Distributed Spatial Data Structures for Big Data

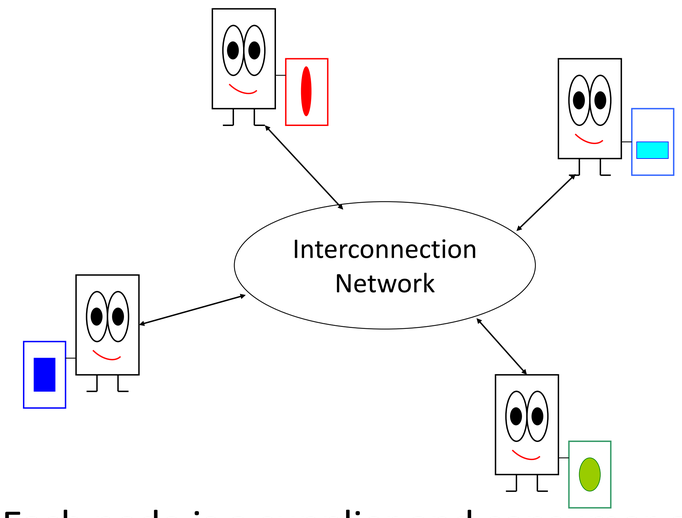
Zahra Miri Kharaji and Bradford G. Nickerson

University of New Brunswick, Faculty of Computer Science

Motivation

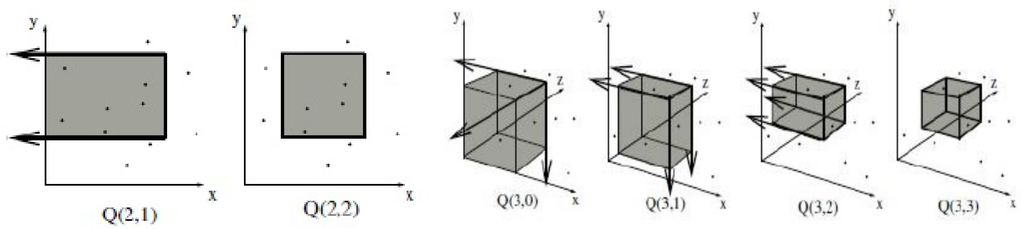
- Supporting orthogonal range search on point data
- Fault tolerance
- Scalability
- Low number of messages
- Dynamic addition and deletion of nodes and data

Peer-to-Peer (P2P) Systems



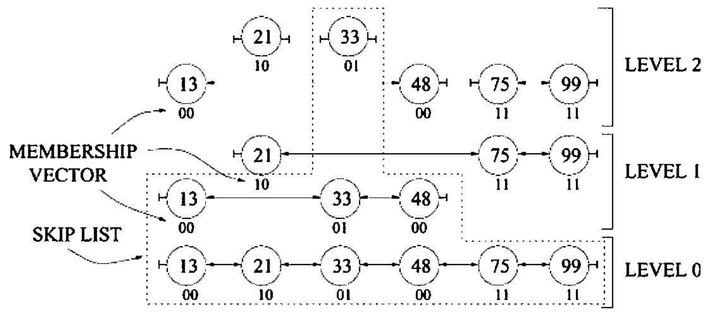
- Each node is a supplier and consumer of data
- Data identified by keys
- Communication is by send and receive message passing

2&3 Dimensional Rectangular Range Search



- Q(d, k): Orthogonal range in d-dimensional space with k dimensions having finite intervals [Afshani et al, 2009]

Skip Graph

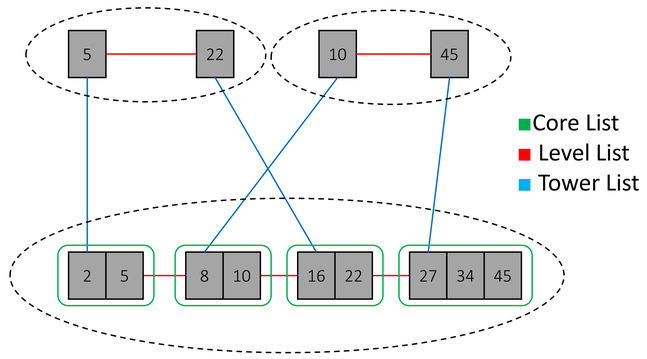


An example of skip graph [Goodrich et al, 2009]

	Expected time	Expected number of messages
Search	$O(\log n)$	$O(\log n)$
Insert	$O(\log n)$	$O(\log n)$
Delete	$O(\log n)$	$O(1)$

Skip graph performance

Rainbow Skip Graph



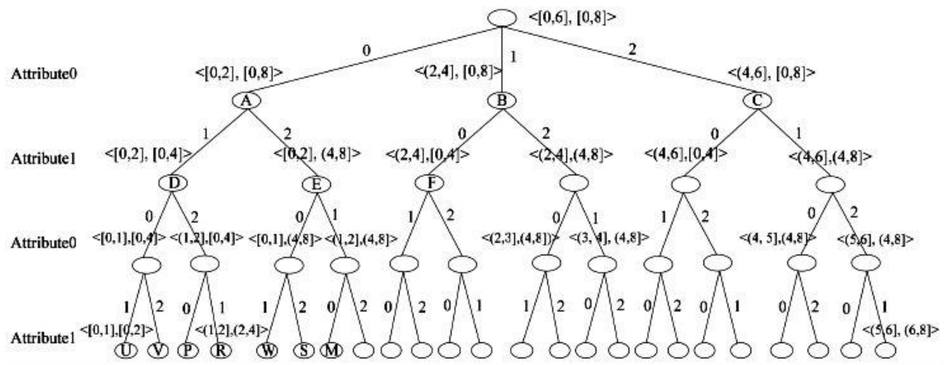
Non-redundant Rainbow Skip Graph [Bisadi et al, 2011]

- Fixed number of pointers per node
- The query cost is O(log n) w.h.p.

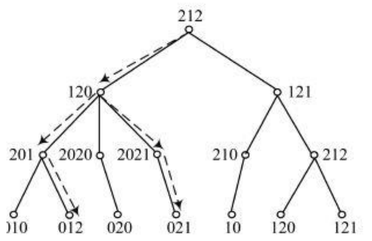
Erasure Resilient Code

Hydra Components

Distributed Hash Table P2P System



Partitioning the space to assign ObjectIDs to objects [Li et al, 2009]



Searching forward routing tree [Li et al, 2009]

- Maximum query delay is less than 2 logN hops in an N-peer P2P system.

Objectives

- Is there a linear space distributed spatial data structure that supports optimal worst case $O(\sqrt{n} + K)$ messages for Q(2,2) range search ?
- Can redundancy be added permitting any one node to be off-line and still answer any 2-d range query?

[1] Li, Dongsheng, et al. "Efficient range query processing in peer-to-peer systems." IEEE Tr. on Knowledge and Data Engineering, 2009.
[2] M. Goodrich, M. Nelson, and J. Z Sun. The rainbow skip graph: A fault-tolerant constant-degree p2p relay structure, 2009.
[3] Bisadi Pouya, and Bradford G. Nickerson. "Orthogonal Range Search using a Distributed Computing Model." *CCCG*. 2011.
[4] P. Afshani, L. Arge, and K. D. Larsen. Orthogonal range reporting in three and higher dimensions. IEEE, 2009.

UNB Faculty of COMPUTER SCIENCE
NSERC CRSNG