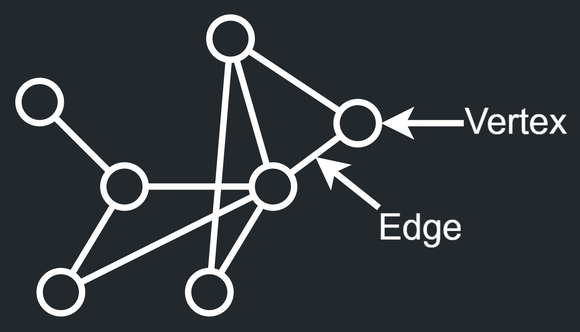
Graph-Optimizer

Performance and Energy Prediction for Graph Processing Workloads

Duncan Bart, Kuan-Hsun Chen, Ana-Lucia Varbanescu

1



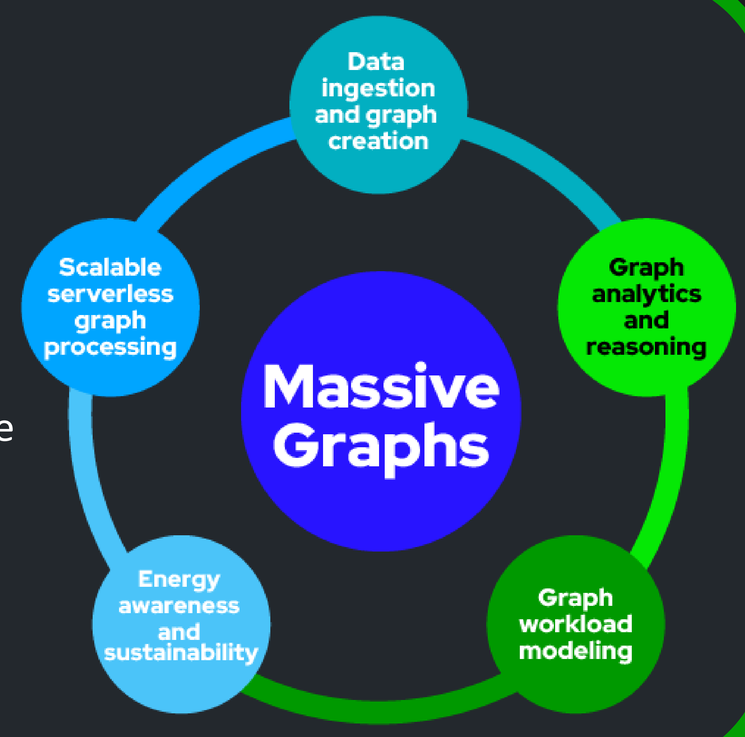
WHAT IS A GRAPH?

Graphs are structures containing entities (called vertices), and relations between these entities (called edges).
Graphs can be used to represent a wide variety of real-world systems, and are used in a wide range of domains. From physics and chemistry, to linguistics and social sciences.
All these domains benefit greatly from graphs, as they have proven to be a vital tool in solving many real-world problems.

2

MASSIVE GRAPH PROCESSING

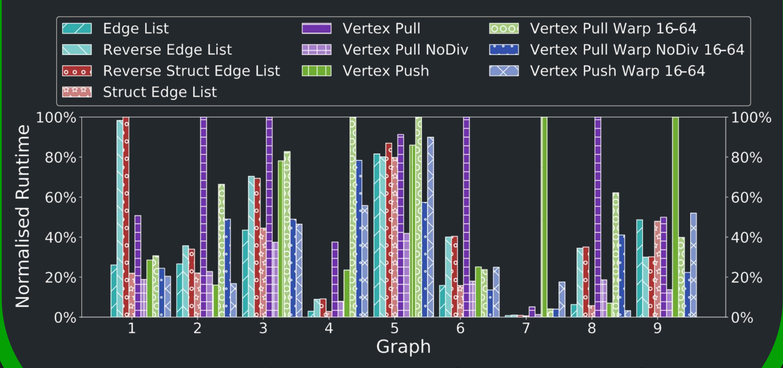
With the increasing scale of data, graph processing becomes a challenge, and existing solutions are no longer viable.
In order for domains to fully benefit from graph processing, it should be:
Easy to use, Scalable, Fast, and Sustainable
The Graph-Massivizer project aims to accomplish this, with the Graph-Optimizer tool playing a vital role in achieving fast and sustainable graph processing.



3

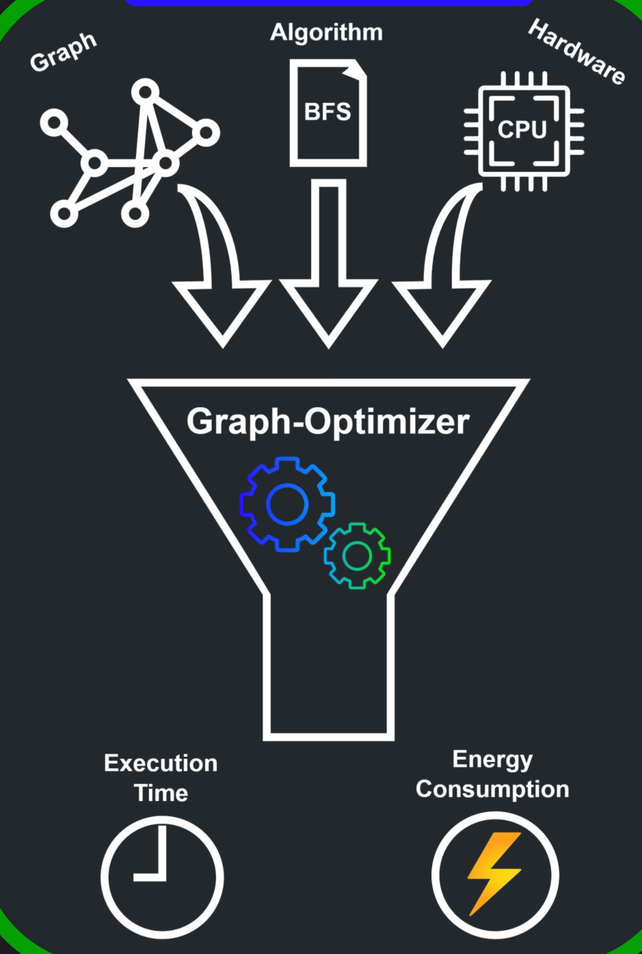
WHY PREDICTION?

For most basic graph operations (BGO), there exist multiple algorithms to solve it. The performance of each algorithm can vary by multiple orders of magnitude depending on the graph topology.
To select the fastest, or most energy efficient algorithm, reliable performance prediction is needed.



Source: Verstraaten, M. (2022) Analysis and prediction of GPU graph algorithm performance

THE SOLUTION



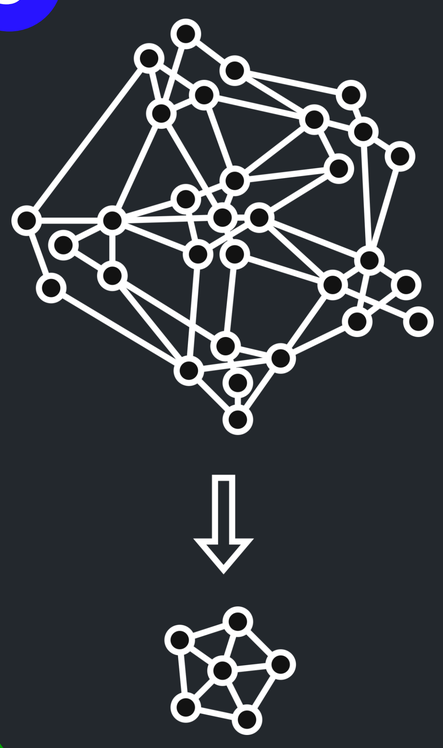
4

HOW IS PREDICTION DONE?

There are three main methods for performance prediction. Analytical modelling, statistical modelling, or simulation. All of them have significant drawbacks when it comes to predicting graph processing workloads. We need a solution that is fast and accurate, and preferably also traceable.

	Fast	Accurate	Traceable
Analytical	✓	✗	✓
Statistical	✓/✗	✓/✗	✗
Simulation	✗	✓	✓

5



HOW WILL WE SOLVE THIS?

We envision two possible solutions to achieve fast and accurate performance prediction for graph processing:

1) Graph summarization

By summarizing the graph, we can significantly reduce its size, while retaining key information about the topology. When executing the algorithms on this reduced graph, we aim to identify the best algorithm for the graph.

2) Hybrid modelling

Analytical modelling has poor accuracy due to its inability to correctly model memory access patterns. However, if we use a fast, lightweight memory simulator, and combine this with the rest of the analytical model, we aim to significantly improve the accuracy of the model.

This project has received funding from the European Union's Horizon Research and Innovation Actions under Grant Agreement № 101093202.

UNIVERSITY OF TWENTE.

GRAPH MASSIVIZER