“Designing a DHT for low latency and high throughput”

Frank Dabek

presented by Besim Avci
Outline

• Background
• Evaluation Methods
• Designing for low latency
• Achieving high throughput
• Conclusion
Background

• Chord
• DHash++
• Synthetic Coordinates
Chord

- Square: successor of the key
- Triangles: lookup path
- Circles: successor’s successor list
- Tick mark: position of key
DHash++

- `<key, value>` system

Synthetic Coordinates

- Prediction of latencies between nodes
- Vivaldi algorithm
System Overview
Evaluation Methods

- Evaluation infrastructure
  - PlanetLab and RON testbeds
- Application Workload
  - Semantic Free Referencing System
  - UsenetDHT
Designing for Low Latency

- Data Layout
- Iterative or Recursive
- Proximity Neighbor Selection
- Coding versus Replication
- Integrating Routing and Fetching
Data Layout

• The first decision to be made about where a DHT should store data is whether it should store data at all.
• Second layout decision is the size of the units of data to store
• Third layout decision is which server should store each block of data
Recursive vs Iterative

- Recursive
  - Low latency
  - Enables TCP

- Iterative
  - Better network management
Proximity Neighbor Selection

- PNS(x) algorithm, x to be idealized
- PNS approximates 1.5 times the average round trip time.
- It uses Vivaldi distances for closeness
Coding vs Replication

• In the case of replication, the originator’s strategy should be to fetch the required data from the successor with lowest latency.
• What is the optimal replication parameter (number of fragments stored) / (number of fragments needed to reconstruct)?
Integrating Routing and Fetching

- Lookup stops as soon as it finds enough number of fragments
- However, this design choice decreases the lookup time at the expense of data fetch latency
Low Latency Design Choices
Achieving High Throughput

- TCP Transport
- STP Transport
  - STP window control
  - Retransmit timers
  - Retransmit policy
Throughput Results
Conclusion

- Techniques can reduce fetch latency by factor of 2 and increase throughput when combined together
- Dynamic network behavior needs to be analyzed.
- Security concerns needs to be addressed.
Questions..