GRAPH/Z: A Key-Value Store Based Scalable Graph Processing System

Tonglin Li1, Chaoqi Ma1, Jiabao Li1, Xiaobing Zhou2, Ke Wang3, Dongfang Zhao4, Iman Sadooghi1, Ioan Raicu1,5

1Illinois Institute of Technology, 2Hortonworks, 3Intel, 4Pacific Northwest National Laboratory, 5Argonne National Laboratory

Abstract

The emerging applications in big data and social networks issue rapidly increasing demands on graph processing. Graph query operations that involve a large number of vertices and edges can be tremendously slow on traditional databases. The state of the art graph processing systems and databases usually adopt master/slave architecture that potentially impairs their scalability. This work describes the design and implementation of a new graph processing system based on Bulk Synchronous Parallel model. Our system is built on top of ZHT, a scalable distributed key-value store, which benefits the graph processing in terms of scalability, performance and persistency. The experiment results imply excellent scalability.

Motivation

- Emerging uses of large graph data sets
- SQL databases don't handle it well
- Large data set can not fit in memory
- Current systems don't allow data change

Contributions

- A BSP model graph processing system on top of ZHT.
- Utilizing data-locality and minimize data movement between nodes.
- Benchmarks up to 16-nodes scales.

Design and Architecture

Design

- A Pregel[2] like graph distributed processing system
- Master: coordinates synchronization
- Use ZHT[3-5] as back end
- Store both intermediate and final result in ZHT

Features

- Handle large data sets
- Don't need to fit all data in main memory
- Dynamic data modification during running
- Load balance
- Fault tolerance
- Support checkpointing

BSP model

BSP[1] model

- Think like a vertex
- Vertices compute
- Edges communicate

Graph/Z system architecture

Experiment setup

- Test bed
  - 2-16 m3.2large Amazon EC2 spot instances
  - 2.5 GHz Xeon, 30GB RAM
- Data set
  - Web-Google from SNAP (Stanford Network Analysis Project)
- 1M vertices and 5M edges

Preliminary results

Reference


Acknowledgement

This work is supported in part by the National Science Foundation grant NSF-1054974. This work used Amazon cloud resources through Amazon research grant.