



Scalable system scheduling for HPC and big data

Albert Reuther^{*}, Chansup Byun, William Arcand, David Bestor, Bill Bergeron, Matthew Hubbell, Michael Jones, Peter Michaleas, Andrew Prout, Antonio Rosa, Jeremy Kepner

Massachusetts Institute of Technology, USA



HIGHLIGHTS

- HPC and Big Data schedulers have many common features.
- HPC schedulers handle parallel jobs better, while Big Data ones have better API.
- Benchmarked schedulers display little overhead for jobs longer than 30 s.
- Some schedulers have significant overhead for jobs shorter than 10 s.
- Job launch overhead can be mitigated through multi-level scheduling.

ARTICLE INFO

Article history:

Received 21 October 2016

Received in revised form 8 May 2017

Accepted 13 June 2017

Available online 8 August 2017

Keywords:

Scheduler

Resource manager

Job scheduler

High performance computing

Data analytics

ABSTRACT

In the rapidly expanding field of parallel processing, job schedulers are the “operating systems” of modern big data architectures and supercomputing systems. Job schedulers allocate computing resources and control the execution of processes on those resources. Historically, job schedulers were the domain of supercomputers, and job schedulers were designed to run massive, long-running computations over days and weeks. More recently, big data workloads have created a need for a new class of computations consisting of many short computations taking seconds or minutes that process enormous quantities of data. For both supercomputers and big data systems, the efficiency of the job scheduler represents a fundamental limit on the efficiency of the system. Detailed measurement and modeling of the performance of schedulers are critical for maximizing the performance of a large-scale computing system. This paper presents a detailed feature analysis of 15 supercomputing and big data schedulers. For big data workloads, the scheduler latency is the most important performance characteristic of the scheduler. A theoretical model of the latency of these schedulers is developed and used to design experiments targeted at measuring scheduler latency. Detailed benchmarking of four of the most popular schedulers (Slurm, Son of Grid Engine, Mesos, and Hadoop YARN) is conducted. The theoretical model is compared with data and demonstrates that scheduler performance can be characterized by two key parameters: the marginal latency of the scheduler t_s and a nonlinear exponent α_s . For all four schedulers, the utilization of the computing system decreases to <10% for computations lasting only a few seconds. Multi-level schedulers (such as LLMMapReduce) that transparently aggregate short computations can improve utilization for these short computations to >90% for all four of the schedulers that were tested.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

The evolution of computing into the hyperscale data center era has created massive computing infrastructures that can be harnessed to solve a wide range of problems. The efficient utilization of these resources is critical to their economical and environmental viability. A 100-megawatt data center that wastes

even 1% of its computing cycles can nullify all the energy-saving measures of a small city. The operating system of a data center is the job scheduler that decides which tasks are executed on which resources. Multiple users submit a wide variety of computational jobs to be processed on computational resources that include various (and sometimes heterogeneous) processing cores, network links, memory stores, storage pools, and software execution licenses (as necessary to run licensed applications). Further, each user's job will follow different parallel execution paradigms from independent process jobs to independently (pleasantly) parallel to synchronously parallel jobs, each of which imposes certain

^{*} Correspondence to: MIT Lincoln Laboratory Supercomputing Center, 244 Wood Street, Lexington, MA 02421, USA.

E-mail address: reuther@mit.edu (A. Reuther).

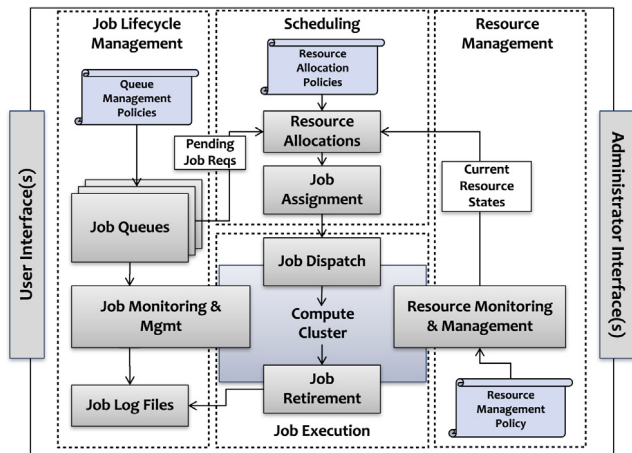


Fig. 1. Key components of cluster schedulers, including job lifecycle management, resource management, scheduling, and job execution.

execution requirements. In other words, heterogeneity is realized in multiple dimensions. Job schedulers are a key part of modern computing infrastructures; if the scheduler is not effectively managing resources and jobs, then the computing capabilities will not perform well.

Job schedulers go by a variety of names, including schedulers, resource managers, resource management systems (RMS), and distributed resource management systems (D-RMS). These terms are used interchangeably. The two main terms – scheduler and resource manager – capture the primary activities of this type of distributed software. At their simplest level, job schedulers match and execute compute jobs from multiple users on multiple computational resources. The users and their jobs usually have different resource requirements and priorities. Similarly, the computational resources have different capabilities and availabilities, and these change dynamically as different jobs are executed. Further, the resources must be managed in such a way that they are best utilized, given the mix of jobs that need to be executed.

In the functional model of schedulers, every job scheduler has an architecture with four key functions: job lifecycle management, resource management, scheduling, and job execution. These functions are shown in Fig. 1 and are described in greater detail as follows:

- The job lifecycle management function receives jobs from users through the user interface and places them in the job queue(s) to wait for execution (even if other jobs are already scheduled). Various resources for the job, including memory, licenses, and accelerators (such as GPUs), are requested through the user interface by the user. The job lifecycle management task is also responsible for prioritizing and sorting candidate jobs for execution by using the queue management policies.
- The resource management function receives availability and resource state information from the compute nodes, aggregates it, and makes it available to the scheduler. It also collects job status information to make available to users and to record in logs.
- The scheduling function allocates resources (including one or more job slots on compute nodes, compute node local memory, accelerator cards, etc.) and assigns the job to the resource(s) if the resources are available to run the job.
- The job execution function is responsible for dispatching/launching the job on the resources. Upon the completion of each job, the job execution function manages closing down the job and sending the job statistics to the job lifecycle management function, which records it in logs.

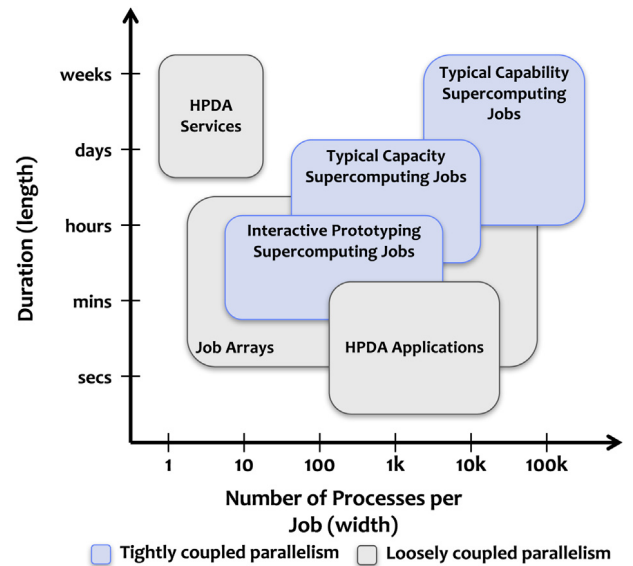


Fig. 2. Characterization of HPC and big data jobs in terms of execution time and parallelism.

This architecture has stood the test of time; it is not much different from those first developed for parallel computing schedulers [39,59]. However, there has been significant evolution in the sophistication of resource management, scheduling, resource allocation, and queue management as resources and jobs have become more diverse. This greater heterogeneity can be captured in a number of dimensions, the most basic being execution time and parallelism. Fig. 2 depicts the characterization of job types from both the traditional high performance computing and big data paradigms. Traditional HPC jobs can be categorized into single-process jobs, a set of single-process jobs that are grouped together and executed asynchronously (job arrays), and sets of processes that are all launched and executed simultaneously and synchronously (parallel jobs). Interactive prototyping parallel jobs have many of the same characteristics of typical capacity supercomputing jobs, but they tend to occupy fewer processors and execute for less time. Most HPC systems are configured to submit jobs to a batch queue, in which they wait to be scheduled from minutes to days. However, there are mechanisms, such as interactive queues, immediate execution flags, and slot reservations, which enable interactive job execution. Big data jobs tend to be in two categories. Some big data jobs are services jobs in which one or a few processes are scheduled and executed to provide services to other jobs. The majority of big data jobs are very short-run, time-critical jobs that execute in a matter of seconds to minutes and are independent of each other. Big data jobs are expected to execute immediately; that is, they tend not to wait in batch queues. High performance data analytics jobs have characteristics of both HPC and big data types of jobs. This paper explores the features that are necessary to execute high performance data analytics jobs because we need the best of both worlds to best accommodate such jobs.

This paper makes the following contributions:

- It introduces the categorization of schedulers into scheduler families and categorizes a set of HPC and big data into these families.
- It enumerates a set of scheduler features that are important for high performance data analysis jobs and compares the support for these features for a representative set of HPC and big data schedulers.

- It recognizes that many features are supported across the schedulers but that a key scheduler feature that has not been explored in much detail is job scheduler latency.
- It derives a mathematical description for job launch latency.
- It presents a benchmarking strategy for measuring job launch latency and benchmarks a representative set of HPC and big data schedulers.
- It measures the marginal latency and nonlinear exponent of job launch latency of the benchmark results of the representative set of HPC and big data schedulers.

The outline of the article is as follows. Section 2 surveys related work. Section 3 overviews the most prominently used schedulers, organizes them into scheduler families (schedulers with similar features), and compares their features. Section 4 introduces latency and utilization models that are important for understanding the job launch latency of schedulers. Section 5 introduces a job- and resource-agnostic job scheduling and execution latency performance benchmark and presents the performance results of several representative schedulers along with a discussion of the results. Section 6 summarizes the paper.

2. Related work

In the world of schedulers, a great deal of research has gone into various components, algorithms, and implementations of job schedulers. Much research has been published on optimizing the placement of jobs onto the resources with a great deal of variety in both the homogeneity and heterogeneity of the jobs and the resources. There has also been significant research in how to characterize and model the execution and resource utilization of various algorithms and applications, both as serial jobs and parallel jobs. A vast number of papers, book chapters, and books have been published on these topics; this Related Work section will overview the most significant literature to differentiate our work on job scheduler features and performance to enable high performance data analytics jobs.

First, it is worthwhile to address the related work in survey and comparison literature. This literature addresses all four functional areas of schedulers, while the feature comparisons most closely address the job lifecycle management functionality. There have been quite a number of such papers, and they were often published as final reports of studies conducted to choose a new scheduler for an organization's supercomputing and/or big data systems. The first papers that brought about comparisons of schedulers came from NASA in the late 1990s when they were replacing the Network Queuing System (NQS) batch scheduler at NASA [45]. This work prompted two reports that catalyzed the development of the Portable Batch System (PBS) [40,41]. Feitelson and Rudolph's 1995 paper [27] established a basis for making comparisons between job schedulers, and as other schedulers were developed by other organizations, more such comparison papers were written, including [2,11,21,70]. In the early 2000s, a number of universities and research organizations were developing technology to share supercomputing resources across their organizations as computing grids [29]. Among the significant papers and books from the research of scheduling for grid computing were [17,46,51]. Similarly among big data schedulers, there have been comparisons between Mesos, MapReduce, and Hadoop YARN [69] and between several Google schedulers [10]. More recently, there have been a few articles that make comparisons between traditional HPC schedulers, grid schedulers, and big data schedulers [31], as well as scheduling in cloud computing [36,64]. In all of these comparison papers, the comparison examines some subset of the features that were considered in service to certain processing workloads, usually real workloads or stochastic models of real workloads. Further, the

comparisons have been within the traditional HPC or big data job scheduling areas, but have not taken candidate schedulers from both.

Next we must address related work within the scheduling function. There are a number of papers that focus on comparing the queue management and scheduling algorithms within schedulers, including discussion of support for first-come, first-served (FCFS) [61], fairshare [62], backfill [48], and gang scheduling [26,53]. These papers include comparisons of how jobs are prioritized within queues and how jobs are matched to resources to most effectively execute them. These articles make comparisons among the various schedulers available within Apache Hadoop YARN [49]; within PBS [34], within the Maui scheduler [38], and among several of the HPC schedulers [24,28,70].

The publications that are covered in the previous paragraphs are all comparisons of actual schedulers as well as the application of queue management and scheduling algorithms on actual workload traces and simulated workload traces based on actual workload characteristics. However, another branch of job scheduler research has approached this topic from a mathematical and statistical operations research vector. Statistical distributions are applied to various aspects of workload characteristics and resource characteristics. First, a taxonomy of workloads, resources/platforms, and mapping and scheduling heuristics is presented in [1]. Workload modeling is extensively treated in [25], while [47] covers a very large set of scheduling algorithms. An extensive study and comparison of mapping and scheduling algorithms was reported in [7], with refinements in [8]. Further refinements of the underlying statistical models and parameter set generation was published in [14,13].

Many of these studies have put much work into crafting a representative mix of heterogeneous resources (i.e., memory, GPUs, cores, compute power) that best represent the systems they are modeling and a representative mix of job workloads that their systems execute. However, these studies primarily stress the ability of the schedulers to match jobs to resources and only test the scheduling and matching efficiency of the schedulers. While that research is extremely important to the effectiveness of job schedulers, job execution is often most important to the short serial and parallel data analytics jobs.

There are several papers that have specifically addressed the efficiency of launching jobs, i.e., measuring the scheduling and job execution functions in schedulers. Two papers compare job launch time of a monolithic (single) scheduler to that of distributed/partitioned job scheduling. [9] explores partitioned parallel job scheduling by modifying the IBM LoadLeveler (a modification of HTCondor) scheduler, while [73] explores distributed resource-allocation techniques by modifying Slurm. Rather than measuring utilization, both papers measure the throughput of how many jobs they can launch through the scheduler per second. As one might expect, their partitioned parallel and distributed scheduling techniques yielded greater job throughput. A white paper by the Edison Group [22] also uses job throughput to compare several modern HPC schedulers. The El-Ghazawi et al. [21] paper measures throughput in jobs per hour and turn-around time of four HPC schedulers –PBS, LSF, Sun Grid Engine/CODINE, and Condor – for small, medium, and large jobs with low, medium, and high submission rates. El-Ghazawi and colleagues randomized the CPU time of their jobs, and the average CPU times per job were 22 s, 7 min and 22 s, and 16 min and 51 s for the small, medium, and large jobs, respectively. Further, each job submission experiment only consisted of 150 jobs. Finally, our team measured launch times of virtual machine jobs onto clusters using a variety of job slot-packing and oversubscription techniques [57]. However, we did not make a distinction between the latency of the scheduler and the latency of the virtual machine launching because we were

reasonably confident that most of the latency was due to virtual machine launching. Clearly, there is a need for more extensive exploration of job launch efficiency, which we address in Sections 4 and 5.

3. Job schedulers

As mentioned in Section 2, a comparison among both traditional HPC or big data job schedulers has yet to be conducted. In this section, we do so. First, we organize these traditional HPC and big data job schedulers into several scheduler families to accentuate common intent and development among them. These families also aid in choosing a representative set of eight representative schedulers for comparison. In this comparison of representative schedulers, we find that there are many commonalities across all of the schedulers. However, we also find that there are several key differences, mainly between the traditional HPC schedulers and big data schedulers.

3.1. Scheduler families

A number of production-quality schedulers have been developed over the past 35 years to address various supercomputing and parallel data analysis computer architectures, network architectures, and software architectures. One of the first full-featured schedulers was the Network Queuing System (NQS) batch scheduler at NASA [45]. From the time of NQS, HPC systems have used job schedulers to manage jobs and resources. These schedulers used a batch queue that kept a backlog of jobs to be run, giving the scheduling function many choices from which to choose the optimal next job on the basis of the current and/or near-future resource availability. However, this batch queue required that jobs wait in the queue, often from minutes to days, in order to begin execution. While most HPC centers continue to use batch queues for scheduling jobs, the MIT Lincoln Laboratory team demonstrated on-demand job launching was possible and desirable for many classes of users [55]. Likewise, the big data community, including Google, Yahoo, Microsoft, and others, found batch-queuing schedulers to be insufficient for their needs. Big data jobs often are short in duration, run independently, and persistently perform data analysis. HPC jobs tend to be long-running, synchronously parallel simulation and modeling jobs.

The MapReduce scheduler, which was among the first big data schedulers, is a very simple scheduler, and it was developed because Google MapReduce [18] (and consequently Hadoop MapReduce [20]) needed a scheduler that could dynamically allocate jobs only to compute nodes that were storing a local copy of the required input files. Subsequently, schedulers like Borg [67] and Mesos [35] were developed because of the impression that HPC schedulers were designed to be used in batch processing modes; that is, their main feature was the capability of optimal scheduling of jobs in batch queues. However, both of these shifts also occurred because of the need for in-language APIs beyond just a command line interface. There was an effort to develop a common in-language API called DRMAA (Distributed Resource Management Application API) for batch schedulers, but the adoption of DRMAA was not widespread because of nontechnical market factors.

Hence, the main division is between HPC and big data schedulers. The HPC scheduler family can be further broken into the traditional HPC and the new HPC sub-families. The traditional HPC schedulers include PBS [34], Grid Engine [65], HTCondor [50], OAR [15], and LSF [72]. The new HPC schedulers include Cray ALPS [52] and Slurm [71]. The big data schedulers can be further broken into commercial and open-source sub-families. The commercial big data schedulers include Google MapReduce [18], Google Borg [67], and Google Omega [60], and components of Borg and Omega

are available as the open-source Kubernetes project [10]. The open-source big data schedulers include Apache Hadoop MapReduce [20], Apache Hadoop YARN [66], and Apache Mesos [35]. There are many plug-ins available for Apache Hadoop and Apache Hadoop YARN, and the two plug-ins that enhance Hadoop YARN functionality the most are Apache Llama [16] and Apache TEZ [58]. Finally, there are many research schedulers that have been developed in academia and industry to explore and evaluate new job scheduler concepts and ideas. Two recent schedulers in this category are Pacora [4,5] and Sparrow [54]. More details about each of the aforementioned schedulers follow.

3.1.1. Traditional HPC family

The Portable Batch System (PBS) Scheduler Family is directly descended from NQS, the first batch scheduler developed under NASA funding. PBS is a fully featured job scheduler that includes a separate queue manager, resource manager, and scheduler, and the Maui scheduler is often used in place of the native PBS scheduler. PBS continues to be developed by Altair Engineering (PBSPro) and Adaptive Computing (TORQUE/Maui/Moab Cluster Suite), which both offer open-source and commercial versions. Recently, scalability challenges have been addressed, and full job array integration has been added.

Grid Engine is a full-featured, very flexible scheduler originally developed and released under the name CODINE in 1993 by Genias Software [30], later acquired and matured by Sun Microsystems and Oracle, and currently offered commercially by Univa. There are also several open-source versions, including Son of Grid Engine and Open Grid Scheduler, though further development of these offerings seems to be waning.

HT Condor (High Throughput Condor) continues to be developed by Prof. Myron Livny's team at the University of Wisconsin. HT Condor was designed to cycle-scamper from distributed heterogeneous desktop computers to managed clusters. It implements a single queue with many resource- and job-distinguishing options, including prioritization and fair execution. It works especially well for many smaller, single-process jobs with execution on a diverse set of computers, clusters, and supercomputers. IBM LoadLeveler was based on Condor [23,42,43].

OAR [15] was developed by the Laboratory of Informatics in Grenoble, France. It is a full-featured scheduler with queues and resources managed through database records and PERL language modules.

IBM Platform Load Sharing Facility (LSF) is a full-featured and high performing scheduler that is very intuitive to configure and use. It was based on research and development of the Utopia job scheduler at the University of Toronto [72]. OpenLAVA [http://www.openlava.org/] is an open-source derivative of LSF that has reasonable feature parity.

3.1.2. New HPC family

Cray Application Level Placement Scheduler (ALPS) [52] was developed in the early 2000s to take advantage of unique Cray architecture features. In 2013–14, it was refactored as a “native” Slurm implementation.

The Simple Linux Utility for Resource Management (Slurm) [71] is an extremely scalable, full-featured scheduler with a modern multithreaded core scheduler and a very high performing plug-in module architecture. The plug-in module architecture makes it highly configurable for a wide variety of workloads, network architectures, queuing policies, scheduling policies, etc. It was inspired by the Quadrics RMS scheduler [71]. To ease transition from other schedulers, translation interfaces are provided. Slurm began development in the early 2000s at Lawrence Livermore National Laboratory to address scalability issues in job schedulers. It is entirely open source, while consulting services and community development leadership are provided by SchedMD.

3.1.3. Commercial big data family

Google MapReduce was introduced by Google through the [18] paper, which detailed how they implemented a map-reduce processing paradigm to process data files stored locally on each compute node. The job scheduler was a simple first-in, first-out scheduler with few features. Subsequent improvements added more features as Google required greater capabilities.

Google Borg is a software container management system [10,67] that evolved from the diversification of workloads at Google that expanded past the map-reduce paradigm of parallel computation. The default execution environment through Borg is Linux containers with static library linking for application resource and execution isolation. The scheduler in Borg is a multithreaded scheduler that accommodates batch and persistent jobs. Further, it is highly scalable; it can manage 10,000 to 100,000 nodes.

Google Omega (Kubernetes/Pilots) [10,60] was based on Borg and was written with improved software architecture and software engineering principles. It is a distributed job scheduler, and it decouples the cluster state from the resource management and scheduling functionalities for greater scalability and flexibility. Some of these features, including distributed scheduling services, were rolled back into Borg.

The Apollo scheduler [6] developed by Microsoft uses distributed estimation-based scheduling to scale to very large (>20,000 node) clusters. Unlike Sparrow, which only takes into account machine queue depth, Apollo also takes into account load and data locality. Boutin and his colleagues were not as concerned with fairness, and they showed task scheduling rates over 10,000 per second.

Microsoft's Mercury scheduler [44] is a hybrid (centralized and distributed) scheduler that allows applications to make trade-offs between the scheduling overhead centralized scheduling and the execution guarantees of distributed scheduling on very large clusters. Mercury was implemented as an extension to Apache Hadoop YARN. Rasler and his colleagues then extended and further refined the work of Sparrow (a distributed research scheduler mentioned below), Apollo, and Mercury with the Yaq-c (centralized) and Yaq-d (distributed) scheduling algorithms with more principled queue management techniques, including queue sizing, worker-side queue reordering, and placement of tasks.

3.1.4. Open-source big data family

Apache Hadoop YARN is an in-memory map-reduce job scheduler that enables scaling map-reduce style job environments to execute efficiently on several thousand servers. It was built to increase the processing speed and scalability of Apache Hadoop MapReduce [20]. It is a monolithic scheduler with a simple API for batch map-reduce jobs, and it does not support microjobs or persistent jobs well. Several YARN add-ons introduce features into the scheduler. These add-ons include YARN Fair Scheduler [3], which enables fair job scheduling; Project Myriad, which enables YARN to run on a Mesos-managed cluster; Apache Llama (Low Latency Application Master), which enables batch and low-latency jobs via a bridge for YARN; and Apache TEZ, which enables job dependencies and directed, acyclic graph (DAG) dependencies in YARN.

Apache Mesos is a two-level scheduler that enables the partitioning of a pool of compute nodes among many scheduling domains. Each scheduling domain has a pluggable scheduler called a Mesos framework (think queues and node pools), and each framework allocates the resources within its domain resources that have been allocated to it by Mesos. It has a rich API for communicating with the scheduler. Frameworks are written to use the Mesos API, thus making application integration on a Mesos cluster straightforward.

Kubernetes is an open-source distributed Docker container management system that includes a scheduler. It is based on

Google Borg and Omega, and it was released as an open-source project in 2015. The scheduler uses a simple first-in, first-out scheme; it has features to launch containers in an on-demand fashion; and it can queue up pending resource requests.

3.1.5. Research family

Pacora [4,5] is a research scheduler developed by Sarah Bird at UC Berkeley for her PhD thesis. She demonstrated the use of dynamic runtime resource utilization of each of the simultaneously executing application to better schedule and load-balance processes onto multicore computers. As future work, she considered extending these concepts to cluster scheduling by feeding back the dynamic state of resource utilization to the job scheduler and using Pacora algorithms to determine job placement more dynamically. Pacora implements resource allocation as a convex optimization.

Among other research schedulers, the focus has been on data-dependent tasks; in other words, they do not schedule synchronously parallel job. A number of these schedulers explored the friction between scalability, placement correctness, and efficiency (including latency), particularly using distributed scheduling architectures.

The Quincy scheduler [37] from Microsoft Research was the first distributed scheduler that tried to optimize the trade-offs between job data locality and fairness using a min-cost flow algorithm. It was purely a research project exploring distributed scheduling, and it is not available.

Sparrow [54] is a research scheduler developed at UC Berkeley that implements decentralized resource management and job scheduling. Sparrow is available as a Spark plug-in with no support.

The Tarcil scheduler [19] developed by Stanford University and MIT is also a distributed scheduler that balances scheduling speed and resource matching quality by employing dynamic job and resource statistics along with dynamic sub-sampling.

Most recently, the Firmament centralized scheduler [32] improves on the min-cost max-flow (MCMF) scheduling optimization of the Quincy scheduler by using relaxation which has much lower scheduling latency.

3.2. Scheduler features

Certain resource manager and scheduler features are key and the primary intent of the large-scale computing system and the nature of the job policies. There are many features that we can compare among these schedulers, and the most important features are briefly explained in this subsection. The features are explained with their impact on both HPC and big data jobs. The features are organized into subsections which will correspond to feature comparison subsections in Section 3.3.

3.2.1. Scheduler metadata

Actively developed assesses whether features are being added to the scheduler code base and if the code is being maintained. This is a bit of a double-edged sword. On the one hand, you want some development to occur, including adding some new features and porting to new operating systems and hardware. But if a project is too active, it can be difficult to keep up with the versions to deliver new features to users in a timely manner.

Cost and licensing conveys whether the scheduler is open source or has a license for which one must pay. While it is nice to have contracted support, it can be very expensive for large clusters. Free is not necessarily free, though. Even if a scheduler is open source and free, it is good to have in-house talent who are responsible for configuring and managing the scheduler. However, some licenses can be simply exorbitant with hefty per-node or even per-core charges.

Operating system support captures the operating systems on which the scheduler runs and on which jobs can be executed. This feature is reasonably straightforward. While all of these schedulers run on Linux, distribution support may not be entirely universal (though you can usually compile the source code of open-source schedulers and possibly incur some porting overhead). Commercial schedulers may have a more limited list of supported Linux distributions, which are usually the more stable, commercially vetted ones like Red Hat Enterprise Linux, CentOS, and Suse, but perhaps not Fedora or Ubuntu.

Language constraints refers to the programming languages in which executed applications can be written. This feature is particularly important because data analysis applications are often written in rapid prototyping/scripting languages such as Matlab, Python, Java, Scala, and R, which are not the typical HPC programming languages. Most of these schedulers can execute jobs in all of the major programming languages. However, with big data scheduler APIs, some languages are favored. Because many data analysis applications are often written with API calls to the scheduler, language support can be very important to data analysis applications.

Access control and security involves user authentication and isolation. Among the HPC schedulers, access control and security have always been parts of the feature set. Early versions of open-source big data schedulers did not have much security or isolation, but that is changing because of greater adoption and scaling to many more users.

3.2.2. Job types

Job types captures the parallel nature and size of the jobs. Among the number of ways to categorize jobs, the primary ways are in the dimensions of time executed (short, medium, long, very long) and the jobs' parallelism or lack thereof. This feature captures the length of jobs that the scheduler is designed to best accommodate. Many data analysis applications and services run as either fast-running small tasks or as persistently executing, long-running service jobs. Schedulers for data analysis need to be able to accommodate both of these job types well.

Parallel and job array support captures the types of parallelism that the scheduler can handle. More specifically, there are single-process jobs; job arrays, in which multiple independent processes are run using a single job identifier with different parameters for each process; and parallel jobs, in which each of the processes are launched simultaneously and communicate during the computation. All schedulers handle single-process jobs. Among the HPC schedulers, all have support for both parallel jobs and job arrays, although in the PBS family of schedulers, the job array commands on the command line are not as well integrated as they are in the other HPC schedulers. Big data schedulers can run job arrays well, and it is possible to run parallel jobs though the synchrony in launching may not be robust.

Parallel and array jobs indicates whether the job scheduler allows synchronous dependent parallel and/or asynchronously independent parallel (array) jobs.

Queue support captures whether the scheduler can manage jobs with different characteristics and resource requirements in separate data structures called queues. Having multiple queues often makes it easier to manage jobs with disparately different requirements; however, having too many queues can make the user experience more confusing than it needs to be by presenting far too many choices of queues into which the users can submit their jobs.

Metascheduling is a feature that enable support for multiple resource managers. In the big data world, the challenge of executing long-running services alongside short-running analytics tasks can cause issues with their schedulers. The solution that

the community settled on was multitiered scheduling in which a metascheduler chose which of one or more schedulers was best for each submitted job. This solution is not that different from the concept of queues among the HPC schedulers. We should note that a large push for metascheduling in the supercomputing research community in the late 1990s and early 2000s enabled supercomputing jobs to be submitted and executed across multiple organizations and supercomputing centers including Globus [29,51], Legion [33], and Condor Flocking [23]. That research has reached maturity and is being used by many supercomputing consortiums. However, these technologies are not being considered in this work because we are addressing data analysis systems that are not deployed across multiple organizations/centers.

To a large extent, the queue support and metascheduler features attempt to solve the same problem from two different perspectives. HPC schedulers added queue support so that different types of jobs, resource policies, priorities, etc., were lumped together to help with prioritization, resource allocation, etc. Each queue can be assigned a certain set of compute nodes (and perhaps other resources) on which its jobs are allowed to run. The metascheduling feature in Mesos tries to solve a very similar problem by enabling users to submit their jobs to Mesos using a single common API.

3.2.3. Job scheduling

Timesharing is the ability to allocate multiple jobs from one or more users to a single compute node. Generally, since modern operating systems have that capability, so do the schedulers that feed jobs to them.

Backfilling is the ability to schedule pending jobs when an executing job finishes early. Usually these backfilled jobs are smaller, both in execution time and number of job slots, than the larger parallel jobs. This is a feature that HPC schedulers have, while big data schedulers tend to schedule jobs on demand.

Job chunking has some relationship with job arrays, but is different. With job chunking, the scheduler searches through a given user's queued jobs to find similar jobs. It then schedules multiple similar jobs simultaneously.

Bin packing is a resource-allocation technique by which the scheduler chooses groups of jobs to launch simultaneously on a node or set of nodes to best utilize the node resources.

Gang scheduling allows a user to submit multiple processes to occupy a single job slot. Only one of the multiple processes can execute at any given time, but they are usually the type of processes that execute sequentially or only part of the time.

Job dependencies and directed acyclic graphs (DAGs) support is a feature that allows users to define execution dependencies between jobs. In data analysis applications, it is common to chain together analytics jobs with dependencies on each other, so support for this feature is essential.

3.2.4. Resource management

Resource heterogeneity refers to the scheduler's ability to accommodate differently configured nodes; i.e., they have different configurations of RAM, CPU, disk space, GPU availability, etc. More sophisticated resource heterogeneity allows administrator-defined resources that can be associated with compute nodes.

Resource allocation policy refers to the ability to write rules about how resources are utilized. Because short-running data analysis applications and long-running services usually require very different resource-usage profiles, it is important that these resource needs can be characterized precisely so that service jobs are provided only the resources that they really need since the allocated resources will be utilized by these long-running services for a long timeframe.

Static and dynamic (consumable) resources refers to how resources are consumed when a job is scheduled to use them. For

instance, the total RAM on a compute node, as well as the CPU cores on a node, are static resources. Conversely, dynamic resources are resources that do change, such as load status and host status.

Network-aware scheduling takes the network topology into account when allocating jobs to nodes. The intent is to minimize network distance between jobs that need to communicate with each other in parallel jobs.

3.2.5. Job placement

Intelligent scheduling captures the fact that the scheduler is using a scheduling policy more sophisticated than first-in, first-out (FIFO).

Prioritization schema assure that resources are allocated to jobs according to priorities. This function can be as simple as submitting jobs with a priority rating, but it can be much more sophisticated, such as having user priorities, user group priorities, and even fairshare policies implemented in the scheduler.

Job replacement and reordering refers to a feature by which a user can replace jobs and/or reorder the execution of jobs that are waiting to be executed in the job queue.

Advanced reservations is a feature that allows a user to request a set of job slots in the future. This feature is often used to enable interactive jobs on systems that use a highly utilized batch queue.

Power-aware scheduling works to minimize the number of nodes that are online and powered up so as to minimize the power being drawn by the system. This function not only involves packing jobs onto nodes effectively but also requires the ability to shut off and reboot nodes when necessary.

User-related job placement enables users to specify whether a job is scheduled (or not) on nodes on which another user's jobs are executing.

Job-related job placement refers to the ability to schedule jobs (or not) where other jobs are executing.

Data-related job placement schedules jobs on nodes on which the required data reside in local storage.

3.2.6. Scheduling performance

Centralized vs. distributed scheduling refers to the architecture with which job scheduling is accomplished. Schedulers need to keep track of jobs that have been submitted and are waiting for execution in a queue and jobs that have been dispatched and are currently executing. Each job has a data structure associated with it that captures a great deal of job information. Hence, most schedulers are centralized so that all of the job information is contained in that central location. In the past few decades, hot backup capabilities have been added to gain higher availability of the scheduler. However, a distributed scheduler architecture would allow for greater resilience but could cost the scheduler in performance. The Mesos scheduler is more distributed than the other schedulers in that it is distributed between its metascheduler role and each of the schedulers that it uses.

Scheduler fault tolerance is related to the centralized or distributed architecture of the scheduler, and it is usually enabled through a hot backup server or a set of distributed servers. It captures whether the scheduler and its jobs continue to execute despite a fault in the scheduler.

Scalability and throughput refer to how many jobs and nodes the scheduler can simultaneously schedule. This performance metric is particularly important to some big data parallel frameworks (particularly MapReduce) in which programmers are encouraged to write their pleasantly parallel software as rather short-duration, simple operations.

Performance and bottlenecks capture what the primary efficiency inhibitor is in scheduling jobs. Because of the complexity of receiving the status of node resources, managing the queue

of pending jobs, performing resource allocations to available resources, dispatching jobs to nodes, and monitoring jobs in execution, it can be quite challenging to pinpoint bottlenecks. Further, the aforementioned list of functions does not even take into account the specific hardware and configurations of the system.

Scheduler latency refers to how quickly a job is dispatched onto job slots/resources once they become available.

3.2.7. Job execution

Prolog/epilog support is a feature that allows scripts to be executed before and/or after job execution. This functionality may include setting up certain execution environments or copying data. We have found this feature quite valuable for copying database datasets and virtual machine images onto and off of the local node storage in dynamic database and virtual machine capabilities.

Data movement/file staging is a built-in scheduler feature that enables the copying of supporting files onto the local storage of nodes. This feature was particularly important in the past when central file systems were less common. Most HPC schedulers have this feature, while big data schedulers generally do not.

Checkpointing allows applications to save the execution state of running jobs so that if a crash occurs, the job can be restarted from the checkpointed state. Most big data schedulers do not handle checkpointing but instead relaunch the job if it fails.

Job migration means that the scheduler can move a job from one compute resource to another while it is executing. Usually this feature requires scheduler-based checkpointing to record the current job state so that the job can be restarted from where it had left off on another node. This feature can be particularly useful in rebalancing long-running or service-oriented jobs.

Job restarting is a feature that restarts jobs if the job is aborted or fails (if so requested by the submitter).

Job preemption allows lower-priority executing jobs to be hibernated in order to execute higher-priority jobs. Usually this action involves using swap space on the compute nodes to hibernate the jobs.

3.3. Feature comparison

Now that we have covered all of the scheduler features, we shall compare how the schedulers stack up against each other on the basis of these features. There are many overlaps and feature parity among the schedulers within each of the families so we will explore a subset of these schedulers. The schedulers must be openly available and have reasonable support and feature development through companies or consulting firms so that medium to large organizations have access to such outside mechanisms (except for the representative scheduler from the research scheduler family). We chose representative schedulers from each scheduler family. We have chosen the following set of schedulers for further evaluation:

- IBM Platform LSF
- OpenLava
- Slurm
- Grid Engine
- Pacora
- Apache Hadoop YARN
- Apache Mesos
- Google Kubernetes

Table 1

Metadata features comparison among job schedulers.

Metadata	LSF	OpenLAVA	Slurm	Grid Engine	Pacora	YARN	Mesos	Kubernetes
Type	HPC	HPC	HPC	HPC	HPC	Big Data	Big Data	Big Data
Actively developed	✓	✓	✓	✓	– ^a	✓	✓	✓
Cost/ Licensing	\$\$\$	Open source	Open source	\$\$\$	N/A	Open source	Open source	Open source
OS support	Linux	Linux, Cygwin	Linux, *nix	Linux, *nix	N/A	Linux	Linux	Linux
Language support	All	All	All	All	N/A	Java, Python ^b	All	All
Access control/ security	✓	✓	✓	✓	–	✓	✓	✓

^a Some further development is ongoing within Microsoft.^b Other languages are supported; Java and Python are the most strongly supported.**Table 2**

Job type features comparison among job schedulers.

Job types	LSF	OpenLAVA	Slurm	Grid Engine	Pacora	YARN	Mesos	Kubernetes
Parallel and array jobs	Both	Both	Both	Both	N/A	Array	Array	Array
Queue support	✓	✓	✓	✓	–	✓	✓	–
Multiple resource managers	–	–	–	–	–	–	✓	–

Table 3

Job scheduling features comparison among job schedulers.

Job scheduling	LSF	OpenLAVA	Slurm	Grid Engine	Pacora	YARN	Mesos	Kubernetes
Timesharing	✓	✓	✓	✓	–	✓	✓	✓
Backfilling	✓	✓	✓	✓	–	–	–	–
Job chunking	–	–	–	✓	–	–	–	–
Bin packing	–	–	✓	–	–	–	–	–
Gang scheduling	–	–	✓	–	–	–	–	–
Job dependencies and DAGs	✓	✓	✓	✓	–	✓	– ^a	–

^a Job dependencies and DAG support in Mesos depends on having the feature in the plugged-in scheduler framework(s).

We will compare the representative schedulers on the features listed above. Please note that since Pacora is a research scheduler, many of the feature comparisons are not available for it.

3.3.1. Metadata features comparison

Table 1 is a comparison of the metadata associated with the representative schedulers. All of the representative schedulers are being actively developed except for the research scheduler, Pacora, which has some further development within Microsoft. While LSF and Grid Engine each have a commercial version, the rest of the schedulers are available as open-source projects. Consulting and support contracts are available for them from various companies. As mentioned in Section 3.1.1, OpenLAVA is an open-source version of LSF with reasonable feature parity, although the two schedulers are reported to not share a common code base. Slurm, LSF, and Grid Engine have been ported to many different operating systems on many different hardware platforms and network topologies over the years. All the production schedulers are supported on all the major server-based Linux distributions. Mesos is written primarily in C++ and should be portable to diverse (non-Linux) environments, if needed. Broad language support is almost universal except for YARN. YARN does support other languages to some extent, but Java and Python are the most strongly supported. While all of the schedulers have some access control and security, there is plenty of online discussion about how YARN, Mesos, and Kubernetes probably should have more. Each of the developer community project teams are addressing these concerns.

3.3.2. Job type features comparison

The comparison of job type features is presented in Table 2. The HPC schedulers have parallel jobs and array jobs, while the big data schedulers support array jobs. YARN, Mesos, and Kubernetes do not have built-in support for tightly coupled parallel jobs. Using the MPI framework with Mesos (or building a custom framework specifically for handling Slurm, PBS, Grid Engine, or LSF schedulers), Mesos can execute parallel jobs. Queue support is reasonably universal, but support for queues in big data schedulers

is generally not as extensive as in the HPC schedulers. In YARN, queues are available in the capacity scheduler. In Mesos, one can think of each of the installed frameworks as a queue for different applications and/or application types. Each of the Mesos frameworks can implement its own set of queues to adjudicate between the jobs of its application flavor. Since the focus of Kubernetes is on container management, the scheduler has no queue support. Mesos is a metascheduler by design, and it allocates shared resources to other schedulers (frameworks). The other schedulers do not support multiple resource managers.

3.3.3. Job scheduling features comparison

The job scheduling features are compared in Table 3. Timesharing is critical for schedulers; one could easily argue that if a scheduler cannot handle scheduling multiple jobs to a compute node, it does not have a full feature set. Other features like backfilling, job chunking, bin packing, and gang scheduling are essential when systems have a very deep set of pending jobs in queues and there are expectations for the systems to be achieving 90% or higher utilization. Similar to timesharing, job dependencies are also essential for job schedulers. All of the HPC schedulers allow job dependencies to be submitted. Mesos can support job dependencies if a given plugged-in framework supports job dependencies. Kubernetes does not support dependencies; in its container-based environment, it is expected that users handle job dependencies at the application level.

3.3.4. Resource management features comparison

Table 4 compares resource management features among the representative job schedulers. Similar to timesharing and job dependencies, resource management and resource allocation policies are essential to the functionality of a modern scheduler. However, how many types of resources a scheduler offers (including site-defined resources) and how optimally jobs are mapped onto the resources is NP-complete and continues to be an open research problem. All of the schedulers support policy-based resource allocation and static and dynamic resources. Static resources are

Table 4

Resource management features comparison among job schedulers.

Resource management	LSF	OpenLAVA	Slurm	Grid Engine	Pacora	YARN	Mesos	Kubernetes
Resource heterogeneity	✓	✓	✓	✓	✓	✓	✓	✓
Resource allocation policy	✓	✓	✓	✓	✓	✓	✓	✓
Static and dynamic resources	Both	Both	Both	Both	Both	Both	Both	Both
Network aware scheduling	✓		✓	✓	–			

Table 5

Job placement features comparison among job schedulers.

Job placement	LSF	OpenLAVA	Slurm	Grid Engine	Pacora	YARN	Mesos	Kubernetes
Intelligent scheduling	✓	✓	✓	✓	–	^a	✓ ^b	
Prioritization schema	✓	✓	✓	✓	–	✓	✓	✓
Job replacement and reordering	✓	✓	✓	✓	–		✓	✓
Advanced reservations	✓		✓	✓	–			
Power-aware scheduling	✓		✓	✓	–			
User-related job scheduling	✓		✓	✓	–			
Job-related job scheduling	✓		✓	✓	–			
Data-related job scheduling					–	✓		

^a Both the Fair Scheduler and Capacity Scheduler in YARN offer some intelligent scheduling.^b Intelligent scheduling depends on intelligent scheduling support in the plugged-in scheduler framework(s).**Table 6**

Scheduling performance features comparison among job schedulers.

Scheduling performance	LSF	OpenLAVA	Slurm	Grid Engine	Pacora	YARN	Mesos	Kubernetes
Centralized vs. distributed	Cent.	Cent.	Cent.	Cent.	Cent.	Cent.	Dist.	Cent.
Scheduler fault tolerance	✓		✓	✓	–	✓	✓	✓
Scalability and throughput	10K+	1K+	100K+	10K+	–	10K+	100K+	100K+

either busy or not (i.e., a job slot), while dynamic resources can be partially used by jobs (i.e., the memory of a compute node). Further, all the schedulers have mechanisms for administrators to set and manage resources via configuration scripts. In YARN, there is some concept of network-aware scheduling in that jobs are dispatched to different parts of the dataset that are stored on each of the Hadoop nodes. Usually files in the Hadoop distributed file system (HDFS) are replicated three times by default. There is the original file, a replica of the file somewhere in the same rack of the data center, and another replica in a different rack of the data center. This distribution scheme is intended to increase the redundancy of each file to provide resilience to failures. However, the data location-oriented job scheduling in Hadoop YARN is not nearly as sophisticated as the network-aware scheduling that is available in LSF, Slurm, and Grid Engine (as well as PBS).

3.3.5. Job placement features comparison

Job placement features are compared in Table 5. Each of the HPC schedulers and Mesos have available scheduling algorithms that are more sophisticated than first-in, first-out. YARN has more sophisticated scheduling algorithms in its Fair Scheduler and Capacity Scheduler, while Kubernetes does not include sophistication beyond first-in, first-out. Every scheduler enables some prioritization scheme though it may be primitive. Many of the other scheduling techniques (advanced reservations, power-aware scheduling, user-related job scheduling, and job-related job scheduling) are more applicable to batch-queue systems with long-running jobs. These techniques are reasonably complex so it is not unreasonable for purely open-source projects like OpenLAVA to not have such a feature. Finally, only YARN supports data-related job scheduling, but one can develop scripts that read location data from distributed file systems like Hadoop distributed file system [63] and annotate dynamically generated submission scripts to have similar effects.

3.3.6. Scheduling performance features comparison

Table 6 compares the scheduling performance of the representative schedulers. All of the schedulers except Mesos employ a centralized scheduler, and all but OpenLAVA and Pacora implement

fault tolerance with one or more spares. Mesos is a distributed scheduler by its metascheduling nature; each of the framework schedulers are distributed. Each of the schedulers is able to handle anywhere from thousands to hundreds of thousands of simultaneously executing job slots. This number depends upon design and implementation of the job and resource data structures and management algorithms. OpenLAVA was just recently reinvigorated, so it would not be surprising for its open-source team to add fault tolerance and greater scalability soon. Performance, bottlenecks, and scheduler latency are addressed in Section 5.

3.3.7. Job execution features comparison

Finally, Table 7 compares job execution features among the representative job schedulers. Prolog and epilog scripting support is fairly well covered among the schedulers while data movement and file staging are supported by the more established HPC schedulers. The big data schedulers leave that functionality to be included in applications. Similarly, HPC schedulers include scheduler-based checkpointing while big data schedulers also leave that functionality to the applications. Data analysis applications need restarting, and migration is very helpful for job fault tolerance; restarting is well supported by the schedulers while migration of jobs is done by the programmers/users in Mesos and Kubernetes. Job preemption support in the scheduler is available in the HPC schedulers and in Mesos and Kubernetes.

3.4. Summary of observations

These comparisons show that there is a great amount of parity among features. Such features common across the majority of traditional HPC and big data schedulers include timesharing, queues, job dependencies, heterogeneous resources, static and dynamic resources, intelligent scheduling, prioritization schema, fault tolerance job restarting, and prolog/epilog support. The depth and complexity of how the features are implemented in each of the schedulers may vary greatly, and one can argue that such details usually only affect the power users of the system and have little to no impact on the novice and average users. There is another set

Table 7

Job execution features comparison among job schedulers.

Job execution	LSF	OpenLAVA	Slurm	Grid Engine	Pacora	YARN	Mesos	Kubernetes
Prolog/epilog support	✓		✓	✓	–		✓	✓
Data movement/file staging	✓		✓	✓	–			
Checkpointing	✓	✓	✓	✓	–			
Job migration	✓	✓	✓	✓	–		?	?
Job restarting	✓	✓	✓	✓	–	✓	✓	✓
Job preemption	✓	✓	✓	✓	–		✓	✓

Table 8

Summary of notation.

Notation	Definition
t_s	Scheduler latency
t	Task time
T_{total}	Total runtime of a job set
T_{job}	Total isolated job execution time
ΔT	Non-execution latency
N	Total tasks
n	Tasks per processor
α_s	Exponent accounting for nonlinear behavior in the scheduler
U	System utilization
U_c	System utilization for constant time tasks
U_v	System utilization for variable time tasks

of more complex features that are generally unique to traditional HPC workloads. These features include synchronously parallel job launches, advanced job scheduling features (backfilling, job chunking, bin packing, and gang scheduling), network-aware scheduling, advanced job placement features, and checkpointing. Finally, big data schedulers stand out in having easy-to-use APIs with which applications are developed, widely shared and distributed, and executed, and that lower the barrier to adoption and utilization. However, the aforementioned DRMAA API is quite similar to these APIs, and it could be utilized to a similar effect.

In the scheduling performance subsection, we postponed addressing performance, bottlenecks, and scheduler latency. These are the key subjects in the next sections.

4. Latency and utilization modeling

In order to better understand scheduler latency, it is worthwhile to construct a mathematical model for this latency. A summary of the notation used in this section is provided in Table 8. Scheduler latency t_s is the overhead associated with executing a task with the scheduler. The scheduler latency is influenced by submission, queue management, resource identification, resource selection, resource allocation, job dispatch, and job tear down. Task time t is the time the task takes to run in complete isolation on a given resource, i.e., compute node. The value of t then falls into one of two cases:

- Case 1: Low scheduler latency impact $-t \gg t_s$
- Case 2: High scheduler latency impact $-t \lesssim t_s$.

Data analysis jobs often execute many short-duration tasks, so they are generally more susceptible to realizing case 2.

The diversity of steps that are performed by a scheduler makes latency difficult to measure. First, it is difficult to position a single observer in the system that will work well with all of the schedulers. Further, there are so many measurements that would need to be taken that the measurements can negatively affect the scheduler latency. However, all is not lost. The runtime model of the overall scheduler latency can be derived from the measured total runtime of a job T_{total} consisting of N tasks running on P processors

$$T_{total}(N, P) = T_{job}(N, P) + \Delta T(N, P)$$

where the isolated job execution time T_{job} is the isolated task times on a processor, and the difference between the total runtime and the isolated job time is ΔT . This equation then leads to the following latency model for each component on the right of the equation. If the job consists of tasks with the same runtime on every processor, then T_{job} and ΔT can be modeled as

$$T_{job} = tn \quad \Delta T = t_s n^{\alpha_s}$$

where $n = N/P$ is the number of tasks that are run on one processor, t_s is the marginal scheduler latency incurred by adding a task to a processor, and $\alpha_s \approx 1$ is the exponent that accounts for nonlinear behavior in the scheduler. (Note, a number of equation formulations were considered for ΔT , and it was this exponential equation that best fit all of the data that we collected and analyzed.) On the basis of the above model, there are three ways to reduce the impact of the scheduler launch latency on the total runtime: reduce n , t_s , and/or α_s .

The above model assumes that all tasks in a job have equivalent runtimes. However, we can account for different runtimes in the following way. If the job consists of tasks with different runtimes t_j , then

$$T_{job}(p) = \sum_{j \in J(p)} t_j \quad \Delta T = t_s n(p)^{\alpha_s}$$

where $J(p)$ is the set of tasks that run on processor p , and $n(p)$ is the number of tasks that run on processor p . Variable task times could impact the scheduler performance if $T_{job}(p)$ and/or $n(p)$ are sufficiently different from their averages.

Now that we have defined the scheduler latency model, we can establish the definition for utilization. Utilization U is the time spent working on a task T_{job} divided by the total time T_{total} that the resource is claimed by the scheduler to execute the task:

$$U = T_{job}/T_{total}.$$

This can be rewritten as

$$U^{-1} = 1 + \Delta T/T_{job}.$$

For constant task times, utilization can be modeled as

$$U_c^{-1} = 1 + (t_s n^{\alpha_s})/(tn).$$

For $\alpha_s \approx 1$, the above formula simplifies to

$$U_c(t)^{-1} \approx 1 + t_s/t.$$

The above model illustrates the strong impact the scheduler latency can have on utilization for tasks similar in duration to t_s :

$$t_s \approx t \Rightarrow U_c^{-1} \approx 2 \Rightarrow U_c \approx 0.5.$$

For variable task times, utilization can be modeled then as

$$U_v(p)^{-1} = 1 + (t_s n(p)^{\alpha_s}) / \left(\sum_{j \in J(p)} t_j \right).$$

For $\alpha_s \approx 1$, this simplifies to

$$U_v(p)^{-1} \approx 1 + (t_s n(p)) / \left(\sum_{j \in J(p)} t_j \right) = 1 + t_s/t(p) = U_c(t(p))^{-1}$$

where $t(p)$ is the average time of the tasks that ran on processor p :

$$t(p) = \left(\sum_{j \in J(p)} t_j \right) / n(p).$$

If the scheduler releases a processor as it completes its work, then the overall utilization is the average of the per-processor utilization

$$U^{-1} \approx P^{-1} \sum_p U_c(t(p))^{-1}.$$

By measuring the constant time task utilization $U_c(t)$ curve, it is possible to estimate the utilization for any distribution of tasks from the average task times on each processor.

There are several takeaways that we can gain from this scheduler latency model. First, scheduler latency can have a large impact on data analytics workloads that have many short-duration tasks. There are three key parameters that impact scheduler performance:

- n , the number of tasks that ran on one processor,
- t_s , the marginal scheduler latency incurred by adding a task to a processor, and
- $\alpha_s \approx 1$, the exponential that accounts for nonlinear behavior in the scheduler.

Scheduler performance can be improved by reducing the value of these parameters. Finally, the utilization curve for jobs with constant time tasks can be used to estimate the utilization of jobs with variable time tasks. In the next section, we will develop a benchmark with constant time tasks and evaluate the utilization and scaling of job scheduling latency with these constant time tasks. This approach allows us to characterize scheduler performance without having to develop more complex variable time task distributions.

5. Performance comparison

If batch scheduling is used, it means that there are consistently (or at least on average) more jobs to be executed than there are computational resources to execute them during a given time interval. With a system that is designed for batch processing large parallel jobs, it is important that the scheduler is able to schedule these large parallel jobs while interspersing shorter, smaller jobs, including job arrays, in order to keep resource utilization high. Thus, HPC schedulers need sophisticated job scheduling algorithms in order to utilize the computational resources most effectively. Conversely, on-demand scheduling is ideal for interactive data analysis jobs because the user does not have to wait for his/her job to percolate through the batch queue before interacting with it. If the system is used mainly to process on-demand interactive jobs such as data analysis jobs, the resource manager must quickly update (with timely periodicity) the status of all of the nodes so that the job scheduler is able to quickly assign new on-demand jobs to resources. This responsiveness is measured by scheduler latency, which involves the latency of submission, queue management, resource identification, resource selection, resource allocation, job dispatch, and job termination.

Building on the model of Section 4, the experiment test plan is depicted in Fig. 3. We schedule n constant time tasks such that each executes in t time on P processor cores. The isolated job execution time is then

$$T_{\text{job}} = t \cdot n.$$

We measure the total time to execute the $n \cdot P$ tasks as T_{total} . We can then calculate the scheduler utilization as

$$U = T_{\text{job}} / T_{\text{total}}.$$

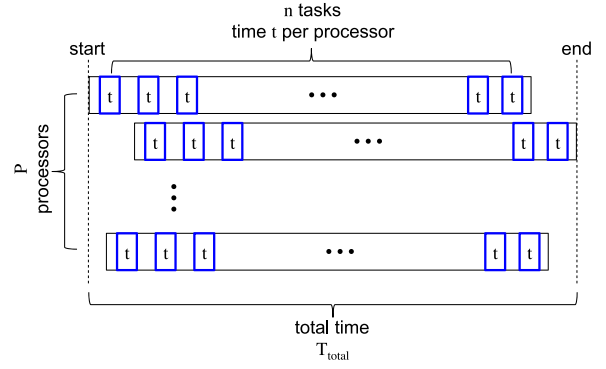


Fig. 3. Schematic of the execution of a job P processors, consisting of n tasks each executing for time t . The total time, T_{total} , includes both the task execution time along with all scheduler overheads.

This test plan allows us to determine the scaling behavior across schedulers for a large number of short-duration jobs by producing curves of utilization U versus task time. As we derived in Section 4, by measuring the constant time task utilization $U_c(t)$, we can estimate the utilization for any distribution of tasks from the average task times on each processor. Hence, we can be confident that any combination of variable task times can be estimated using these constant task time results.

One must keep in mind that we are evaluating the performance of the scheduler from the scheduler perspective. That is, we are executing constant time tasks to occupy a given job slot for a set amount of time. So for that job slot, the scheduler does not need to reevaluate its assignment until that occupation time has expired. The constant time tasks are inserting delays after which the scheduler must schedule and dispatch a new task to a given resource. There are naturally other considerations such as slowdown of task execution due to resource contention that affect the amount of time it takes to complete the task. However, such considerations are an issue of scheduling functionality effectiveness rather than scheduler latency.

5.1. Benchmarking environment

To minimize uncertainty, all measurements were carried out with our MIT SuperCloud cluster [56]. This cluster was fully isolated from extraneous processes in order to deliver consistent results. Forty-five nodes were used: one to serve as the scheduler node and forty-four as compute nodes (1408 cores total). All of the nodes were connected via a 10 GbE interface to a Lustre parallel storage array.

To compare the launch latency for the schedulers, we chose four representative schedulers from across the scheduler landscape: Slurm, Son of Grid Engine, Mesos, and Hadoop YARN. Each of the four schedulers was installed on the scheduler node and compute nodes, and the daemons of only the scheduler under test were running; i.e., only one scheduler was running at a time. Each of the schedulers had numerous parameters that needed to be tuned for optimal performance on short tasks. All four were configured to achieve optimal performance on short-duration jobs.

The version designators and scheduler configuration details are as follows. We used Slurm version 15.08.6 with the following parameter settings:

```
ProctrackType = proctrack/cgroup
SchedulerType = sched/builtin
SelectType = select/cons_res
SelectTypeParameters = CR_Core_Memory
PriorityType = priority/basic
```

Table 9

Parameter sets and runtimes used to measure scheduler latency as a function of job task time. Each task set was run three times for each scheduler, and all three times are listed. Note: The Hadoop YARN Rapid Task trial set was not executed because execution time was exceedingly long.

Configuration	Rapid Tasks	Fast tasks	Medium tasks	Long tasks
Task time t	1 s	5 s	30 s	60 s
Job time per processor T_{job}	240 s	240 s	240 s	240 s
Tasks per processor n	240	48	8	4
Processors P (cores)	1408	1408	1408	1408
Total tasks N	337,920	67,584	11,264	5632
Total processor time	93.7 h	93.7 h	93.7 h	93.7 h
Runtimes (sec)				
Slurm	2774, 2787, 2790	622, 603, 606	280, 278, 255	287, 264, 300
GE	3057, 3073, 3082	622, 634, 623	278, 279, 277	275, 281, 274
Mesos	1794, 1795, 1792	366, 364, 367	280, 280, 281	306, 306, 305
Hadoop YARN	–	2013, 1798, 1710	479, 472, 510	342, 445, 347

DefMemPerCPU = 2048

For the Grid Engine testing, we used Son of Grid Engine 8.1.8 and enabled the high-throughput configuration. With Mesos, we used Mesos version 0.25.0 with a single-node master and single ZooKeeper running on the master. Finally, with Apache Hadoop YARN, we used version 2.7.1 with a single-node master running NameNode and ResourceManager (YARN) daemons and compute nodes running DataNode and NodeManager daemons.

5.2. Latency and utilization measurements

The jobs that were launched on the 1408 cores were all sleep jobs of 1, 5, 30, or 60 s. The total number of tasks N and the number of tasks per processor n were set so that the total processing time was always the same: 93.7 h (337,920 s). The four task sets used for comparison are given in Table 9. For each task set and each scheduler, three trials were executed, and the results are the average of the three trials. The Hadoop YARN trials for rapid tasks were abandoned because it took too much time to execute because of excessive scheduler latency. With all four schedulers, the jobs were submitted as job arrays because they introduce much less scheduler latency than is introduced by submitting the same workload as individual jobs.

Each of the task sets is composed of only one duration length of jobs. As we determined from the scheduler latency model of the previous section, heterogeneous mixes of job durations can be composed of combinations of these task sets (and other in-between duration values); hence, the trends that we learn from these task sets can easily be used to deduce the scheduler latency performance of more heterogeneous mixes of job durations.

First, we analyzed the distribution of ΔT for each of the four schedulers under test as n , the number of tasks per processor, was increased. These results are shown in Fig. 4. Note that these plots are on log–log axes. Fig. 4(a) shows that Slurm has low scatter among trials at higher values of n , and the model is a good fit at higher values of n . At lower values of n , shot noise impacts the model. For Grid Engine, there is low scatter among trials at all values of n , and the model is a good fit at most values of n , as seen in Fig. 4(b). Similar to Slurm, shot noise impacts the model at lower values of n . In Fig. 4(c) for Mesos, there is very low scatter among trials at all values of n , and the model is a reasonable fit at most values of n , though there is some curvature away from the power law. There is some shot noise impact on the model at lower values of n . Finally, in Fig. 4(d) for the Hadoop YARN scheduler, there is some scatter among trials at all values of n , and the model is a reasonable fit at all values of n . At the highest values of n , large marginal latency limits the ability to conduct experiments at all.

It is also interesting to examine the model fit parameters for each of the schedulers. The model fit parameters are shown in Table 10, where the second column is the y -axis crossing points and

Table 10

Measured model fit parameter comparison for scheduler latency model.

Scheduler	Marginal latency t_s (smaller is better)	Nonlinear exponent α_s (smaller is better)
Slurm	2.2 s	1.3
Grid Engine	2.8 s	1.3
Mesos	3.4 s	1.1
Hadoop YARN	33 s	1.0

the third column is the angle of the fit line in the log–log plot. For this set of experiments, Slurm has the best marginal latency, while Grid Engine and Mesos have acceptable marginal latency. This result suggests that these three schedulers do not use excessive overhead to launch jobs. Mesos and Hadoop YARN have the best nonlinear exponent, suggesting that they have more uniform per-task launch latency overhead.

Fig. 5 shows the results of launch latency as measured by the utilization defined along with the approximate and more exact utilization models derived previously. In both cases, the models qualitatively agree with the data and indicate they can be used as reasonable approximations in theoretical utilization analyses. Since we have a fixed total amount of work, all additional time that it takes the cluster to execute the total set of jobs decreases the utilization of the cluster. Simply put, any time the scheduler consumes time in managing job execution (submission, queue management, resource identification, resource selection, resource allocation, job dispatch, and job termination) will decrease utilization. As we can see in Fig. 5, all of the schedulers do well with 60-second tasks with the exception of Hadoop YARN; all of the schedulers launch jobs reasonably fast. Hadoop YARN has greater overhead for each job, including launching an application master process for each job [68]. Slurm, Grid Engine, and Mesos perform similarly with 1-, 5-, and 30-second tasks. However, Hadoop YARN is much less efficient among these schedulers; it was so inefficient that the 1-second task trials were prohibitive to run. Each of the schedulers loses efficiency as it must manage large queues of pending jobs, determine the best candidate job to launch next, and manage the launching and finishing of executing jobs. Besides the poor performance from Hadoop YARN, Slurm, Grid Engine, and Mesos show much lower utilization rates for 1- and 5-second jobs. This lackluster performance beg for a solution to realize better utilization.

5.3. Multilevel scheduling

The key to increasing the utilization for 1- and 5-second jobs is to decrease the job launch latency or not launch as many jobs overall while still getting all of the work done. With pleasantly parallel jobs, which most data analysis jobs are, the analysis code can be changed slightly to allow processing of multiple datasets

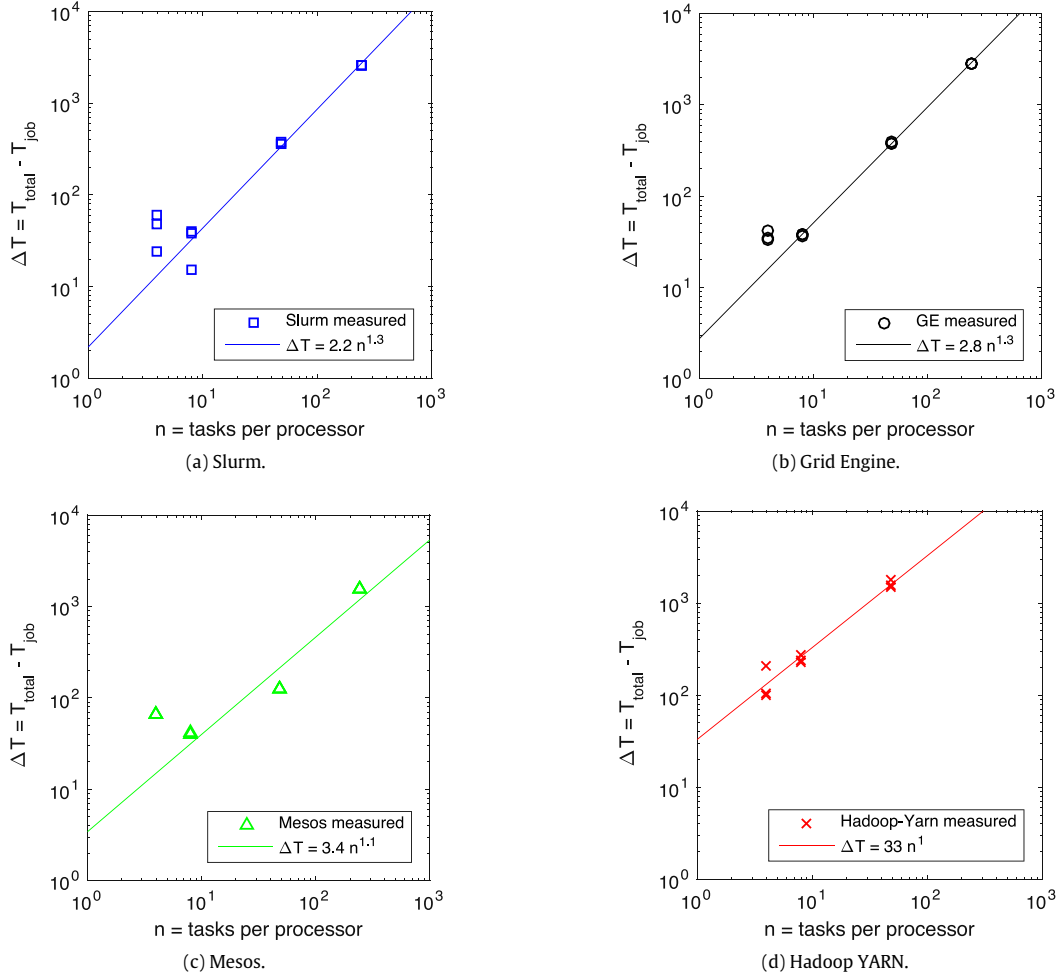


Fig. 4. Comparison of scheduler latency versus number of tasks per processor.

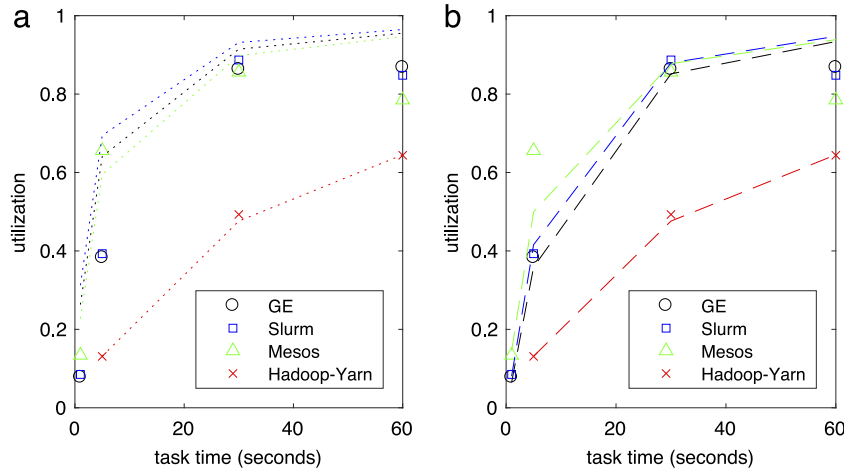


Fig. 5. Scheduler utilization as a function of varying task times. Points denote measured times. (a) Dotted lines show approximate utilization model given by $U_c(t)^{-1} \approx 1 + t_s/t$. (b) Dashed lines show utilization model given by $U_c^{-1} = 1 + (t_s n^{qs})/(tn)$.

or files with a single job execution. We can then use a tool like LLMaReduce [12] to efficiently run the jobs on the computing resource. This technique is referred to as multilevel scheduling.

LLMaReduce is a tool that greatly simplifies the setup and execution of processing many data files with the same program. It employs the map-reduce parallel programming model, which consists of two user-written programs: a Mapper and a Reducer.

The input to the Mapper is a file, and the output is another file. The input to the Reducer is the set of Mapper output files. The output of Reducer is a single file. Launching consists of starting many Mapper programs, each with a different input file. When the Mapper programs all have completed, the Reduce program is run on the Mapper outputs. By default, LLMaReduce expects that the map application takes a single input and single output path

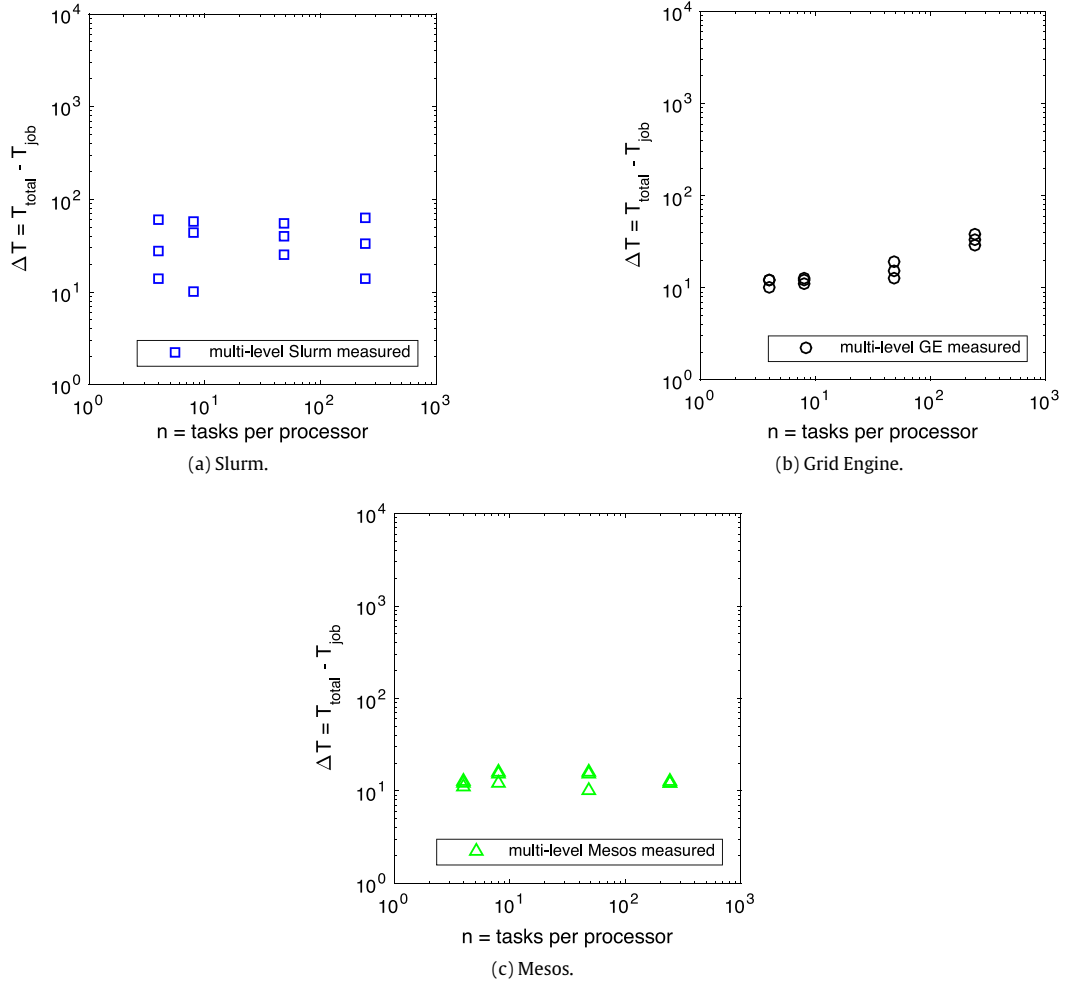


Fig. 6. Comparison of scheduler latency versus number of tasks per processor with multilevel scheduling.

(siso). However, this siso approach will incur overhead associated with repeated startups of the map application. Some applications, such as MATLAB codes, can save significant overhead cost with the minor change of having the map application start only once and process many input/output file pairs. Using the multiple input and multiple output mode requires a (mildly) modified map application that will read the input file with the multiple lines of input/output filename pairs [12].

Again, we first analyzed the distribution of ΔT for each of the four schedulers under test as n , the number of tasks per processor, is increased. These results are shown in Fig. 6. Again, note that these plots are on log–log axes. Fig. 6(a) shows significant scatter among trials at all values of n with the Slurm scheduler with no apparent ΔT dependence upon n . Further, multilevel scheduling with Slurm reduces ΔT to below 100 s for all values of n , and there is a $30\times$ reduction in ΔT at the largest values of n when compared to not using multilevel scheduling. Grid Engine shows low scatter among trials at all values at n in Fig. 6(b), and there is a small increase in ΔT as n increases. Here, multilevel scheduling reduces ΔT to below 40 s at all values of n ; there is a $40\times$ reduction for ΔT at the largest values of n when compared to non-multilevel scheduling. Finally for multilevel scheduling with Mesos, Fig. 6(c) shows that there is low scatter among trials at all values of n , and there is no apparent ΔT dependence on n . Multilevel scheduling in Mesos reduces ΔT to below 20 s at all values of n , and there is a $100\times$ reduction in ΔT at the largest values of n when compared to non-multilevel scheduling. All of these results show that if pleasantly parallel

computational jobs are amenable to using multilevel scheduling techniques, the effort to cast them as such is well worth it.

The same four task sets from Table 9 in Section 5.2 were used with the LLMapReduce tool to test the utilization of multilevel scheduling on Slurm, Grid Engine, and Mesos. Fig. 7 shows the same results as Fig. 5 and includes the results of the use of multilevel scheduling. Slurm, Grid Engine, and Mesos all have high utilization with this technique. The figure shows that multilevel scheduling brings the utilization rates for all three schedulers around 90%, which is on par with the 30- and 60-second jobs.

6. Summary

Schedulers are a primary component of a scalable computing infrastructure because they control all of the work on the system and directly impact overall system effectiveness. Schedulers have many features, and certain features can enhance the execution of certain workloads. In this paper, we have compared a number of scalable compute cluster schedulers and developed the concept of scheduler families based on their lineage and features. We then compared and contrasted a number of key features and their impact on high performance data analytic workloads. We found that, in general, there are many common features across all the schedulers that we examined.

We then focused on scheduler latency, creating a scheduler latency model and validating the model with a set of timing measurements on a subset of schedulers: Slurm, Grid Engine, Hadoop

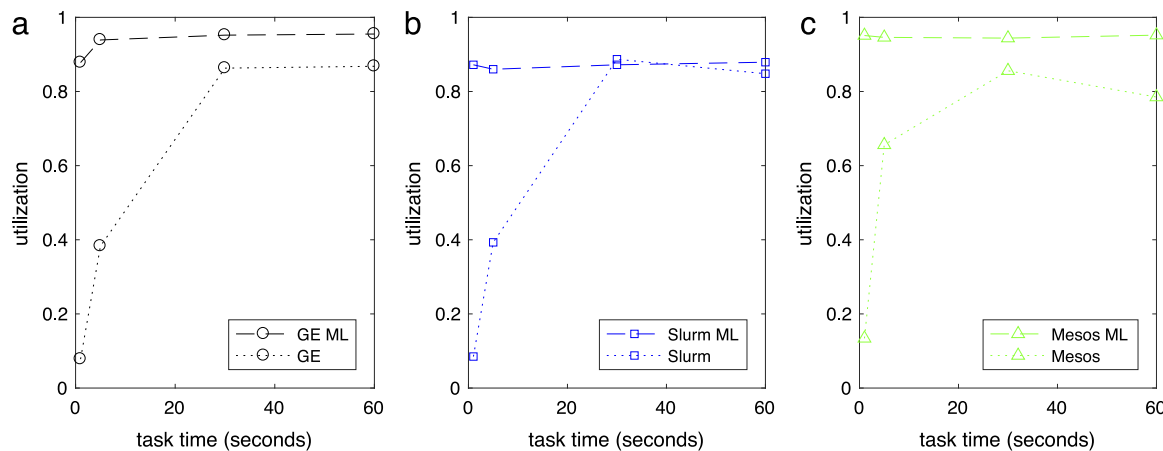


Fig. 7. Utilization as function of tasks for regular and multilevel scheduling for (a) Grid Engine (b) Slurm, and (c) Mesos. In all cases multilevel scheduling significantly improves utilization.

YARN, and Mesos. We found that Hadoop YARN performed worse than the other three schedulers and that the other three performed similarly. We then used the multilevel scheduling technique with the LLMaReduce tool to increase the overall utilization from Slurm, Grid Engine, and Mesos.

Acknowledgments

The authors wish to acknowledge the following individuals for their contributions: Alan Edelman, Justin Brukardt, Steve Pritchard, Chris Clarke, Sterling Foster, Paul Burkhardt, Victor Roytburd, Dave Martinez, Vijay Gadepally, Anna Klein, Lauren Milechin, Julie Mullen, Sid Samsi, and Chuck Yee.

This material is based upon work supported by the National Science Foundation under Grant No. DMS-1312831. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

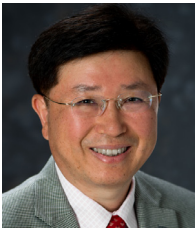
References

- [1] S. Ali, T.D. Braun, H.J. Siegel, A.A. Maciejewski, N. Beck, L. Bölöni, M. Maheswaran, A.I. Reuther, J.P. Robertson, M.D. Theys, et al., Characterizing resource allocation heuristics for heterogeneous computing systems, *Adv. Comput.* 63 (2005) 91–128.
- [2] M. Baker, G.C. Fox, H.W. Yau, A Review of Commercial and Research Cluster Management Software. Tech. Rep. Paper 19, Northeast Parallel Architecture Center, 1996.
- [3] A.A. Bhattacharya, D. Culler, E. Friedman, A. Ghodsi, S. Shenker, I. Stoica, Hierarchical scheduling for diverse datacenter workloads, in: *Proceedings of the 4th Annual Symposium on Cloud Computing, SOCC '13*, ACM, New York, NY, USA, 2013, pp. 4:1–4:15. URL <http://doi.acm.org/10.1145/2523616.2523637>.
- [4] S.L. Bird, B. Smith, Pacora: Performance aware convex optimization for resource allocation, in: *Proceedings of the 3rd USENIX Workshop on Hot Topics in Parallelism, HotPar: Posters*, 2011, pp. 1–6.
- [5] S.L. Bird, Spring, Optimizing Resource Allocations for Dynamic Interactive Applications, (Ph.D. thesis), University of California, Berkeley, 2014.
- [6] E. Boutin, J. Ekanayake, W. Lin, B. Shi, J. Zhou, Z. Qian, M. Wu, L. Zhou, Apollo: Scalable and coordinated scheduling for cloud-scale computing, in: *11th USENIX Symposium on Operating Systems Design and Implementation, OSDI 14*, USENIX Association, Broomfield, CO, 2014, pp. 285–300. URL <https://www.usenix.org/conference/osdi14/technical-sessions/presentation/boutin>.
- [7] T.D. Braun, H.J. Siegel, N. Beck, L.L. Bölöni, M. Maheswaran, A.I. Reuther, J.P. Robertson, M.D. Theys, B. Yao, D. Hensgen, R.F. Freund, A comparison of eleven static heuristics for mapping a class of independent tasks onto heterogeneous distributed computing systems, *J. Parallel Distrib. Comput.* 61 (6) (2001) 810–837. URL <http://www.sciencedirect.com/science/article/pii/S0743731500917143>.
- [8] T.D. Braun, H.J. Siegel, A.A. Maciejewski, Y. Hong, Static resource allocation for heterogeneous computing environments with tasks having dependencies, priorities, deadlines, and multiple versions, *J. Parallel Distrib. Comput.* 68 (11) (2008) 1504–1516.
- [9] D. Brelford, G. Chochia, N. Falk, K. Marthi, R. Sure, N. Bobroff, L. Fong, S. Seelam, Partitioned parallel job scheduling for extreme scale computing, in: W. Cirne, N. Desai, E. Frachtenberg, U. Schwiegelshohn (Eds.), *Job Scheduling Strategies for Parallel Processing: 16th International Workshop, JSSPP 2012*, Shanghai, China, May 25, 2012. Revised Selected Papers, Springer, Berlin, Heidelberg, 2013, pp. 157–177. URL http://dx.doi.org/10.1007/978-3-642-35867-8_9.
- [10] B. Burns, B. Grant, D. Oppenheimer, E. Brewer, J. Wilkes, Borg, Omega, and Kubernetes, *Commun. ACM* 59 (5) (2016) 50–57. URL <http://doi.acm.org/10.1145/2890784>.
- [11] C. Byun, C. Duncan, S. Burks, Fall, A comparison of job management systems in supporting hpc clustertools, in: *Technical Forum of Sun Users Performance Group-SUPERG*, 2000.
- [12] C. Byun, J. Kepner, W. Arcand, D. Bestor, W. Bergeron, V. Gadepally, M. Hubbell, P. Michaleas, J. Mullen, A. Prout, A. Rosa, C. Yee, A. Reuther, LLMaReduce: Multi-level Map-Reduce for high performance data analytics, in: *High Performance Extreme Computing Conference, HPEC, 2016*, IEEE, 2016, pp. 1–6.
- [13] L.-C. Canon, L. Philippe, On the Heterogeneity Bias of Cost Matrices when Assessing Scheduling Algorithms. Tech. Rep. RR-FEMTO-ST-8663, FEMTO-ST, 2015.
- [14] L.-C. Canon, L. Philippe, On the heterogeneity bias of cost matrices when assessing scheduling algorithms, in: *European Conference on Parallel Processing*, Springer, 2015, pp. 109–121.
- [15] N. Capit, G.D. Costa, Y. Georgiou, G. Huard, C. Martin, G. Mounie, P. Neyron, O. Richard, A batch scheduler with high level components, in: *CCGrid 2005 IEEE International Symposium on Cluster Computing and the Grid*, Vol. 2, 2005, pp. 776–783.
- [16] Cloudera, Llama –low latency application master, 2016. URL <http://cloudera.github.io/llama/index.html>.
- [17] K. Czajkowski, I. Foster, N. Karonis, C. Kesselman, S. Martin, W. Smith, S. Tuecke, A resource management architecture for metacomputing systems, in: *Workshop on Job Scheduling Strategies for Parallel Processing*, Springer, 1998, pp. 62–82.
- [18] J. Dean, S. Ghemawat, MapReduce: simplified data processing on large clusters, *Commun. ACM* 51 (1) (2008) 107–113.
- [19] C. Delimitrou, D. Sanchez, C. Kozyrakis, Tarcil: Reconciling scheduling speed and quality in large shared clusters, in: *Proceedings of the Sixth ACM Symposium on Cloud Computing, SoCC '15*, ACM, New York, NY, USA, 2015, pp. 97–110. URL <http://doi.acm.org/10.1145/2806777.2806779>.
- [20] J. Dittrich, J.-A. Quiané-Ruiz, Efficient big data processing in Hadoop MapReduce, *Proc. VLDB Endowment* 5 (12) (2012) 2014–2015.
- [21] T. El-Ghazawi, K. Gaj, N. Alexandridis, F. Vroman, N. Nguyen, J.R. Radzikowski, P. Samipagdi, S.A. Suboh, A performance study of job management systems, *Concurr. Comput. Pract. Exp.* 16 (13) (2004) 1229–1246.
- [22] M. Elkourie, HPC Workload Management Tools: A Competitive Benchmark Study. Tech. Rep., The Edison Group, 2014.
- [23] D.H.J. Epema, M. Livny, R. van Dantzig, X. Evers, J. Pruyne, A worldwide flock of condors: Load sharing among workstation clusters, *Future Gener. Comput. Syst.* 12 (1) (1996) 53–65. URL [http://dx.doi.org/10.1016/0167-739X\(95\)00035-Q](http://dx.doi.org/10.1016/0167-739X(95)00035-Q).
- [24] Y. Etsion, D. Tsafir, A Short Survey of Commercial Cluster Batch Schedulers. Tech. Rep. 44221: 2005–13, School of Computer Science and Engineering, The Hebrew University of Jerusalem, 2005.
- [25] D.G. Feitelson, Workload modeling for performance evaluation, in: *IFIP International Symposium on Computer Performance Modeling, Measurement and Evaluation*, Springer, Berlin, Heidelberg, 2002, pp. 114–141.

- [26] D. Feitelson, M. Jette, Improved utilization and responsiveness with gang scheduling, in: *Job Scheduling Strategies for Parallel Processing*, Springer, Berlin/Heidelberg, 1997, pp. 238–261.
- [27] D.G. Feitelson, L. Rudolph, Parallel job scheduling: Issues and approaches, in: *Workshop on Job Scheduling Strategies for Parallel Processing*, Springer, 1995, pp. 1–18.
- [28] D.G. Feitelson, L. Rudolph, U. Schwiegelshohn, Parallel job scheduling—a status report, in: *Workshop on Job Scheduling Strategies for Parallel Processing*, Springer, Berlin, Heidelberg, 2004, pp. 1–16.
- [29] I. Foster, C. Kesselman, *The Grid 2: Blueprint for a New Computing Infrastructure*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2003.
- [30] W. Gentzsch, Codine, Computing in Distributed Networked Environments, User's Guide and Reference Manual. Genias Software GmbH, Erzgebirgstr. 2B, D-93073 Neutraubling, Germany, 1994.
- [31] V. Getov, et al., Workload schedulers—genesis, algorithms and comparisons, *Int. J. Comput. Sci. Softw. Eng.* 4 (6) (2015) 141–155.
- [32] I. Gog, M. Schwarzkopf, A. Gleave, R.N.M. Watson, S. Hand, Firmament: Fast, centralized cluster scheduling at scale, in: *12th USENIX Symposium on Operating Systems Design and Implementation*, OSDI 16, USENIX Association, CA, 2016, pp. 99–115. URL <https://www.usenix.org/conference/osdi16/technical-sessions/presentation/gog>.
- [33] A.S. Grimshaw, W.A. Wulf, The legion vision of a worldwide virtual computer, *Commun. ACM* 40 (1) (1997) 39–45.
- [34] R.L. Henderson, Job scheduling under the portable batch system, in: *Workshop on Job Scheduling Strategies for Parallel Processing*, Springer, 1995, pp. 279–294.
- [35] B. Hindman, A. Konwinski, M. Zaharia, A. Ghodsi, A.D. Joseph, R.H. Katz, S. Shenker, I. Stoica, Mesos: A platform for fine-grained resource sharing in the data center, in: *NSDI*, Vol. 11, 2011, pp. 295–308.
- [36] H. Hussain, S.U.R. Malik, A. Hameed, S.U. Khan, G. Bickler, N. Min-Allah, M.B. Qureshi, L. Zhang, W. Yongji, N. Ghani, J. Kolodziej, A.Y. Zomaya, C.-Z. Xu, P. Balaji, A. Vishnu, F. Pinel, J.E. Pecero, D. Kliazovich, P. Bouvry, H. Li, L. Wang, D. Chen, A. Rayes, A survey on resource allocation in high performance distributed computing systems, *Parallel Comput.* 39 (11) (2013) 709–736. URL <http://www.sciencedirect.com/science/article/pii/S016781911300121X>.
- [37] M. Isard, V. Prabhakaran, J. Currey, U. Wieder, K. Talwar, A. Goldberg, Quincy: fair scheduling for distributed computing clusters, in: *Proceedings of the ACM SIGOPS 22nd Symposium on Operating Systems Principles*, SOS '09, ACM, New York, NY, USA, 2009, pp. 261–276. URL <http://doi.acm.org/10.1145/1629575.1629601>.
- [38] D. Jackson, Q. Snell, M. Clement, Core algorithms of the maui scheduler, in: *Workshop on Job Scheduling Strategies for Parallel Processing*, Springer, 2001, pp. 87–102.
- [39] J.P. Jones, NAS Requirements Checklist for Job Queuing/Scheduling Software. Tech. Rep. NAS-96-003, NAS High Performance Processing Group, NASA Ames Research Center, Moffett Field, CA, 1996.
- [40] J.P. Jones, Evaluation of Job Queuing/Scheduling Software: Phase 1 Report. Tech. Rep. NAS-96-009, NAS High Performance Processing Group, NASA Ames Research Center, 1996.
- [41] J.P. Jones, C. Brickell, Evaluation of Job Queuing/Scheduling Software: Phase 1 Report. Tech. Rep. NAS-97-013, NAS High Performance Processing Group, NASA Ames Research Center, 1997.
- [42] S. Kannan, M. Roberts, P. Mayes, D. Brelsford, J.F. Skovira, IBM LoadLeveler: User's Guide. No. SH26-7226-00 in IBM Redbooks, IBM Corporation, 1993.
- [43] S. Kannan, M. Roberts, P. Mayes, D. Brelsford, J.F. Skovira, Workload Management with LoadLeveler, IBM Corporation, 2001.
- [44] K. Karanasos, S. Rao, C. Curino, C. Douglas, K. Chaliparambil, G.M. Fumarola, S. Heddaya, R. Ramakrishnan, S. Sakalanaga, Mercury: Hybrid centralized and distributed scheduling in large shared clusters, in: *USENIX Annual Technical Conference*, 2015, pp. 485–497.
- [45] B.K. Kingsbury, The Network Queueing System Tech. Rep. NASA-CR-177433, NASA, 1986.
- [46] K. Krauter, R. Buyya, M. Maheswaran, A taxonomy and survey of grid resource management systems for distributed computing, *Softw. - Pract. Exp.* 32 (2) (2002) 135–164.
- [47] J. Leung, L. Kelly, J.H. Anderson, *Handbook of Scheduling: Algorithms, Models, and Performance Analysis*, CRC Press, Inc., Boca Raton, FL, USA, 2004.
- [48] D.A. Lifka, The ANL/IBM SP scheduling system, in: *Workshop on Job Scheduling Strategies for Parallel Processing*, Springer, 1995, pp. 295–303.
- [49] J.-C. Lin, M.-C. Lee, Performance evaluation of job schedulers on Hadoop YARN, *Concurr. Comput. Pract. Exp.* 28 (9) (2016) 2711–2728. cPE-15-0190. R1URL <http://dx.doi.org/10.1002/cpe.3736>.
- [50] M.J. Litzkow, M. Livny, M.W. Mutka, Condor—a hunter of idle workstations, in: *8th International Conference on Distributed Computing Systems*, IEEE, 1988, pp. 104–111.
- [51] J. Nabrzyski, J.M. Schopf, J. Weglarz (Eds.), *Grid Resource Management: State of the Art and Future Trends*, Kluwer Academic Publishers, Norwell, MA, USA, 2004.
- [52] T. Newhouse, J. Pasquale, ALPS: An application-level proportional-share scheduler, in: *HPDC*, 2006, pp. 279–290.
- [53] J.K. Ousterhout, *Scheduling Techniques for Concurrent Systems*, 1982, pp. 22–30.
- [54] K. Ousterhout, P. Wendell, M. Zaharia, I. Stoica, Sparrow: distributed, low latency scheduling, in: *Proceedings of the Twenty-Fourth ACM Symposium on Operating Systems Principles*, ACM, 2013, pp. 69–84.
- [55] A.I. Reuther, T. Currie, J. Kepner, H.G. Kim, A. McCabe, P. Michaleas, N. Travinin, Technology requirements for supporting on-demand interactive grid computing, in: *2005 Users Group Conference*, DOD-UGC'05, 2005, pp. 320–327.
- [56] A. Reuther, J. Kepner, W. Arcand, D. Bestor, B. Bergeron, C. Byun, M. Hubbell, P. Michaleas, J. Mullen, A. Prout, A. Rosa, LLSuperCloud: Sharing HPC systems for diverse rapid prototyping, in: *High Performance Extreme Computing Conference*, HPEC, 2013, IEEE, 2013, pp. 1–6.
- [57] A. Reuther, P. Michaleas, A. Prout, J. Kepner, HPC-VMs: Virtual machines in high performance computing systems, in: *High Performance Extreme Computing*, HPEC, 2012 IEEE Conference on, 2012, pp. 1–6.
- [58] B. Saha, H. Shah, S. Seth, G. Vijayaraghavan, A. Murthy, C. Curino, Apache Tez: A unifying framework for modeling and building data processing applications, in: *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*, ACM, 2015, pp. 1357–1369.
- [59] W. Saphir, L.A. Tanner, B. Traversat, Job management requirements for NAS parallel systems and clusters, in: *Workshop on Job Scheduling Strategies for Parallel Processing*, Springer, 1995, pp. 319–336.
- [60] M. Schwarzkopf, A. Konwinski, M. Abd-El-Malek, J. Wilkes, Omega: flexible, scalable schedulers for large compute clusters, in: *Proceedings of the 8th ACM European Conference on Computer Systems*, ACM, 2013, pp. 51–364.
- [61] U. Schwiegelshohn, R. Yahyapour, Analysis of first-come-first-serve parallel job scheduling, in: *SODAm*, Vol. 98, Citeseer, 1998, pp. 629–638.
- [62] U. Schwiegelshohn, R. Yahyapour, Fairness in parallel job scheduling, *J. Sched.* 3 (5) (2000) 297–320.
- [63] K. Shvachko, H. Kuang, S. Radia, R. Chansler, The Hadoop distributed file system, in: *Proceedings of the 2010 IEEE 26th Symposium on Mass Storage Systems and Technologies*, MSST, MSST '10, IEEE Computer Society, Washington, DC, USA, 2010, pp. 1–10. URL <http://dx.doi.org/10.1109/MSST.2010.5496972>.
- [64] S. Singh, I. Chana, A survey on resource scheduling in cloud computing: Issues and challenges, *J. Grid Comput.* 14 (2) (2016) 217–264.
- [65] P. Slapničar, U. Seitz, A. Bode, I. Zoraja, Resource management in message passing environments, *J. Comput. Inf. Technol.* CIT 9 (1) (2001) 43–54.
- [66] V.K. Vavilapalli, A.C. Murthy, C. Douglas, S. Agarwal, M. Konar, R. Evans, T. Graves, J. Lowe, H. Shah, S. Seth, et al., Apache hadoop YARN: Yet another resource negotiator, in: *Proceedings of the 4th Annual Symposium on Cloud Computing*, ACM, 2013, p. 5.
- [67] A. Verma, L. Pedrosa, M. Korupolu, D. Oppenheimer, E. Tune, J. Wilkes, Large-scale cluster management at google with borg, in: *Proceedings of the Tenth European Conference on Computer Systems*, ACM, 2015, p. 18.
- [68] T. White, *Hadoop: The Definitive Guide*, fourth ed., O'Reilly, Sebastopol, CA, 2015.
- [69] M.G. Xavier, M.V. Neves, C.A.F. De Rose, A performance comparison of container-based virtualization systems for mapreduce clusters, in: *2014 22nd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing*, IEEE, 2014, pp. 299–306.
- [70] Y. Yan, B. Chapman, Comparative Study of Distributed Resource Management Systems—SGE, LSF, PBS Pro, and LoadLeveler Tech. Rep. En: Technical Report, 2008.
- [71] A.B. Yoo, M.A. Jette, M. Grondona, Slurm: Simple Linux utility for resource management, in: *Workshop on Job Scheduling Strategies for Parallel Processing*, Springer, 2003, pp. 44–60.
- [72] S. Zhou, X. Zheng, J. Wang, P. Delisle, Utopia: a load sharing facility for large, heterogeneous distributed computer systems, *Softw. - Pract. Exp.* 23 (12) (1993) 1305–1336.
- [73] X. Zhou, H. Chen, K. Wang, M. Lang, I. Raicu, Exploring Distributed Resource Allocation Techniques in the Slurm Job Management System, Technical Report, Illinois Institute of Technology, Department of Computer Science, 2013.



Dr. Albert I. Reuther is the Manager of the MIT Lincoln Laboratory Supercomputing Center (LLSC). He brought supercomputing to Lincoln Laboratory through the establishment of LLGrid, and he leads the LLSC computational science and engineering team. He developed gridMatlab and is the system architect of the MIT Supercloud along with numerous interactive supercomputing clusters based on Supercloud including those in the LLSC. His areas of research include interactive high performance computing, parallel signal processing, machine learning, and graph analytics. Dr. Reuther earned a dual BS Comp/EE, MSEE, and Ph.D. ECE (1994, 1996, 2000), all from Purdue University. He also earned an MBA degree (2001) from the Collège des Ingénieurs in Paris, France, and Stuttgart, Germany.



Dr. Chansup Byun is a Computational Scientist in the MIT Lincoln Laboratory Supercomputing Center. He primarily supports LLSC users to be successful and productive in using the LLSC resources. He is also involved in maintaining, developing, and porting the schedulers, pMatlab, gridMatlab, and LLMapReduce software in addition to other open-source software for various computing environments. He has a Ph.D. in Aerospace Engineering from Virginia Tech (1991) and M.S. (1986) and B.S. (1984) in Aerospace Engineering from Seoul National University in South Korea.



William Arcand is a Systems Engineer in the MIT Lincoln Laboratory Supercomputing Center. His primary role at the LLSC is to design and maintain the various high speed parallel file system and storage assets for all LLSC related clusters. He also assists in the general system administration and support for all LLSC clusters, as well as providing expert advice on designing and maintaining high speed parallel file systems for various application system teams at Lincoln.



David Bestor is a Systems Engineer in the MIT Lincoln Laboratory Supercomputing Center. He primarily handles maintenance for the various storage systems in the LLSC. He is also involved with system administration of the TX-Green HPC cluster. He received a Bachelor of Science in Information Technology from University of Massachusetts at Lowell in 2005 and a Master of Science in Information Technology from the University of Massachusetts at Lowell in 2014.



Bill Bergeron is a Systems Engineer in the MIT Lincoln Laboratory Supercomputing Center. Since 2007, Mr. Bergeron has been supporting the LLSCs effort to provide High Performance Computing for research at MIT LL. His primary task has been to build and maintain systems for unclassified, classified, external, development, satellite, and MIT campus environments. He has been an integral part of the development of the LLSC software stack and system imaging platform providing consistency and performance integrity across all supported systems. He also is the primary support for the research computing collaborative software and code repository systems. He received a Bachelors degree in Mechanical Engineering from the University of Massachusetts at Amherst in 1991.



Matthew L. Hubbell is a Systems Engineer with the Lincoln Laboratory Supercomputing Center. Since joining the Laboratory in 2006, his activities include building High Performance Computing (HPC) Systems as member of the LLSC Team. He currently oversees operations of the Holyoke Field site, the home of TX-Green, Lincoln Laboratories general purpose Lab-wide HPC system located in Holyoke, MA. His other activities include the development of immersive 3-D environments leveraging video game technology to provide a multi-player platform to achieve situational awareness in diverse environments. He received a Bachelor's degree business management from Gettysburg College in 2002 and a Master's in Business Administration from the University of Massachusetts at Lowell in 2010.



Michael Jones is a Systems Engineer with the Lincoln Laboratory Supercomputing Center. Since joining the Laboratory in 2015, he has focused on the design and architecture of High Performance Computing (HPC) systems, developing innovative tools and techniques for managing supercomputing resources and leveraging novel prototype hardware to further the needs of the research community and maintain the LLSC's leadership position in the area of interactive supercomputing. His other interests include supporting efforts to develop and integrate novel hardware and algorithms to aid in understanding the complex relationships between diverse data feeds and advance the state of the art in graph analytics on extremely large datasets.



Peter Michaleas is the Deputy Manager of the Lincoln Laboratory Supercomputing Center. His focus is on systems architecture and execution of the LLSC vision. He leads LLSC's award-winning operations team. Peter has overseen the growth of LLSC computing capabilities from a small number of computers to the largest number of computers in the region spread across several world-class data centers.

Mr. Michaleas joined the Laboratory in 2004 and has been involved in many mission areas including Advanced Networks, Embedded Systems, Cyber Security, and High Performance Computing. He has co-authored over a dozen peer reviewed papers on supercomputing and big data systems.



Andrew Prout is a Systems Engineer in the MIT Lincoln Laboratory Supercomputing Center. He developed the dynamic database management system, dynamic virtual machine system, dynamic web application portal, and user-based firewall technologies used to provide advanced supercomputing capabilities to Lincoln Laboratory staff. He has contributed to many open source projects and is experienced with low-level systems and kernel programming.

He holds a CISSP certification and a Bachelor of Science in Information Technology Security.



Antonio Rosa is a Systems Engineer in the MIT Lincoln Laboratory Supercomputing Center with over 20 years of experience in Information Technology. He has a strong background in Computer Information Management, Administration, Analysis and Security. He has achieved various certifications in the IT field including CISSP (Certified Information Systems Security Professional), Network+, and CIW (Certified Internet Webmaster). Antonio Rosa acquired a B.S. Computer Information Systems from the State University of New York and is involved in the operations and logistics of the Holyoke, MA EcoPOD site.



Dr. Jeremy Kepner is an MIT Lincoln Laboratory Fellow who brought supercomputing to Lincoln Laboratory through the establishment of LLGrid and who led the establishment of the Massachusetts Green High Performance Computing Center. He developed a novel database management language and schema (Dynamic Distributed Dimensional Data Model, D4M). He is the architect of pMatlab (Parallel Matlab Toolbox), which is used by hundreds of Lincoln Laboratory staff and tens of thousands of scientists and engineers worldwide.

Dr. Kepner holds a BA degree in astrophysics from Pomona College and a Ph.D. degree in astrophysics from Princeton University.