Orthogonal Range Search using a Distributed Computing Model

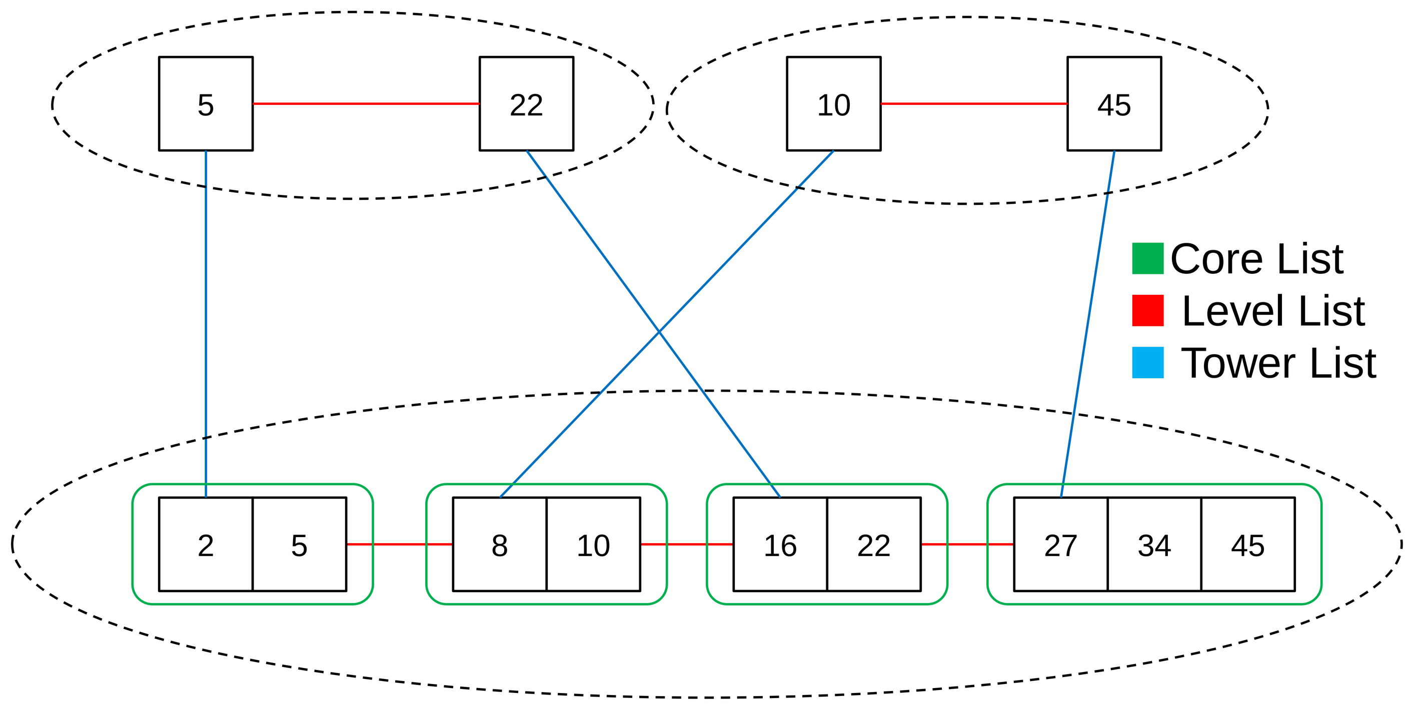
Pouya Bisadi and Bradford G. Nickerson

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 backup

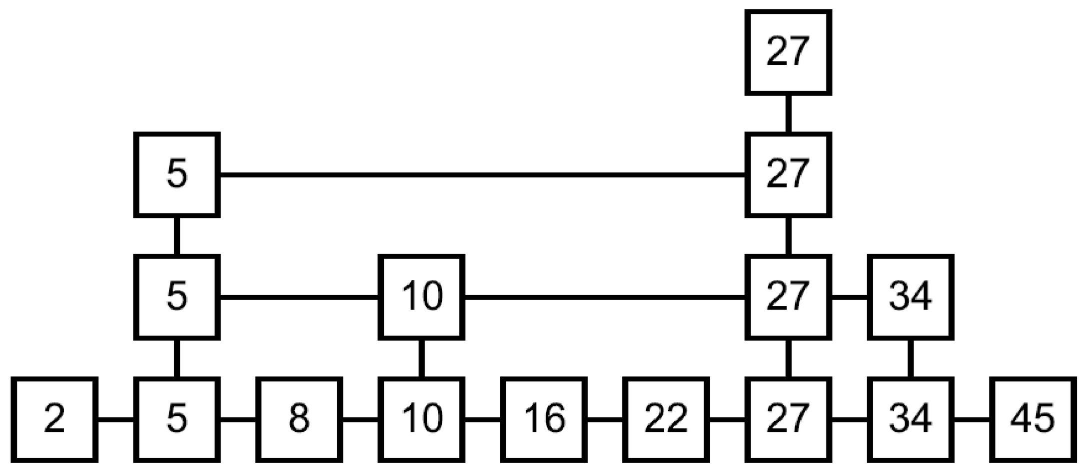
Non-redundant Rainbow Skip Graph[1]



- 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

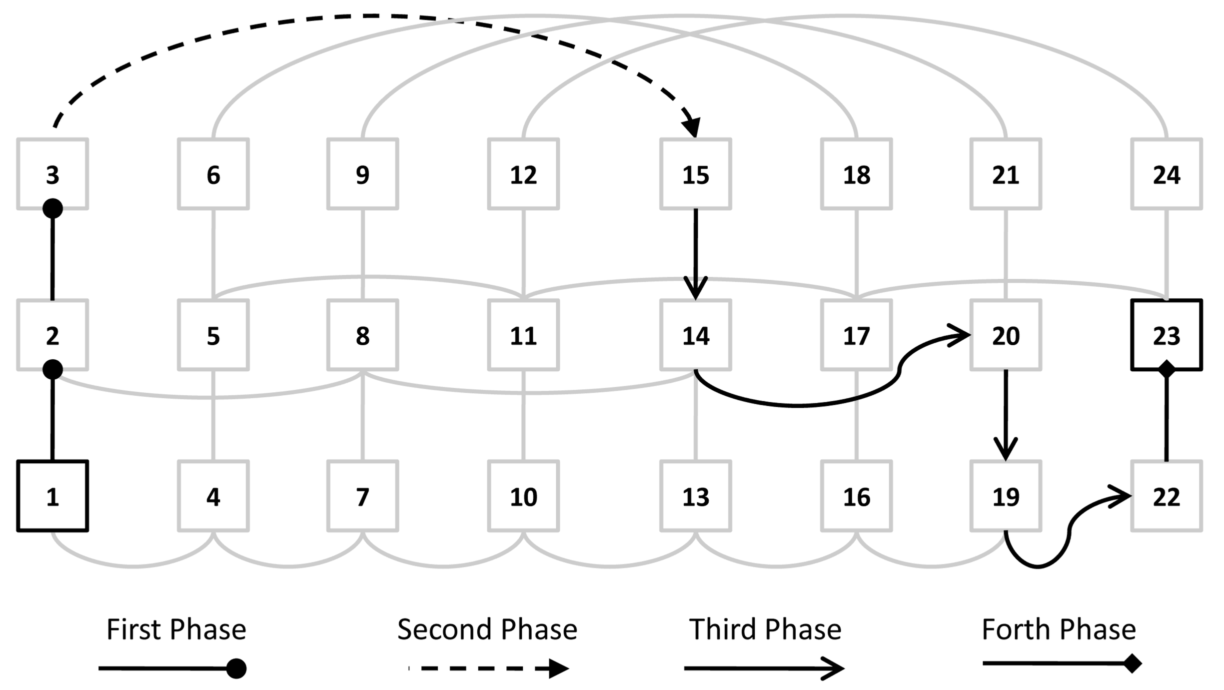
Total order binary relation (≤) should be definable on the set of keys

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)*

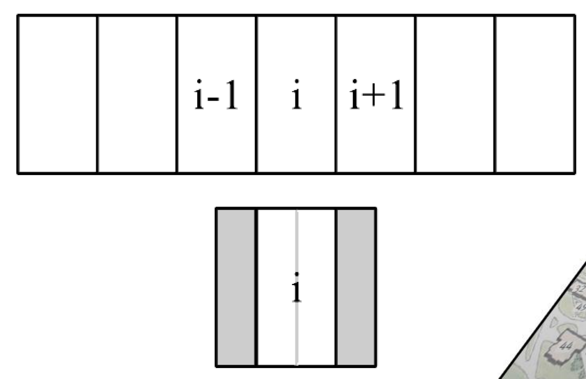


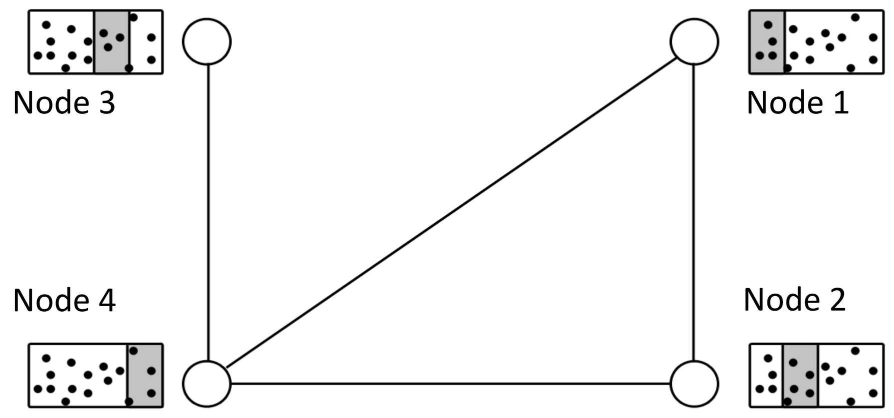
Search algorithm in Rainbow Skip Graph

(1) Sending the query to the topmost level (2) using the hops in the topmost level. (Largest possible hops) (3) Finding the target supernode by sending it to lower levels and using hops in each level. (4) Passing the query through the core list

Data Distribution and replication

Half of the data on previous and next nodes are stored as backup





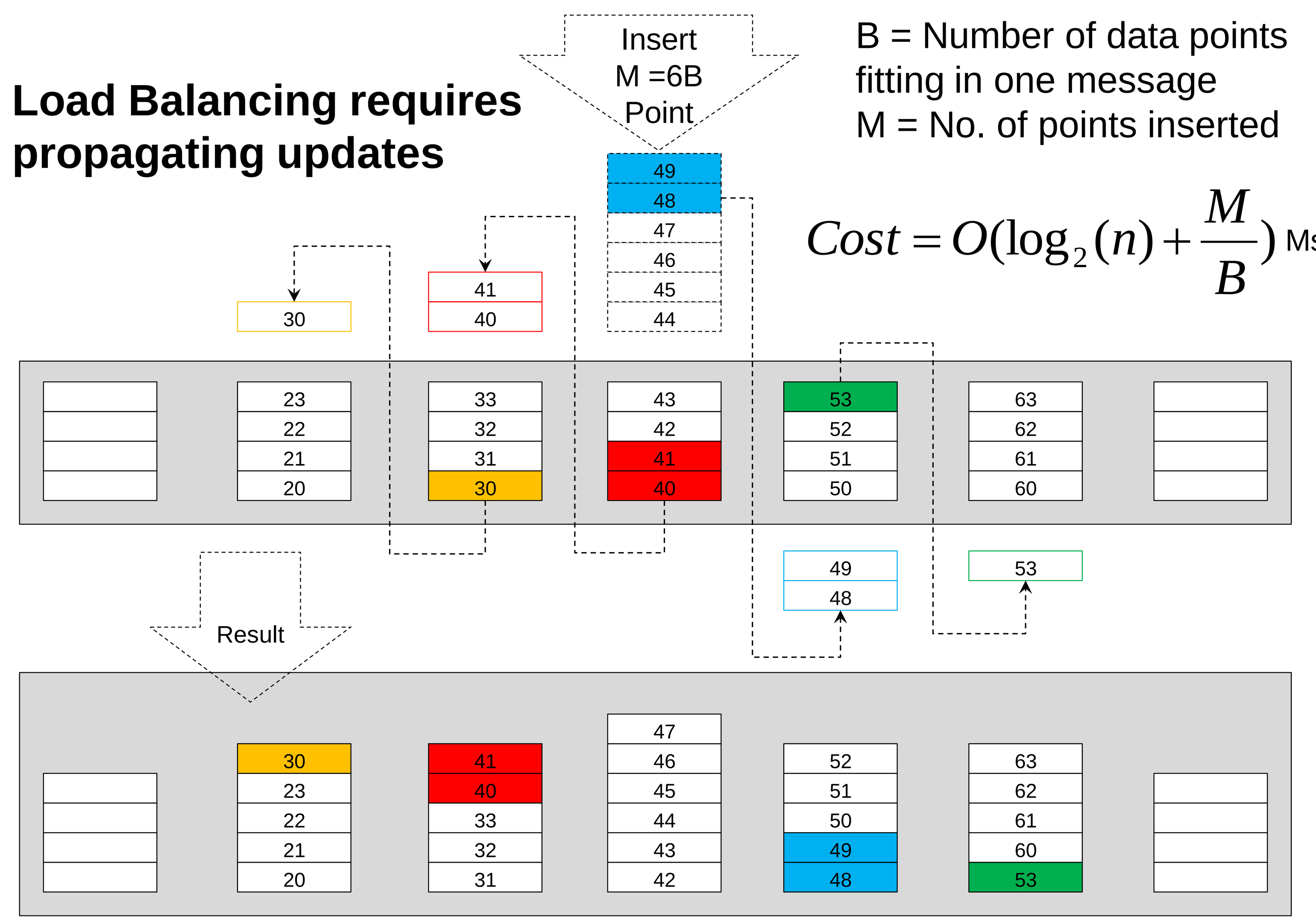
Distribution by X coordinate

Overlapped data replication

Load Balancing requires propagating updates

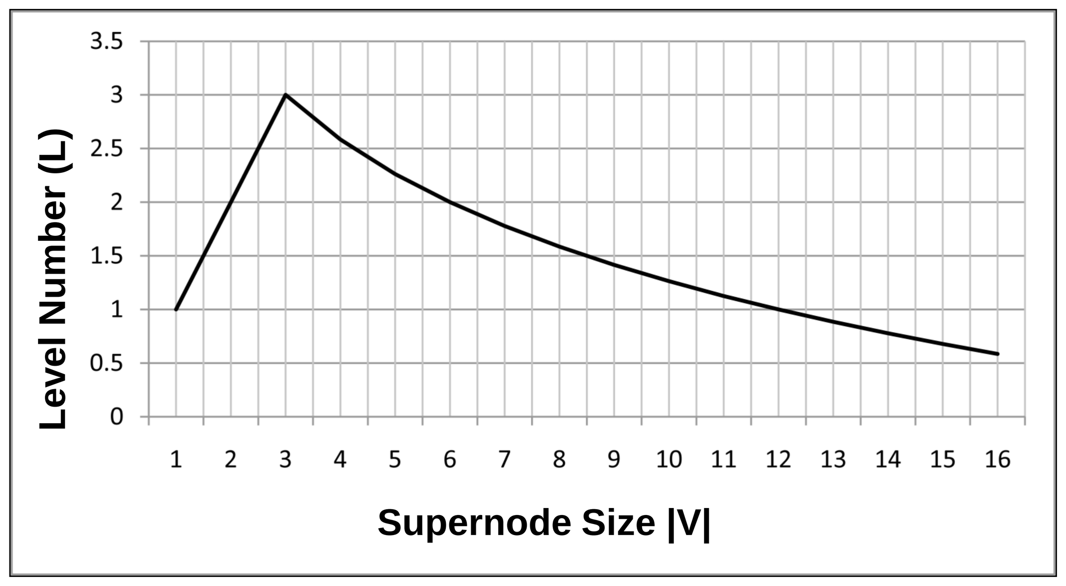
B = Number of data points fitting in one message
M = No. of points inserted

$$Cost = O(\log_2(n) + \frac{M}{B})$$ Msg.



Number of levels in Rainbow Skip Graph

The number of levels L is always the minimum of $|V|$ and $\log_2 \frac{n}{|V|}$

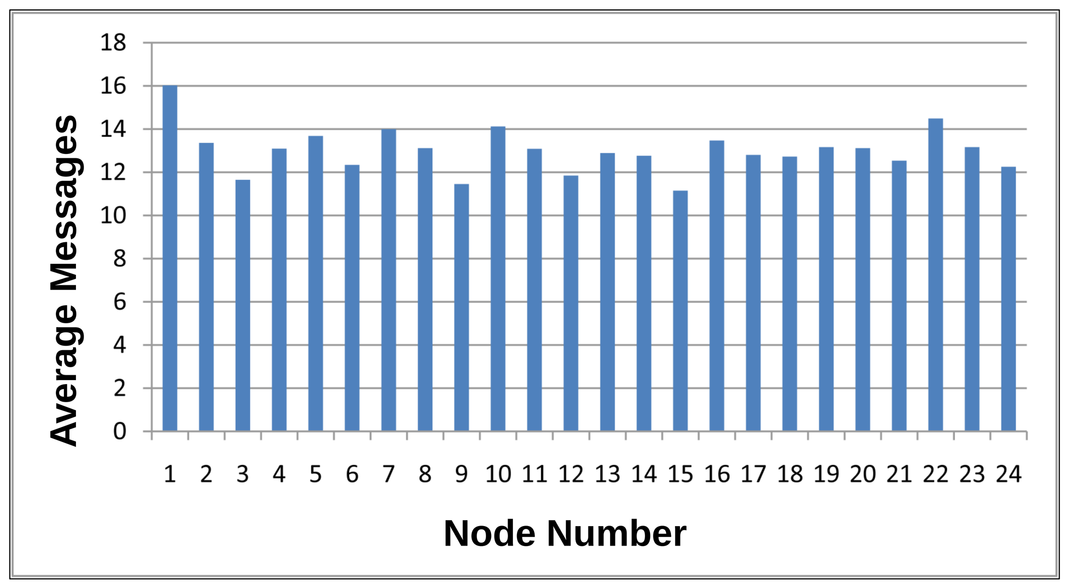


Theorem: The maximum number of levels L for a rainbow skip graph is

$$L = \frac{W(n \ln 2)}{\ln 2}$$

where W is the *lambertW* function $z = W(z)e^{W(z)}$

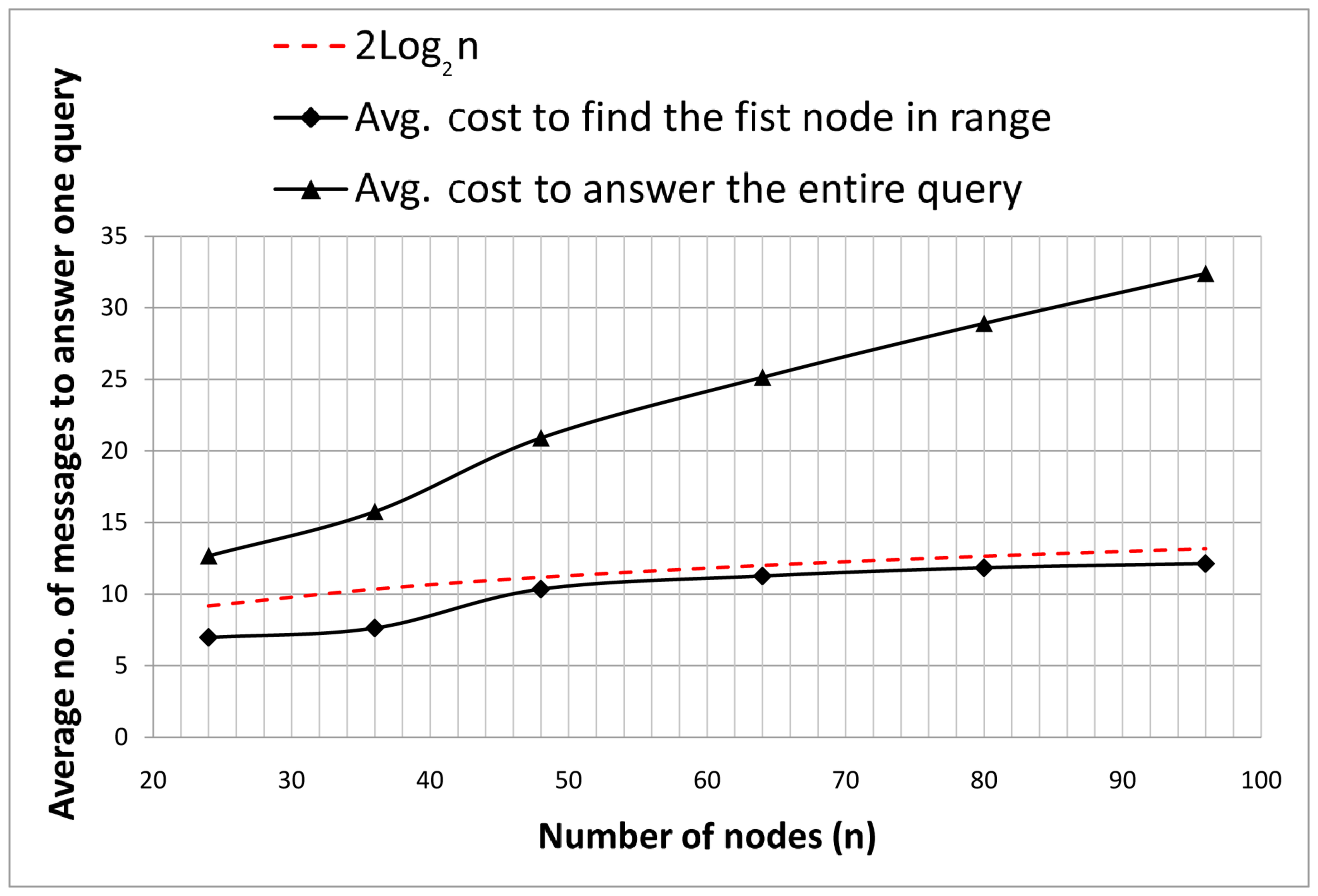
Average Message per Query



The average number of messages for answering 24,000 random point queries on 24 nodes of fundy.cs.unb.ca

Point Query: $O(\log_2 n)$ Messages

Range Query: $O(n + \frac{K}{B})$ Messages



[1] M. Goodrich, M. Nelson, and J. Sun. The rainbow skip graph: A fault-tolerant constant-degree p2p relay structure. pages 384{393, New York, NY, USA, 2009. ACM.

Sponsored by: UNB Faculty of COMPUTER SCIENCE NSERC CRSNG