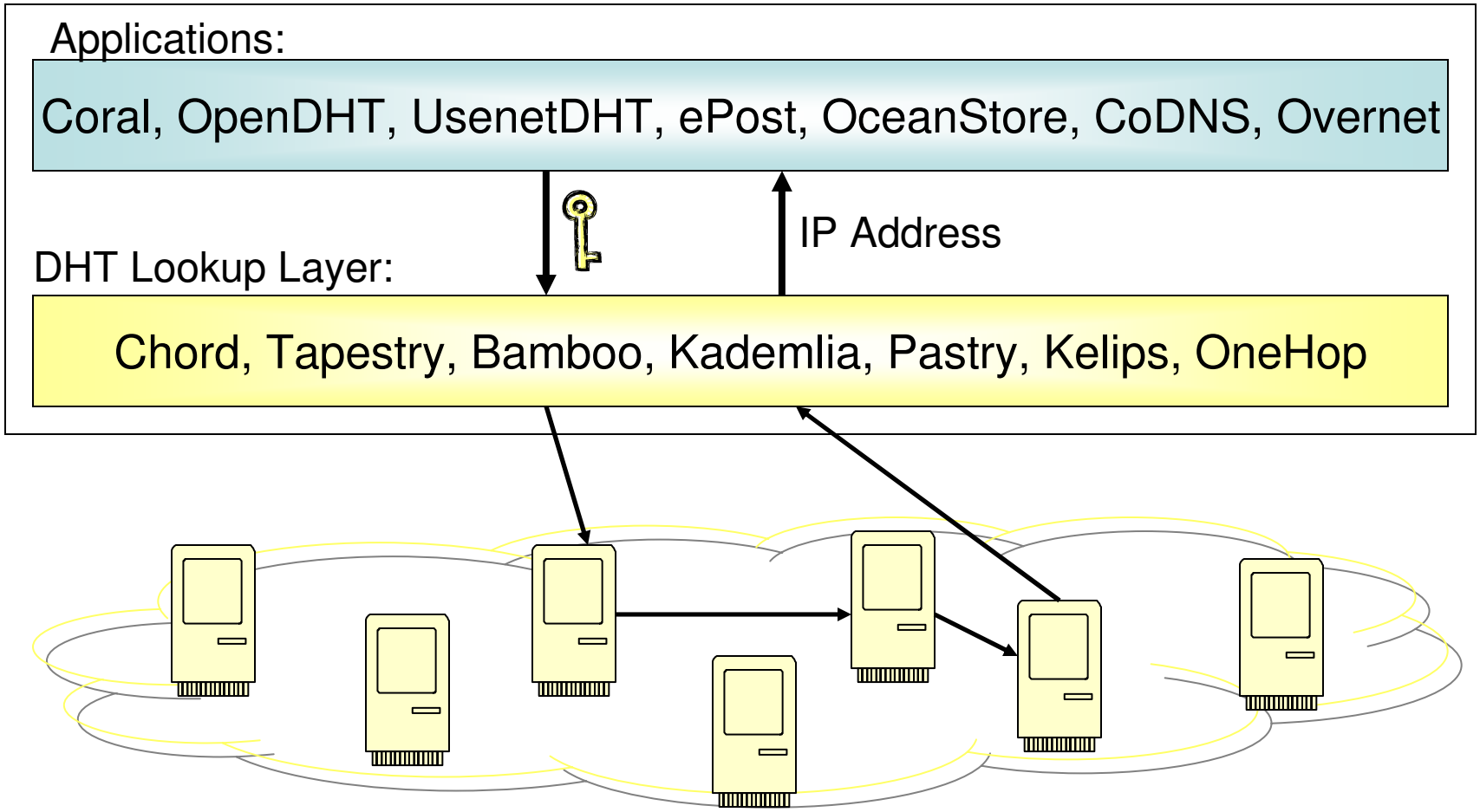


Jinyang Li, **Jeremy Stribling**,
Robert Morris, Frans Kaashoek, Thomer Gil

MIT Computer Science and Artificial Intelligence Laboratory

Distributed Hash Tables



Static DHT Comparison

DHT	Hop Count	Table Size
Chord	log n	log n
Tapestry	log n	log n
Bamboo	log n	log n
Kademlia	log n	log n
Pastry	log n	log n
Kelips	2	√n
OneHop	1	n

Static performance metrics don't reflect *cost*

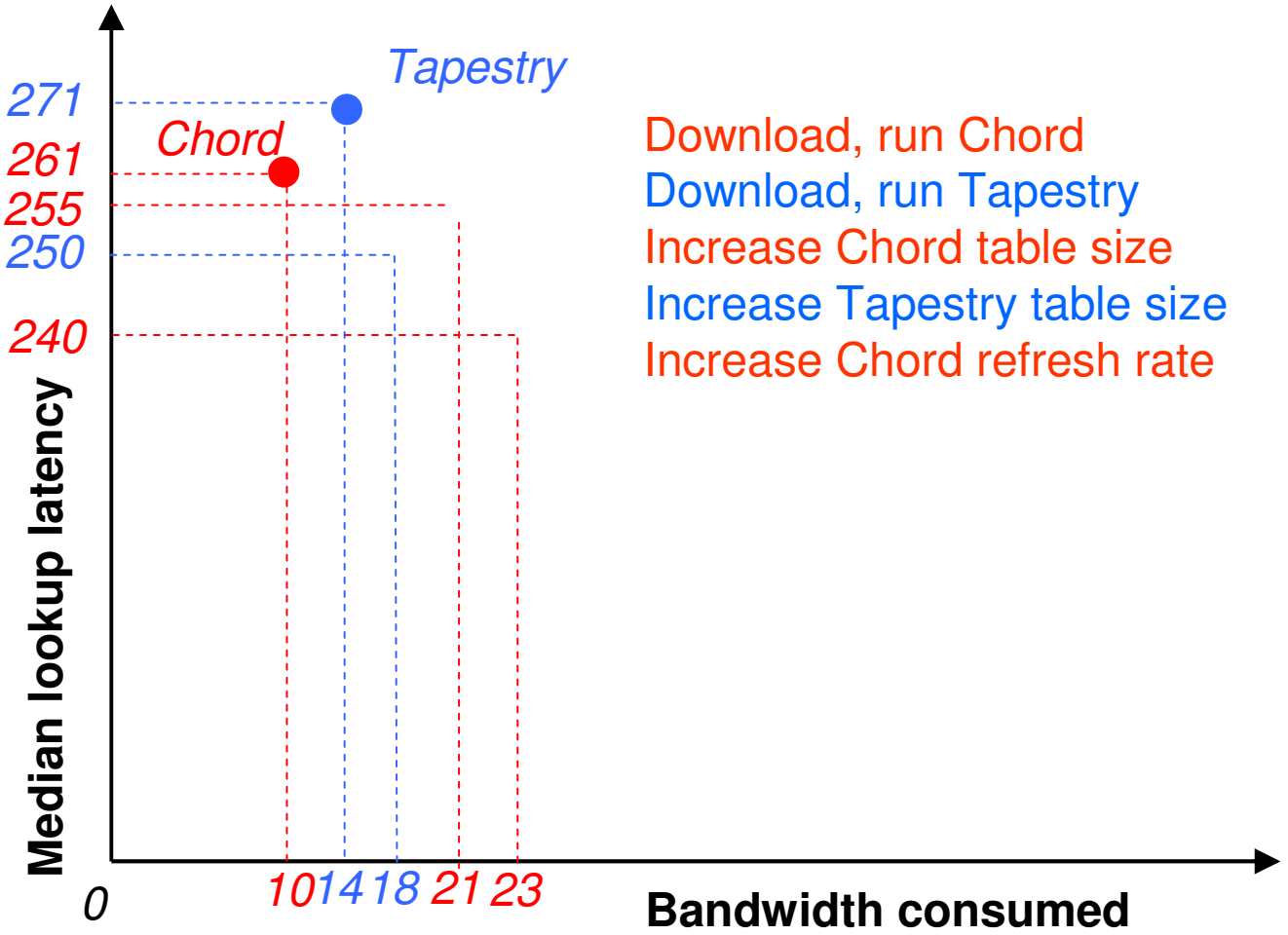
Goal: Compare DHTs Under Churn

- Protocols differ in *efficiency* of their mechanisms
- Example mechanism: Extra state
  - Reduce hop count
  - Increased table maintenance bandwidth
- **PVC**: A performance vs. cost framework
  - Performance: median lookup latency
  - Cost: average bandwidth consumed / node

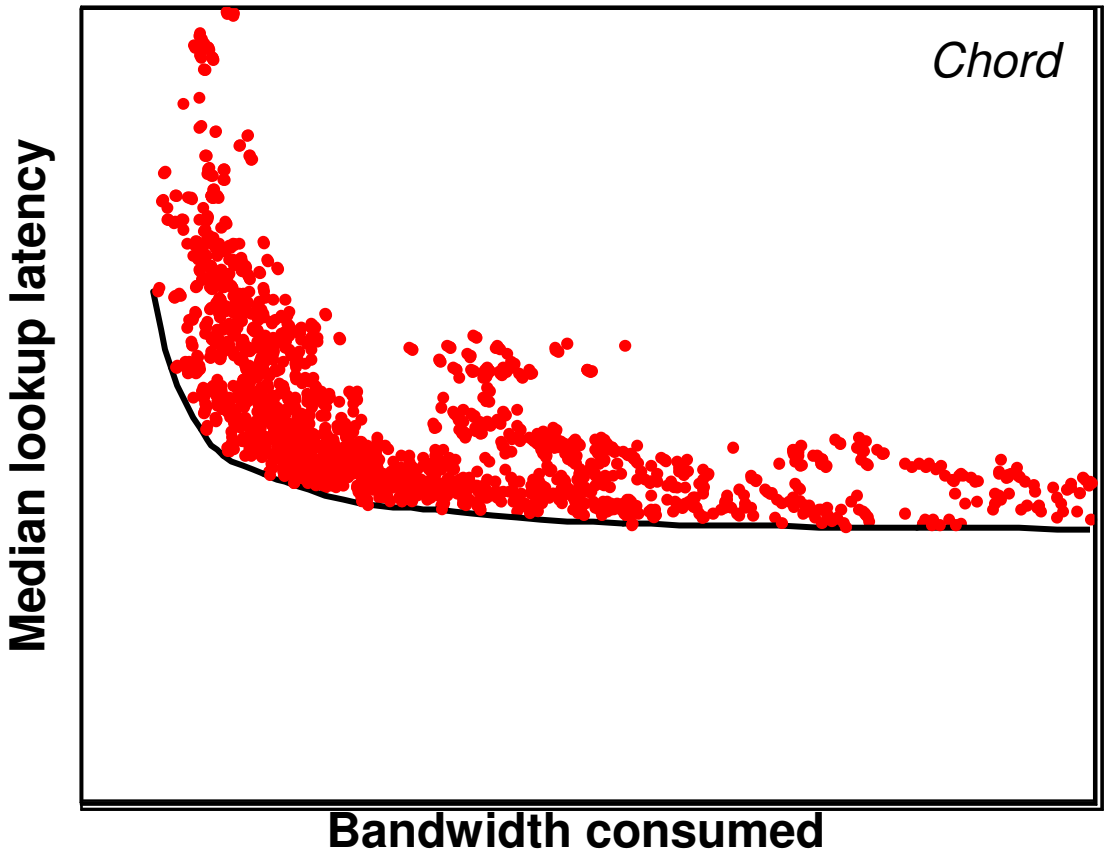
Why Is Comparing DHTs Hard?

- Protocols have lots of parameters:
  - Routing table size
  - Refresh interval
  - Parallelism degree
- How do we ensure a fair comparison?

Parameters Affect Comparison

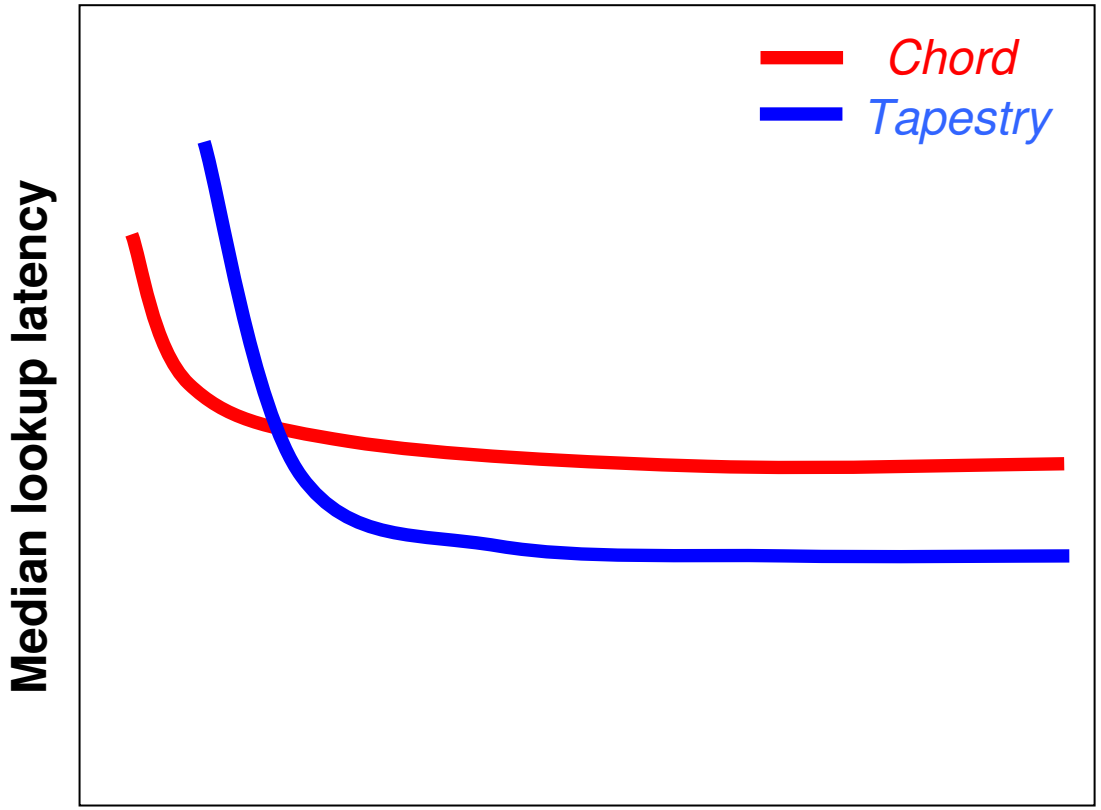


PVC Parameter Exploration

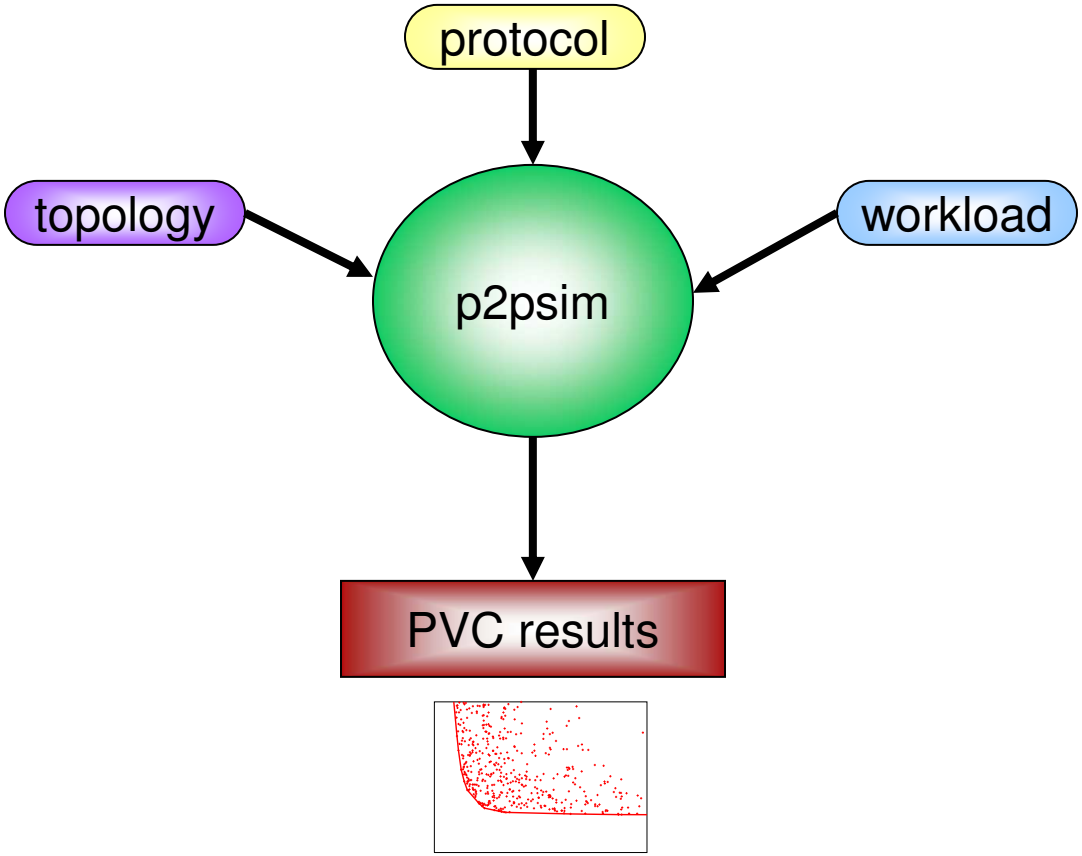


- Convex hull outlines the most efficient tradeoffs

Comparing DHTs with PVC



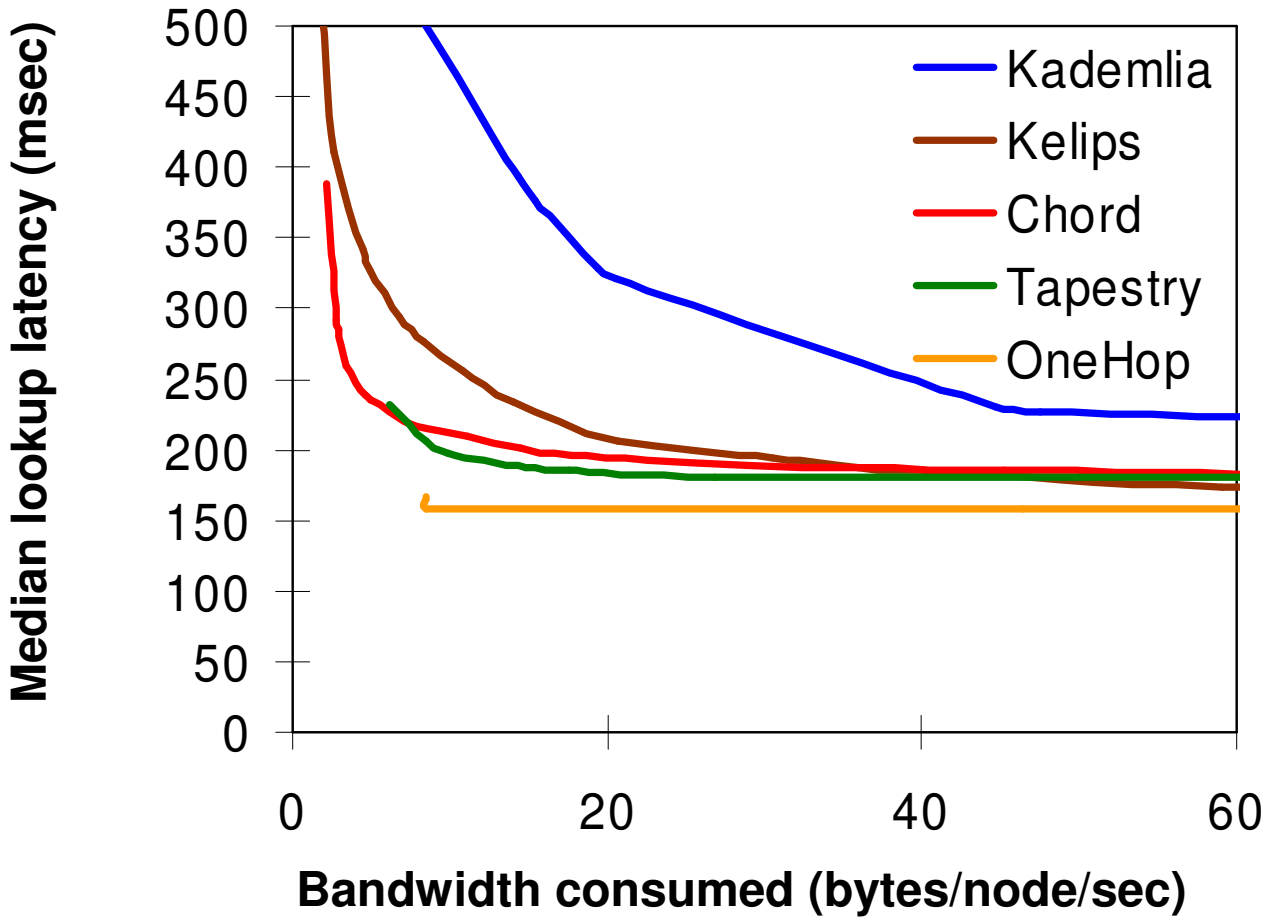
Implementation



Experimental Setup

- Chord, Tapestry, Kademlia, Kelips, OneHop
- Use measured topology of DNS servers [Gummadi et al., IMW '02]
  - 1,024 nodes
  - Median RTT: 156 ms
- Workload:
  - Each node joins/crashes with mean of 1 hour
  - Each node issues lookups with mean of 10 min

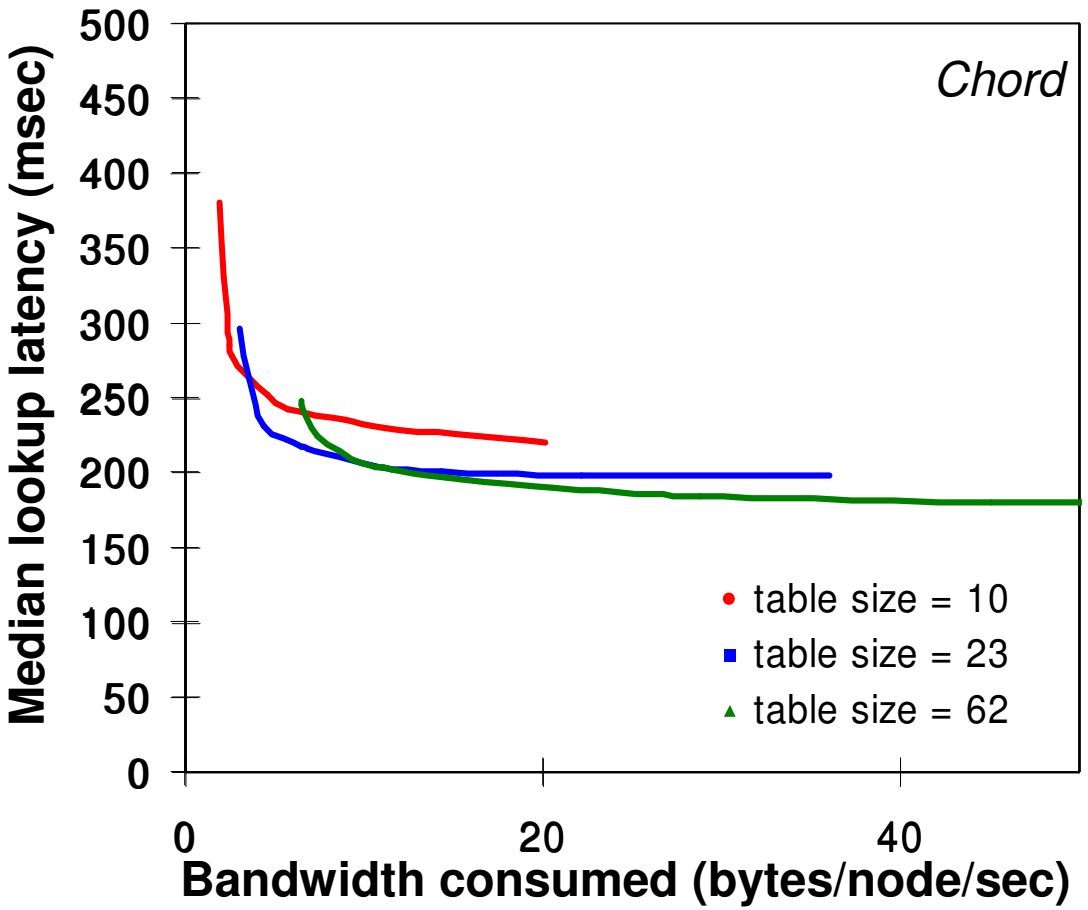
DHT Comparison



DHT Mechanisms and Parameters

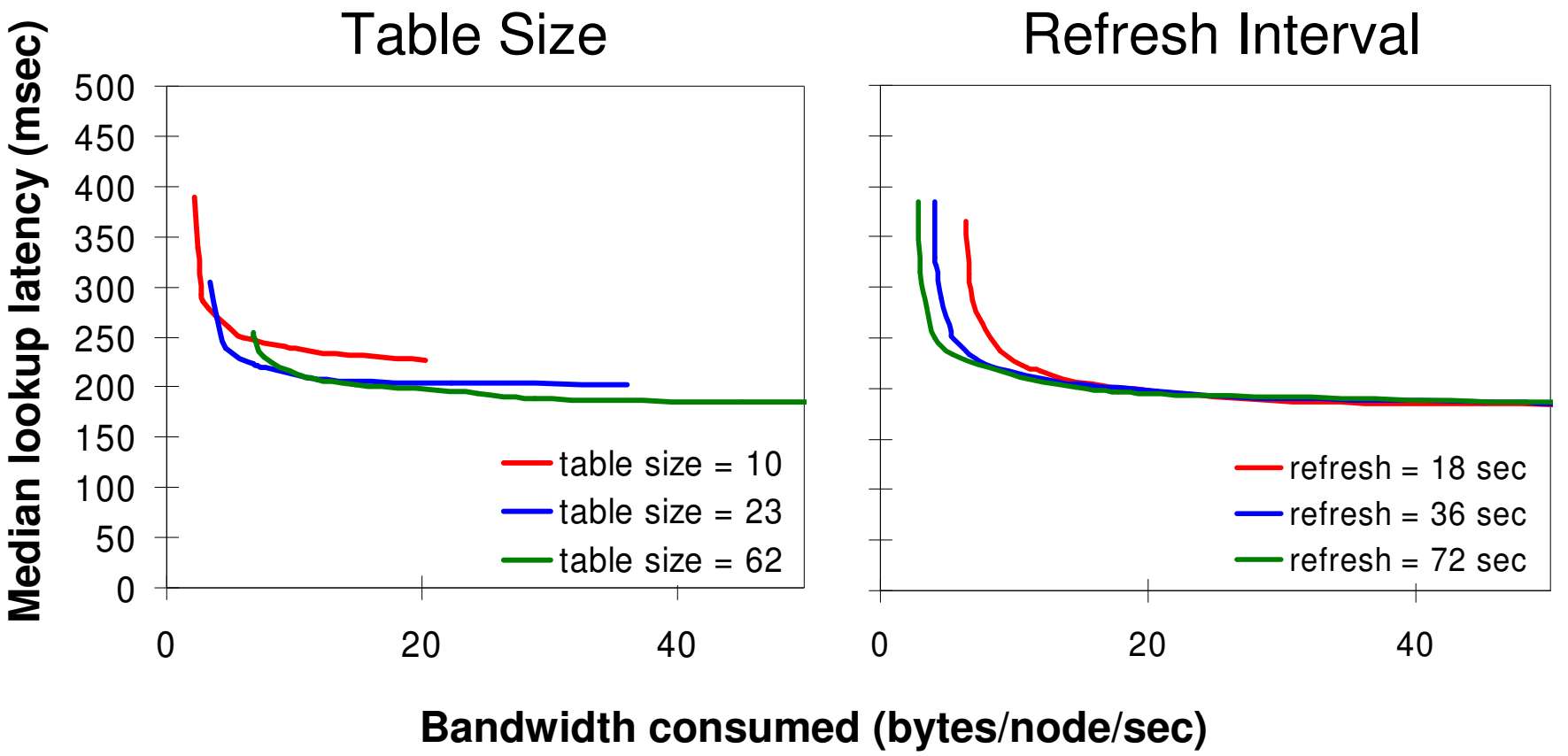
- How can we use PVC to study mechanisms?
- Observation: Parameters control mechanisms
  - Extra state → table size parameter
  - Table freshness → refresh interval

PVC Parameter Convex Hulls



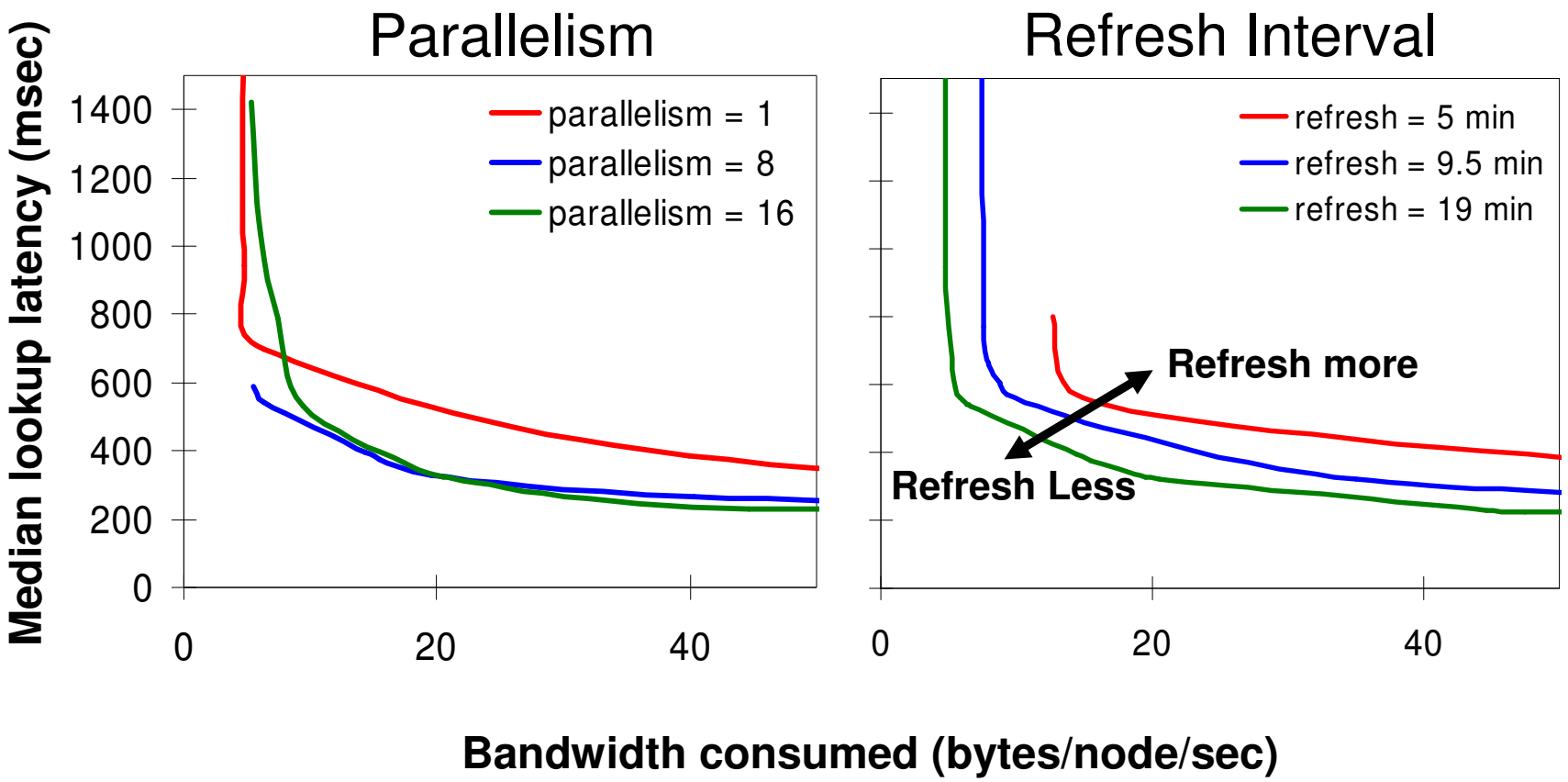
Extra State Is More Efficient than Maintaining Freshness

Chord



Parallel Lookup Is More Efficient than Maintaining Freshness

Kademlia



DHT Design Insights with PVC

Section	Insights
V	Minimizing lookup latency requires complex workload-dependent parameter tuning.
V-A	The ability of a protocol to control its bandwidth usage has a direct impact on its scalability and performance under different network sizes.
V-B	DHTs that distinguish between state used for the correctness of lookups and state used for lookup performance can more efficiently achieve low lookup failure rates under churn.
V-C	The strictness of a DHT protocol's routing distance metric, while useful for ensuring progress during lookups, limits the number of possible paths, causing poor performance under pathological network conditions such as non-transitive connectivity.
V-D	Increasing routing table size to reduce the number of expected lookup hops is a more cost-efficient way to cope with churn-related timeouts than stabilizing more often.
V-E	Issuing copies of a lookup along many paths in parallel is more effective at reducing lookup latency due to timeouts than faster stabilization under a churn intensive workload.
V-F	Learning about new nodes during the lookup process can essentially eliminate the need for stabilization in some workloads.
V-G	Increasing the rate of lookups in the workload, relative to the rate of churn, favors all design choices that reduce the overall lookup traffic. For example, one should use extra state to reduce lookup hops (and hence forwarded lookup traffic). Less lookup parallelism is also preferred as it generates less redundant lookup traffic.

TABLE I

INSIGHTS OBTAINED BY USING PVC TO EVALUATE A SET OF PROTOCOLS (CHORD, KADEMLIA, KELIPS, ONEHOP AND TAPESTRY).

the impact of different design decisions on performance under churn. Section VI relates this paper's study to previous studies. Finally, Section VII summarizes our findings.

II. PVC: A PERFORMANCE VS. COST FRAMEWORK

The goal of PVC is to address two challenges in evaluating lookup protocols for DHTs. First, most protocols can be tuned to have low lookup latency by including features such as aggressive membership maintenance, faster routing state

state storage costs (*e.g.*, the size of each node's routing table) because communication is typically far more expensive than storage. The main cost of state is often the communication cost necessary for maintaining the correctness of that state.

In PVC, nodes try to forward lookups to the node responsible for the lookup key. The identity of the responsible node is returned to the sender as the result of the lookup. A lookup is considered failed if it returns the wrong node among the current set of participating nodes (i.e. those that have completed

Related Work

- Protocols designed for churn:
  - Bamboo [Rhea et al., USENIX '04]
  - MSPastry [Castro et al., DSN '04]
- Churn theory:
  - Statistical theory of Chord under churn [Krishnamurthy et al., IPTPS '05]
  - Chord half-life [Liben-Nowell et al., PODC '02]
- Churn metrics:
  - K-consistency [Lam and Liu, SIGMETRICS '04]

Summary/Future Work

- A unified performance/cost framework for DHT evaluation (PVC)
- We used the lessons from our PVC-based study to design **Accordion** [Li et al., NSDI '05]
  - One parameter: bandwidth
  - Self-tunes all others

http://pdos.csail.mit.edu/p2psim