Fast and Efficient Update Handling for Graph H²TAP

Muhammad Attahir Jibril ▪ Hani Al-Sayeh ▪ Alexander Baumstark ▪ Kai-Uwe Sattler

(EDBT 2023)

Sixteenth LDBC TUC Meeting
24.06.2023
Introduction

- **Graph workloads:**
  - Transactions e.g. Inserts, deletes, lookups
  - Analytics e.g. BFS, SSSP, PR

- **Graph HTAP:** concurrent mix of transactional and analytical workloads
Design Space

- **Single Instance CPU**
  - Pros: easier to manage
  - Cons: no workload-specific optimizations, no accelerated analytics

- **Single Instance GPU**
  - More complex and problematic than single instance CPU

- **Separate Instances**
  - Pros: workload-specific optimizations, leverage GPU
  - Cons: synchronizing the instances
Heterogeneous HTAP (H^2TAP)

Leveraging each processor for the most suitable workload
Challenges

1. Propagating updates

2. Meeting freshness requirements

3. Guaranteeing consistency

Update Handling
Update Handling: Overview

Phase 1: Update storage (in CPU)

Phase 2: Update propagation from CPU to GPU
Evaluation

Setup

- **Machines:**
  - GPU server
  - PMem server

- **Data:**
  - LDBC SNB dataset
  - Graph 500 dataset

- **Workload:**
  - LDBC SNB-like insert and delete queries
  - LDBC Graphalytics
Evaluation

Transaction Time

(a) INSERT NODE  (b) DELETE NODE  (c) INSERT RELATIONSHIP  (d) DELETE RELATIONSHIP  (e) MIXED
Evaluation

Update Propagation Time

- **DELTA\(^I\)\(_{LoDeg}\)**
- **DELTA\(^I\)\(_{HiDeg}\)**
- **DELTA\(^F\)E\(_{LoDeg}\)**
- **DELTA\(^F\)E\(_{HiDeg}\)**

(a) **INSERT NODE**
(b) **DELETE NODE**
(c) **INSERT RELATIONSHIP**
(d) **DELETE RELATIONSHIP**
(e) **MIXED**
Evaluation

Update Propagation Time Components

(a) **Total Time**

(b) **Scan and Merge Times**

(c) **Delta Merge**
## Evaluation

### H²TAP and Single Instance HTAP (CPU)

<table>
<thead>
<tr>
<th></th>
<th>BFS</th>
<th>PR</th>
<th>SSSP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sortledton</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analytics on CPU</td>
<td>1.48</td>
<td>21.34</td>
<td>57.30</td>
</tr>
<tr>
<td><strong>DELTA\textsuperscript{FE}</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Update Propagation</td>
<td>5.38</td>
<td>5.38</td>
<td>5.38</td>
</tr>
<tr>
<td>Analytics on GPU</td>
<td>0.07</td>
<td>0.30</td>
<td>0.13</td>
</tr>
<tr>
<td>Sum</td>
<td>5.45</td>
<td>5.68</td>
<td>5.51</td>
</tr>
</tbody>
</table>
Evaluation

Direct Relational Approach

(a) Transactional Update Time
(b) Delta Store Scan
Conclusion

- Fast and efficient approach for handling updates in a graph $H^2TAP$ setting
  - minimal overhead
  - freshness
  - consistency

- Outperforms existing approaches
Thank You