Graph Prefetching Using Data Structure Knowledge

Sam Ainsworth and Timothy M. Jones

Computer Laboratory
Graph500 Search Performance

Stall Rate (%)

Miss Rate (%)

Edge factor 15
Edge factor 10
Edge factor 5

Scale
Current Prefetching Techniques

- Stride

- Software
### Work List

- 5
- 4
- 1
- 2
- 3
- 7
- ...

### Vertex List

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>#</td>
<td>#</td>
<td>3</td>
<td>5</td>
<td>#</td>
<td>#</td>
</tr>
</tbody>
</table>

### Edge List

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>#</td>
<td>#</td>
<td>#</td>
<td>0</td>
<td>#</td>
<td>#</td>
</tr>
</tbody>
</table>

### Visited

- False
- True
- True
- True
- True
- False
- True
Problems

• Need address bounds of data structures

• Need to schedule prefetches

• Need to react to variable latency loads
Problems

• Need address bounds of data structures
  • Configure them in software!

• Need to schedule prefetches

• Need to react to variable latency loads
Problems

- Need address bounds of data structures
  - Configure them in software!
- Need to schedule prefetches
  - Use observation hardware – EWMAs.
- Need to react to variable latency loads
Problems

- Need address bounds of data structures
  - Configure them in software!

- Need to schedule prefetches
  - Use observation hardware – EWMAs.

- Need to react to variable latency loads
  - React to arrival of prefetches, not loads!
Graph Prefetcher

Main Memory
- Work List
- Vertex List
- Edge List
- Visited List

L2 Cache

Dcache
- To / From L2 Cache
- Snoops
- Prefetch Reqs
- Prefetched Data
- EWMA Calculator
- Address Generator
- Request Queue
- Prefetcher
- Config

Core

EWMA Calculator
Address Generator
Request Queue
Prefetcher
Config
Graph Prefetcher: Microarchitecture

- Snoops & Prefetched Data From L1 Cache
- To DTLB & L1 Cache
- Prefetch Request Queue
- Address Filter
- Address Bounds Registers:
  - Work List Start
  - Work List End
  - Vertex List Start
  - Vertex List End
  - Edge List Start
  - Edge List End
  - Visited List Start
  - Visited List End
- EWMA Unit:
  - Work List Time EWMA
  - Data Time EWMA
  - Ratio Register

- Prefetch Address Generator
Results – Graph500

![Graph500 Results](image)
Results – Boost Graph Library

![Speedup Comparison Graph]

The graph compares the speedup for different algorithms (BFS, BC, ST) across different datasets (amazon, web, road). The x-axis represents the algorithms, and the y-axis represents the speedup. The datasets are color-coded: red for amazon, blue for web, and green for road.
Results – Sequential Iteration

![Graph showing speedup for PR and SC with categories 'amazon', 'web', 'road', and respective speedup values.]
Generalized Prefetching - Databases

Hash Table

Bucket (43, 2, ptr)

Bucket

12
43
ptr

13
87
null

Key

12
62
43

Lookahead by striding in the key list

Hash(43) = 3
Graph Prefetching Using Data Structure Knowledge

Sam Ainsworth and Timothy M. Jones
sam.ainsworth@cl.cam.ac.uk timothy.jones@cl.cam.ac.uk

For more information, see our paper from ICS 2016!