Mosaic: Processing a Trillion-Edge Graph on a Single Machine
Steffen Maass, Changwoo Min, Saridhya Kashyap, Woonhak Kang, Mohan Kumar, Taesoo Kim
Georgia Institute of Technology

Heterogenous Systems
- Powerful, Heterogenous Systems become commodity:
  - Terabytes of RAM
  - Powerful Co-processors (GPGPU, Xeon Phi)
  - Large-capacity, high-throughput Non-Volatile Memory (NVMe)

Large-scale Graph Processing
- Large-scale Graph Processing is ubiquitous
  - Machine Learning
  - Web (Social Networks, Search, ...)
  - Genome Analysis
  - ...

Our Approach: Employ sub-graph centric Encoding for Locality and Compression

Architecture

API

Key Abstraction: Hilbert-ordered Tiles

Performance Comparison

Cache Locality

Future Work
- Adaption for Evolving Graphs
  - Web graphs, Social Networks, ...
- New Abstractions for Algorithms?
  - Explore new classes of Algorithms
- Deep Learning on Graphs?
- Extension to Distributed Systems
  - Load balancing issues, ...

Take Away:
Organization of Graph Structure matters when processing a Trillion Edges on a Single Machine
Check our paper at EuroSys’17 for more details!