Deepak Ajwani
Max-Planck-Institute for Informatics

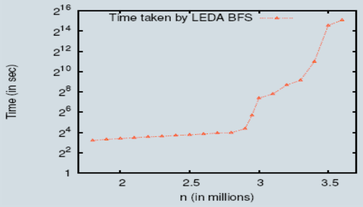
External Memory Breadth First Search

Problem: Given a massive sparse undirected graph, compute its BFS level decomposition

Problem

Internal memory solution: Simple linear time algorithm for computing BFS

RAM model BFS algorithm perform very poor on real architecture

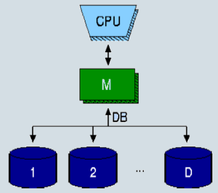


Cause

- RAM model does not capture the cost of I/Os, which becomes the bottleneck on real architecture
- Remembering visited nodes may cause $\Omega(m)$ I/Os
- Unstructured accesses to adjacency lists may cause $\Omega(n)$ I/Os

Solution

Need to design and analyze the BFS algorithm in the I/O model



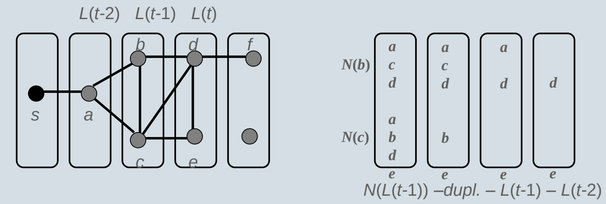
- External memory: D disks
- Data transfer in blocks of size B (~1M)
- Up to $D \cdot B$ elements transfer in one I/O

Goal: Minimize total number of I/O steps

BFS algorithm by Munagala and Ranade

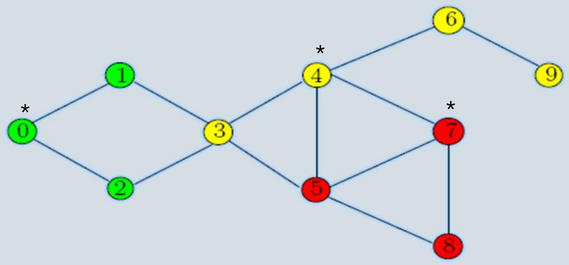
Observation

In an undirected graph, the edges from BFS level t-1 go only to levels t-2, t-1 and t



Mehlhorn and Meyer BFS algorithm - Randomized clustering

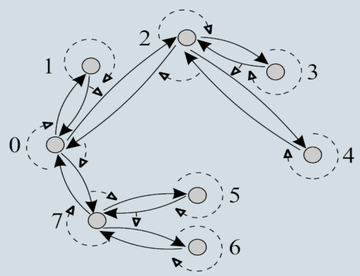
Randomly select master nodes (marked *) with uniform probability and expand around them “in parallel”

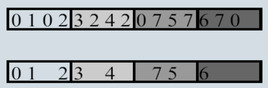


Mehlhorn and Meyer BFS algorithm - Deterministic clustering

Key Idea

Cluster the graph and re-arrange the graph layout on the disk to make accesses to adjacency lists structured





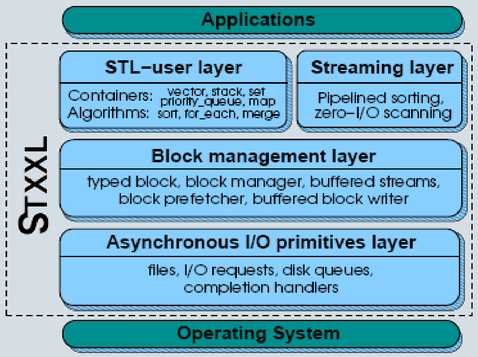
Steps involved in deterministic clustering:

- Compute a minimum spanning tree of the input graph
- Compute an Euler tour around the spanning tree
- Sort nodes in the order in which they appear in the Euler tour
- Chop the ordered list into blocks of size $\sqrt{B}$ and remove duplicates

Implementation

Key features

- Open source software freely available
- Uses pipelining to reduce constant factors in I/Os
- Can exploit disk parallelism
- Maximizes overlap between I/O and computation
- Based on external memory library STXXL which has the same interface as the standard template library, STL



Graph class	n	m	MR_BFS	MM_BFS_R	MM_BFS_D
Random	2^{28}	2^{30}	**1.4**	7x	6x
Web Graph	$\sim 1.4 \cdot 2^{28}$	$\sim 1.2 \cdot 10^9$	**2.6**	3.5x	2x
Grid (2^{14} x 2^{14})	2^{28}	2^{29}	2.5x	1.25x	**21**
Grid (22^1 x 2^7)	2^{28}	$\sim 2^{29}$	> 100x	> 10x	**4.0**
Grid (2^{27} x 2)	2^{28}	$\sim 2^{28} + 2^{27}$	> 500x	> 25x	**3.8**
Simple Line	2^{28}	2^{28} - 1	**0.4**	7x	7x
Random Line	2^{28}	$2^{28} - 1$	> 1300x	> 75x	**3.6**
Max			~ ½ year	~ 1 week	< 1 day