Orthogonal Range Search using a Distributed Computing Model

Pouya Bisadi and Bradford G. Nickerson

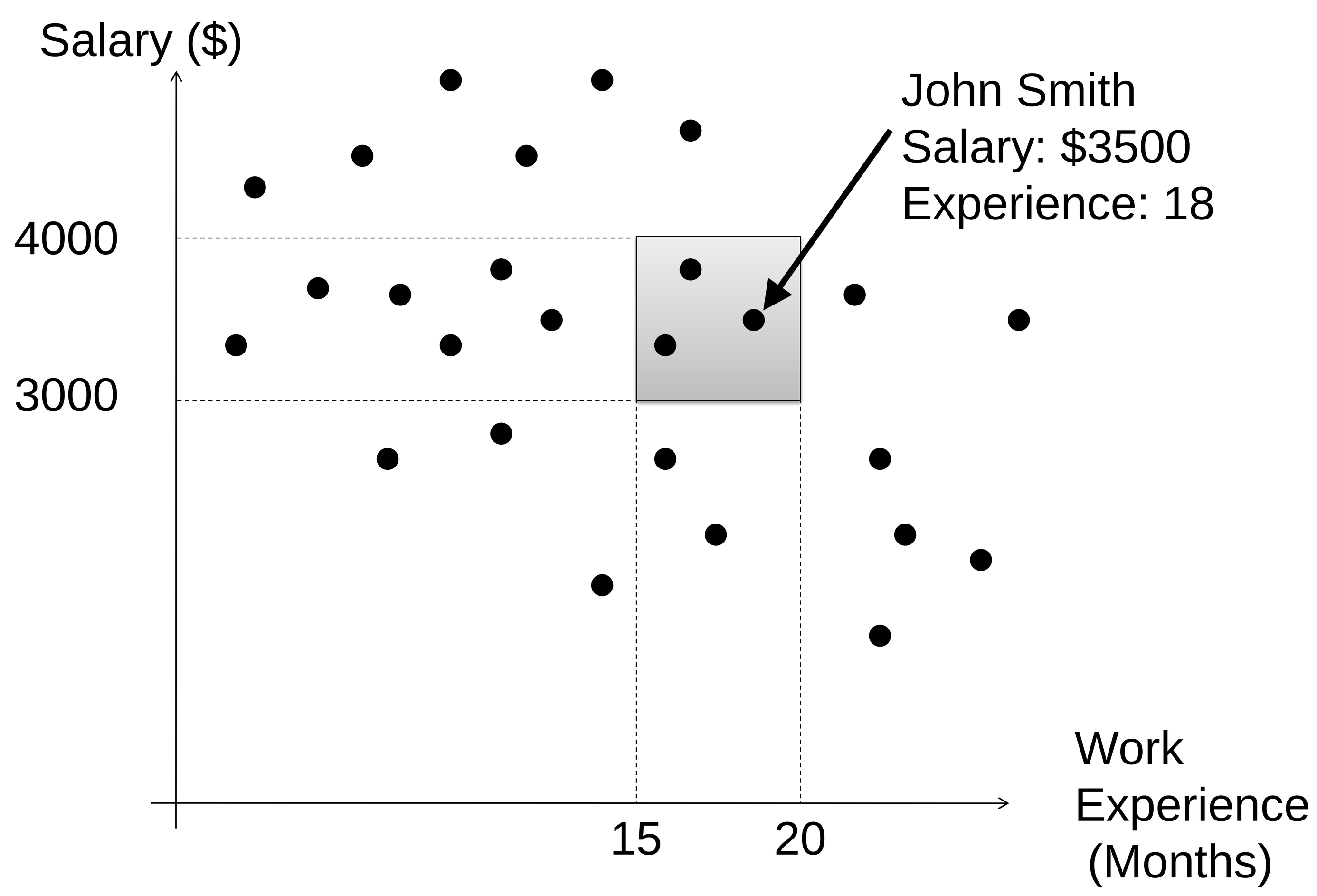
Faculty of Computer science, University of New Brunswick, Fredericton, New Brunswick, Canada

Motivation

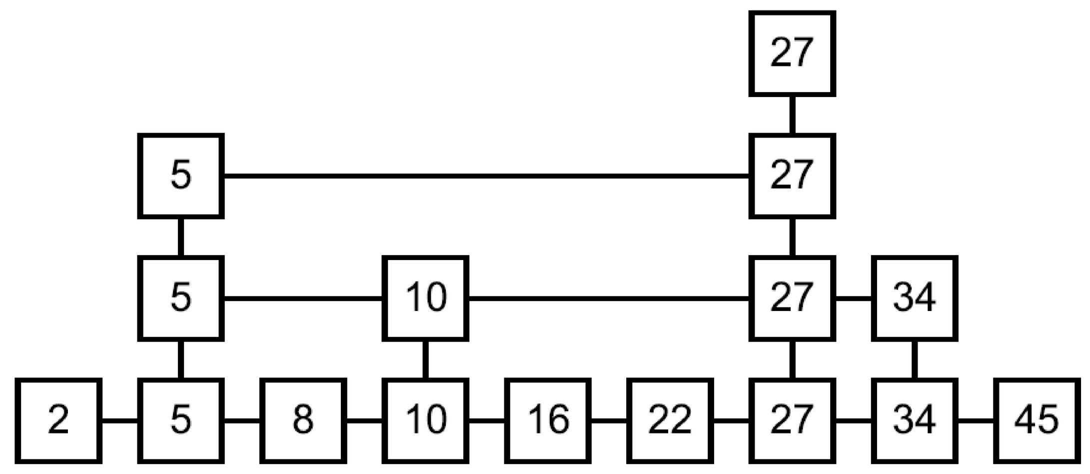
- Reliability
- Low Congestion per host
  - O(log n / n) for n= # of random queries
- Improved data access in a P2P network
- Geographical distribution of data
- Multiple party access control
- Automated data replication for back

Orthogonal Range Search

- Basis of database engines

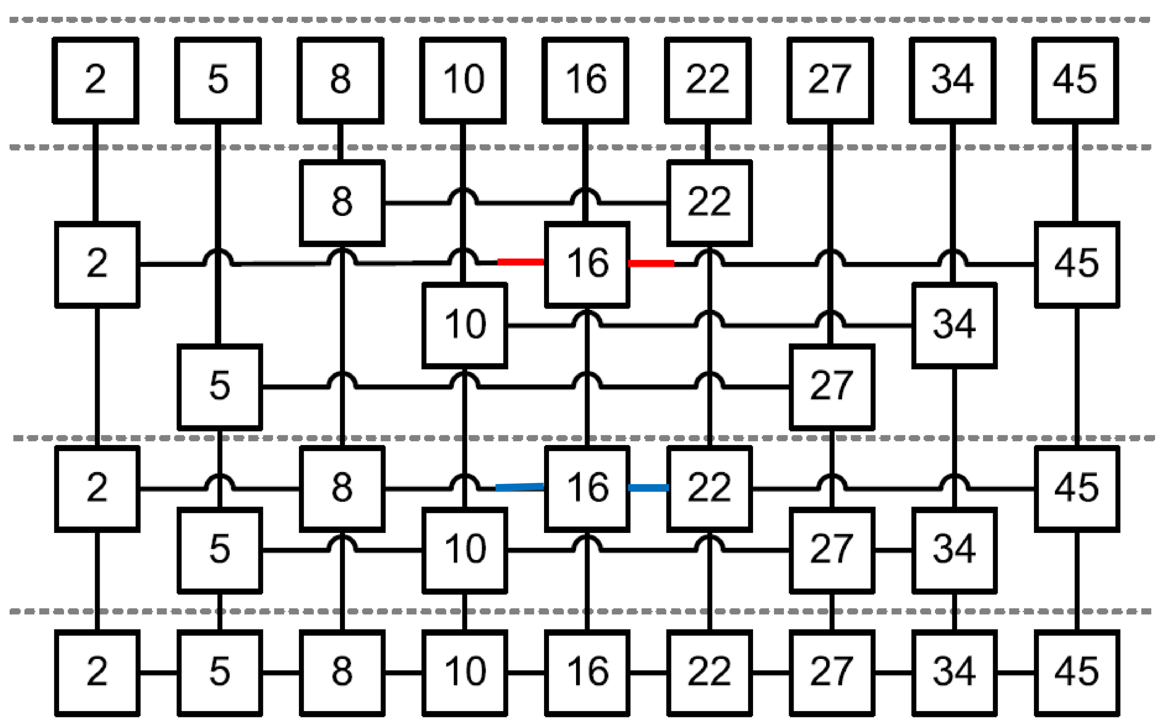


Data Structure on n nodes



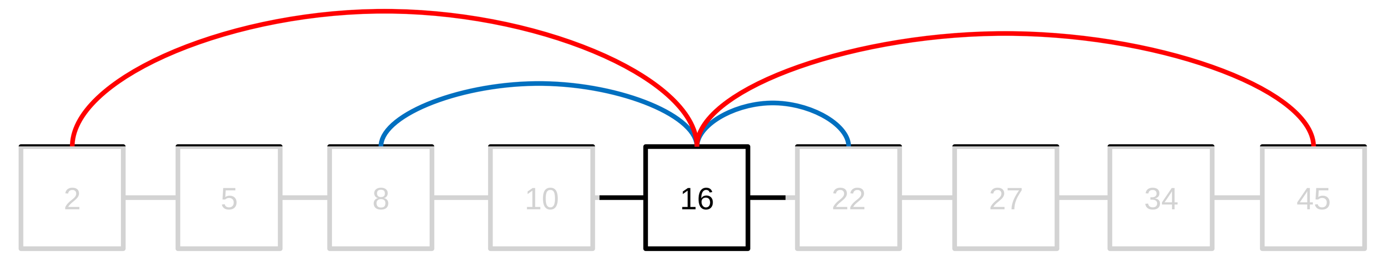
Skip List [1]

- Random data structure
- Searches can start from any element
- Query messages (W.H.P): $O(\log n + (k/B))$
  - k = # of point in range and
  - B = # of points in one message
- Space: $O(n)$

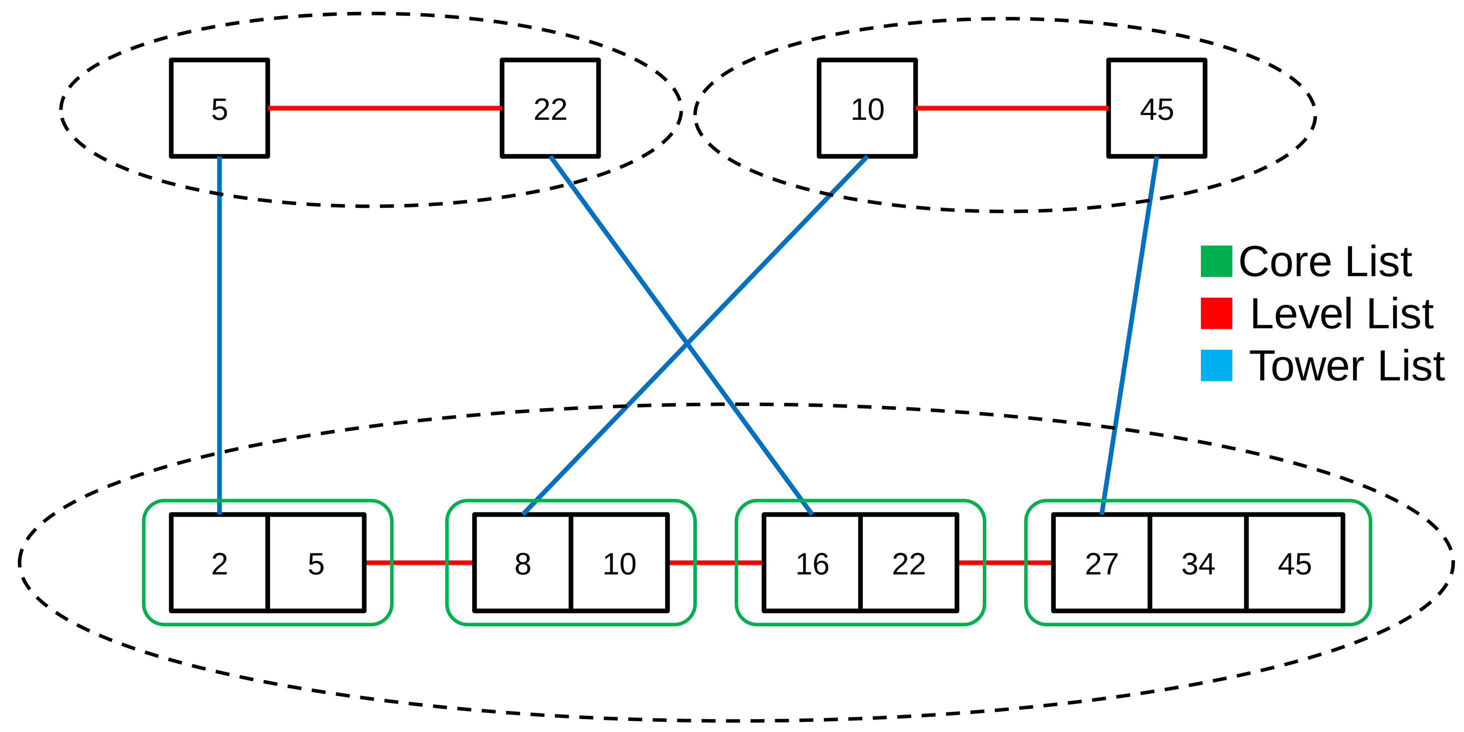


Skip Graph [1]

- Distributed extension of skip lists.
- Set of increasingly sparse lists



Multiple list membership of each node

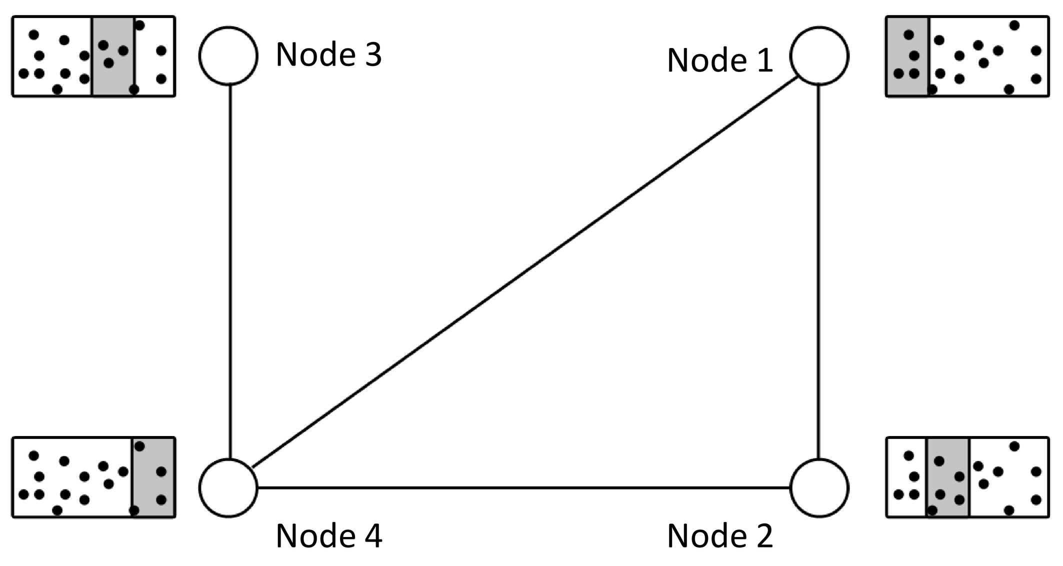


Non-redundant Rainbow Skip Graph

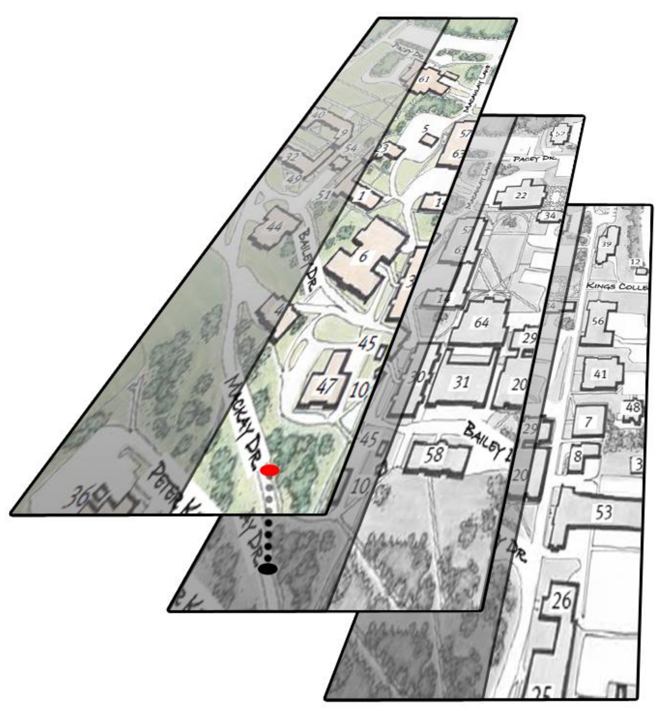
- A skip graph on θ(n/log n) Supernodes
- A Supernode consists of θ(log n) nodes
- Constant number of pointers
  - (vs. Log(n) for skip graph

In all 3 data structures, **Total order** binary relation (≤) should be definable on the set of keys

Data Distribution



Distribution by X coordinate



Overlapped data replication

Planned Experimental Validation

- 10^9 uniform random generated points ϵ $[0,1]^2$
- 100 nodes physically located on ACEnet
- 3000 size-limited axis-aligned random queries
- Queries will be asked from different nodes

[1] Michael T. Goodrich, Michael J. Nelson, and Jonathan Z. Sun, The rainbow skip graph: a fault-tolerant constant-degree distributed data structure, SODA '06: Proceedings of the seventeenth annual ACM-SIAM symposium on Discrete algorithm (New York, NY, USA), ACM, 2006, pp. 384{393.

Sponsored by: