On Zone Balancing of Peer to Peer Networks: Analysis of Random Node Join

Xiaoming Wang, Yueping Zhang, Xiafeng Li, and Dmitri Loguinov

Computer Science
Texas A&M University
College Station, TX 77843
Motivation

- Structured P2P systems construct DHTs (Distributed Hash Tables) for efficient routing
  - Chord, CAN, de Bruijn

- Data objects are hashed into some virtual coordinate spaces

- Each user holds a zone in the DHT space
  - Stores data objects within its zone and answers queries for these objects

An instance of zone partition
Motivation 2

• Notice that the amount of user load is proportional to zone size
  – Imbalance can lead to “hotspots” and lower performance

• In addition, graph structure is unbalanced
  – Which leads to increased diameter, smaller node degree, lower bisection width

• Our paper studies how zone-balancing decisions during node join affect the resulting zone sizes
  – We derive the probability bounds on the maximum and minimum zone sizes
Basics

• Consider a system with $n$ users
  – Assume a sequential join process

• Define two metrics for load balancing:

  \[
  f_{\text{min}} = \frac{\text{avg}}{\text{min}} \quad f_{\text{max}} = \frac{\text{max}}{\text{avg}}
  \]

• We focus on the bounds of these two metrics that hold with probability $1 - n^{-\varepsilon}$ ($\varepsilon > 0$)
Random Join Process

- Each new user randomly samples one or more existing peers and splits one of their zones
- The join decision includes two factors:

  - **Splitting**
    - Random
    - Center
  - **Sampling**
    - Single-Point
    - Multi-Point
Random Join Process 2

• We will compare these algorithms in terms of $f_{max}$ and $f_{min}$
  – The optimal bound for the two metrics is 2
  – No method can achieve better load-balancing

• Due to the time limit, we skip the single-point algorithms
  – Summary for random and center splits:

<table>
<thead>
<tr>
<th>Random</th>
<th>Center</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_{max} \leq (1 + \varepsilon) \log n$</td>
<td>$f_{max} \leq (1 + \varepsilon) \log n$</td>
</tr>
<tr>
<td>$f_{max} \leq n^{1+\varepsilon}$</td>
<td>$f_{max} \leq 3.246^{\sqrt{\log n}}$</td>
</tr>
</tbody>
</table>
A Big Map

Motivation

Single-point
- Random Split
- Center Split

Multi-point
- Random \(d\)-sampling
- Deterministic \(d\)-sampling

P2P Simulations

Conclusion
Multi-Point Center-Split

• Next we examine multi-point schemes
  – We use center-split for the rest of the talk

• Greedy methods
  – Motivated by the “power of two choices”

• Idea: extend the center-split model to sample $d$ random points before the actual join

• Intuitive observation:
  – The more points sampled, the better the graph is balanced, but what are the actual bounds?
Multi-Point Center-Split 2

- The extreme case is to sample every peer
  - The resulting $f_{max}$ is always optimal and concentrates on the ideal value 2
- However, this method will suffer from huge traffic overhead
- Thus, the tradeoff is between:
  - The balancing performance of the algorithm, and
  - The amount of sampling traffic
- Next we study two multi-point schemes and present our analysis of this problem
A Big Map

Motivation

Single-point
- Random Split
- Center Split

Multi-point
- Random \(d\)-sampling
- Deterministic \(d\)-sampling

P2P Simulations

Conclusion
Purely Random $d$-sampling

- The method samples $d$ independent uniformly random points $X_1, X_2, \ldots, X_d$
  - Splits the largest zone among the $d$ choices

- How does the performance improve as a function of $d$?

- Based on the “balls-into-bins” model, we derive an asymptotic bound on $f_{\text{max}}$
  - The analysis is intractable when applying this model to $f_{\text{min}}$
  - We leave this direction for future work
Purely Random $d$-sampling 2

- **Theorem 1**: Under $d$-point sampling and center-splits, the following bound holds with probability at least $1 - \frac{c}{n}$

  $$f_{\text{max}} \leq 2 + \frac{(1 + \varepsilon) \log n}{d} - \frac{\Theta(\log(d + \log n))}{d}$$

- For $d = 1$, it reduces to the single-point model

  $$f_{\text{max}} \leq (1 + \varepsilon) \log n - \Theta(\log \log n)$$

- For $d \geq 2$, the term $(1 + \varepsilon) \log n$ is scaled down by a factor of $d$
  
  - The “power of two choices” bound $\log \log n / \log d$ is not achieved here
Simulating Random $d$-sampling

- Each of the following simulations is run for 1,000 graphs with 30,000 nodes each.
Further Discussion

• For $d = c \log n$,

$$f_{\text{max}} \leq 2 + \frac{1 + \varepsilon}{c} - o(1)$$

• For $c \to \infty$, the second term goes to zero
  – And $f_{\text{max}}$ is bounded by 2 with high probability

• Recall from the single-point method,
  – $f_{\text{max}} \leq 28$ for $n = 10^6$ with probability $1 - 1/n$

• The improvement is significant
  – But results in additional traffic overhead
Reducing Traffic Overhead

• How to reduce the join overhead?
  – While keeping the graph balanced

• Idea:
  – Randomly sample a peer
  – Then deterministically sample its neighbors
  – Subsequently walk along the edges of the graph to find additional peers to sample

• Two walking strategies:
  – Random walk selects arbitrary (random) neighbors
  – Biased walk selects the largest neighbors
Reducing Traffic Overhead 2

• Intuition: "larger" nodes are more likely to know additional "large" nodes

• This reduces the join overhead by a factor of \( \Theta(kD_{av}) \)
  – \( k \) is graph degree and \( D_{av} \) is the average distance

• The exact analysis is nontrivial since the walk process depends on the state of peers
  – We leave the exact model for future work

• Instead, we study a similar deterministic model
  – According to our analysis, it provides a lower bound on the performance of the other \( d \)-walk models
A Big Map

Motivation

Single-point
- Random Split
- Center Split

Multi-point
- Random d-sampling
- Deterministic d-sampling

P2P Simulations

Conclusion
Deterministic $d$-sampling

• The model samples a random point $X_1$
  – Then checks $d - 1$ additional points according to a simple deterministic rule
  – Points $X_2, \ldots, X_i, \ldots, X_d$ are obtained by adding $i/d$ of the total size of the DHT space to $X_1$

• An example of $d = 4$
  – $X_1$ is the first random sample
  – The points $X_2, X_3, X_4$ are found by adding $1/4, 1/2, \text{ and } 3/4$ of the circle’s circumference to $X_1$
Deterministic $d$-sampling 2

- **Theorem 2**: In deterministic sampling, the following bound holds with probability at least $1 - n^{-\varepsilon}$

$$f_{max} \leq 2 + \frac{(1 + \varepsilon) \log n}{d} + \eta - \frac{\Theta(\log \log n)}{d}$$

where $\eta = \log \left(1 + \frac{1+\varepsilon}{c} + \log \left(1 + \frac{1+\varepsilon}{c} + \ldots\right)\right)$

- This result differs from that of random $d$-sampling by a constant $\eta$

- Notice that $\eta$ is positive
  - Thus, the deterministic model is worse than the random model
  - But how much is the difference?
Simulating Deterministic Sampling

- The model is conservative on some points
  - Round-off errors at $d$ not powers of 2
Purely Random vs Deterministic

- With the previous results on $f_{max}$, we compare the two multi-point models ($\varepsilon = 0.22$)

$n = 10^6$

$n = 10^7$
Purely Random vs Deterministic

- Further question:
  - How many samples does the deterministic model need to approximate the random model?

- **Theorem 3**: Assuming that the random method samples $c_1 \log n$ points and the deterministic method samples $c_2 \log n$ points, the corresponding upper bounds on $f_{\text{max}}$ are equal if

\[
  c_2 = \frac{(1 + \varepsilon)c_1}{1 + \varepsilon - c_1 \log(1 + \frac{1+\varepsilon}{c_1})}
\]
Pure Random vs Deterministic 3

• For $c_1 = 1$ ($f_{max} \leq 4$) and $\varepsilon = 1$ (probability $1 - 1/n$), the two methods are equivalent if
  – The deterministic model samples 2.2 times more points than the random model

• For $c_1 = 2$ ($f_{max} \leq 3.5$) and $\varepsilon = 2$ (probability $1 - 1/n^2$), the difference is by a factor of 5.1

• In summary:
  – Each model has its benefits (low overhead vs. performance)

• What about graph properties?
A Big Map

Motivation

Single-point
- Random Split
- Center Split

Multi-point
- Random $d$-sampling
- Deterministic $d$-sampling

P2P Simulations

Conclusion
P2P Simulations

• We next compare the performance of multi-point methods in P2P simulations
  – Our main metric of interest is the degree distribution

• Three models
  – Purely random $d$-sampling
  – Random walk
  – Biased walk

• De Bruijn DHT (based on ODRI, SIGCOMM 2003) with $n = 30,000$ nodes and degree $k = 8$
Degree Distribution - CDF

- Single-point, center-split scheme sets the basis for comparison

- 100 iterations
- Largest degree 81
- 5.7% of all nodes have degree 1
- 13% with degree 1 or 2
Degree Distribution - CDF

- Multi-point schemes perform much better

- Purely random
  - \( d_1 = 11 \)

- Deterministic
  - \( d_2 = 24 = 2.2d_1 \)

- 40% of the nodes have the ideal degree 8
Degree Distribution - PDF

- Overhead: random 55 messages per join and deterministic 7 per join, but performance is similar
A Big Map

Motivation

Single-point
随机分割 中心分割

Multi-point
随机分割 $d$-采样 定制分割 $d$-采样

P2P 模拟

结论
Conclusion

Legend:
- Single-point center split
- Deterministic
- Purely random

Performance vs. Overhead

Chord
CAN

Naor et al. SPAA 2003
D2B PODC 2003
ODRI SIGCOMM 2003
Adler et al. STOC 2003