A Revised Implementation of the GRAPH/Z Graph Processing System

Ballmer, Alexander  
alexandersballmer@gmail.com

Moudgalya, Shreyas  
smoudgal@hawk.iit.edu

The emerging applications in big data science and social increasing demands on large-scale graph-structured processing has led to the development of several graph-parallel abstractions including GRAPH/Z. However, in the case of using ZHT, a scalable distributed key-value store as building block, the distribution is forced to use hash-based (random) partitioning which has potentially impaired their locality. The goal of this proposal would be to develop a graph partitioning scheme that can be used in the loading process of GRAPH/Z, a scalable graph processing system, to overcome several problems of its performance.

2

# Background

Graph partition problems are very common to run on systems like GRAPH/Z. Partitioning on graphs falls under the category of NP-hard problems. Solutions to these problems are generally derived using heuristics and approximation algorithms. However, uniform graph partitioning or a balanced graph partition problem can be shown to be NP-complete to approximate within any finite factor. Even for special graph classes such as trees and grids, no reasonable approximation algorithms exist, unless P=NP.  
There have been a lot of work to handle big datasets of both commercial and scientific applications, including work flow systems, data streaming management systems and graph databases. These systems typically save the data in distributed file systems(such as Hadoop HDFS and FusionFS), SQL databases(such as Oracle and DB2), and NoSQL database(such as Cassandra and ZHT).  
Graph related query is tremendously slow on the traditional relational database, which makes it even more challenging to fully reveal and utilize the scientific and commercial value from the continuously increasing graph data sets. An ideal solution for this problem is to replace the traditional data infrastructure with a graph-centric model, including storage and computing, thus to better serve graph-based applications in terms of performance and programmability.  
However, since natural graphs are difficult to partition, the original GRAPH/Z system used ZHT’s hash function to physically distribute all the vertices. As a result, vertices and edges, which are the basic elements of graphs, are spread around different nodes. Due to the random partitioning scheme, when each vertex needs to get access to its edge list, it will statistically communicate to another node and make the communication cost become tremendously higher. Therefore, using hash function to partition the graph data can be considered as a worst case, which is determined by the characteristic of graph structure and the hash function. The original use case of hashing partition is dispersing the data to avoid hot spot, but it doesn’t help in the case of graph processing system.

# Problem

We will also design and implementation a partitioning scheme that works well on the large graph datasets on Graph/Z. The design and implementation of the algorithm must include a fast, balanced and lightweight graph partition scheme based on graph traversal.

# Related Work

Various graph partitioning algorithms exist. Since graph partitioning is a NP-hard problem, practical solutions are based on heuristics. There are two broad categories of methods used. They are local and global. Well known local methods are the Kernighan–Lin algorithm, and Fiduccia-Mattheyses algorithms, which were the first effective 2-way cuts by local search strategies. Their major drawback is the arbitrary initial partitioning of the vertex set, which can affect the final solution quality. Other Global approaches rely on properties of the entire graph and do not rely on an arbitrary initial partition. The most common example is spectral partitioning, where a partition is derived from the spectrum of the adjacency matrix.  
A multilevel graph partitioning algorithm works by applying one or more stages. Each stage reduces the size of the graph by collapsing vertices and edges, partitions the smaller graph, then maps back and refines this partition of the original graph. A wide variety of partitioning and refinement methods can be applied within the overall multi-level scheme. In many cases, this approach can give both fast execution times and very high quality results. One widely used example of such an approach is METIS, a graph partitioner, and hMETIS, the corresponding partitioner for hypergraphs.

# Proposed Solution

Our main work to solve our problems of scalability is to backtrack and try to achieve good scaling on a single node. We believe that the problem with GRAPH/Z lies in bad data locality with ZHT. Our main goal is to achieve some level of competitiveness against GraphLab on a single node, eventually expanding to multi-node scaling through ZHT or another distributed datastore. We will measure performance from weak scaling with larger datasets, and profiling tools such as valgrind/callgrind. This main method will also apply to the implementation of a graph partitioning algorithm.

# Evaluation

We will use a graph partitioning algorithm defined as:  
Given a graph and a a set of partitions, , an output partition such that

1. are disjoint
2. are roughly balanced
3. let   
   Minimize

Results will come from from load balancing tests with other partitioning algorithms such as the Dynamic Cut-Cluster Algorithm, EvoPartition Algorithm, and the Kernighan–Lin Algorithm will be used for comparison. We would also profile the running time of the algorithms based on increase in graph vertices and increase in graph edges.

# Deliverables

We will test the data on an instance of AWS EC2 using the dataset from Stanford Network Analysis Project.

# Conclusion

[placeholder]