Adaptive Partitioning of Large-Scale Dynamic Graphs

Luis Vaquero, Félix Cuadrado, Dionysios Logothetis, Claudio Martella

http://eecs.qmul.ac.uk/~fcuadrado

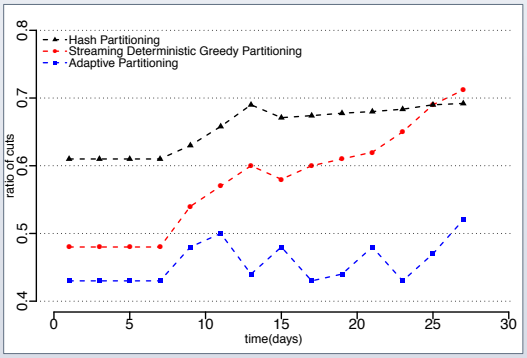
Queen Mary University of London

Telefónica

VU

Problem Statement

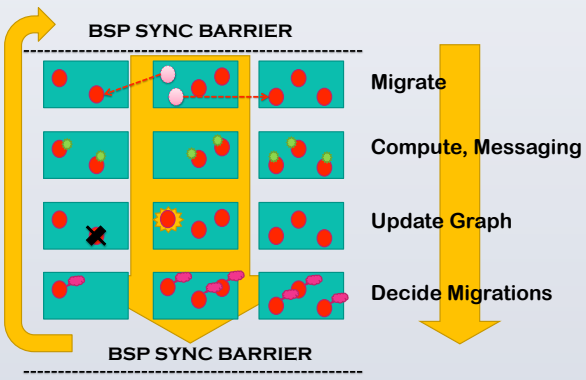
- Large-Scale Dynamic Graph Processing
- Graph partitioning impacts computation performance
- Graph changes degrade quality of partitioning



Evolution of partitioning on a graph from mobile CDR data

Approach

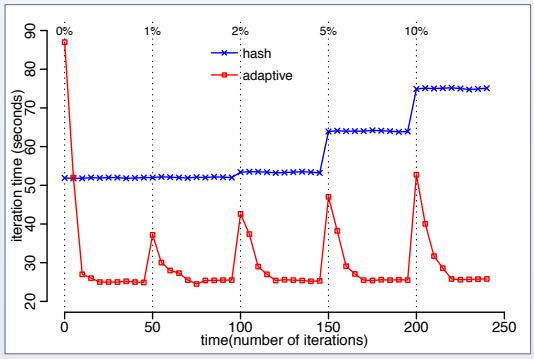
- Iterative partitioning optimization
- Vertex-level decision, greedy migration
- Adaptive partitioning: Trigger migration decisions as the graph changes
- Computation and partitioning joint execution loop



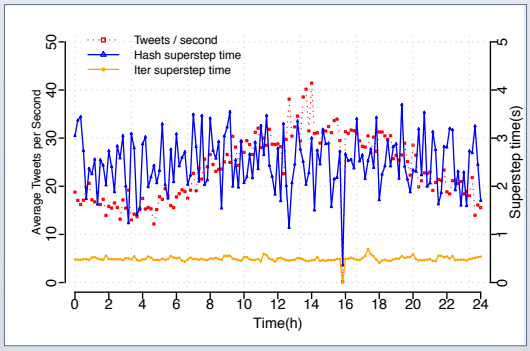
Adaptive Optimization on BSP Execution Loop

Experiments

- Performance comparison with hash partitioning



Livejournal graph expanded with bursts of new nodes/edges



Twitter Stream analytics (TunkRank)

Conclusions

- Migration overhead noticeable on initial steps
- Computation time stabilizes on frequent graph changes
- Vertex migrations improve computation performance

References

[1] R. Cheng, J. Hong, A. Kyrola, Y. Miao, X. Weng, M. Wu, F. Yang, L. Zhou, F. Zhao, and E. Chen. Kineograph: taking the pulse of a fast-changing and connected world. In EuroSys , 2012.

[2] G. Malewicz, M. H. Austern, A. J. Bik, J. C. Dehnert, I. Horn, N. Leiser, and G. Czajkowski. Pregel: asystem for large-scale graph processing. In PODC , 2009.

[3] I. Stanton and G. Kliot. Streaming graph partitioning for large distributed graphs. In KDD , 2012.

[4] L. Vaquero, F. Cuadrado, D. Logothetis, and C. Martella. xDGP: A dynamic graph processing system with adaptive partitioning. http://arxiv.org/abs/1309.1049 , 2013.