PROBLEMS WITH DHTS

1. Assignment Violation – Multiple nodes want to own the same key.
2. Routing Violation – A node with incorrectly route messages due to a failure to recognize a new node configuration.
GOALS OF SCATTER

1. Consistency
2. Scalability
3. Adaptability
BASIC BUILDING BLOCK IS A GROUP. NOT A NODE.

- Groups use Paxos which means..
- They have an elected leader
- There are ways to add and drop members
- Use quorum to decide on operations
- Maintain internal integrity – Scatter can assume consistency within groups!
GROUP INTERACTION – STATIC GROUPS NOT ENOUGH

Groups maintain knowledge of proceeding and following groups (i.e. their members)

1. Merge – the Union of Two Groups
2. Migrate – Move members around
3. Split – Break one group into two
4. Repartition – change way keys are divided among two groups
HOW DO GROUPS COMMUNICATE? NESTED CONSENSUS

- They keep on saying two-phase commitment (2PC)
- Consensus within group
- Agreement between groups
THAT’S A LOT OF STEPS!
STORING STUFF

- Each member of a group is the “primary” for some subset of keys the group owns.
- Uses UDP – so no guarantees anything gets anywhere in any order.
<table>
<thead>
<tr>
<th>Policy</th>
<th>Load Balance</th>
<th>Latency</th>
<th>Resilience</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low churn, uniform latency</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low churn, non-uniform latency</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>High churn, non-uniform latency</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
EVALUATION

Looked at three cases:

1. Low churn/uniform latency – 150 nodes, all homogenous.

2. Low churn/non-uniform latency – 50 nodes from California, Virginia, and Utah. (homogenous?)

3. High churn/non-uniform latency – 840 nodes from PlanetLab (a random group of computers meant for research)
LATENCY

(a) Single Cluster

(b) Multi-Cluster
EFFECT OF THE LATENCY-JOIN POLICY

AT PLANET LAB! SO THE HIGH CHURN/NON-UNIFORM LATENCY SITUATION
THROUGHPUT

(a) Single Cluster

(b) Multi-cluster
GROUP OPERATIONS

Figure 11: Scatter group throughput in single cluster and multi-cluster settings.
SHOWING THEIR STUFF - CHIRP

• Twitter-like application
  • Create a user name.
  • Username is associate with two keys – updates and follows (your tweets and who you follow)

• Deployed it through PlanetLab.
• Compare it to OpenDHT
PERFORMANCE – POISSON

(a) Performance

(b) Availability

(c) Consistency
PERFORMANCE - PARETO

(a) Performance
(b) Availability
(c) Consistency
LOAD BALANCING

(a) Latency

(b) Node Load
SCALABILITY – THE WHOLE POINT!
SCATTER

- Fault tolerant groups
- Groups self organize (membership, load balancing)
- Scalable AND reliable