

Fast Processing of Large Graph Applications Using Asynchronous Architecture

Michel A. Kinsy, Rashmi S. Agrawal and Hien D. Nguyen
Adaptive and Secure Computing Systems (ASCS) Laboratory
Department of Electrical and Computer Engineering, Boston University
8 Saint Mary's Street, Boston MA 02215 - Email: mkinsy@bu.edu

Abstract—Graph algorithms and techniques are increasingly being used in scientific and commercial applications to express relations and explore large data sets. Although conventional or commodity computer architectures, like CPU or GPU, can compute fairly well dense graph algorithms, they are often inadequate in processing large sparse graph applications. Memory access patterns, memory bandwidth requirements and on-chip network communications in these applications do not fit in the conventional program execution flow. In this work, we propose and design a new architecture for fast processing of large graph applications. To leverage the lack of the spatial and temporal localities in these applications and to support scalable computational models, we design the architecture around two key concepts. (1) The architecture is a multicore processor of independently clocked processing elements. These elements communicate in a self-timed manner and use handshaking to perform synchronization, communication, and sequencing of operations. By being asynchronous, the operating speed at each processing element is determined by actual local latencies rather than global worst-case latencies. We create a specialized ISA to support these operations. (2) The application compilation and mapping process uses a graph clustering algorithm to optimize parallel computing of graph operations and load balancing. Through the clustering process, we make scalability an inherent property of the architecture where task-to-element mapping can be done at the graph node level or at node cluster level. A prototyped version of the architecture outperforms a comparable CPU by 10~20x across all benchmarks and provides 2~5x better power efficiency when compared to a GPU.

I. INTRODUCTION

Advances in mobile computing, coupled with the proliferation of online social networks, have given rise to a new class of applications and computing challenges [1]. These applications tend to be relational by nature. In other words, they express or encode relations, communications, connectivity and interactions between people, places, objects or systems. As such, the data of interest in these applications are often best represented in the form of graphs. Graph-based applications range from social network analyses to anomaly detections [2]. For computing purposes, graphs are commonly represented in one of two forms: (1) as adjacency matrix or (2) as adjacency list. Adjacency Matrix works well for densely connected graphs, i.e., the number of edges in the graph is close to the maximal number of edges. In general, computing on dense graphs can be easily parallelized and GPU and SIMD architectures have proven to be the platform of choice for executing such graph-based applications [2]. Unfortunately, the vast majority of large graph-based applications are sparse. For the efficient storage of large sparse graphs, adjacency list or other compressed representation schemes are used. Memory

access and load balancing are some of the key bottlenecks to the efficient processing of large sparse graph algorithms and applications [1]. The memory access patterns often lack spatial and temporal localities resulting in high cache miss rates. Current cache-based processor architectures are simply not well suited for the computational flow of graph processing. In addition to the storage problem, computing on large sparse graphs currently presents a number of challenges including effective programming abstractions and models of computation that leverage the graph structure in the application. In this work, we present a domain-specific architecture tailored to graph-based algorithms and applications.

II. PROPOSED GRAPH PROCESSOR ARCHITECTURE

Figure 1 shows an illustration of the proposed architecture. The three key modules of the architecture are (1) the graph processor, (2) the co-processor and (3) the main memory. The graph processor (1) module has a Memory Interface unit (1a) to coordinate batch accesses to the main memory or external memory units, a Dispatch Logic (1b) to perform scatter operations on data from the main memory, an Output Logic (1c) to gather output data from the graph processor, and a systolic array of simple processing elements called *Node Arithmetic Logic Engines* (NALEs) (1d) to carry out the actual graph computations. The co-processor (2) performs three key functions. It (1) executes non-graph parts of the application, (2) schedules the graph part of the application and (3) monitors the execution flow of the graph.

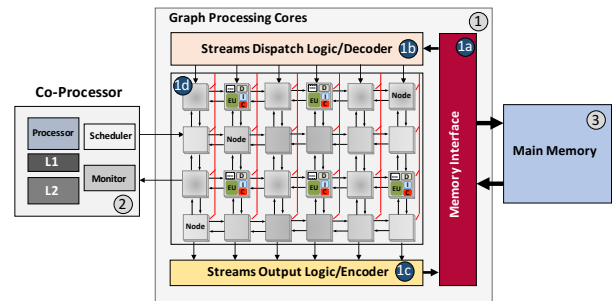


Fig. 1. Proposed graph processor system architecture.

Graph Processor Micro-Architecture: Figure 2 shows the micro-architecture of a NALE. The NALE is optimized for fast MAC (Multiply-And-Accumulate) operations with a three-state output comparator for fast node value sorting. It has two FIFO structures, one to communicate with neighbors and one internal FIFO to emulate multiple graph nodes (node cluster mode execution).

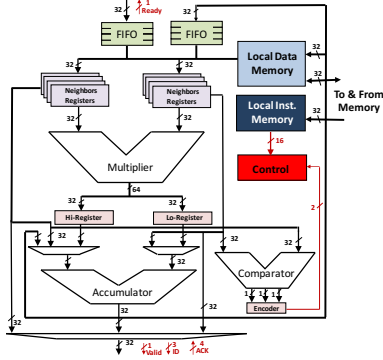


Fig. 2. Micro-architecture of a Node Arithmetic Logic Engine (NALE).

Each NALE operates independently of others depending on the readiness of inputs. Communicating through FIFOs only allows each NALE to run on its own clock speed. Furthermore, this approach allows us to adopt a GasP asynchronous [3] design methodology that can seamlessly scale to hundreds of thousands of NALEs. Figure 3(a) illustrates the clockless handshake logic between NALEs. In addition to the scalability benefits, the absence of a global clock allows for the underlining data dependencies to dictate application execution time. Figure 3(b) shows the synthesizable equivalent of the GasP circuit.

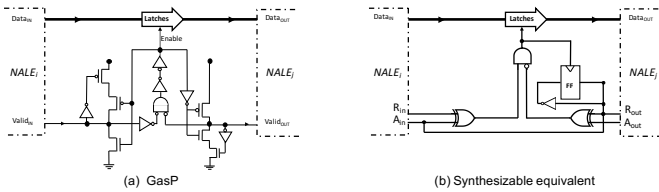


Fig. 3. GasP asynchronous communication circuit between NALEs.

Model of computation and compilation : An asynchronous model of computation is adopted to fully take advantage of the graph processor. Given a graph application specification and a number of available NALEs for its computation, the execution preprocessing flow follows five key steps. Figure 4 illustrates these steps. In the first step, the application is profiled to extract the graph topology, followed by the clustering of nodes, clusters dependency analysis, placement and finally the compilation step.

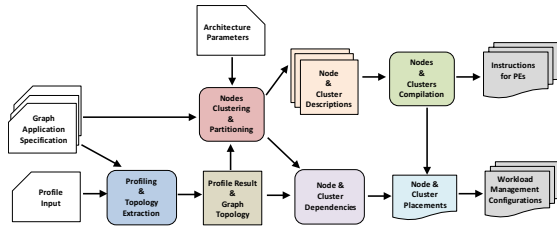


Fig. 4. Compilation process steps for the graph processor.

III. ARCHITECTURE EVALUATION

Experimental setup: To get high-fidelity performance and power measurements for the proposed architecture, we prototype it alongside a conventional CPU and a GPU with comparable complexity in FPGA. The Xilinx Virtex7-XC7VX980T FPGA device is used for our prototyping platform. We implement a synthesizable RTL version of the graph processor. We

use the 7-stage RISC core in the Heracles [4] RTL simulator for the CPU. We adopte the MIAOW open-source general-purpose graphics processor (GPGPU) based on the AMD Southern Islands ISA [5] for the GPU architecture. The three architectures are evaluated based on their execution time and power.

Graph algorithms and applications: For the evaluation, we consider a set of representative graph algorithms, namely, Single Source Shortest Path (SSSP), Breadth First Search (BFS), Depth First Search (DFS), PageRank (PR), Minimal Enclosing Triangles (MiniTri), and Connected Components (CC). We use three difference graph applications: (1) California road network (CA) which has 1,965,206 vertices, 2,766,607 edges and an average degree of 1.41, (2) Facebook social network (FB) with 2,937,612 vertices, 41,919,708 edges and an average degree of 14.3 and (3) Livejournal social network (LJ) with 4,847,571 vertices, 85,702,475 edges and 17.6 average degree.

Results: Figures 5 and 6 present the execution time in terms of number of cycles and power usage for each platform for the different applications and graph algorithms.

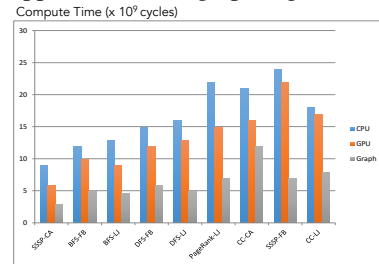


Fig. 5. Performance in terms of execution for the three architecture types on the different graph applications.

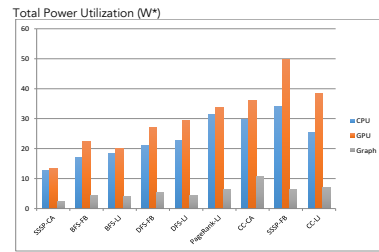


Fig. 6. Efficiency in terms of power usage for the three architecture types on the different graph applications.

REFERENCES

- [1] A. Kyrola, G. Blelloch, and C. Guestrin, "Graphchi: Large-scale graph computation on just a pc," in *Proceedings of the 10th USENIX Conference on Operating Systems Design and Implementation*, ser. OSDI'12. Berkeley, CA, USA: USENIX Association, 2012, pp. 31–46.
- [2] A. Lumsdaine, D. Gregor, B. Hendrickson, and J. Berry, "Challenges in parallel graph processing," *Parallel Processing Letters*, vol. 17, no. 01, pp. 5–20, 2007.
- [3] M. Roncken, S. M. Gilla, H. Park, N. Jamadagni, C. Cowan, and I. Sutherland, "Naturalized communication and testing," in *21st IEEE International Symposium on Asynchronous Circuits and Systems*, May 2015, pp. 77–84.
- [4] M. A. Kinsy, M. Pellauer, and S. Devadas, "Heracles: A tool for fast rtl-based design space exploration of multicore processors," in *Proceedings of the ACM/SIGDA International Symposium on Field Programmable Gate Arrays*, ser. FPGA 2013, pp. 125–134.
- [5] R. Balasubramanian et al., "Enabling gpgpu low-level hardware explorations with miaow: An open-source rtl implementation of a gpgpu," *ACM Trans. Archit. Code Optim.*, vol. 12, no. 2, pp. 21:25, Jun. 2015.