

Measurement and analysis of GPU-accelerated applications with HPCToolkit

Keren Zhou, Laksono Adhianto, Jonathon Anderson, Aaron Cherian, Dejan Grubisic, Mark Krentel, Yumeng Liu, Xiaozhu Meng, John Mellor-Crummey*

Department of Computer Science, Rice University, Houston, TX, United States of America

ARTICLE INFO

Keywords:

Supercomputers
High performance computing
Software performance
Performance analysis

ABSTRACT

To address the challenge of performance analysis on the US DOE's forthcoming exascale supercomputers, Rice University has been extending its HPCToolkit performance tools to support measurement and analysis of GPU-accelerated applications. To help developers understand the performance of accelerated applications as a whole, HPCToolkit's measurement and analysis tools attribute metrics to calling contexts that span both CPUs and GPUs. To measure GPU-accelerated applications efficiently, HPCToolkit employs a novel wait-free data structure to coordinate monitoring and attribution of GPU performance. To help developers understand the performance of complex GPU code generated from high-level programming models, HPCToolkit constructs sophisticated approximations of call path profiles for GPU computations. To support fine-grained analysis and tuning, HPCToolkit uses PC sampling and instrumentation to measure and attribute GPU performance metrics to source lines, loops, and inlined code. To supplement fine-grained measurements, HPCToolkit can measure GPU kernel executions using hardware performance counters. To provide a view of how an execution evolves over time, HPCToolkit can collect, analyze, and visualize call path traces within and across nodes. Finally, on NVIDIA GPUs, HPCToolkit can derive and attribute a collection of useful performance metrics based on measurements using GPU PC samples. We illustrate HPCToolkit's new capabilities for analyzing GPU-accelerated applications with several codes developed as part of the Exascale Computing Project.

1. Introduction

In recent years, compute nodes accelerated with Graphics Processing Units (GPUs) have become increasingly common in supercomputers. In June 2020, six of the world's ten most powerful supercomputers employ GPUs [1]. Each of the US DOE's forthcoming exascale systems Aurora, Frontier, and El Capitan are based on GPU-accelerated compute nodes.

While GPUs can provide high performance, without careful design GPU-accelerated applications may underutilize GPU resources by idling compute units, employing insufficient thread parallelism, or exhibiting poor data locality. Moreover, while higher-level programming models such as RAJA [2], Kokkos [3], OpenMP [4], and DPC++ [5] can simplify development of HPC applications, they can increase the difficulty of tuning GPU kernels (routines compiled for offloading to a GPU) for high performance by separating developers from many key details, such as what GPU code is generated and how it will be executed.

To harness the full power of GPU-accelerated nodes, application developers need tools to identify performance problems. Performance tools for GPU-accelerated programs employ *trace* and *profile* views. A trace view presents events that happen over time on each process, thread, and GPU stream. A profile view aggregates performance metrics

over the time dimension. Most performance tools that support GPUs [6–14] only provide trace and profile views with the name of each GPU kernel. For large-scale GPU-accelerated applications, it is often difficult to understand how performance problems arise without associating the cost of GPU kernels with their CPU calling contexts. Manually associating the performance of GPU kernels with their CPU calling contexts is difficult when a kernel is called from many contexts or when the name of a kernel is the result of a C++ template instantiation.

Since 2015, NVIDIA GPUs support fine-grained measurement of GPU performance using Program Counter (PC) sampling [15]. Intel's GT-Pin [16] and NVIDIA's NVBit [17] provide APIs to instrument GPU machine code to collect fine-grained metrics. While tools such as MAP [8], nvprof [14], Nsight-compute [13], TAU [7], and VTune [11] use PC sampling or instrumentation to associate fine-grained metrics with individual source lines for GPU code, they do not associate metrics with loop nests or calling contexts for GPU device functions, which are important to understand the performance of complex GPU kernels. For example, a template-based dot product kernel in the RAJA performance suite [18] yields 25 different GPU functions that implement the computation.

* Corresponding author.

E-mail address: johnmc@rice.edu (J. Mellor-Crummey).

To address these challenges, we are extending Rice University's HPCToolkit performance tools [19] to support scalable measurement and analysis of GPU-accelerated applications running on NVIDIA, AMD, and Intel GPUs. HPCToolkit collects call path profiles and presents them with a graphical user interface that provides both profile and trace views. After our initial extensions to support GPU-accelerated programs, HPCToolkit has the following capabilities:

- It uses a GPU-independent measurement framework to monitor and attribute performance of GPU code.
- It employs wait-free queues for efficient coordination between application, runtime, and tool threads.
- It supports measurement and attribution of fine-grained metrics using PC sampling and instrumentation.
- It employs compact sparse representations of performance metrics to support efficient collection, storage, and inspection of performance metrics within and across processes, threads, and GPU streams.
- It employs a combination of distributed-memory parallelism and multithreading to aggregate global performance metrics across a large number of profiles.
- It provides useful information to guide performance optimization, including heterogeneous calling contexts, derived metrics, and idleness analysis.

We present some early experiences with codes from the Exascale Computing Project to illustrate HPCToolkit's attribution of fine-grained measurements to heterogeneous calling contexts on multiple GPU platforms and its capability to measure and analyze executions across hundreds of GPUs.

This rest of the paper is organized as follows. Section 2 reviews related work and highlights HPCToolkit's features. Section 3 describes HPCToolkit's workflow for GPU-accelerated applications. Section 4 describes the design of HPCToolkit's measurement framework for collecting GPU performance metrics. Section 5 discusses analysis of GPU binaries for performance attribution. Section 6 presents scalable algorithms for aggregating performance data from parallel programs. Section 7 describes HPCToolkit's profile and trace views for analyzing measurements of GPU-accelerated applications. Section 8 illustrates HPCToolkit's capabilities with views of several codes from the Exascale Computing Project. Section 9 reflects on our experiences and briefly outlines some future plans.

2. Related work

Developing performance tools for GPU-accelerated applications has been the focus of considerable past and ongoing work. NVIDIA provides tools [12–14] to present a trace view of GPU kernel invocations and a profile view for individual kernels. Intel's VTune [11] monitors executions on both CPUs and GPUs. AMD provides ROCProfiler [20] to monitor GPU-accelerated applications. In addition, a collection of third-party performance tools have been developed for GPU-accelerated applications. Malony et al. [21] describe early tools for collecting kernel timings and hardware counter measurements for CUDA and OpenCL kernels. Welton and Miller [22] investigated hidden performance issues that impact several HPC applications but are not reported by tool APIs. Kousha et al. [23] developed a tool for monitoring communications on multiple GPUs. Unlike the aforementioned tools, HPCToolkit collects call path profiles and shows calling context information in both trace and profile views. Early work on HPCToolkit [24] describes using GPU events and hardware counters for kernel-level monitoring on NVIDIA GPUs to compute profiles that blame CPU code for associated GPU idleness.

With the increased complexity of GPU kernels, fine-grained measurement of performance metrics within GPU kernels are critical for providing optimization insights. At present, only NVIDIA GPUs support using PC sampling [15] to collect fine-grained instruction stall

information. NVIDIA's nsight-compute collects data using PC sampling and provides performance information at the GPU kernel level. CUDA-Blamer [25] was a proof-of-concept prototype that collects PC samples and reconstructs static call paths on GPUs with information from LLVM-IR. Unlike CUDABlamer, HPCToolkit reconstructs GPU calling context trees by analyzing GPU binaries and distributes costs of GPU functions based on PC sample counts.

Several vendor tools support instrumentation of GPU kernels. NVIDIA's NVBit [17] and Sanitizer API [26], as well as Intel's GT-Pin [16] provide callback APIs to inject instrumentation into GPU machine code. Tools can use these APIs to collect fine-grained metrics. For example, Goroshov et al. [27] use instrumentation to measure basic block latency and detect hot code regions. GVProf [28] instruments GPU memory instructions to profile value redundancies. In HPCToolkit, we use GT-Pin to measure instruction counts within GPU kernels.

Scalable analysis of performance measurements will be critical for gaining insight into executions on forthcoming exascale platforms. NVIDIA's NVProf [14] and Intel's VTune [11] record measurements as traces. To our knowledge, these tools lack support for scalable analysis of measurement data. Scalasca, TAU [7], and Vampir [6] present data gathered by the Score-P measurement infrastructure [29]. At execution finalization, Score-P aggregates profile data in parallel into the CUBE storage format. To date, there has only been a preliminary study exploring the addition of sparsity to CUBE [30]; for GPU-accelerated applications, sparsity is essential. MAP [8] selects a user-defined subset of collected samples at runtime to limit the amount of measurement data collected per thread. This is effective for scalable overview analysis, however this does not retain sufficient data for in-depth analyses.

Past research has used trace analysis to identify performance bottlenecks within and across compute nodes. Wei et al. [31] describe a framework that diagnoses scalability losses in programs using multiple MPI processes and CPU threads. Choi et al. [32] analyze traces from simulators to estimate performance on GPU clusters. Schmitt et al. [33] use Vampir [6]'s instrumentation of MPI primitives to gather communication traces. From these traces, they construct a dependency graph and explore dependencies between communication events and GPU computations. Unlike other tracing tools, HPCToolkit gathers CPU traces using sampling rather than instrumentation, which has much lower overhead.

3. Overview

Fig. 1 shows HPCToolkit's workflow to analyze programs running on GPUs. HPCToolkit's *hpcrun* measurement tool collects GPU performance metrics using profiling APIs from GPU vendors or custom hooks with LD_PRELOAD. *hpcrun* can measure programs that employ one or more GPU programming models, including OpenMP, OpenACC, CUDA, HIP, OpenCL, and DPC++. As GPU binaries are loaded into memory, *hpcrun* records them for later analysis. For GPUs that provide APIs for fine-grained measurement, *hpcrun* can collect instruction-level characterizations of GPU kernels using hardware support for sampling or binary instrumentation. *hpcrun*'s output includes profiles and optionally traces. Each profile contains a calling context tree in which each node is associated with a set of metrics. Each trace file contains a sequence of events on a CPU thread or a GPU stream with their timestamps.

hpcstruct analyzes CPU and GPU binaries to recover static information about procedures, inlined functions, loop nests, and source lines. There are two aspects to this analysis: (1) recovering information about line mappings and inlining from compiler-recorded information in binaries, and (2) analyzing machine code to recover information about loops.

hpcprof and *hpcprof-mpi* correlate performance metrics for GPU-accelerated code with program structure. *hpcprof* employs a multi-threaded streaming aggregation algorithm to quickly aggregate profiles, reconstruct a global calling context tree, and relate measurements

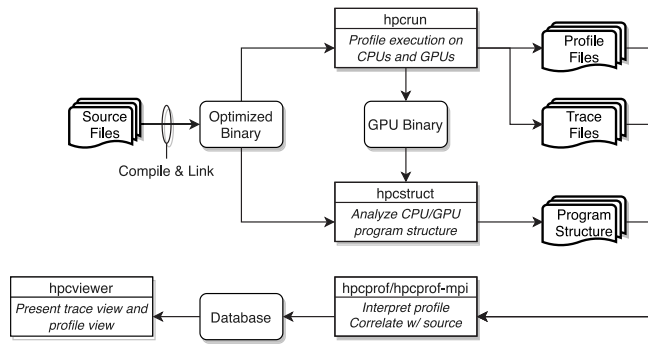


Fig. 1. HPCToolkit's workflow for analysis of GPU-accelerated applications.

associated with machine instructions back to CPU and GPU source code. To accelerate analysis of performance data from extreme-scale executions, *hpcprof-mpi* additionally employs distributed-memory parallelism for greater scalability. Both *hpcprof* and *hpcprof-mpi* write sparse representations of their analysis results in a database.

Finally, *hpcviewer* interprets and visualizes the database. In its profile view, *hpcviewer* presents a heterogeneous calling context tree that spans both CPU and GPU contexts, annotated with measured or derived metrics to help users assess code performance and identify bottlenecks. In its trace view, *hpcviewer* identifies each CPU or GPU trace line with a tuple of metadata about the hardware (e.g., node, core, GPU) and software constructs (e.g., rank, thread, GPU stream) associated with the trace. Automated analysis of traces can attribute GPU idleness to CPU code.

4. Performance measurement on GPUs

HPCToolkit's *hpcrun* collects GPU performance metrics and associates them with calling context at every GPU API invocation. Section 4.1 describes HPCToolkit's unified infrastructure for collecting and attributing performance metrics on AMD, Intel, and NVIDIA GPUs. Section 4.2 describes how HPCToolkit collects fine-grained metrics using hardware instruction sampling or binary instrumentation. Section 4.3 describes support for measuring GPU kernel executions with hardware counters. Section 4.4 describes how HPCToolkit employs performance measurement substrates from GPU vendors. Section 4.5 describes how HPCToolkit collects metrics at runtime and computes derived metrics during post-mortem analysis. Section 4.6 describes HPCToolkit's use of sparse representations of performance metrics at runtime and as the products of post-mortem analysis. Section 4.7 explains the utility of combining measurements from multiple runs.

4.1. Infrastructure

Fig. 2 illustrates how *hpcrun* monitors the execution of GPU-accelerated applications. As application threads offload computations to GPUs, HPCToolkit employs a GPU monitor thread to asynchronously process measurement data from the GPUs. If tracing is enabled, *hpcrun* creates one or more tracing threads to record an activity trace for each GPU stream.

When an application thread performs an invocation I of a GPU operation (e.g., a kernel or a data copy), *hpcrun* unwinds the application thread's call stack to determine the CPU calling context of I , inserts a placeholder P representing the operation in that context, communicates I and P to the monitor thread, and initiates the GPU operation after tagging it with I . The monitor thread collects measurements consisting of one or more GPU activities $\mathcal{A}_1, \dots, \mathcal{A}_n$ associated with I and sends them back to the application thread for attribution below P to form a heterogeneous calling context. When tracing is enabled, the monitor thread separates GPU activities by their associated stream and

sends each stream of activities to a tracing thread. Each tracing thread records one or more GPU streams of activities and their timestamps into trace files. For efficient inter-thread communication, HPCToolkit uses bidirectional channels, each consisting of a pair of wait-free single-producer and single-consumer queues [34]. The precise instantiation of HPCToolkit's monitoring infrastructure is tailored to each GPU vendor's software for monitoring GPU computations.

When using NVIDIA's CUPTI [35] and AMD's ROCTracer [36] libraries for monitoring GPU activities, a monitor thread created by these libraries receives measurements of GPU activities via a buffer completion callback. Each application thread \mathcal{T} shares two channels with the GPU monitor thread, including an activity channel C_A , from which \mathcal{T} receives information about GPU activities associated with operations it invoked, and an operation channel C_O on which \mathcal{T} enqueues GPU operation tuples of (I, P, C_A) , representing an invocation I , its associated placeholder P , and its activity channel C_A . Every time the GPU monitor thread receives a buffer completion callback, it drains its incident operation channels prior to processing a buffer full of GPU activities. The GPU monitor thread matches each GPU activity \mathcal{A} , tagged with its invocation I , with its associated operation tuple (I, P, C_A) . The monitor thread enqueues a pair (\mathcal{A}, P) into activity channel C_A to attribute the GPU activity back to \mathcal{T} .

When using OpenCL [37] and Level Zero [38], depending upon the GPU operation invoked, either an application thread or a runtime thread will receive a completion callback providing measurement data. At each GPU API invocation I by an application thread \mathcal{T} , *hpcrun* provides a `user_data` parameter [39], which includes a placeholder node P for the invocation I and \mathcal{T} 's activity channel C_A . The OpenCL or Level Zero runtime will pass `user_data` to the completion callback associated with I . At each completion callback, some thread receives measurement data about a GPU activity \mathcal{A} . Using information from its `user_data` argument, the completion callback correlates \mathcal{A} with placeholder P and then enqueues an operation of (\mathcal{A}, P, C_A) for the monitor thread in its operation channel C_O . The monitor thread enqueues an (\mathcal{A}, P) pair in \mathcal{T} 's activity channel C_A . If the thread receiving the callback enqueued (\mathcal{A}, P) pairs directly into \mathcal{T} 's activity channel C_A , C_A would need to be a multi-producer queue since more than one thread may receive completion callbacks for \mathcal{T} . Our design, which employs a GPU monitor thread created by *hpcrun*, replaces the need for a multi-producer queue with several wait-free single producer queues.

When tracing is enabled, the monitor thread checks the GPU stream id S of each GPU activity and enqueues the activity and its placeholder P into a trace channel for S . One or more tracing threads handle the recording of traces. Each tracing thread handles a set of trace channels by polling each channel periodically and processing its activities. For each activity in a trace channel for stream S , the tracing thread records its timestamp and placeholder in a trace file for S . Depending on the number of application threads used, the number of tracing threads can be adjusted by users to balance tracing efficiency with tool resource utilization.

4.2. Fine-grained performance measurement

On NVIDIA GPUs, HPCToolkit uses PC sampling to monitor both instruction execution and stalls in GPU kernels. On Intel GPUs, HPCToolkit uses Intel's GT-Pin to instrument GPU kernels to collect fine-grain, instruction-level measurements. AMD GPUs currently do not support either instrumentation-based or hardware-based fine-grained measurement.

If PC sampling is used, the monitor thread receives a buffer full of PC samples in a completion callback. Each PC sample for a kernel includes an instruction address, a stall reason, and a count of the times the instruction was observed. The monitor thread enqueues an instruction measurement record into the activity channel of the application thread that launched the kernel. When an application thread receives an instruction measurement record, it creates a node in its

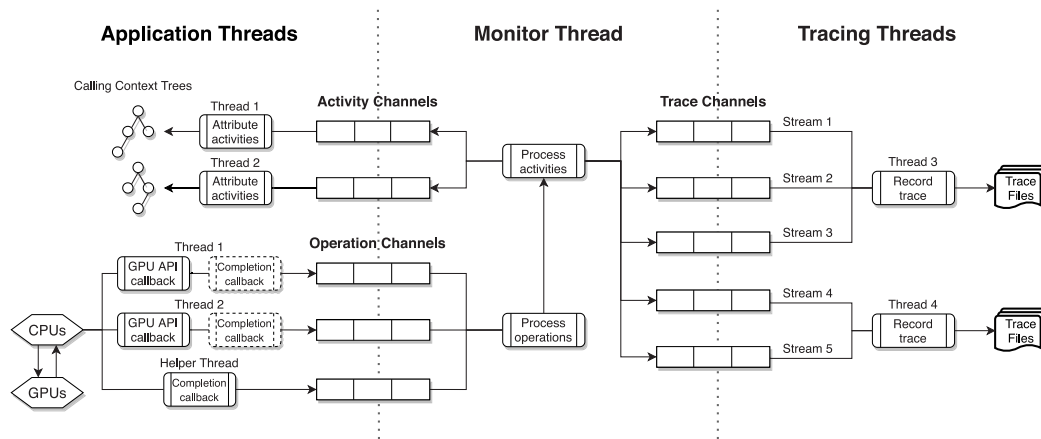


Fig. 2. HPCToolkit's infrastructure for coordinating application threads, monitor thread, and tracing threads.

calling context tree representing the GPU instruction as a child of the placeholder node for the corresponding kernel invocation.

If instrumentation is used, when a GPU binary is loaded, HPCToolkit injects code into each GPU kernel to collect measurements. Measurement data is collected on a GPU and provided to HPCToolkit in a completion callback. On Intel GPUs, HPCToolkit instruments a GPU kernel to count the execution frequency of each basic block. In a completion callback following kernel execution, HPCToolkit iterates over each basic block and propagates its execution count to each instruction in the block. Information about each instruction and its count is sent to the monitor thread in an operation channel. The monitor thread passes the information back to the application thread that launched the kernel using an activity channel. The application thread processes the instruction measurement like a PC sample.

4.3. Measuring performance with hardware counters

HPCToolkit uses hardware performance counters to observe how an application interacts with an accelerated compute node. CPUs and GPUs each provide a collection of programmable hardware counters that can be configured to measure device metrics (e.g., temperature and power), functional unit utilization, memory hierarchy activity, inefficiency, and more.

On CPUs, HPCToolkit uses the Linux `perf_event` interface [40] to configure hardware counters with events and thresholds. HPCToolkit unwinds the call stack to attribute a metric to a call path each time a counter reaches a specified threshold. With appropriately chosen event thresholds, such measurement has low overhead.

On GPUs, HPCToolkit uses the University of Tennessee's PAPI [41] as a vendor-independent interface to measure GPU activity using hardware counters. PAPI supports hardware counter-based measurement on NVIDIA, AMD, and Intel GPUs. At present, the only way tools can associate hardware counter measurements with individual GPU kernels using existing vendor APIs is to serialize kernels and read data from counters before and after kernel execution. Serializing kernels may both slow execution and alter execution behavior.

4.4. Interaction with measurement substrates

While developing HPCToolkit's GPU measurement infrastructure, we encountered a few problems using each vendor's measurement substrate(s). This section describes some of the difficulties encountered and how they were handled.

Each GPU vendor and/or runtime system provides different levels of monitoring support. NVIDIA's CUPTI [35] supports both coarse-grained and fine-grained measurements for CUDA programs. AMD's ROCTracer [36] only supports coarse-grained measurements for HIP

programs. Both of these monitoring frameworks enable a tool to register a callback function that will be invoked at every GPU API invocation. These callbacks can be used to gather information about an invocation, such as its calling context. Intel's GT-Pin enables tools to add instrumentation for fine-grained measurement of GPU kernels; however, neither its OpenCL [37] or Level Zero [38] runtimes provide APIs for collecting coarse-grained metrics. As a result, HPCToolkit wraps Intel's OpenCL and Level Zero APIs using `LD_PRELOAD` to collect custom information in each API wrapper. Wrapping APIs is sensitive to changes in APIs as the runtimes evolve (interfaces in Level Zero have changed over the last few months) and may not provide access to all information of interest, e.g., implicit data movement associated with kernel arguments in OpenCL.

As a GPU program executes, vendor runtime and/or tool APIs typically create helper threads. For example, if PC sampling is used, CUPTI creates a short-lived helper thread each time the application launches a kernel. Thus, in a large-scale execution that launches kernels millions of times, CUPTI will create millions of short-lived threads. Similarly, several components in AMD's software stack create threads, including the HIP runtime, ROCm debug library, and ROCTracer. To reduce monitoring overhead, HPCToolkit wraps `pthread_create` to check if a thread is created by a GPU runtime or its tool library. If yes, HPCToolkit avoids monitoring the thread.

In CUPTI and ROCTracer, a single helper thread in each process handles GPU measurement data using the buffer completion callback. However, for OpenCL and Level Zero, an event completion callback may be asynchronous, as described for OpenCL [39]. Hence, it is *hpcrun's* responsibility to ensure that callbacks gather and report information in a thread-safe fashion. To avoid races reporting data back to an application thread, *hpcrun* first delivers measurement data from the thread that receives the callback, which might be the application thread, to a monitoring thread using a point-to-point operation channel between the threads. The monitoring thread then delivers measurement data back to the proper application thread using a point-to-point activity channel.

While CUPTI and ROCTracer typically order activities within each stream, the order in which GPU activities are reported is undefined for OpenCL [39]. On Power9 CPUs, we have even observed overlapping intervals on a stream using CUPTI. Rather than taking extreme measures to order each stream's activities in *hpcrun*, we simply record each stream into a trace file and note if any activity is added out of order. If so, HPCToolkit sorts the trace stream to correct the order during post-mortem analysis.

4.5. Measuring and computing metrics

As a GPU-accelerated program executes, HPCToolkit collects performance metrics and associates them with heterogeneous calling contexts. HPCToolkit supports several strategies for measuring and computing metrics. A raw CPU or GPU metric for a heterogeneous calling context in an application thread is simply the sum of all measured values of a specific kind associated with that context. For instance, raw metrics for GPU data copies associated with a context include the operation count, total bytes copied, and total copy time.

To facilitate analysis, HPCToolkit also computes two types of *derived* metrics. The first type of derived metrics is computed during post-mortem analysis by HPCToolkit's *hpcprof*. Built-in derived metrics for combining metrics from different thread profiles during post-mortem analysis include *sum*, *min*, *mean*, *max*, *std. deviation*, and *coefficient of variation*. With the exception of *sum*, these metrics can provide insight into imbalances across threads. The second type of derived metrics is computed in HPCToolkit's *hpcviewer* user interface. HPCToolkit uses *hpcviewer* to compute GPU metrics including GPU utilization and GPU theoretical occupancy.

Computing some GPU metrics requires a bit of creativity. For instance, NVIDIA's CUPTI reports static information about a kernel's resource consumption (e.g., registers used) each time a kernel is invoked. To avoid the need for a special mechanism for collecting such metrics, HPCToolkit simply records raw metrics such as the sum of the count of registers used over all kernel invocations in a particular calling context and the count of kernel invocations in that context. After summing these raw metrics over threads and MPI ranks, *hpcviewer* computes the ratio of these two values to recover the number of registers used.

4.6. Sparse representation of metrics

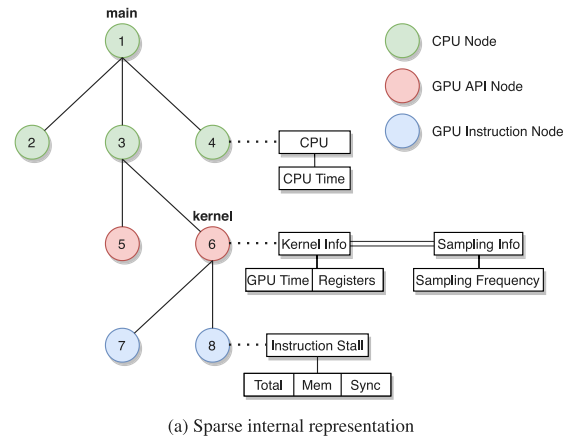
hpcrun maintains a Calling Context Tree (CCT) for each CPU thread or GPU stream it measures. In a CCT, each node represents the address of a machine instruction in a CPU or GPU binary as a (load module, offset) pair. When a CCT node is allocated, it receives a companion metrics array to store associated performance metrics. In HPCToolkit, well over 100 metrics can be measured; some for CPUs and some for GPUs. When measuring the performance of GPU-accelerated programs, many CCT nodes have CPU metrics only; all of their many GPU metrics are zero. Storing zero values for all unused metrics at a CCT node would waste considerable memory.

To reduce storage during measurement, *hpcrun* partitions metrics into *kinds*, such as *GPU kernel info* kind, *GPU instruction stall* kind, and *CPU time* kind. Each CCT node is associated with a metric kind list, and each metric kind represents an array of one or more metrics. For example, when measuring an execution with PC sampling, the CCT node representing a GPU kernel has *GPU kernel info* kind and *GPU instruction sampling info* kind. The kernel kind includes kernel running time, register usage, and shared memory usage, among others.

Fig. 3(a) illustrates the sparse representation of metrics associated with CCT nodes as *hpcrun* measures the performance of a GPU-accelerated application. In the figure, each CCT node is categorized as a CPU node, a GPU API node, or GPU instruction node. Each type of CCT node is associated with different metric kinds.

In addition to representing metrics sparsely in memory, *hpcrun* also writes profiles to the file system using a sparse format to save space. The output format of each profile file has the following sections.

- A *Load Modules* section that contains all the libraries loaded in execution.
- A *CCT* section that depicts the structure of a CCT, including each node's module id, offset, and parent fields.
- A *Metrics* section that contains the index and name of each performance metric, as well as some properties.



(a) Sparse internal representation

```

Load Modules
module 2: {path: /home/code/add.gpubin}
...
CCT
node 6: {load module: 1, pc: 0x100, parent: 3}
node 7: {load module: 2, pc: 0x80, parent: 5}
node 8: {load module: 2, pc: 0x70, parent: 6}
...
Metrics
metric 0: {name: Total Stall, properties}
metric 1: {name: Mem Stall, properties}
metric 2: {name: Sync Stall, properties}
...
Metric Values
index 5: {value: 2, metric id: 0}
index 6: {value: 1, metric id: 1}
index 7: {value: 1, metric id: 2}
...
CCT Metric Values
node 7: {metric value index: [5, 8]}
node 8: {metric value index: [8, 10]}
...

```

(b) Sparse output format

Fig. 3. *hpcrun*'s sparse representation of a CCT and its metrics in memory and on the disk.

- A *Metric Values* section and a *CCT Metric Values* section that indicates the metric values associated with each CCT node.

To generate the *Metric Values*, *hpcrun* iterates through the metrics kind list of each CCT node, counts the number of non-zero metrics N , and records their values. In the *CCT Metric Values* section, a CCT node with an index range $[I, N)$ indicates that it has metrics in the *Metric Value* section at positions from I to $I + N - 1$. Profiles produced by *hpcrun* employ this scheme to represent only non-zero metrics.

Fig. 3(b) illustrates the sparse representation of metrics in *hpcrun*'s output files. In the *CCT Metric Values* section of the figure, node 7 has three metrics—metric index 5, metric index 6, and metric index 7. We locate metric index 5's value (2) by at index (5) in the *CCT Metric Values* section. Further, we can retrieve the metrics's name (Total Stall) by looking up its *metric id* (0) in the *Metrics* section.

4.7. Minimizing measurement errors

There are many ways that HPCToolkit can measure the performance of a GPU-accelerated application: sampling on the CPU using timers or hardware counters, reading GPU hardware counters before and after GPU operations, profiling and/or tracing timings of GPU operations, and using PC sampling or instrumentation to collect fine-grained measurements of GPU kernels. Performing all of these measurements in a single execution is unwise. For instance, fine-grained instrumentation of GPU kernels can distort GPU profiles and traces recorded in the same execution. To minimize the distortion in measurements, it is best to collect each kind of measurements in a separate run, if one can afford

multiple runs. HPCToolkit’s post-mortem analysis can combine performance measurements from multiple runs to produce a comprehensive representation of an application’s performance.

5. Program structure recovery

Relating performance metrics associated with a node in a calling context tree to source code requires understanding the mapping from machine instructions to source lines, loops, functions, and inlined code. HPCToolkit’s *hpcstruct* analyzes both CPU and GPU binaries to determine this mapping. This strategy works for any statically-compiled language, including C, C++, and Fortran.

Binaries for GPU-accelerated applications are rather unusual. A host application typically has GPU binaries embedded within. Furthermore, NVIDIA and Intel GPU binaries are unlike any CPU binaries we have ever seen. While each of the GPU vendors rely in part on ELF representations for GPU code, their formats for GPU binaries are quite different.

NVIDIA’s GPU binaries sometimes have device functions embedded inside a global GPU function. Each function is in a separate text segment and all functions start with offset 0 in its symbol table. To aid measurement and analysis, we relocate each function to a unique address—its offset in the binary’s section table. We split any overlapping GPU functions into disjoint address ranges. Then we use NVIDIA’s *nvdiasm* binary tool to analyze GPU machine code and dump a control flow graph for each function in a binary. Because *nvdiasm* renders merged basic blocks in control flow graphs, we split superblocks into basic blocks to fit analysis.

Intel’s GPU binaries are embedded in a “fat” binary in which each section contains information about a single GPU kernel. At runtime, we dump the fat binary to the file system. During post-mortem analysis, *hpcstruct* reads a fat binary, iterates through all kernels, and processes the ELF binary for each kernel separately. It uses Intel’s IGA API to decode each kernel binary to obtain its control flow graph.

AMD’s GPU binaries are also embedded in a “fat” binary. The locations of AMD’s GPU binaries inside the fat binary are specified by Uniform Resource Identifiers (URIs) reported by ROCm debug API and ROCTracer. AMD’s GPU binary uses a conventional ELF representation. We use the Dyninst [42] binary analysis framework to decode AMD GPU instructions and construct control flow graphs. The Dyninst team has developed preliminary support for analyzing AMD GPU binaries, including support for decoding scalar instructions and resolving direct control flow; support for decoding vector instructions and resolving indirect control flow is a work in progress.

Control flow graphs extracted from Intel and NVIDIA GPU binaries are converted to a uniform representation and injected into Dyninst, which analyzes loop nests they contain. We classify GPU instructions by their operation types, instruction length, and kind of memory they access. As Dyninst can analyze AMD GPU binaries, the aforementioned analyses are performed directly in Dyninst.

hpcstruct uses Dyninst to read DWARF sections in a GPU binary to attribute instructions to source lines and recover inlined functions. For AMD, Intel and NVIDIA GPU binaries, it is worth noting their information about inlined code is not perfect. None of the GPU binaries provides information about the call site of an inlined function, even when they are compiled with debug information (e.g., `-lineinfo`) and optimization (e.g., `-O3`).

If *hpcrun* records many GPU binaries at for a complex execution, *hpcstruct* analyzes GPU binaries in parallel.

6. Scalable performance analysis

This section describes our approach to improve the scalability of our performance analysis, to handle the additional data gathered from GPU-accelerated applications running at the extreme scales supported by forthcoming systems. Section 6.1 describes our new approach for

scalable performance analysis of measurement data from extreme-scale executions. Section 6.2 presents the structure of the new sparse data format HPCToolkit uses to address the sparsity in measurements and analysis results for heterogeneous applications. Section 6.3 describes an algorithm to reconstruct approximate GPU calling context trees for complex GPU kernels.

6.1. Streaming aggregation

Forthcoming exascale compute platforms pose significant challenges for performance tools. Performance issues in applications or system software may only become apparent during executions at very large scales, requiring performance tools to support measurement of applications executing on systems with tens of thousands of compute nodes equipped with multicore processors and GPUs. The aggregation of measurement data into statistics suitable for human inspection becomes increasingly expensive as the scale of executions increases, requiring tools to employ well-designed algorithms for data aggregation with good scaling to large inputs.

To address these issues, we developed a novel algorithm for performance analysis that exploits both process-based (via MPI ranks) and thread-based parallelism. Exploiting thread-level parallelism reduces communication overheads and the memory footprint during analysis. The key to our algorithm is to process profiles of individual application threads or GPU streams as concurrent parallel tasks, sharing a limited set of data structures between profiles containing the resulting aggregated values. In effect, we “stream” the input profile data in parallel to the appropriate destination, retaining only values required for the final “aggregation” of all inputs, hence the name. Our algorithm then can be described as a sequence of operations applied in parallel on the profile inputs, as follows:

Input Acquisition: Before any other processing is performed, the set of input profiles is acquired and initial offsets are prepared to facilitate later data reads. A profile contains a tree of call paths and measurements attributed to nodes in call paths as described in Section 4.6, and may additionally include an execution trace referencing individual call paths. These profiles are then distributed evenly across the available ranks, where they continue to be processed in parallel using a dynamic scheduling algorithm.

Call Path Profile Unification: Once the profiles have been distributed across the available ranks, we unify the tree of call paths from each profile into a single tree available at the root of a collective reduction operation. This can be done optimally utilizing a reduction tree of the same arity as the number of threads available in each rank t , as the rank has sufficient capacity to handle data from t children at once.

Calling Context Expansion: As call path nodes are received in parallel at the root rank, both from allocated profiles and child ranks, calling context nodes are created based on program structure information for each instruction in a call path. Basic program structure information (lines and inlined code) can be generated from line map and inlining information recorded in a binary by a compiler. If information about loops is of interest, then program structure files generated by *hpcstruct*, as described in Section 5, can be imported. Once the unified calling context tree is fully constructed on the root rank, the complete conversion mapping from call paths nodes to calling contexts is broadcast to all other ranks, which use it to construct compatible subsets of the calling context tree represented by their contributions. Compared to our previous implementation of post-mortem analysis using MPI-everywhere parallelism, our new implementation has a smaller memory footprint when using the same number of cores because multiple threads in each MPI rank share a single copy of the unified calling context tree.

Statistic Generation: At this stage, the measurements present within each profile are read in parallel, propagating values up the calling context tree to generate per-thread metrics for every referenced calling context. On each rank these metrics are then composed into a set of accumulators present on each calling context, which are used to

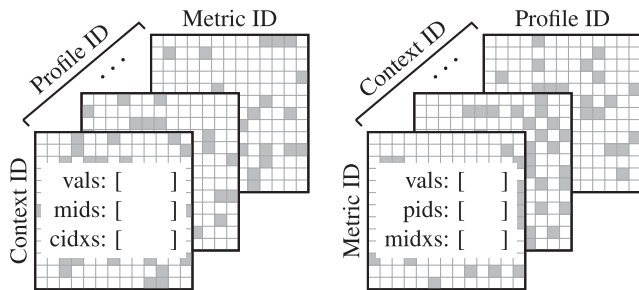


Fig. 4. Visualization of Profile Major Sparse format (left) and CCT Major Sparse format (right). Each plane uses a modified CSR. Gray cells represent non-zeros on each plane.

generate metric statistics global to all threads in the application's execution. At this point the per-thread metrics can be written to the output database immediately, while the accumulators are aggregated by a second reduction operation similar to the first.

Trace and Final Outputs: If a profile has a companion execution trace, at this stage its trace sequence is converted one sample at a time to reference calling contexts instead of the call paths present at runtime and then written directly into the output database. After all profiles have been processed, the final unified calling context tree and global statistics are written to the database by the root rank.

6.2. Sparse representations of analysis results

Inspired by *Compressed Sparse Row* (CSR) format [43] for storing a sparse matrix, we designed and implemented a pair of compact sparse formats to store HPCToolkit's performance metrics for threads and GPU streams. These sparse formats can save storage and ensure efficient access of values needed for presentation and inspection.

Fig. 4 shows a high-level visualization of our design: *Profile Major Sparse* (PMS) format and *CCT Major Sparse* (CMS) format. As mentioned earlier, one of our goals is to access the data efficiently, which means accessing the data both within one profile and between profiles. Therefore, we designed PMS and CMS to compare performance within and across threads or GPU streams, respectively, so we only need to open one file for all comparisons as long as they belong to the same type.

Unlike CSR, we need to locate a value with three indices: a metric ID, a context ID, and a profile ID. If we consider the matrix represented by CSR a sparse plane, then our formats represent sparse cubes. The filled cells represent non-zero values and the three arrays show how to represent each plane. PMS consists of a vector of profile offsets, one per profile, that indicate the start of each plane of (profile, context, metric) triples. CMS consists of a vector of context offsets, one per calling context, that indicate the start of each plane of (context, metric, profile) triples. For brevity, in the rest of Section 6.2, we only discuss the design details of CMS.

As Fig. 4, each plane can be seen as a CSR. *row* in CSR is a dense array, however, after some experiments, we found many repeating values in *midxs* for CMS, which means there are no non-zero values for these metrics for the specific context, and we call them *empty* metrics. As discussed previously, we expect many empty metrics in calling context trees for heterogeneous programs. To exploit this sparsity, we make *midxs* a sparse array: each entry is a pair of metric ID and this metric's starting index (*midx*) in *pids* and *vals*.

With x non-zeros, m non-empty metrics in an average profile, and c contexts, then CMS uses $\mathcal{O}(c \times (2x + m + 1))$ space. This saves space when $(2x + m + 1) < MP$, where M is the number of metrics and P is the number of profiles. $(2x + m + 1)$ can even be slightly larger than MP since the dense version uses a consistent number of bytes for each data, but CMS can use fewer for some data whenever appropriate, for instance, metric ID. To access a specific value, it takes constant time to locate a context plane and $\mathcal{O}(\log m)$ time to binary search for the metric index

midx. To compare performance between profiles for a specific context and metric, go to *pids[midx]* and then compare all of the values until the start of the next metric. To access one specific value, let p be the number of profiles associated with the context and the metric, then it takes $\mathcal{O}(\log m + \log p)$ time in total. Therefore, CMS can save storage and ensure efficient access to values.

To construct our sparse formats using both distributed-memory and multithreaded parallelism, *hpcprof-mpi* uses MPI for communication between ranks and shared memory for communication between OpenMP threads. Both CMS and PMS use the same general idea to generate the file: find the right offset of each plane, collect the related bytes, write the bytes to the offset, and record the location in the offsets section. *hpcprof-mpi* uses *exscan* operations within and between ranks to find the right offsets, and then all threads can mostly finish the remaining three steps concurrently without communication.

For PMS, *hpcprof-mpi* splits work within and between ranks based on the number of profiles to ensure load balance. However, since different calling contexts may have huge differences in the number of associated metrics, for CMS, *hpcprof-mpi* partitions the work within and between ranks based on the number of non-zeros for contexts to ensure load balance. As a performance tool for exascale applications, *hpcprof-mpi* does not assume it has enough memory to store all contexts or profiles at the same time. For that reason, it processes the data in an out-of-core fashion. For example, for CMS, *hpcprof-mpi* has a pre-set maximum memory that it can use for one round, and it processes the data in multiple rounds if necessary.

6.3. GPU calling context tree

To analyze the performance of complex GPU kernels that call device functions, it is useful to organize performance data into GPU calling context trees. Using binary instrumentation to collect instruction traces, we could reconstruct a GPU calling context tree on the CPU as an application executes. However, this method would have high overhead with frequent communication between CPUs and GPUs to copy traces [28].

To address this, we designed a method to reconstruct approximate GPU calling context trees offline from fine-grain instruction-level measurements gathered using instrumentation or PC samples. This approach can be applied in any environment where collecting precise call stacks is expensive.

Our method reconstructs an approximate GPU calling context tree for each GPU kernel invocation with the following four steps.

1. Construct a GPU static call graph based on function symbols and call instructions. Initialize weights on call edges using exact counts of call instructions or call instruction sample counts.
2. *For call graphs based on samples:* if a function has samples and none of its incoming call edges has a non-zero weight, we assign each of its incoming call edges a weight of one; we repeat this propagation through callers until at least one call edge of a function has a non-zero weight.
3. Identify strongly connected components (SCCs) using Tarjan's algorithm [44]. In the call graph, add an SCC node to represent the set of nodes in each SCC. Link external calls to functions inside an SCC to the SCC node and remove call edges between functions in the SCC.
4. Finally, build a calling context tree by splitting the call graph. Like Gprof [45], assume that every invocation of a function takes the same time. Apportion the number of samples or instructions in each function among its call sites using ratios of calls from each call site to the total number of calls from all call sites.

Fig. 5 shows the reconstruction process for a small synthetic example. We first construct a static call graph in Step 1. In Step 2, we assign the edge between A to B one call sample because B does not have a sampled call site. If executed instructions are collected, each edge representing a call that executed should have a non-zero weight. In Step

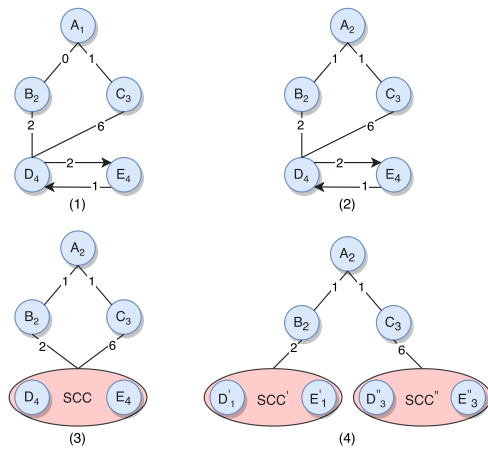


Fig. 5. An example of splitting a call graph into a calling context tree. A–D denote functions. Each subscript denotes the number of samples associated with the function. Each edge is annotated with its number of call samples.

3, we identify an SCC that contains D and E. Finally, we apportion the number of samples of the SCC using ratios of weights of calls from each call site to the total weight of calls from all call sites.

7. User interfaces

hpcviewer is HPCToolkit’s graphical user interface (GUI), built on top of the Eclipse Rich Client Platform (RCP), which enables us to support multiple platforms including Windows, macOS, Linux x86/64, and Linux ppcle64. We plan to support Linux Arm in the near future; at present, we are awaiting a forthcoming release of Eclipse for Arm. hpcviewer is designed to support huge application performance databases by using sparse data representations, loading GUI components lazily, and employing multiple threads to read and load data dynamically, as needed.

The database generated by hpcprof consists of 4 dimensions: profile, time, context, and metric. We employ the term profile to include any logical threads (such as OpenMP, pthread and C++ threads), and also MPI processes and GPU streams. The time dimension represents the

timeline of the program’s execution, context depicts the path in calling-context tree, and metric constitutes program measurements performed by hpcrun such as cycles, number of instructions, stall percentages and ratio of idleness.

To simplify performance data visualization, hpcviewer displays only two dimensions at a time: the profile viewer displays pairs of (context, metric) or (profile, metric) dimensions; and the trace viewer visualizes the behavior of threads or streams over time. Section 7.1 describes HPCToolkit’s profile viewer and Section 7.2 describes its trace viewer.

7.1. Profile viewer

HPCToolkit’s profile viewer provides a code-centric user interface for interactive examination of performance databases. As shown in Fig. 6, it displays pairs of (context, metric) dimensions, which enables users to view context-sensitive performance metrics correlated to program structure and mapped to a program’s source code, if available. It can also present an arbitrary collection of performance metrics gathered during one or more runs or compute derived metrics.

Measurements can be analyzed in a variety of ways: top-down in a calling context tree, which associates costs with the full calling context in which they are incurred; bottom-up in a view that apportions costs associated with a function to each of the contexts in which the function is called; and in a flat view that aggregates all costs associated with a function independent of calling context. This multiplicity of code-centric perspectives is essential to understanding a program’s performance for tuning under various circumstances. hpcviewer also supports a thread-centric perspective, which enables one to see how a performance metric for a calling context differs across threads. The viewer can plot a graph of metric values associated with the selected node in CCT for all processes or threads. This functionality enables users to display pairs of (profile, metric) dimensions.

To collect all necessary metrics, GPU performance tools often employ multiple runs and merge all metrics together. For example, Nsight-compute runs nine passes to collect all of its default metrics for a small GPU kernel; this approach is infeasible for a large-scale application. As mentioned in Section 4.5, hpcviewer lets users compute derived metrics based on measurements gathered using a single pass to identify optimization opportunities. A derived metric is a spreadsheet-like formula composed from existing metrics, operators, functions, and numerical constants. For example, for each GPU kernel, based on the number of total PC samples (S) and stalled PC samples (S_{stall}), we can estimate the Warp Issue Rate (W) of schedulers as $W = \frac{S - S_{stall}}{S}$.

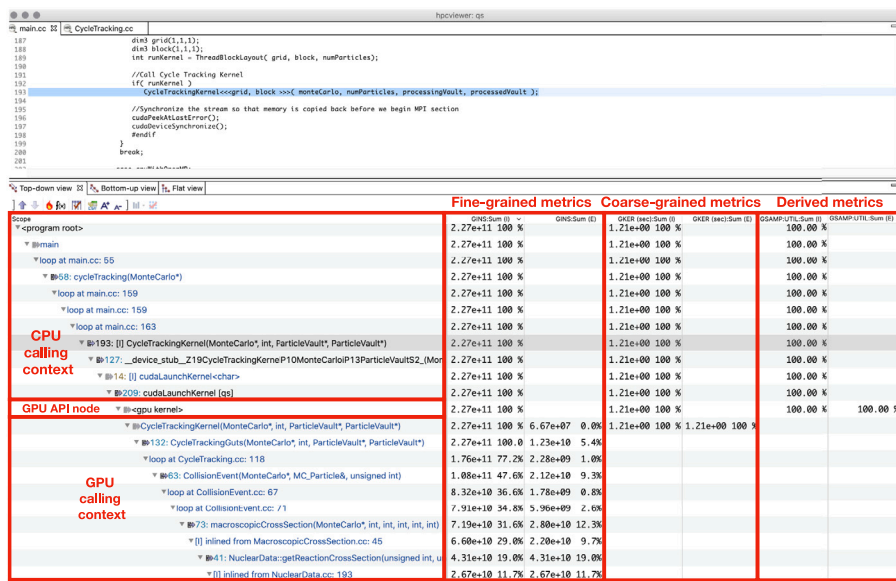


Fig. 6. Performance measurements of an optimized, GPU-accelerated execution of LLNL’s Quicksilver proxy application on a Power9 and an NVIDIA GPU. Fine-grained measurements on the GPU were collected using PC sampling.

Table 1
Experimental platforms.

Node	Hardware	Software
OLCF Summit	2 × IBM POWER9 CPUs 6 × NVIDIA V100 GPUs	GCC-6.4.0 CUDA-11.0.3
Argonne JLSE Iris	1 × Intel E3-1585Lv5 CPU 1 × Intel Iris Pro P580 GPU	GCC-8.3.1 DPC++, oneAPI beta 10
Local AMD	2 × AMD EPYC 7402 CPUs 2 × AMD MI50 GPUs	GCC-8.3.1 Rocm-3.8

7.2. Trace viewer

HPCToolkit’s trace viewer provides a time-centric user interface for the interactive examination of a sample-based time series (hereafter referred to as a trace) view of program execution. It is designed to interactively present traces of large-scale execution across both CPUs and GPUs, relating activity to both hardware contexts (e.g., nodes, GPUs, cores) and software contexts (e.g., MPI ranks, threads, GPU contexts, and streams).

As shown in Fig. 9, the trace viewer’s main pane shows (profile, time) dimensions, for each available call-stack depth. By changing call-stack depth, a user can change the granularity of trace lines, and gain insight into an execution at different levels of abstraction. For the routine pointed by the cursor, functions from the call stack are listed on the right to the main pane. Each routine is uniquely identified with a specific color, while idle activity is assigned the color white.

The trace viewer’s *Statistics* and *GPU Idleness Blame* tabs analyze the information in traces and offer some high-level characterizations of what the traces show. The *Statistics* tab calculates the percentage of the area occupied by each routine in the main pane and lists routines in descending order according to their percentage of the area. The *GPU Idleness Blame* tab employs blame analysis in an attempt to help application developers understand the causes of GPU idleness. To do so, it identifies times when all GPU streams are idle and at least one CPU thread is active. In such cases, it partitions the cost of GPU idleness among routines being executed by active CPU threads. The *GPU Idleness Blame* tab then presents normalized blame associated with each CPU function in sorted in descending order. CPU routines associated with high GPU idleness may be candidates for optimization.

8. Case studies

We tested HPCToolkit’s support for analyzing GPU-accelerated applications on the Summit supercomputer with NVIDIA GPUs, the JLSE Iris testbed which consists of nodes equipped with Intel Skylake processors that have integrated GPUs, and a local machine with AMD GPUs. The hardware and software specification of the three platforms are shown in Table 1. We evaluated HPCToolkit with three HPC applications described below:

- Quicksilver [46] is an ECP proxy application that solves a simplified dynamic Monte Carlo particle transport problem, representing some elements of the workload of LLNL’s Mercury [47] radiation transport code.
- PeleC [48] is an application using adaptive-mesh compressible hydrodynamics for reacting flows. We studied its Premixed Flame (PMF) and Taylor–Green Vortex (TG) examples.
- Nyx [49] is a massively parallel code computes compressible hydrodynamic equations on a grid with particles of dark matter. We ran Nyx using up to 128 GPUs on Summit.

All three applications were compiled with `-O3`. In the remaining section, we describe insights based on the analysis result of HPCToolkit and its overhead on a single GPU and multi-GPU execution.

8.1. Measurement overhead

On NVIDIA GPUs, HPCToolkit and nvprof introduce 2.24× and 2.20× profiling overhead for PeleC’s TG example. To trace Nyx on 128 ranks and 640 GPU streams, HPCToolkit and nvprof introduce 1.85× and 1.42× overhead. It is worth noting that unlike nvprof, HPCToolkit collects CPU call stack information for every GPU API invocation. On Intel GPUs, HPCToolkit’s overhead for PeleC’s TG application with coarse-grained and fine-grained measurements is 1.81× and 2.23× respectively. AMD’s ROCm software stack is evolving rapidly, and as a result, we have not assessed measurement overhead. We have not yet invested much effort in analyzing and tuning HPCToolkit’s measurement overhead for any of the GPU-accelerated platforms.

8.2. Analysis overhead

We compared the size of measurement data and analysis results for a GPU-accelerated version of Nyx using our original dense formats with that of our new sparse formats. The data in sparse formats is consistently much smaller than that using the original dense formats. The size of the measurement results in the sparse format is 167.1 MB, which is 22× smaller than the dense format. The size of the analysis results in sparse format is 153.3 MB, which is 3701× smaller than the dense format.

To measure analysis overhead, we used *hpcprof-mpi* to aggregate and analyze measurements of a GPU-accelerated execution of LAMMPS [50] on 167 nodes and 1002 GPUs. Using thread-level parallelism and streaming aggregation, *hpcprof-mpi* analyzed the 85 GB of measurement data in 91 s on 48 nodes of Summit using 48 MPI ranks with 42 threads per rank (one thread per Power9 core). This is over 3.6× faster than using the original *hpcprof-mpi*, which employs only inter-process parallelism, to analyze the data using the same resources.

8.3. Quicksilver

Fig. 6 shows the profile view for Quicksilver, an ECP proxy application for LLNL’s Mercury radiation transport code. The bottom left pane shows a detailed heterogeneous calling context that spans both CPU and GPU. The GPU calling context inside the `CycleTrackingKernel` kernel is reconstructed from flat PC samples using the algorithm described in Section 6.3. The bottom right pane shows two measured metrics (GPU instructions and GPU kernel time), and a derived metric (GPU utilization). Note that the code was compiled with `-O3`. Although many device functions are inlined, others are not due to function size or register limitations. With HPCToolkit’s approximate GPU calling context tree reconstruction, an application developer can readily understand where GPU code is costly and/or inefficient. (Understanding where a GPU code is inefficient requires fine-grained stall metrics, which were collected for Quicksilver although they are not shown in Fig. 6.)

8.4. PeleC

We studied PeleC’s PMF problem on NVIDIA’s GPUs only because there are not yet stable HIP or DPC++ implementations. We ran PeleC’s TG problem on the three platforms shown in Table 1 and compared the performance characteristics to understand relative performance of different GPUs and how efficiently the software maps to each.

	DIFF
1660: amrex::AmrLevel::FillPatch(amrex::AmrLevel&, amrex::MultiFab&, int, double, int, int, int)	1.43e+04
2209: [] FillPatchIterator	1.08e+04
712: amrex::FillPatchIterator::Initialize(int, double, int, int, int)	1.08e+04
998: [] amrex::FabArray<amrex::FArrayBox>::setDomainBndry<amrex::FArrayBox, C	3.59e+03
Case 1 1474: amrex::FabArray<amrex::FArrayBox>::setDomainBndry<amrex::FArrayBox,	3.59e+03
1496: amrex::MFilter::~MFilter()	3.59e+03
993: amrex::MultiFab::define(amrex::BoxArray const&, amrex::DistributionMapping	3.59e+03
589: amrex::MultiFab::define(amrex::BoxArray const&, amrex::DistributionMapp	3.59e+03
Case 2 601: amrex::FabArray<amrex::FArrayBox>::define(amrex::BoxArray const&, ar	3.59e+03
1269: [] synchronize	3.59e+03
Case 3 loop at AMReX_AmrLevel.cpp: 1003	3.59e+03
2211: [] ~FillPatchIterator	3.59e+03

Fig. 7. An unnecessary synchronization calling context in PeleC.

8.4.1. NVIDIA GPU

We profiled PeleC's PMF example with its default input using PC sampling. Using HPCToolkit's heterogeneous calling context, we were able to identify the hot kernel quickly. We noted that the GPU utilization for kernel `pc_exp1_reactions` was only 2.5% on average, indicating low SM utilization. By reducing the number of threads per block from 256 to 128, we increased the number of blocks for this kernel from 16 to 32 and improved its performance by 1.14×.

Next, we profiled PeleC's TG example with its default input. `hpcviewer` shows that the time spent on GPU synchronization is close to GPU kernel execution time. We found an unusual phenomenon: in some cases, the number of GPU synchronization API invocations exceeds the number of GPU kernel launches. We computed a derived metric: $diff = sync_count - kernel_count$ in `hpcviewer` to find where synchronizations are unnecessary. Fig. 7 shows three calling contexts where $diff$ is high. In the first calling context, no kernel launch occurs in the loop while there are synchronization invocations. We found that an `MFilter` object is created for the loop, and the synchronization is called in the object's destructor to synchronize computations in the loop. We optimized the code by not invoking synchronization if no computation is performed in the loop. In the second calling context, the synchronization is needed only when there are multiple MPI ranks. Thus, we can conditionally invoke synchronization by checking the number of MPI ranks. In the third calling context, the synchronization is redundant because a copy function immediately before always synchronizes all GPU streams. These three code changes reduced the number of synchronization invocations by 38% and improved end-to-end performance by 1.05×.

8.4.2. Intel and AMD GPUs

We compiled the DPC++ code of PeleC's TG example and ran it on an Intel Gen9 GPU using Intel's OpenCL backend. We captured four coarse-grained metrics, including kernel execution time, memory transfer time, memory transfer bytes, and memory allocation bytes. For fine-grained measurement, HPCToolkit collected GPU instruction counts using instrumentation added to GPU kernels using Intel's GT-Pin library.

The profile view in Fig. 8 enables one to quickly identify the most costly GPU kernels with respect to GPU instruction count: `pc_compute_hyp_mol_flux` and `pc_compute_diffusion_flux`, which execute 37.6% and 29.6% of the GPU instructions respectively. Measurement and attribution of instruction counts within each kernel revealed an interesting detail: roughly one third of the instructions in each kernel are for array index calculation. It is worth noting that the fraction of kernel execution time measured for this kernel while collecting instruction level measurements may not be accurate since kernel instrumentation inflates kernel execution time.

In addition, we noted 28.8E6 bytes of GPU memory was allocated, but no memory transfers were measured. In this case, the memory transfers are implicit; thus, wrapping public OpenCL APIs may miss observing internal memory transfers performed by a vendor's implementation. To observe implicit memory transfers, we either need either

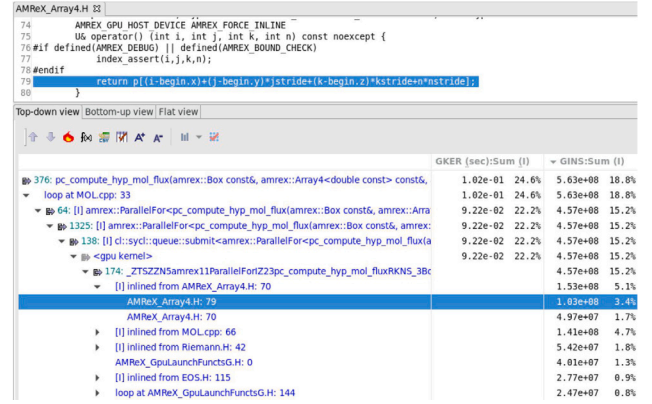


Fig. 8. Fine-grained measurements of PeleC's TG benchmark collected with binary instrumentation on an Intel Gen9 GPU. PeleC is implemented using DPC++ and executing GPU atop Intel's OpenCL runtime.

vendors to use only public APIs for data transfers or we need a better OpenCL tool API.

In experiments with PeleC on AMD GPUs, HPCToolkit can collect heterogeneous call stacks that attribute costs to GPU operations, including kernel launches and data copies. Without support for fine-grained measurements, we are unable to measure or analyze performance inside GPU kernels.

8.5. Nyx

Fig. 9 shows a trace view of Nyx executing on Summit using 640 streams across 128 GPUs. The trace view shows that this execution consists of five phases. By inspecting the call stacks in the trace view, we can determine what the phases do, e.g., initializing particle and dark matter data, performing hydrodynamic calculations, calculating gravity, etcetera.

For each phase, we employed blame analysis to understand why GPUs are idle. In the first phase, 58.01% of idleness is caused by a call to `cuCtxSynchronize`, which synchronizes all streams on the GPU. Because only a single stream is used in this phase and a synchronized memory copy always follows the call to `cuCtxSynchronize`, we can safely remove the synchronization call and reduce the running time by 0.6 s.

In the second phase, we easily identified that the major cause of idleness is JIT compilation at runtime. By providing the specific GPU architecture flag to the compiler and recompiling the program, we reduced the running time by 0.2 s with this single optimization.

In the following three phases, we identified that idleness is caused by calls to `MPI_Waitall`, which suggests that there may be opportunities for improving performance by optimizing communication.

The insights provided by blame shifting analysis reduced the GPU running time of Nyx from 10.6 s to 9.8 s, achieving a 1.08× speedup with 640 GPU streams. While the problems and improvements that we describe here are small, the important part is that HPCToolkit provides developers insights to identify even small problems.

9. Conclusions and future work

Based on our experiences extending Rice University's HPCToolkit performance tools for GPU-accelerated platforms, we offer some observations about our perceptions of the hardware and software challenges faced by performance tools.

Measurement. To understand in detail the performance of complex GPU kernels, hardware support for measuring GPU performance is essential.

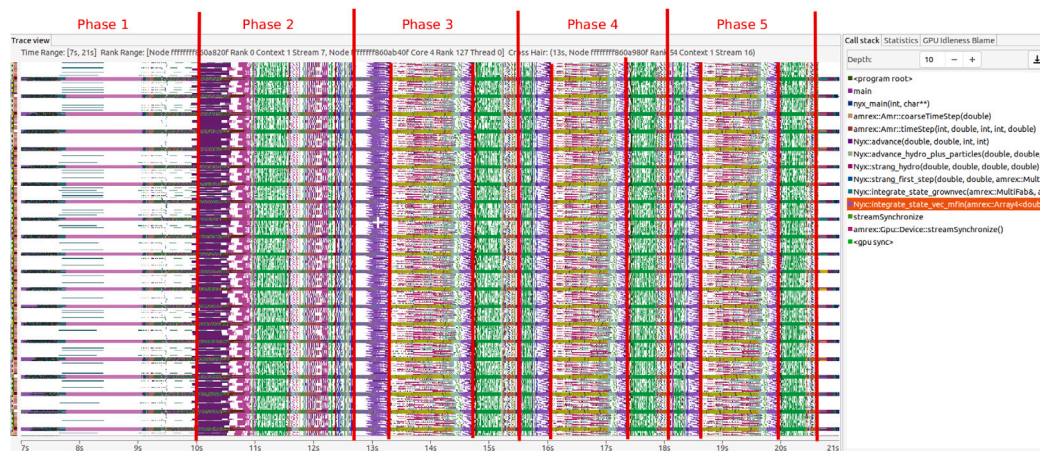


Fig. 9. Nyx' trace view running across 22 nodes using 640 GPU streams and 128 processes.

Today, NVIDIA is the only vendor whose GPUs provide support for fine-grained measurement using PC sampling. Without hardware support for fine-grained performance measurement, application developers must use the “guess and check” strategy for performance optimization, which significantly complicates the process of analysis and optimization of large-scale applications.

Today, NVIDIA's PC sampling measurement serializes execution of GPU kernels to simplify performance measurement. This approach both slows the execution of applications when under observation by performance tools and distorts the performance of applications available for tools to observe.

Another problem tools face is that not all GPU activities are readily observable. As described in 8.4.2, there is no portable mechanism that a tool can use to monitor implicit buffer transfers in OpenCL. In addition, NVIDIA purposefully avoids having CUPTI monitor implicit synchronization invoked by primitives such as `cudaMemcpy` as well as driver calls and other operations called by NVIDIA libraries such as `cuBLAS` [22]. The inability to monitor all data transfers and synchronization gives tools an incomplete picture of an application's behavior. For tools to be maximally effective, all GPU operations must be observable.

Attribution. To help application developers understand the performance of optimized GPU kernels that are generated from or employ inlined templates or functions, tools need compilers to generate high-quality DWARF information that describes the provenance of every machine instruction in each GPU kernel. While line mappings generated by today's compilers relate most machine instructions back to an associated source line, they often fail to relate machine instructions back to any inlined call chain that caused it to be included. In our experience, even the latest compilers for AMD, Intel, or NVIDIA GPUs do not generate sufficiently precise DWARF for attributing the performance of each GPU instruction to its full calling context in the presence of inlined functions and templates.

Analysis. Measuring and attributing the performance of GPU kernels with hardware support for fine-grained measurement is only the first step towards understanding the performance of complex applications. Within GPU kernels, understanding the causes of performance problems and opportunities for improvement requires understanding the interplay between a wide range of factors, including block-level parallelism, thread-level parallelism within blocks, data sizes, data alignment and how it maps into the memory hierarchy, details of generated code (e.g., resource consumption, use of type conversions, use of special functional units, and instruction sequencing), and reasons for stalls that result in uncovered latency.

While our work developing a framework for scalable analysis of performance data for extreme-scale executions is an important building

block, it is clear that our tool could use further development to better exploit the capabilities of parallel file systems found on supercomputers. Also, it would be beneficial if we could exploit accelerators in our analysis.

Presentation. The most common performance problem we have observed in emerging ECP applications and libraries is that GPUs are idle too much of the time. Understanding the root causes of idleness in large-scale applications composed of many layers of software is difficult for application developers using existing user interfaces. Our initial prototype support for automated analysis of GPU traces offers some help in this regard, but our early experiences have shown that attributing performance at the proper level to generate insight will require new design insights and implementation.

Declaration of competing interest

One or more of the authors of this paper have disclosed potential or pertinent conflicts of interest, which may include receipt of payment, either direct or indirect, institutional support, or association with an entity in the biomedical field which may be perceived to have potential conflict of interest with this work. For full disclosure statements refer to <https://doi.org/10.1016/j.parco.2021.102837>.

The work was funded in part by Advanced Micro Devices, though software support for their GPUs is only briefly mentioned in the paper.

We are expecting to receive some modest funding from Intel to support continued work with Intel GPUs, but funds have not been received yet.

Acknowledgments

This research was supported in part by the Exascale Computing Project (17-SC-20-SC)—a collaborative effort of the U.S. Department of Energy Office of Science and the National Nuclear Security Administration, United States Argonne National Laboratory, United States (DE-AC02-06CH11357), DOE Tri-labs (Lawrence Livermore National Laboratory (LLNL), United States Subcontract B639429), and Advanced Micro Devices, Inc., United States

This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725, and Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357.

References

- [1] E. Strohmaier, J. Dongarra, H. Simon, M. Meuer, June 2020 | Top500 super-computer sites, 2020, [Accessed Oct. 4, 2020], <https://www.top500.org/lists/top500/2020/06/>.
- [2] R.D. Hornung, J.A. Keasler, The RAJA Portability Layer: Overview and Status, Tech. rep., Lawrence Livermore National Lab. (LLNL), Livermore, CA (United States), 2014.
- [3] H.C. Edwards, C.R. Trott, D. Sunderland, Kokkos: Enabling manycore performance portability through polymorphic memory access patterns, *J. Parallel Distrib. Comput.* 74 (12) (2014) 3202–3216.
- [4] C. Bertolli, S.F. Antao, A.E. Eichenberger, K. O'Brien, Z. Sura, A.C. Jacob, T. Chen, O. Sallénave, Coordinating GPU threads for OpenMP 4.0 in LLVM, in: 2014 LLVM Compiler Infrastructure in HPC, IEEE, 2014, pp. 12–21.
- [5] Intel Corporation, DPC++, 2020, [Accessed Oct. 24, 2020], <https://spec.oneapi.com/versions/latest/elements/dpcpp/source/index.html>.
- [6] A. Knüpfer, H. Brunst, J. Doleschal, M. Jurenz, M. Lieber, H. Mickler, M.S. Müller, W.E. Nagel, The Vampir performance analysis tool-set, in: *Tools for High Performance Computing*, Springer, 2008, pp. 139–155.
- [7] S.S. Shende, A.D. Malony, The TAU parallel performance system, *Int. J. High Perf. Comput. Appl.* 20 (2) (2006) 287–311.
- [8] C. January, J. Byrd, X. Oró, M. O'Connor, Allinea MAP: Adding energy and OpenMP profiling without increasing overhead, in: *Tools for High Performance Computing 2014*, Springer, 2015, pp. 25–35.
- [9] M. Schulz, J. Galarowicz, D. Maghrak, W. Hachfeld, D. Montoya, S. Cranford, OpenSpeedShop: An open source infrastructure for parallel performance analysis, *Sci. Program.* 16 (2–3) (2008) 105–121.
- [10] Lawrence Berkeley National Laboratory, Cray Reveal, 2020, [Accessed Oct. 4, 2020], <https://docs.nersc.gov/development/performance-debugging-tools/reveal>.
- [11] J. Reinders, VTune Performance Analyzer Essentials, Intel Press, 2005.
- [12] NVIDIA Corporation, NVIDIA Nsight systems, 2020, [Accessed Oct. 4, 2020], <https://developer.nvidia.com/nsight-systems>.
- [13] NVIDIA Corporation, NVIDIA Nsight compute, 2020, [Accessed Oct. 4, 2020], <https://developer.nvidia.com/nsight-compute>.
- [14] NVIDIA Corporation, The user manual for NVIDIA profiling tools for optimizing performance of CUDA applications, 2020, [Accessed Oct. 4, 2020], <https://docs.nvidia.com/cuda/profiler-users-guide>.
- [15] NVIDIA Corporation, PC sampling, 2020, [Accessed Oct. 4, 2020], https://docs.nvidia.com/cupti/cupti/r_main.html#r_pc_sampling.
- [16] M. Kambadur, S. Hong, J. Cabral, H. Patil, C.-K. Luk, S. Sajid, M.A. Kim, Fast computational GPU design with GT-Pin, in: 2015 IEEE International Symposium on Workload Characterization, IEEE, 2015, pp. 76–86.
- [17] O. Villa, M. Stephenson, D. Nellans, S.W. Keckler, NVBit: A dynamic binary instrumentation framework for NVIDIA GPUs, in: *Proceedings of the 52nd Annual IEEE/ACM International Symposium on Microarchitecture*, ACM, 2019, pp. 372–383.
- [18] Lawrence Livermore National Laboratory, RAJA performance suite, 2020, [Accessed Oct. 4, 2020], <https://github.com/LLNL/RAJAPerf>.
- [19] L. Adhianto, S. Banerjee, M. Fagan, M. Krentel, G. Marin, J. Mellor-Crummey, N.R. Tallent, HPCToolkit: Tools for performance analysis of optimized parallel programs, *Concurr. Comput.: Pract. Exper.* 22 (6) (2010) 685–701.
- [20] Advanced Micro Devices Corporation, AMD ROCm Profiler, 2020, [Accessed Oct. 4, 2020], https://rocm.docs.amd.com/en/latest/ROCM_Tools/ROCM-Tools.html.
- [21] A.D. Malony, S. Biersdorff, S. Shende, H. Jagode, S. Tomov, G. Juckeland, R. Dietrich, D. Poole, C. Lamb, Parallel performance measurement of heterogeneous parallel systems with GPUs, in: 2011 International Conference on Parallel Processing, IEEE, 2011, pp. 176–185.
- [22] B. Welton, B.P. Miller, Diogenes: looking for an honest CPU/GPU performance measurement tool, in: *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, 2019, pp. 1–20.
- [23] P. Kousha, B. Ramesh, K.K. Suresh, C.-H. Chu, A. Jain, N. Sarkauskas, H. Subramoni, D.K. Panda, Designing a profiling and visualization tool for scalable and in-depth analysis of high-performance GPU clusters, in: 2019 IEEE 26th International Conference on High Performance Computing, Data, and Analytics (HiPC), IEEE, 2019, pp. 93–102.
- [24] M. Chabbi, K. Murthy, M. Fagan, J. Mellor-Crummey, Effective sampling-driven performance tools for GPU-accelerated supercomputers, in: *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*, ACM, 2013, p. 43.
- [25] H. Zhang, J. Hollingsworth, Understanding the performance of GPGPU applications from a data-centric view, in: 2019 IEEE/ACM International Workshop on Programming and Performance Visualization Tools (ProTools), 2019, pp. 1–8, <http://dx.doi.org/10.1109/ProTools49597.2019.00006>.
- [26] NVIDIA Corporation, NVIDIA Compute Sanitizer, 2020, [Accessed Oct. 5, 2020], <https://docs.nvidia.com/cuda/compute-sanitizer/index.html>.
- [27] A.V. Gorshkov, M. Berezalsky, J. Fedorova, K. Levit-Gurevich, N. Itzhaki, GPU instruction hotspots detection based on binary instrumentation approach, *IEEE Trans. Comput.* 68 (8) (2019) 1213–1224.
- [28] K. Zhou, et al., GVProf: A value profiler for GPU-based clusters, 2020, [Accessed Oct. 4, 2020], <https://github.com/Jokeren/GVProf>.
- [29] A. Knüpfer, et al., Score-p: A joint performance measurement run-time infrastructure for periscope, scalasca, TAU, and vampir, in: *Tools for High Performance Computing 2011*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012, pp. 79–91.
- [30] B.U. Wuppertal, Parallel MPI I/O in Cube: Design & Implementation, Jülich Forschungszentrum, 2018.
- [31] L. Wei, J. Mellor-Crummey, Using sample-based time series data for automated diagnosis of scalability losses in parallel programs, in: *Proceedings of the 25th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*, 2020, pp. 144–159.
- [32] J. Choi, D.F. Richards, L.V. Kale, A. Bhatele, End-to-end performance modeling of distributed GPU applications, in: *Proceedings of the 34th ACM International Conference on Supercomputing*, 2020, pp. 1–12.
- [33] F. Schmitt, R. Dietrich, G. Juckeland, Scalable critical-path analysis and optimization guidance for hybrid MPI-CUDA applications, *Int. J. High Perform. Comput. Appl.* 31 (6) (2017) 485–498.
- [34] K. Zhou, M.W. Krentel, J. Mellor-Crummey, Tools for top-down performance analysis of GPU-accelerated applications, in: *Proceedings of the 34th ACM International Conference on Supercomputing*, in: ICS '20, Association for Computing Machinery, New York, NY, USA, 2020, <http://dx.doi.org/10.1145/3392717.3392752>.
- [35] NVIDIA Corporation, CUPTI User's guide DA-05679-001_v10.1, 2020, [Accessed Oct. 5, 2020], https://docs.nvidia.com/cuda/pdf/CUPTI_Library.pdf.
- [36] Advanced Micro Devices Corporation, ROC-Tracer, 2020, [Accessed Oct. 5, 2020], <https://github.com/ROCm-Developer-Tools/roctracer>.
- [37] J.E. Stone, D. Gohara, G. Shi, OpenCL: A parallel programming standard for heterogeneous computing systems, *Comput. Sci. Eng.* 12 (3) (2010) 66–73.
- [38] Intel Corporation, oneAPI Level Zero: 1.0.4.46, 2020, [Accessed Oct. 24, 2020], <https://spec.oneapi.com/level-zero/index.html>.
- [39] Khronos Group, clSetEventCallback(3) Manual Page, <https://www.khronos.org/registry/OpenCL/sdk/2.2/docs/man/html/clSetEventCallback.html>.
- [40] V.M. Weaver, Linux perf_event features and overhead, in: *The 2nd International Workshop on Performance Analysis of Workload Optimized Systems*, FastPath, Vol. 13, 2013, p. 5.
- [41] D. Terpstra, H. Jagode, H. You, J. Dongarra, Collecting performance data with PAPI-C, in: *Tools for High Performance Computing 2009*, Springer, 2010, pp. 157–173.
- [42] University of Wisconsin-Madison, Dyninst, 2020, [Accessed Oct. 4, 2020], <https://github.com/dyninst/dyninst>.
- [43] S.C. Eisenstat, M.C. Gursky, M.H. Schultz, A.H. Sherman, Yale sparse matrix package I: The symmetric codes, *Internat. J. Numer. Methods Engrg.* 18 (8) (1982) 1145–1151, <http://dx.doi.org/10.1002/nme.1620180804>.
- [44] R. Tarjan, Depth-first search and linear graph algorithms, *SIAM J. Comput.* 1 (2) (1972) 146–160.
- [45] S.L. Graham, P.B. Kessler, M.K. Mckusick, Gprof: A call graph execution profiler, in: *Proceedings of the 1982 SIGPLAN Symposium on Compiler Construction*, in: SIGPLAN '82, ACM, New York, NY, USA, 1982, pp. 120–126, <http://dx.doi.org/10.1145/800230.806987>.
- [46] Lawrence Livermore National Laboratory, Quicksilver, 2020, [Accessed Oct. 28, 2020], <https://github.com/LLNL/Quicksilver>.
- [47] Lawrence Livermore National Laboratory, Mercury, 2020, [Accessed Oct. 28, 2020], <https://wci.llnl.gov/simulation/computer-codes/mercury>.
- [48] PeleC, 2020, [Accessed Oct. 6, 2020], <https://github.com/AMReX-Combustion/PeleC>.
- [49] Nyx, 2020, [Accessed Oct. 6, 2020], <https://github.com/AMReX-Astro/Nyx>.
- [50] S. Plimpton, Fast parallel algorithms for short-range molecular dynamics, *J. Comput. Phys.* 117 (1) (1995) 1–19.