Load balancing techniques for CHAOS

May 15, 2016

Author:
Vlad Ioan Haprian

Supervisors:
Laurent Bindschaedler
Prof. Willy Zwaenepoel

1 Abstract

Chaos scales graph processing from secondary storage to multiple machines in a cluster. The preprocessing time is minimized by dividing the graph in partitions that contain the same number of vertices. This implies that the edge sets of different partitions have very different sizes thus work stealing is needed in order to reduce the imbalance between machines. Even if chaos achieves perfect balance, the work stealing comes with an overhead as the stealer need to copy additional informations.

Our project aims to exploit new partitioning modes in order to balance the computation and I/O between machines by trading the work stealing option against one more pass over the graph in the preprocessing phase.