Literature Review

Cheng Liu

2016-11-03

1 GraphMat

On top of the vertex programming model, this work [3] focuses on transforming graph processing to sparse matrix computing. With the well-studied matrix computing on HPC, GraphMat achieves very good performance and scales on multi-core systems.

With this framework, it is possible to develop graph processing accelerator over spare matrix accelerator which is also well-studied.

Although a few classical graph algorithms including BFS and SSSP can be transformed to spare matrix computing, there is no guarantee for general graph problems.

2 In-memory Graph Database for Web-scale Data

In this work [1], Graph Database Engine for Multithreaded System (GEMS) is developed for implementing Resource Description Framework (RDF) database on distributed memory high-performance cluster. In this framework, SPARQL queries will be compiled by SPARQL-to-C++ compiler. With a SGLib, the queries will be eventually converted to graph pattern matching operations. This seems to be a good application of Graph processing on RDF database, but intensive background knowledge needs to be investigated to get better understanding of this work.

3 Distributed Graph Engine for Web Scale RDF Data

In this work, Trinity.RDF [4] is developed for a distributed system. Previous work usually relies on join operations for processing SPARQL queries and a large amount of useless data will be generated. To avoid this problem, this work models and stores the RDF data with native graph data. The SPARQL query is represented as a query graph and the SPARQL query processing problem is transformed to be subgraph matching. Trinity.RDF is part of the Trinity project and more details can be found [2].

References

[1] V. G. Castellana, A. Morari, J. Weaver, A. Tumeo, D. Haglin, O. Villa, and J. Feo. In-memory graph databases for web-scale data. *Computer*, 48(3):24–35, Mar 2015.

[2] Microsoft. Trinity. `https://www.microsoft.com/en-us/research/project/trinity/`, 2013. [Online; accessed 10-November-2016].

[3] Narayanan Sundaram, Nadathur Satish, Md Mostofa Ali Patwary, Subramanya R Dulloor, Michael J Anderson, Satya Gautam Vadlamudi, Dipankar Das, and Pradeep Dubey. Graphmat: High performance graph analytics made productive. *Proceedings of the VLDB Endowment*, 8(11):1214–1225, 2015.

[4] Kai Zeng, Jiacheng Yang, Haixun Wang, Bin Shao, and Zhongyuan Wang. A distributed graph engine for web scale rdf data. In *Proceedings of the VLDB Endowment*, volume 6, pages 265–276. VLDB Endowment, 2013.