

Presentation Abstract

Experimental Comparison of Vertex Scheduling Policies for Distributed Asynchronous Graph Processing Systems

HAYATO WATAHIKI^{1,a)} SHIGEYUKI SATO¹ KENJIRO TAURA¹

Presented: July 30, 2020

Vertex-centric graph processing is known to be useful for distributed processing of large-scale graphs and its asynchronous variants enable us to implement various vertex scheduling policies that have a big impact on experimental performance. Although many vertex scheduling policies were used in frameworks proposed over the past decade, an apple-to-apple comparison of their characteristics has not been done enough, and therefore it is still unclear how to select and design schedulers for specific applications. To deepen the knowledge on vertex scheduling, we implemented representative vertex schedulers, which were used in GraphLab and Maiter, into the same system and measured their performance impacts on five common iterative graph algorithms. Our experiments demonstrate that the number of messages sent among parallel workers is the dominant factor of the performance improvement by schedulers, and show what the most suitable scheduling policy is for each type of application. Furthermore, we investigate what aspect of each scheduling policy affects the number of messages and provide some suggestions for designing a new scheduling policy that suppresses the communication among workers, and also present the limitations of the performance improvement by vertex scheduling. Our investigation should help both users and developers of asynchronous vertex-centric graph processing systems to make better use of the potential of vertex scheduling.

This is the abstract of an unrefereed presentation, and it should not preclude subsequent publication.

¹ Graduate School of Information Science and Technology, The University of Tokyo, Bunkyo, Tokyo 113-8656, Japan

^{a)} watahiki@eidos.ic.i.u-tokyo.ac.jp