

04

Fast graph analysis by efficient CPU utilization

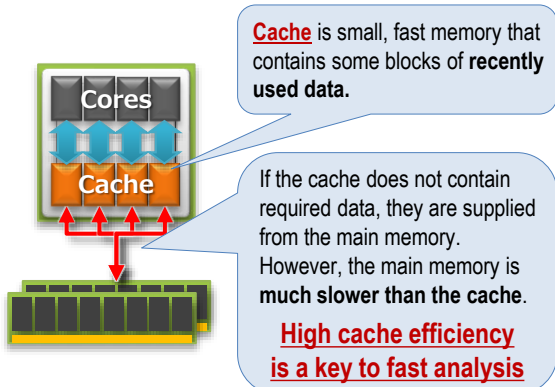
~Scalable parallel graph processing by reordering~

Abstract

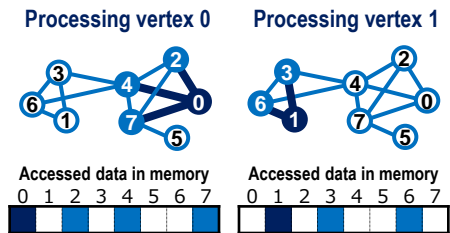
We are working towards improving the performance of big-data analysis. Particularly, our focus is on the analysis of *graphs*, which represent network structures such as hyperlinks between web pages and personal relationships on social networking services. Analysis of these graphs offers a wide range of knowledge. For example, we can find valuable Web pages, communities of closely related people, and items recommendable for customers. Our novel method lets graph analysis algorithms efficiently utilize multicore CPUs and improves their performance by up to 17 times. Fast graph analysis enables accurate knowledge discovery from large-scale data, a quick response to queries for recommendation or searches for a person, and so on.

Our method optimizes memory access patterns of various graph analysis algorithms (e.g., PageRank and clustering) and improves their performance by up to 17 times.

Structure of CPU and memory



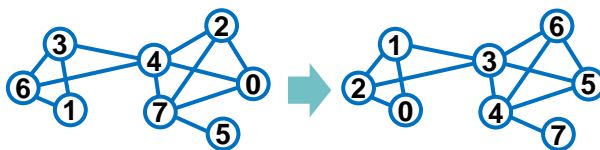
Problem: Irregular memory access



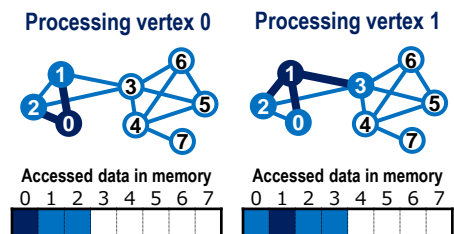
- Accessed data are scattered in memory
- Few data are reused between vertices

Poor cache efficiency

Solution: Optimization of the vertex order



Reorder vertices so that adjacent vertices are co-located in memory



Co-located data & many data reuses

High cache efficiency

- Reordering improves the performance of graph analysis (Personalized PageRank) by up to 12 times (twice better than the existing method)
- Additionally, reordering with graph compression improves the performance by **up to 17 times**

Related works

[1] Junya Arai, Hiroaki Shiokawa, Takeshi Yamamuro, and Makoto Onizuka. "Scalable Parallel Graph Processing by Optimizing Vertex Order." In *Proc. the 6th Forum on Data Engineering and Information Management*, 2015. (In Japanese)

Contact

Junya Arai Software Innovation Center
E-mail : arai.junya(at)lab.ntt.co.jp

