GPU-Accelerated Graph500 Benchmarks on Frontier

Piyush Sao, Research Scientist
Oak Ridge National Laboratory
Phone: 865-341-0265; saopk@ornl.gov

Abstract

This project is aimed at improving the performance of Single-Source Shortest Path (SSSP) and Breadth-First Search (BFS) benchmarks on the Frontier supercomputer, which is one of the most powerful systems worldwide. We plan to accomplish this by optimizing and innovating the RMAT graph generation and preprocessing stages to utilize the massive parallelism provided by Frontier's GPU architecture. The project will be carried out in phases, beginning with the transfer of existing algorithms to the GPU, followed by performance profiling to establish a baseline. Subsequently, we will create and implement advanced algorithms designed for distributed GPU execution, with the goal of achieving significant performance enhancements. Our final deliverables will include optimized codebases, a set of performance models, and a research article intended for a top-tier supercomputing conference or journal. We expect that by the end of the project, we will have set new standards for Graph500 benchmarks, strengthening Frontier's position in the global supercomputing arena.

1 Optimizing Graph500 Benchmark Performance on Frontier

The Graph500 benchmark is a crucial tool for evaluating the performance and efficiency of supercomputers when dealing with data-intensive applications that involve large-scale graphs. These benchmarks simulate operations like navigating social networks, decoding genetic sequences, and mapping the intricate connections of neural networks, which are essential in various scientific and social fields.

Frontier, one of the most advanced supercomputers in the world, represents a remarkable combination of computational power and potential. However, improving its Graph500 benchmarks presents challenges due to the intrinsic complexities associated with graph analytics operations. These challenges include:

- **High Communication Cost**: The need for frequent data exchanges to represent graph structures can slow down performance, especially in distributed environments where communication costs are significant.
- **Limited Bandwidth and High Synchronization Costs**: These issues worsen communication challenges, leading to poor computational resource utilization due to bandwidth constraints and the necessity for synchronization across computing nodes.
- **Inefficient Mapping to GPU Architecture**: Graph algorithms struggle on GPU architectures, which favor high data parallelism. The irregular data patterns and branching inherent in graph processing do not align well with GPUs.
- **Sparse and Irregular Memory Access**: Graph data structures cause non-sequential and indirect memory accesses, which GPUs, optimized for uniform memory operations, handle poorly.
- **Complex Branching Structures**: The decision-making in graph algorithms involves branching, causing warp divergence on GPUs. This divergence leads to inefficiency as threads in a warp follow different execution paths.

These factors collectively contribute to a scenario where, despite the significant floating-point operation (FLOP) capabilities provided by GPUs in supercomputing systems, few Graph500 submissions exploit this power. The underutilization is primarily due to the challenges mentioned above, which this project aims to overcome. The primary goal of this project is to optimize Graph500 benchmark performance on Frontier and make better use of GPU resources in supercomputing graph analytics. By addressing the fundamental obstacles of communication, memory access, and algorithmic efficiency, we will unlock new capabilities for Frontier and set a precedent for GPU-accelerated graph processing on heterogeneous supercomputing systems.

2 Scope of Work for Optimizing Graph500 Benchmarks on Frontier

The Scope of Work (SoW) for optimizing Graph500 benchmarks on the Frontier supercomputer outlines a structured approach to enhance its graph processing capabilities. This project comprises several interrelated tasks:

- **GPU Acceleration of Graph Generation**: The first phase of optimizing the Graph500 involves refining the RMAT graph generation and preprocessing code for effective GPU performance. This is crucial because the as the current graph generation step significantly lags behind graph traversal in terms of speed—sometimes by an order of magnitude. Although the graph generation and preprocessing are not part of the benchmark's main metrics, it is still important to optimize it for a comprehensive evaluation of different configurations and performance profiling. Our focus is on improving the data structures and algorithms to match the processing capabilities and memory hierarchy of the GPU. This will help us the optmization process streamlined and suitable for extensive experimentation. We will also compare the GPU-accelerated graph generation with traditional CPU-based methods to ensure the accuracy and reliability of the generated graphs. This validation is essential to make sure that subsequent experiments are based on precise and dependable data.
- **Design Space Exploration for Optimizing Graph500 for Frontier**:Our goal is to thoroughly examine algorithmic and architectural designs to find the best configurations for improving Graph500 benchmark performance on the Frontier supercomputer. We will investigate ways to optimize the execution of SSSP and BFS algorithms on large graphs, focusing on algorithmic choices, architecture-specific optimizations, software availability, and promising configuration implementations. Our exploration will include arranging MPI processes, combining MPI with OpenMP, using communication routines like broadcast algorithms, and strategies for data distribution, graph partitioning, and load balancing. We will also assess Frontier's architecture to identify optimization opportunities for Graph500 performance and examine software tools and libraries for high-performance graph processing on Frontier. Our process aims to select configurations and optimizations that potentially enhance Graph500 performance significantly. Ultimately, we seek to create a foundation for targeted optimizations, focusing our efforts on the most promising strategies to improve Frontier's Graph500 benchmarks.
- **Performance Profiling & Modeling**: After exploring algorithmic and architectural designs, we move to profiling and modeling. This step is crucial for understanding the performance characteristics of key components, algorithms, and configurations. It helps us see how factors like problem size and computing nodes count impact performance. **Profiling** involves running key components alone and within the full benchmark to measure their actual throughput. By comparing these results with theoretical expectations, we can identify performance gaps. This reveals optimization opportunities by highlighting differences between expected and actual performance. **Modeling** is about creating empirical models to predict performance based on various parameters. These models assist in choosing the best parameter mix, determining the smallest viable problem size, and deciding on the necessary number of nodes for certain performance targets.

  Through these performance models, we aim to set achievable performance goals, identify where performance lags, and locate optimization needs. This structured approach will help us enhance the performance of essential algorithms and configurations on the Frontier supercomputer, ensuring focused and efficient optimization efforts.
- **New Algorithm Development and Optimization**: Our project's pivotal element involves creating new graph processing algorithms tailored for GPU performance. We aim to develop optimized distributed GPU algorithms for Single Source Shortest Path (SSSP) and Breadth-First Search (BFS), primarily focusing on reducing communication overhead. Additionally, we plan to incorporate a single broadcast mechanism to boost both efficiency and scalability on GPUs.
- **Graph500 Package and Documentation:** Our efforts will culminate in submitting an optimized Graph500 code package specifically tailored for Frontier, accompanied by a detailed guide outlining setup, execution processes, and expected performance outcomes. In addition, we will provide a detailed guide outlining the setup, execution processes, and expected performance outcomes. We will document challenges, solutions, and critical decisions. Additionally, we will author a research article detailing our methodologies and their impact on Frontier's Graph500 results, targeting a leading supercomputing conference or jour-

nal. This emphasizes our contribution to high-performance computing and graph processing.

This SoW sets a clear roadmap for addressing the challenges of graph processing on Frontier, aligning each task with strategic goals to achieve significant performance enhancements.

Timeline

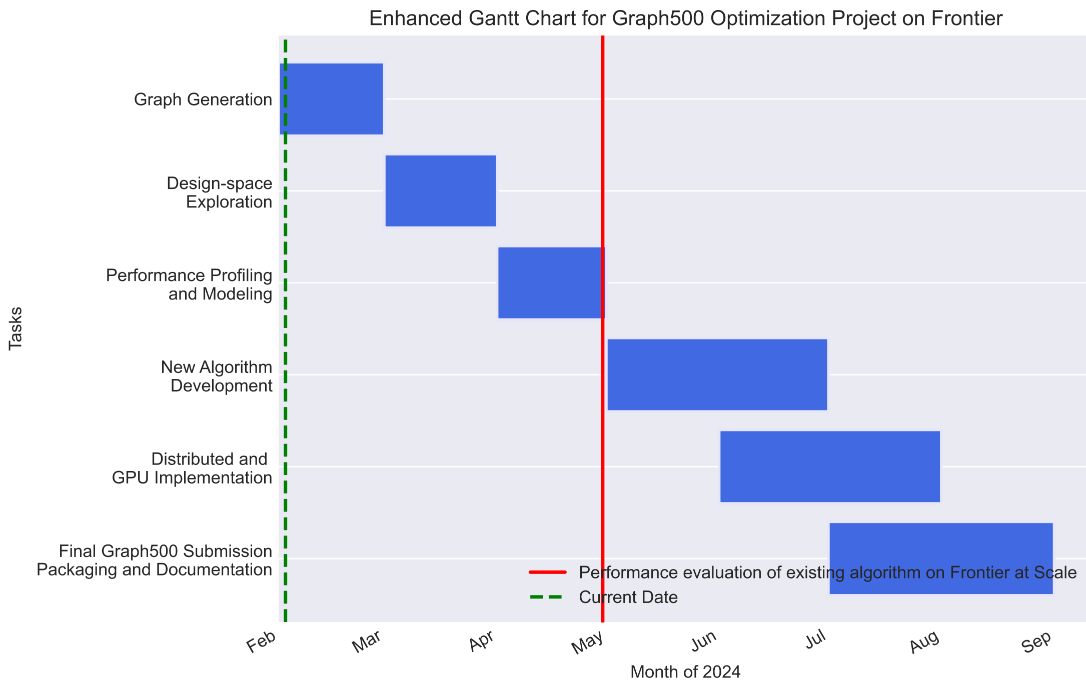

Figure 1: Gantt chart for the project timeline.

3 Deliverables

The deliverables for this project are designed to underscore the advancements at each stage, ultimately enhancing the Frontier supercomputer's Graph500 benchmarks. The project will produce the following key outcomes:

- **GPU-Accelerated RMAT Code and Documentation**: This deliverable includes an enhanced RMAT graph generation code optimized for GPU efficiency, complete with source code and executable binaries. Additionally, detailed documentation will cover the optimizations, usage instructions, and considerations unique to Frontier, ensuring thorough understanding and ease of use.
- **Profiling Tools and Performance Reports**: A suite of profiling tools will be developed specifically for Frontier's architecture, enabling detailed analysis of graph processing performance. Performance reports will provide a baseline for SSSP and BFS algorithms on Frontier, including an analysis of existing bottlenecks and opportunities for further improvements.
- **Performance Models and Analysis Reports**: Performance models for RMAT graph generation and key algorithms (SSSP and BFS) will be created, applicable across various graph sizes and computational conditions. Analysis reports will discuss these models to identify optimal graph sizes and problem scales for peak performance and efficient profiling.
- **Optimized SSSP and BFS Algorithms**: The project aims to develop advanced GPU-optimized SSSP and BFS algorithms that focus on achieving high performance with minimal communication overhead. The deliverable will include source code and comprehensive documentation, highlighting innovations such as single-broadcast techniques that enhance scalability and efficiency.

- **Optimized Graph500 Code Submission for Frontier**: The culmination of this project will be the final, optimized Graph500 code package specifically prepared for Frontier, ready for benchmark submission. A detailed guide will accompany this package, outlining the setup, execution processes, and expected performance metrics, providing a clear framework for evaluation.
- **Research Article for Conference or Journal Submission**: A research article, prepared for submission to a prestigious supercomputing conference or journal, will detail the methodologies, findings, and contributions of the project to the field of GPU-based high-performance computing. This article aims to share insights and advancements with the broader scientific community, promoting further innovation and research.

These deliverables, serving as tangible evidence of the project's impact, are meticulously planned to highlight our achievements and foster further innovation and research within the scientific and computational community.