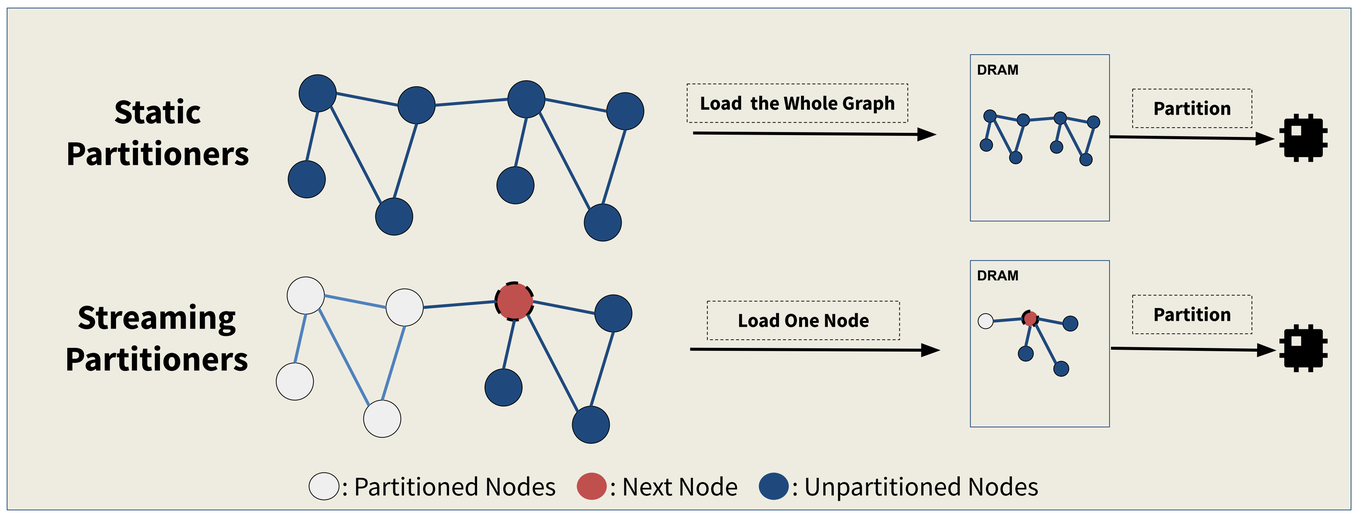
Department of Computer Science

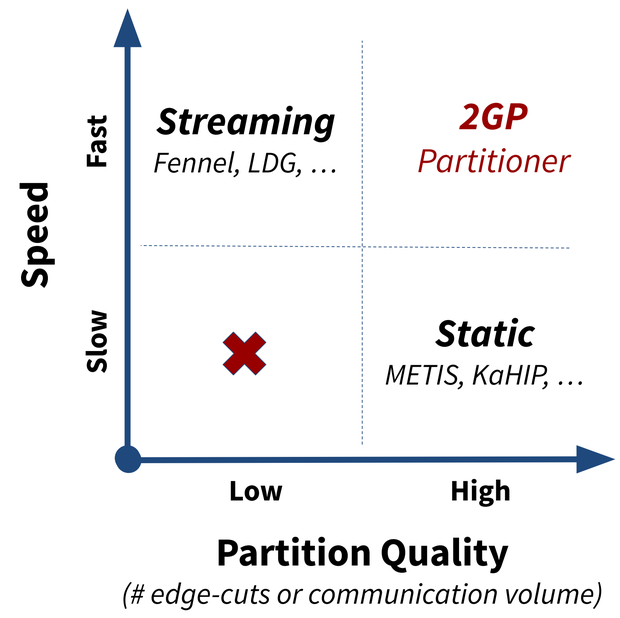
Two-Phase Graph Partitioner (2GP)

Hadi Sinaee (sinaee@cs.ubc.ca), Margo Seltzer (mseltzer@cs.ubc.ca)

Problem: Partitioning of Large Graphs Is Challenging!

- Graph partitioning enables **parallel** and **distributed processing**
- As graphs increase in **size** and **scale**, graph partitioning becomes the only viable means for their processing
- Good partitioning requires global knowledge of the graph



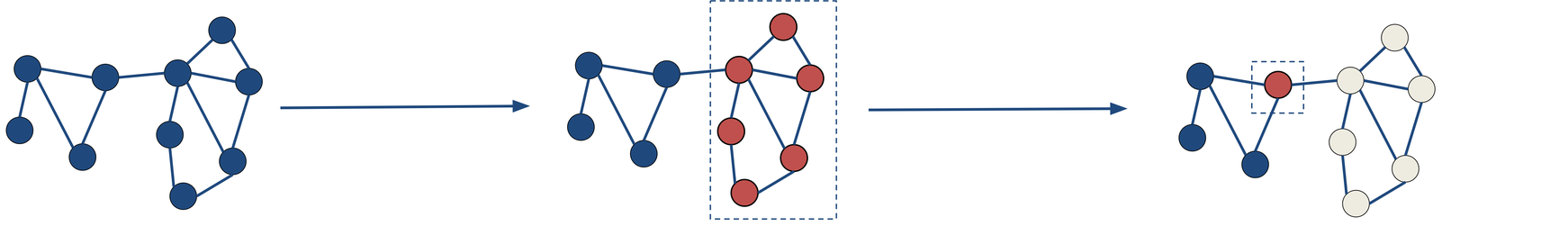


How can we achieve a good partitioning quality without sacrificing speed for large graphs?

Solution: Two-Phase Graph Partitioner (2GP)

We introduce **two-phase graph partitioner (2GP)** that combines the advantages of both static and streaming partitioners.

2GP Overview



Pre-processing
(reorder the graph)

Phase 1: Static Partitioner (METIS)
Partition a subgraph
(determined by available resources)

Phase 2: Stream Partitioner (FENNEL)
Partition the rest of the graph one node at time.

2GP Configurable Parameters

Input Graph Ordering	BFS/DFS, Highest Degree Order, …
Partitioners	Static: METIS, KaHIP, … Streaming: FENNEL, LDG, …
Sub-graph Size	25% of edges

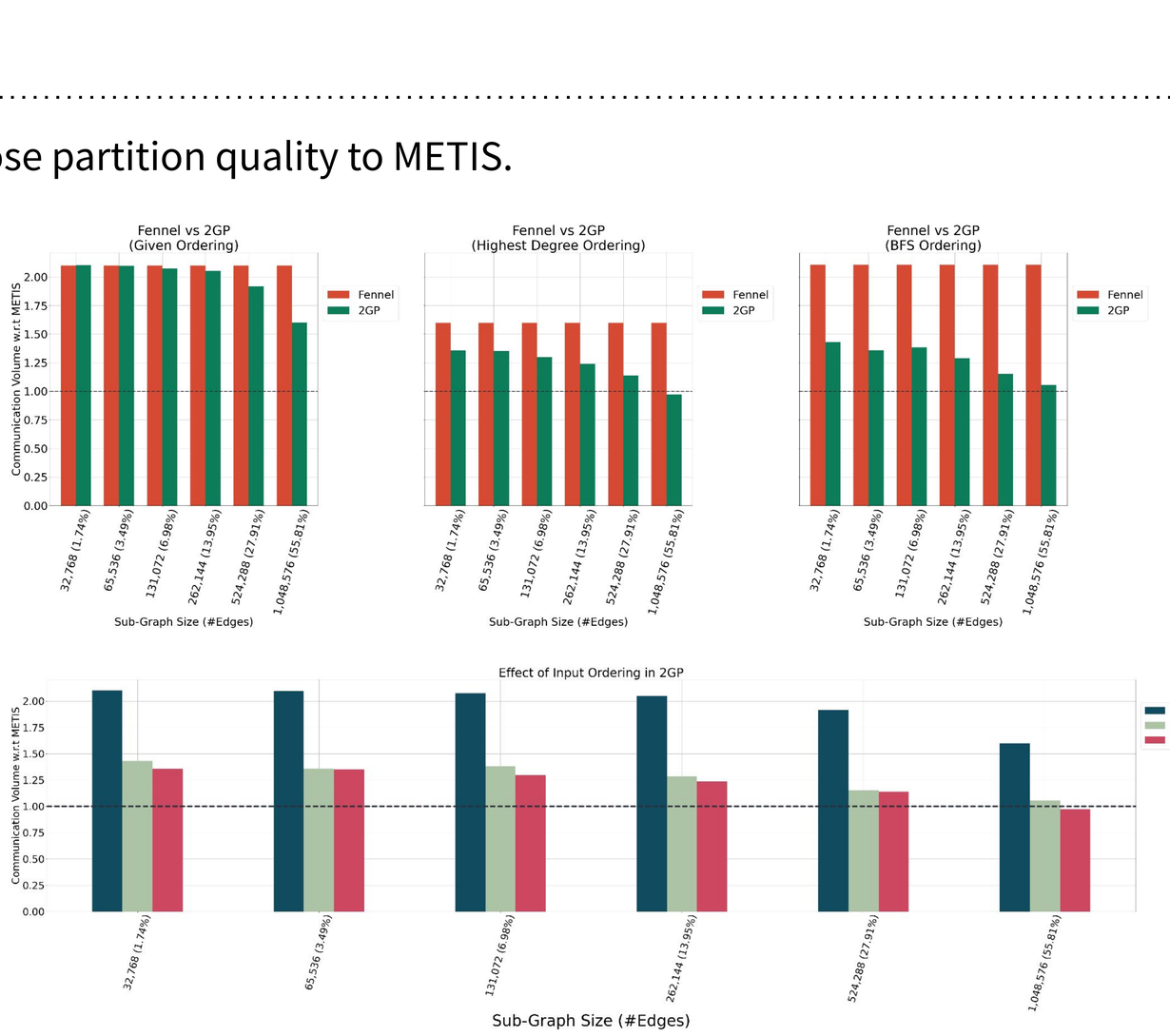
Evaluation: Preliminary Results

2GP has **35% better** partition quality than Fennel and close partition quality to METIS.

Dataset: Yahoo! Messenger Graph (~1.8M nodes, ~4M edges)

Communication Volume(**CV**): Fennel vs 2GP
(Normalized by Metis)

- **2GP** has **35% better** CV than Fennel
- **2GP** has close CV to Metis in *"Highest Degree Ordering"*



Effect of Ordering in 2GP
(Normalized by Metis)

- "*BFS Ordering*" and "*Highest Degree Ordering*" have close CV in **2GP**
- On average, "*Highest Degree Ordering*" performs better than other orderings