

HPC Cloud for Scientific and Business Applications: Taxonomy, Vision, and Research Challenges

MARCO A. S. NETTO, IBM Research

RODRIGO N. CALHEIROS, Western Sydney University

EDUARDO R. RODRIGUES and RENATO L. F. CUNHA, IBM Research

RAJKUMAR BUYYA, University of Melbourne

High performance computing (HPC) clouds are becoming an alternative to on-premise clusters for executing scientific applications and business analytics services. Most research efforts in HPC cloud aim to understand the cost benefit of moving resource-intensive applications from on-premise environments to public cloud platforms. Industry trends show that hybrid environments are the natural path to get the best of the on-premise and cloud resources—steady (and sensitive) workloads can run on on-premise resources and peak demand can leverage remote resources in a pay-as-you-go manner. Nevertheless, there are plenty of questions to be answered in HPC cloud, which range from how to extract the best performance of an unknown underlying platform to what services are essential to make its usage easier. Moreover, the discussion on the right pricing and contractual models to fit small and large users is relevant for the sustainability of HPC clouds. This article brings a survey and taxonomy of efforts in HPC cloud and a vision on what we believe is ahead of us, including a set of research challenges that, once tackled, can help advance businesses and scientific discoveries. This becomes particularly relevant due to the fast increasing wave of new HPC applications coming from big data and artificial intelligence.

CCS Concepts: • **Computing methodologies** → **Massively parallel and high-performance simulations**; • **Computer systems organization** → **Cloud computing**; • **Networks** → *Cloud computing*; • **Software and its engineering** → Massively parallel systems;

Additional Key Words and Phrases: HPC cloud, high performance computing, parallel applications, resource allocation, charging models, advisory systems, big data

ACM Reference format:

Marco A. S. Netto, Rodrigo N. Calheiros, Eduardo R. Rodrigues, Renato L. F. Cunha, and Rajkumar Buyya. 2018. HPC Cloud for Scientific and Business Applications: Taxonomy, Vision, and Research Challenges. *ACM Comput. Surv.* 51, 1, Article 8 (January 2018), 29 pages.
<https://doi.org/10.1145/3150224>

1 INTRODUCTION

In the early 1990s, clusters of computers (Buyya 1999; Sterling 2002) became popular in high performance computing (HPC) environments due to their low cost compared to traditional supercomputers and mainframes. Computers with high processing power, fast network connections,

This work was partially supported by FINEP/MCTI under grant 03.14.0062.00.

Authors' addresses: M. A. S. Netto, E. R. Rodrigues, and R. L. F. Cunha, IBM Research, Rua Tutoia 1157, Sao Paulo, Brazil; emails: {mstelmar, edrodri, renatoc}@br.ibm.com; R. N. Calheiros, Western Sydney University, Locked Bag 1797 Penrith, NSW 2150 Australia; email: R.Calheiros@westernsydney.edu.au; R. Buyya, University of Melbourne, Doug McDonnell Building, Melbourne, VIC 3010, Australia; email: rbuyya@unimelb.edu.au.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 ACM 0360-0300/2018/01-ART8 \$15.00

<https://doi.org/10.1145/3150224>

and Linux were fundamental for this shift to occur. To this day, these clusters can handle complex computational problems in industries such as aerospace, life sciences, finance, and energy. They are managed by *batch schedulers* (Feitelson et al. 1997) that receive user requests to run jobs, which are queued whenever resources are under heavy use. Considering that service-level agreements (SLAs) are usually not in place in these environments, users have no visibility or concerns on costs of running jobs. However, large clusters do incur expenses and, when not properly managed, can generate resource wastage and poor quality of service.

Motivated by the different utilization levels of clusters around the globe and by the need to run even larger parallel programs, in the early 2000s, grid computing became relevant for the HPC community. Grids offer users access to powerful resources managed by autonomous administrative domains (Foster and Kesselman 2003; Foster et al. 2001). The notion of monetary costs for running applications was soft, favoring a more collaborative model of resource sharing. Therefore, quality of service was not strict in Grids, having users relying on best-effort policies to run applications.

In the late 2000s, cloud computing (Armbrust et al. 2010; Mell and Grance 2011; Buyya et al. 2009) was quickly increasing its maturity level and popularity, and studies started to emerge on the viability of executing HPC applications on remote cloud resources. These applications, which consume more resources than traditional cloud applications and usually are executed in batches rather than 24x7 services range from parallel applications written in Message Passing Interface (MPI) (Gropp et al. 1996, 1999) to the newest big data (Reed and Dongarra 2015; Assunção et al. 2015; Bahrami and Singhal 2015; Dean and Ghemawat 2008) and artificial intelligence applications—the latter mostly relying on deep learning (Coates et al. 2013; Krizhevsky et al. 2012). Cloud then came up as an evolution of a series of technologies, mainly on virtualization and computer networks, which facilitated both workload management and interaction with remote resources, respectively. Apart from software and hardware, cloud offers a business model where users pay for resources on demand. Compared to traditional HPC environments, in cloud computing users can quickly adjust their resource pools, via a mechanism known as *elasticity*, due to the size of the platforms managed by large cloud providers.

HPC cloud refers to the use of cloud resources to run HPC applications. Parashar et al. (2013) break down the usage of cloud computing for HPC into three categories: (i) “HPC in the cloud,” which focuses on moving HPC applications to cloud environments; (ii) “HPC plus cloud,” in which users use the cloud to complement their HPC resources (a scenario known as cloud bursting to handle peak demands (de Assuncao et al. 2009)); and (iii) HPC as a service (HPCaaS), which exposes HPC resources via cloud services. These categories are related to how resources are allocated and abstractions to simplify the use of the cloud.

HPC cloud still has various open issues. To exemplify, the abstraction of the underlying cloud infrastructure limits the tuning of HPC applications. Moreover, most cloud networks are not fast enough for large-scale tightly coupled applications—those with high interprocessor communication. The business model of HPC cloud is also an open field. Cloud providers stack several workloads on the same physical resources to explore economies of scale—an approach not always appropriate for HPC applications. In addition, although small companies benefit from fast access to resources from public clouds with no advance notice, this is usually not true for large users. The market forces that operate at large scales of cloud computing are the same as for other products and services. If one wants large amounts of resources, other methods of delivery are more suitable, such as private clouds, customized long-term contracts (e.g., strategic outsourcing), or even multiparty contracts.

Although several advances happened in past years in the cloud space, there is still a lot to be done in HPC cloud. Studies have shown some challenges on HPC cloud (Vecchiola et al. 2009; Gentzsch and Yenier 2013, 2014; U.S. Department of Energy Office of Science 2011; Mauch et al.

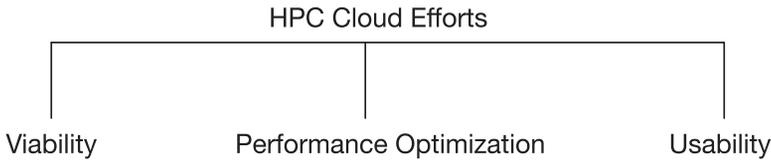


Fig. 1. Classification of HPC cloud main research efforts.

2013; Richter 2016; Yang et al. 2014; Gantikow et al. 2015; Sterling and Stark 2009; Galante et al. 2016); however, they do not present a comprehensive view of findings and challenges in the area. Therefore, this article aims at helping users, research institutions, universities, and companies understand solutions in HPC cloud. These are organized via a taxonomy that considers the viability of HPC cloud, existing optimizations, and efforts to make this platform easier to be consumed. Moreover, we provide a vision with directions for the research community to tackle open challenges in this area.

2 TAXONOMY AND SURVEY

The main difficulties in using cloud computing to execute HPC applications come from their properties in comparison to those from traditional cloud services such as standard enterprise and Web applications (Varghese and Buyya 2017). High Performance Computing (HPC) applications tend to require more computing power than cloud services. Such computing requirements come not only from CPUs but also from the amount of memory and network speeds to support their proper execution. In addition, such applications have a particular execution mechanism compared to cloud services that run 24x7. HPC applications tend to be executed in batches. Users run a set of jobs, which are instances of the application with different inputs, and wait until results are generated to decide whether new jobs need to be executed. Therefore, moving HPC applications to cloud platforms requires not only special care on resource allocation and optimizations in the infrastructure but also care on how users interact with this new environment. Therefore, proper understanding of all of these aspects is necessary to bring HPC users to cloud platforms.

Research in the area of HPC cloud can be classified in three broad categories, as depicted in Figure 1: (i) viability studies on the use of the cloud over on-premise clusters to execute HPC applications; (ii) performance optimization of cloud resources to execute HPC applications; and (iii) services to simplify the use of HPC cloud, in particular for non-IT-specialized users.

For research in the first category, analyses were carried out by executing HPC benchmarks (mostly NAS parallel benchmark (NPB) (Bailey et al. 1991)—and Intel MPI benchmarks (IMB) (Intel 2017)), microbenchmarks (Barbosa et al. 2009), and HPC user applications—all focusing on CPU, memory, storage, and disk performance. A few studies utilized other types of applications, such as scientific workflows (Kwok and Ahmad 1999) and parameter sweep applications (Casanova et al. 2000). For cluster infrastructure comparisons, some studies utilized clusters interconnected via high bandwidth/low latency networks including the well-known Myrinet (Boden et al. 1995) and InfiniBand (Ruivo et al. 2014; Vienne et al. 2012), whereas other studies investigated the performance of clusters interconnected via commodity Ethernet networks, with and without system virtualization.

Still in the first category, on the cloud side, most of the studies utilized Amazon Web Services (AWS) (Amazon 2017a). This happened because AWS provided credits for researchers to utilize the cloud infrastructure and Amazon was the first player in the market that simplified the use of cloud resources for individuals and small organizations. More recent works compare different public cloud providers—an analysis that would not have been possible in the early days of cloud

computing. Another effect of advances in the cloud technology over time in HPC research is the availability of HPC-optimized cloud resources (Amazon 2017b). The first generation of such machines were made available by AWS in 2010, and thus earlier research did not investigate their performance. In this category, the *Magellan Report* (U.S. Department of Energy Office of Science 2011) is the most comprehensive document in terms of range of architectures (i.e., number of different HPC clusters, virtualized clusters, and cloud providers), number of workloads (benchmarks and applications), and metrics. In fact, the report is a compilation of several studies sponsored by the U.S. Department of Energy (DOE) Office of Advanced Scientific Computing Research (ASCR) to study the viability of clouds to serve the department's computing needs, which were at the time met mostly by on-premise HPC clusters. In addition, most of the viability studies are focused on public clouds because most large-scale workloads from industry are not published in academic papers.

The second category (i.e., on optimizing the performance of HPC clouds), targeted either the infrastructure level or the resource management level. In the former, networking has been the main target, as it was established that networking accounted for most of the inefficiencies of executing HPC workloads in cloud. In the latter, scheduling policies that are aware of application and platform characteristics were proposed. In the optimization of resource allocation for HPC cloud, we observe a series of efforts on platform selectors due to hybrid cloud and multiple cloud choices, and studies aligned with specific features in cloud environments such as spot instances and elasticity. All of these optimizations benefit from resource usage and performance predictions.

Efforts in the third category focused on abstracting away the infrastructure from HPC users. One of the goals is to create HPCaaS platforms where HPC applications are executed in the cloud without requiring users to have any understanding of the underlying cloud infrastructure. Users just submit the application, relevant parameters, and quality-of-service (QoS) expectations, such as deadline, via a Web portal, and a middleware takes care of resource provisioning and application scheduling, deployment, and execution. This category of studies is relevant to increase the adoption of HPC, especially for new users with no expertise in system administration and configuration of complex computing environments. This category also highlights efforts on moving legacy applications to software-as-a-service (SaaS) deployments. This changes the user workflow from submitting jobs that wait in cluster scheduler queues to a cloud environment where resources are provisioned on demand according to user needs.

In the following sections, we detail the work in each category, having a large body of the work on HPC cloud viability. After the survey, we introduce our vision on the main missing components to enhance the capabilities of HPC cloud environments and their respective research challenges and opportunities.

2.1 Viability: Performance and Cost Concerns

There are four main aspects that were considered in viability studies, as depicted in Figure 2: (i) metrics used to evaluate how viable it is to use HPC cloud; (ii) resources used in the experiments; (iii) computational infrastructure; and (iv) software, which comprised well-known HPC benchmarks and user applications.

Gupta et al. (2013a) ran experiments using benchmarks and applications on various computing environments, including supercomputers and clouds, to answer the following question “Why and who should choose the cloud for HPC, for what applications, and how should the cloud be used for HPC?” They also considered thin virtual machines (VMs),¹ OS-level containers (Soltesz et al. 2007; Felter et al. 2015), and hypervisor- and application-level CPU affinity (Love 2003). They

¹VMs written specifically to run on top of hypervisors with the objective of reducing overhead.

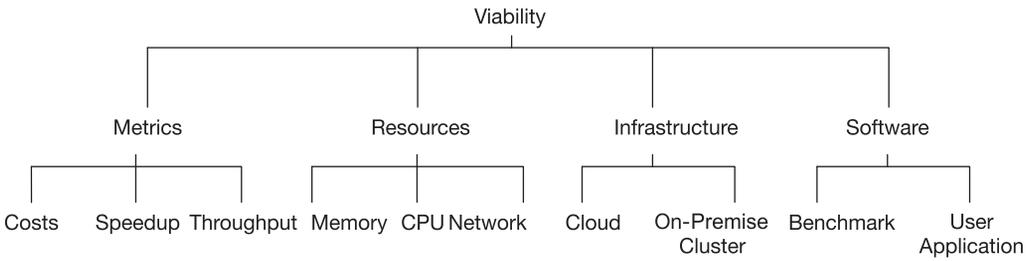


Fig. 2. Classification of HPC cloud viability studies.

concluded that public clouds are cost effective for small-scale applications but can complement supercomputers (i.e., HPC plus cloud (Parashar et al. 2013)) using cloud bursting and application-aware mapping. They also mentioned that network latency is a key limitation for scalability of applications in the cloud. For the experiments, based on their analyses, they found out that a cost ratio between two times and three times is a proper approximation to capture the differences between both cluster and cloud environments. This cost ratio reflects the shift from CAPEX to OPEX. The in-house HPC environment cost includes hardware acquisition, facility, power, cooling, and maintenance (Eubank 2003), which are not directly managed by the user in the cloud environment. However, IT support may still be present in the cloud, as most HPC applications are user specific. For further details, Kashaf and Altmann (2011) present a cost model for hybrid clouds itemizing all elements that go into the in-house and cloud environments.

Gupta and Milojicic et al. (2011) highlighted that the cloud computing can be suitable for only a subset of HPC applications; and for the same application, the choice of the environment may depend on the number of processors used. Gupta et al. (2012) also remarked that obtaining application signatures (or characterization) is a challenging problem, but with substantial benefits in terms of cost savings and performance. They evaluated the performance of HPC benchmarks across clusters, grids, and a private cloud. The analysis confirmed that HPC applications can have performance degradation in the cloud if they are communication intensive, but they can achieve a good performance otherwise. Furthermore, if the cost of running an HPC infrastructure is taken into consideration, the cost-performance ratio can be in favor of the cloud for HPC applications that do not demand high-performance networks.

The study from Gupta et al. (2014) was further expanded for more platforms—including public cloud providers and public HPC-optimized clouds—and more applications. They identified different classes of applications, considering cloud scalability, driven by different communication patterns and the ratio between the number of messages and message sizes. The cause for the differences in scalability has been identified as network virtualization, multitenancy, and hardware heterogeneity. Based on these findings, the authors identified two general strategies for countering performance limitations of clouds called *cloud-aware HPC* and *HPC-aware clouds*. The first is about decomposing work units, setting up optimal problem size, and tuning network parameters to improve computation/communication ratio; the second is about using lightweight virtualization, setting up CPU affinity, and handling network aggregation to reduce the overhead of the underlying virtualization platform. Their study also identified that the more CPU cores an application requires, the more likely an HPC platform offers the best value for the money, although startups and small- and medium-size companies, which are usually sensitive to CAPEX, might still be better off using clouds rather than clusters.

Napper and Bientinesi (2009) used High-Performance LINPACK (Dongarra et al. 2003) to evaluate whether the cloud could potentially be included in the TOP500 list (TOP500 2017)—the

list of the most powerful computers worldwide. The experiments were conducted using Amazon Elastic Compute Cloud (EC2). Their results showed that the raw performance of EC2 instances is compatible with resources from current on-premise systems. However, memory and network performance were not sufficient to scale the application. They also investigated giga floating-point operations per second (GFLOPS) and GFLOP per dollar to evaluate the trade-off of costs and performance when running user applications in remote cloud resources. On the cloud, these metrics behaved differently from traditional HPC systems. In November 2011, it was announced that an Amazon EC2 cluster reached position 42 in the TOP500 ranking.²

Belgacem and Chopard (2015) investigated the use of hybrid HPC clouds to run multiscale large parallel applications. Their motivation was the lack of memory in their local cluster to run their application. They showed that using proper strategies to balance the load between local and remote sites had a considerable influence on the overall performance of this large parallel application, thus leading to the conclusion that HPC hybrid clouds are relevant for large applications.

Marathe et al. (2013) compared HPC clusters against top-of-the-line EC2 clusters using two metrics: (i) turnaround time and (ii) total cost for executions. The turnaround time includes the expected queue wait time determined by the cluster management system, which is a commonly ignored but highly important factor for viability studies considering QoS. They showed that although the clusters produced superior raw performance, EC2 was able to produce better turnaround times. The results showed turnaround times more than four times longer in HPC on-premise resources than in the cloud, even with much faster executions on local clusters. They also highlighted that applications should be properly mapped to clusters, which are not generally the most powerful ones. And the choice between the cloud and local HPC clusters is complicated—which relies on application scalability and the goal of optimizing cost or turnaround time. Therefore, having tools that help in making resource allocation decisions is fundamental to save money and time.

Expósito et al. (2013) studied the performance of HPC applications in Amazon EC2 resources and focused mainly on I/O and scalability aspects. In particular, they compared CC1 instances against the CC2 instances released in 2011 and used up to 512 cores. They also investigated the cost benefit of using these instances to run HPC applications. Their conclusions were that although CC2 provides more raw and point-to-point communication performance, collective-based communication-intensive applications performed worse compared to using CC1. They also concluded that using multilevel parallelism—one level of message passing and another with multithreading—generated a scalable and cost-effective alternative for user applications in Amazon EC2 CC instances. Such instances were also investigated by Sadooghi and Raicu (2013) using the high-performance LINPACK benchmark, and they found instabilities in network latency compared to their on-premise HPC environment.

Carlyle et al. (2010) conducted an experiment to identify if it would be worthwhile, costwise, using Amazon EC2 cluster instances rather than their two clusters at Purdue University. The motivation of their study was the lack of cost-aware studies between community clusters and public clouds. Their conclusion was that the community cluster had a better cost benefit than using the cloud, especially given the high utilization of their clusters. However, they acknowledged that for low utilization clusters or small and underfunded projects, cloud computing could be a cost-effective alternative.

Ostermann et al. (2009) presented a detailed study on the performance of EC2 for scientific computing. They evaluated EC2 instances using multiple benchmarks and considered various aspects, including time to acquire and release resources, computing performance, I/O performance,

²<https://www.top500.org/list/2011/11/>.

memory hierarchy performance, and reliability. Although they found performance and reliability to be the major limitations for scientific computing, they found cloud computing as attractive for scientists in need of resources immediately and temporarily.

Zaspel and Griebel (2011) reported performance results of a parallel application in the area of computational fluid dynamics. They ran experiments on Amazon EC2 instances using both CPUs and GPUs to evaluate the application scalability. The application was executed with up to 256 CPU cores and 16 GPUs, and their findings indicated that the application scaled well until 64 CPU cores and 8 GPUs.

Berriman et al. (2010) relied on three applications with workflow models to study the benefit of a cluster at the National Center for Supercomputing Applications (NCSA) against Amazon EC2 cloud offers. The workflows are Montage³ from astronomy, Broadband from seismology, and Epigenome from biochemistry. They concluded that, for their experiments, the cloud can provide resources that are more powerful and cheaper compared to their on-premise cluster, particularly considering applications that are CPU and memory intensive. However, for applications that manipulate large volumes of data, clusters with parallel file systems and low-latency networks provide better performance. They also highlighted the growing efforts on creating academic clouds, which, according to them, may have different levels of services than those provided by commercial clouds.

Zhai et al. (2011) described in the Supercomputing 2011 conference a detailed study comparing the cloud and an in-house cluster considering tightly coupled applications. They investigated the EC2 cluster compute instances released by Amazon in 2010. They used NAS benchmarks and three parallel applications with different characteristics. Similar to the other studies, they observed the network as a bottleneck for tightly coupled applications. However, they also evaluated which types of messages were not suitable for the 10GB Ethernet offered by these instances. They concluded that for large messages with few processors, the network performance was comparable to their in-house cluster. They also compared the costs of both environments with a detailed analysis on the required utilization level a cluster should have to be considered cheaper or more expensive than running HPC applications in the cloud. They observed that depending on the application, clusters should have at least a utilization of 8.5% to 31.1% to be beneficial compared to the cloud—as highlighted by the authors, these numbers were calculated with a set of assumptions that brought benefit to the in-house cluster.

The *Magellan Report* (U.S. Department of Energy Office of Science 2011), commissioned by the DOE, is one of the most comprehensive documents in the area of adoption of clouds for scientific HPC workloads. Extensive evaluation of diverse cloud architectures and systems, and comparison with HPC clusters, have been carried out as part of the report's activities. Certain workloads from the DOE suffered slowdown of up to 50 times when utilized in clouds, due to the particular patterns of communication between application tasks. The report highlights that latency-limited applications (where numerous small point-to-point messages are exchanged) are the most penalized ones by the lack of high performance networks in clouds, whereas bandwidth-limited applications (exchanging few large messages, or performing collective communication) are less penalized. Another obstacle for HPC cloud, noted in the report, is the eventual absence of support, in the hypervisor level, to specialized instruction sets. When such instructions are enabled at the hypervisor, no performance loss is incurred during computation. Furthermore, when users cannot enforce a certain CPU architecture (as in the case in most public infrastructure-as-a-service (IaaS) providers), CPU set-specific compiler optimizations cannot be utilized, what also limit the potential increase in performance of applications. On the positive side for cloud adoption, the report notes that high

³<http://montage.ipac.caltech.edu>.

performance cluster queuing systems do not provide proper support to embarrassingly parallel applications, and thus this class of applications can benefit from clouds.

Roloff et al. (2012) described a study on HPC cloud considering three well-known cloud offers (Microsoft Azure, Amazon EC2, and Rackspace) and analyzed three aspects—deployment of HPC applications, performance, and costs—utilizing NAS benchmarks. A few highlights from their study are the following: (i) there is no single clear provider that best meets all three aspects analyzed; (ii) in various scenarios, the cloud showed an interesting alternative for an on-premise cluster to run HPC applications considering both raw performance and monetary costs; and (iii) the lack of information on network interconnection among the cloud resources is still an issue for all cloud providers, particularly for communication-intensive applications. The authors also described in detail the deployment strategies offered by each provider.

Egwutuoha et al. (2013) compared IaaS against not an on-premise cluster but a cloud provider that offers bare metal machines, termed *hardware as a service* (HaaS). By using an HPL benchmark and an application called *ClustalW-MPI* for bioinformatics sequence alignment, they were able to show that HaaS can save up to 20% of the cost to run the applications compared to a traditional IaaS provider.

Evangelinos and Hill (2008) reported their experience using the cloud to run their atmosphere-ocean climate model parallel application. Their experiment started by running NAS benchmarks to understand the cloud performance. Next, they tested various MPI implementations and highlighted the issue of not being able to control in which subnet the VMs were provisioned. Having the environment set up, they tested their application and concluded that EC2 instances were a feasible alternative for their experiments, which considered only performance aspects.

He et al. (2010) ran experiments to verify the performance of the NPB, LINPACK, and an application on climate and numerical weather prediction application using three cloud environments: Amazon EC2 Cloud, GoGrid Cloud, and IBM Cloud. They concluded that with a few changes from a cloud perspective, mainly in relation to network and memory, the cloud could be an alternative from on-premise clusters for HPC applications. They highlighted that different from traditional on-premise HPC platforms, in which FLOPS is the main optimization criterion, in cloud FLOPS per dollar is an important metric. From their experiments, one of the examples was that changing an execution from one environment to another produced a gain of 30% of performance but paying four times more. Thus, users can get to a point where they will accept having 30% slower executions but with a set of resources that is four times cheaper.

Jackson et al. (2010) evaluated the performance of the NERSC benchmarking framework, which contains application internals from several fields including astrophysics, climate, and materials science. They compared Amazon EC2 against three on-premise clusters. They also used the Integrated Performance Monitoring framework (Borrill et al. 2005), which helps determine the different application phases (i.e., computing and communication). Their main conclusion is that the cloud is suitable for various applications but not, in particular, for tightly coupled ones. They found a considerable relationship between the amount of time an application spends in communication and its performance when using cloud resources. Similar findings were discussed by Ekanayake and Fox (2009) for MapReduce technologies, by Church and Goscinski (2011) for bioinformatics and physics applications, and by Hill and Humphrey (2009) with STREAM memory bandwidth benchmark and Intel's MPI benchmark.

Hassan et al. (2015) investigated the performance of the IMB suite and NPBs on Microsoft Azure on 16 VMs. They were more interested in analyzing scalability of these benchmarks considering different point-to-point communication approaches using both MPICH and OpenMPI. They obtained more promising results when running experiments on single VMs due to the shared memory communication, as internode communication showed to be a key bottleneck.

Table 1. Overview of the Related Work on Viability of HPC Cloud

Aspect	References	Key Takeaways
Cost	(Gupta et al. 2013a; Napper and Bientinesi 2009; Carlyle et al. 2010; Roloff et al. 2012)	Related to financially related decisions. When compared to low-utilization clusters and when running small applications, cloud environments can be preferred over on-premise clusters.
Throughput	(Marathe et al. 2013; Carlyle et al. 2010; Ostermann et al. 2009)	Related to job turnaround times. For running single applications with low communication requirements, cloud can have higher throughput due to the lack of queueing systems present in on-premise clusters.
Resources	(Gupta et al. 2014; Napper and Bientinesi 2009; Expósito et al. 2013; Sadooghi and Raicu 2013; Zhai et al. 2011; Egwutuoha et al. 2013; Hassan et al. 2015)	Related to how different resource types impact performance of the applications. Network virtualization and hardware heterogeneity are main causes of poor performance of HPC in cloud; single-node performance is comparable between environments; improvements in infrastructure affect HPC applications positively.
Network	(U.S. Department of Energy Office of Science 2011; Gupta and Milojevic 2011; Gupta et al. 2012; Napper and Bientinesi 2009; Sadooghi and Raicu 2013; Zhai et al. 2011; Jackson et al. 2010; He et al. 2010; Evangelinos and Hill 2008)	Related to the impact of network speeds considering different communication models of parallel applications. Cloud is suitable when loosely coupled or embarrassingly parallel applications are used.

Hassani et al. (2014) implemented a version of the Radix sorting algorithm using OpenMPI and tested it on both their on-premise cluster and Amazon EC2 extra large instances. They varied the input size of data to be sorted and concluded that the cloud environment generated 20% faster completion time compared to their on-premise cluster using up to eight nodes. After that point, they highlight that network bandwidth could limit the performance if instances were not reserved in advance.

Summary and takeaways. Overall, all of these studies showed that the cloud has a great potential to host HPC applications, with network performance as one of the main bottlenecks at the moment. Fortunately, this network bottleneck may not be a problem for several scientific and business applications that are CPU intensive. Most of the studies still rely on standard HPC benchmarks to stress different resource requirements of HPC applications. In addition, users and institutions should avoid limiting their cloud versus an on-premise cluster decision based only on raw application performance. It is important to understand turnaround time—that is, execution time plus the time to access resources, costs, and resource demands. A mix of on-premise cluster and the cloud seems to be a proper environment to balance this equation. The key takeaways of the works referenced in this section are summarized in Table 1. In Table 2, we provide a list of

Table 2. Recommendation to Use Cloud

Application Type	Recommendation
Large-scale tightly coupled	They are typical MPI applications that use thousands of cores and require a high-performance network, such as weather, seismic, geomechanical, and computational fluid dynamics models (timestepped applications). Any virtualization bottleneck and high latency network will have a negative impact on application performance. Therefore, it is recommended to use these applications in traditional supercomputing centers, or on private clouds featuring bare metal machines and high-speed networks.
Mid-range tightly coupled	They utilize several cores ranging from tens to hundreds and have lower performance requirements than the large-scale tightly coupled type. Consequently, they are more tolerant of virtualization and traditional networks. Event-driven simulation is an example of this type, but also other timestepped applications whose jobs are less deadline sensitive. It is recommended that these applications explore the benefits of fast resource access to the cloud, especially when lightweight virtualization (i.e., containers) are becoming pervasive.
High throughput	They are composed of independent tasks that require little or no communication, popular in Monte Carlo simulations and many other bag-of-tasks and Map-Reduce applications, and can benefit from a variable number of available resources and are tolerant of resource heterogeneity. It is recommended the use of cloud for these applications especially by exploring elasticity mechanisms. They can also benefit from HPC hybrid cloud environments by spreading out tasks on both on-premise clusters and public clouds.

recommendations on when the cloud could be used depending on application types defined in the *Magellan Report* (U.S. Department of Energy Office of Science 2011).

2.2 Performance Optimization: Resource Allocation and Job Placement

As depicted in Figure 3, most of the work on performance optimization for HPC cloud lies on areas related to resource management and job placement systems. These efforts concern where and how jobs should be placed, inside a cloud or in a hybrid environment (HPC in the cloud and HPC plus cloud according to Parashar et al. (2013) definitions), how to leverage cheaper instances, how to use elasticity of cloud resources, and how to use prediction systems for job placement decisions.

Gupta et al. (2013b) introduced an HPC-aware scheduler for cloud platforms that considers topology requirements from HPC applications. Their scheduler utilizes benchmarking information, which classifies the type of network requirement of the application and how its performance is affected when resources are shared with other applications. Their experiments, using three applications and NAS benchmarks, showed performance improvements of up to 45% by using their scheduler compared to an HPC-agnostic scheduler. In a related topic, Gupta et al. (2012) developed a tool to extract characteristics from HPC applications and then map them to the most suitable computing platform, including both clusters and clouds.

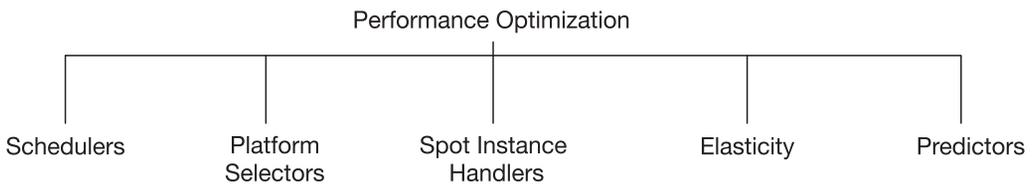


Fig. 3. Classification of HPC cloud performance optimization studies.

Church et al. (2015) addressed the issue of resource selection in the Uncinus framework. The selection is based on prepopulated information on available HPC applications and resources, which can be clouds and clusters. Resource selection is then carried out with the use of historical information on cloud resource usage and availability of resources at request time. The approach relies on a broker that mediates access to the cloud (via credentials, resource discovery, and resource selection) and enables sharing of resources and applications. Their goal is to help non-IT specialized users deploy their applications in the cloud.

Gupta et al. (2014) proposed a set of heuristics for the problem of choosing a platform (including clusters and clouds) for a stream of jobs. The objective is to improve makespan and job completion time. They tested a combination of static and dynamic heuristics that are application aware and application agnostic. Results demonstrated that dynamic heuristics that consider application characteristics and their performance on each particular platform generate better throughput than other heuristics.

Ashwini et al. (2013) introduced a framework to allocate cloud resources for HPC applications. Their motivation is the heterogeneity of physical cloud resources, and their framework can select resources based on similar performance levels. They used processing power and point-to-point communication bandwidth to select clusters of VMs to run HPC applications.

Marathe et al. (2014) designed and implemented techniques to reduce costs when running HPC applications in the cloud. They developed techniques for both determining bid prices for Amazon EC2 spot instances and scheduling checkpoints of user applications. They were able to obtain gains of 7x compared to traditional on-demand instances.

Somasundaram and Govindarajan (2014) developed a framework to schedule HPC applications in cloud platforms. Their goal was to meet users' deadline and at the same time reduce costs to run their applications. The scheduler is based on particle swarm optimization, and their experiments were based on simulations and on two real applications and a testbed setup with Eucalyptus Cloud middleware (Nurmi et al. 2008).

Netto et al. (2015) discussed the challenges users have when making job placement decisions in HPC hybrid clouds, particularly with respect to uncertainties coming from execution time and job waiting time predictions. As a follow-up to this problem, Cunha et al. (2017) presented a detailed implementation of an advisor that relies on job runtime and wait time predictions and the confidence level of such predictions to make job placement decisions. On the direction of using performance predictions, Shi et al. (2012) applied Amdahl's law to predict performance of NAS benchmarks over a private HPC cloud environment.

Network is another type of resource that requires proper management. Mauch et al. (2013) investigated the use and configuration of InfiniBand (InfiniBand Trade Association 2000) in virtualized environments. Their motivation is the limitation of the Ethernet technology used by cloud providers for HPC workloads. Their HPC cloud architecture based on InfiniBand allows customers to allocate virtual clusters isolated from other tenants in the physical network. This type of platform, augmented with the capacity of establishing elastic *virtual clusters* for users and achieving

network isolation, was shown to incur only a small overhead in the order of microseconds. They also envisioned as future directions how to use InfiniBand to allow live migration of VMs.

Marshall et al. (2013) provided an overview of their work on HPC cloud having elasticity as their main cloud functionality to be explored for HPC workloads. They developed a prototype using Torque and Amazon EC2 in which jobs would receive machines in the cloud if not enough resources were available at the job submission moment. They raised a series of questions that they believe are relevant for HPC workloads related to when cloud machines should be provisioned—for instance, if they should be provisioned at the moment new jobs arrive or once they stay stuck in queue, if instances should be deleted once jobs are completed or should remain until the hour completes as new jobs may arrive, and if autoscaling should be done in a reactive or proactive manner.

Mateescu et al. (2011) proposed a concept they called *Elastic Cluster*, which aims at combining the strengths of clouds, grids, and clusters by offering intelligent infrastructure and workload management that are aware of different types and locations of resource and performance requirements of applications. Performance guarantees are obtained via a combination of statistical reservation and dynamic provisioning. The architecture is complemented by the capacity of setting *personal virtual clusters* for execution of user workflows and by the capacity of combining resources from multiple elastic clusters for execution of workflows. Elasticity was also explored by da Rosa Righi et al. (2016), who created a software at the platform as a service (PaaS) to support this functionality for HPC applications. The aim of their project is to offer elasticity with no need of application source code access.

Zhang et al. (2016) studied the use of container-based virtualization and its relationship with the performance of MPI applications. The authors found out that the overhead on the computation stage of applications is negligible, although performance degradation can be observed in the communication stage, even when all containers hosting application processes are executed on the same physical server. To tackle such an issue, the authors proposed a method to enable the MPI runtime to detect processes that share the same physical host and use in-memory communication among them. The approach, combined with optimization of configuration parameters of the MPI runtime, improved by 900% and 86% the performance of point-to-point and collective communications, respectively.

Fan et al. (2012, 2014) proposed a framework that considers application internal communication patterns to deploy scientific applications. This topology-aware deployment framework consists of clustering cloud machines that have communication affinity. Experiments were conducted using the PlanetLab (Chun et al. 2003) environment with 100 machines and relied on NPBs and the IMB. They were able to reduce execution times by 30% to 35% compared to not considering application and topology and intermachine communication performance.

Summary and takeaways. All efforts described in this section indicate that even with challenges in HPC cloud environments, optimization techniques and technologies have helped improve the performance of user HPC applications running in remote cloud resources. These optimizations influence allocation of different types of resources, including CPUs, memory, and computing networks. Performance prediction plays a key role to help in resource allocation decisions to make the best use of cloud platforms. In addition, users have to be careful with application placement decisions, particularly trying to explore possibilities of using powerful single nodes whenever possible. Most of these efforts were motivated by network and virtualization issues from current cloud platforms. Even with new virtualization technologies and high-performance networks, such optimizations will still bring benefits to users. Table 3 summarizes the efforts discussed in this section.

Table 3. Overview of the Related Work on Performance Optimization for HPC Cloud

Aspect	References	Key Takeaways
Scheduler	(Gupta et al. 2013b; Ashwini et al. 2013; Somasundaram and Govindarajan 2014; Fan et al. 2012, 2014)	Related to how scheduling decisions impact job performance. HPC-aware schedulers indeed improve performance of HPC applications in cloud environments, as they exploit both HPC application and infrastructure properties.
Platform selectors	(Gupta et al. 2014; Netto et al. 2015; Cunha et al. 2017; Church et al. 2015)	Related to the impact of environment selection on job performance. Automated selection of execution environments eases transition to the cloud as users may be overloaded with many infrastructure configuration choices.
Spot instance handlers	(Marathe et al. 2014)	Related to the use of novel pricing models in cloud. Exploiting different cloud pricing models transparently reduces costs in the cloud because users have different resource consumption patterns.
Elasticity	(Mateescu et al. 2011; Mauch et al. 2013; Zhang et al. 2016; Marshall et al. 2013; da Rosa Righi et al. 2016)	Related to dynamic allocation of resources. Cloud elasticity adds more flexibility to HPC applications; VMs placed on same host improve performance.
Predictors	(Shi et al. 2012; Netto et al. 2015; Cunha et al. 2017)	Related to collect information of future resource consumptions. Prediction of expected runtime and wait time helps on job placement decisions due to a proper match of resource configuration and resource consumption.

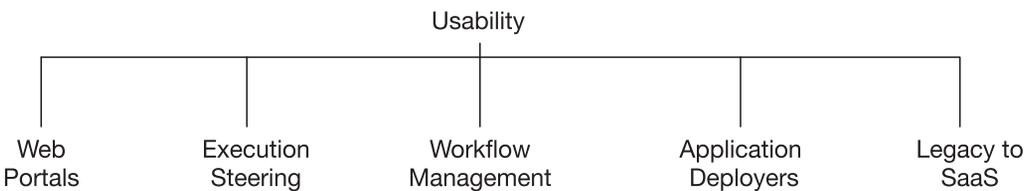


Fig. 4. Classification of HPC cloud usability efforts.

2.3 Usability: User Interaction and High-Level Services

Understanding the cost benefit of moving HPC workloads to the cloud and optimizing the execution of these workloads are the main efforts of researchers working in the area. However, all of these efforts will have limited applicability if HPC cloud is not easy to use. Figure 4 depicts the main research areas concerning the usability of HPC cloud. Most of the work relies on Web portals and how users interact with their applications and the cloud infrastructure. There are also

various efforts to support the creation of easy-to-use SaaS based on legacy applications and offer HPC resources as a service (HPCaaS as defined by Parashar et al. (2013)).

Belgacem and Chopard (2015) reported their experience of porting a parallel 3D simulation in the computational fluid dynamics domain to a hybrid cloud consisting of a cluster located in Switzerland and AWS EC2 resources located in the United States. To facilitate the integration between clusters and clouds, the authors utilized a methodology called *distributed multiscale computing* (DMC), which provides abstractions that enable phenomena to be described in different scales. Each scale becomes a parallel application (in the presented work, written in MPI) to be executed in a physical computing resource. These different scales can be processed in parallel or can hold dependencies among themselves, leading to a workflow model for application execution. Results demonstrated that utilization of a hybrid infrastructure only improves execution time if the application is tuned to adapt to the difference in CPU speeds between clusters and clouds and, more importantly, to adjust to the difference in speeds between local networks and WAN connections. This can be done, for example, by adapting the application to allow for overlap between communication and computation, thus reducing the time tasks are waiting for data to compute.

In the area of computational steering, SciPhy (Ocaña et al. 2011) is a tool to help scientists in the area of phylogeny/phylogenomics run experiments for drug design. The tool, based on SciCumulus (de Oliveira et al. 2010), uses cloud resources and is an example on how a service can be created to facilitate the use of the cloud for non-IT specialists. We believe that several efforts on computational steering (Mattoso et al. 2015) applied to grid and cluster computing will be leveraged by cloud users.

By using a middleware called *Aneka*, Vecchiola et al. (2009) showed two case studies on HPC cloud: (i) gene expression data classification and (ii) fMRI brain imaging. Both applications were executed in Amazon EC2. They highlighted research directions on helping users in performance/cost trade-offs and decision-making tools to specify accuracy level of the experiment or parts of data to be processed according to predefined SLAs.

Huang (2014) created a mineral physics SaaS, called *Fonon*, for on-premise HPC clusters. Usage of *Fonon* allowed users of different backgrounds to submit jobs more easily and to obtain more useful results. The benefits achieved come from users submitting and analyzing jobs by means of a Web-based application. From his analysis, although he noticed the virtualization overhead and network limitations when experimenting the application in a public cloud, he also identified a demand to facilitate the use of the application through the creation of a service. Therefore, his SaaS encapsulated several activities of the users, such as execution of preprocessing and postprocessing scripts, elimination of data movement by creating figures and tables in the cloud provider site, and allocation of cluster resources. In a similar direction, Buyya and Barreto (2015) described the use of *Aneka* to help users run and deploy HPC applications on multicloud environments. They used BLAST as an example of a parameter sweeping application that could benefit from *Aneka* application deployment system.

In 2012, Church et al. (2012) discussed the difficulties that users, especially those with no IT expertise, had to execute applications using remote cloud resources. Having this motivation, they started to develop a framework to translate HPC applications into services. Over the years, the same group (Church et al. 2015) presented novel use cases of their framework in the area of genomics. Church and Goscinski (2014) also presented a survey on technologies available to help researchers working with mammalian genomics run their experiments. They highlighted the technical difficulties that these researchers have when using IaaS. They then developed a system to encapsulate several activities related to resource allocation, data management, images containing software stacks, and graphical user interfaces.

AbdelBaky et al. (2012) described a system to enable HPCaaS using an application that models oil reservoir flows as a use case. Users are able to easily interact with the application, which has access to IBM BlueGene/P systems and can provide dynamic resource allocation (e.g., elasticity). The system relies on both DeepCloud and CometCloud, which provide IaaS and PaaS functionalities, respectively. Also in the context of CometCloud, AbdelBaky et al. (2014) proposed an SaaS abstraction for scientists to run experiments using distributed resources; as a use case, they considered an application for experimental chemists to use mobile devices to run dissipative particle dynamics experiments.

Wong and Goscinski (2013) created a software to configure, deploy, and offer HPC applications in the cloud as services. The approach also includes mechanisms for discovery of already deployed applications—thus avoiding duplication of the endeavor. Once deployed, the application becomes available for end users via a Web-based portal, and the only input required in this case are input parameters, which are supplied via Web forms. Petcu et al. (2014) proposed a different framework with similar objectives, with the added difference of including support to hardware accelerators, such as GPUs.

Balis et al. (2017) presented a methodology for porting HPC applications to a cloud environment using a multifrontal solver as a case study. The methodology focuses on a task agglomeration heuristic to increase task granularity while ensuring that there is enough memory to run them, a task scheduler to increase data locality, and a two-level storage to enable in-memory storage of intermediate data.

Another effort to simplify the usage of HPC comes from Bunch et al. (2011), who created a domain-specific language, called *Neptune*, to deploy HPC applications in the cloud. Neptune offers support to applications written with various packages, including MPI, X10, and Hadoop. It can also be used to add and remove resources to the underlying cloud platform and to control how applications are placed across multiple cloud infrastructures.

Summary and takeaways. As described in Table 4, there has been a growing interest in creating services to facilitate the use of the cloud for HPC applications and to transform legacy applications into cloud services. HPC is still not easy for non-IT experts to use, and having cloud services, even if in private environments, can help HPC become more popular. With these services, elasticity, which is a key functionality of cloud computing, could be embedded and explored more easily by end users. There is a considerable engineering effort to improve the usability of HPC cloud, such as the creation of Web portals or simplification of procedures to allocate and access remote resources. However, research can also make contributions. For instance, when transforming an HPC application into SaaS, the amount of computing resources to meet user expected QoS needs to be properly defined. This may also be an opportunity of collaboration with researchers working with a human-computer interface. In the area of workload management, languages for non-IT experts could also be beneficial to facilitate the use of cloud resources. Several of these efforts could leverage the work done in grid computing; however in HPC cloud, cost management is a crucial aspect. Another relevant aspect of usability is to increase user productivity. It may be more valuable for several users to reduce their time TO setup an experiment rather than optimize resources to run jobs. The value of proper usability technologies and practices is to reduce turnaround times and minimize costs whenever possible.

3 VISION AND RESEARCH CHALLENGES

As presented in the previous section, plenty of work has been done in HPC cloud. However, clients and cloud providers can benefit more from this platform if additional modules/functionalities become available. Here we discuss a vision and research challenges for HPC cloud and relate them to what was presented in the previous section.

Table 4. Overview of the Related Work on the Usability of HPC Cloud

Aspect	References	Key Takeaways
Web portals	(Huang 2014; Church et al. 2012; Church and Goscinski 2014; Church et al. 2015; AbdelBaky et al. 2012; Wong and Goscinski 2013; Petcu et al. 2014)	Related to creation of easy-to-use interfaces. Users have a hard time selecting cloud resources; developing portals that abstract details can increase user productivity.
Execution steering	(Ocaña et al. 2011; de Oliveira et al. 2010)	Related to automatic creation of jobs with different input parameters. Facilitation of parameter sweeping experiments for non-IT specialists.
Workflow management	(Vecchiola et al. 2009; Belgacem and Chopard 2015; Church et al. 2015; Church and Goscinski 2014; Wong and Goscinski 2013)	Related to proper mapping of multiple activities, possibly including dependencies. Frameworks that have knowledge of cloud pricing models can reduce time and costs for deploying and exposing HPC applications in the cloud.
Application deployers	(Wong and Goscinski 2013; Bunch et al. 2011; Buyya and Barreto 2015)	Related to HPC application software stack dependencies. HPC applications have a different software stack than traditional cloud applications; systems that are aware of HPC application requirements reduce time to solution in clouds.
Legacy-to-SaaS	(Balis et al. 2017; Petcu et al. 2014; Huang 2014)	Related to transform legacy applications running in traditional computing platforms to cloud environments. Direct ports of legacy applications to cloud environments might have inefficiencies due to different infrastructure assumptions; principled methodologies can overcome such inefficiencies.

Figure 5 illustrates our vision of an HPC cloud architecture. The architecture contains three major players: the Internet, the client, and the cloud provider. Sensors and social media networks are two relevant sources of data generation that serve as input for various HPC applications, especially for the increasing workloads coming from big data, artificial intelligence, and sensor-based stream computing. Data from these sources can come from places that are both internal and external to client and cloud provider environments. The other two players are the client who needs data to be processed and the cloud provider, who offers HPC-aware services. The architecture comprises components already discussed in Section 2, including those from performance optimization, which are more related to infrastructure, and those from HPC cloud usability, which are closer to the user. In the following sections, we discuss a set of modules that we believe require further development, and we split them in two categories: one specific for HPC workloads and the other general for cloud but that can bring benefits to HPC cloud.

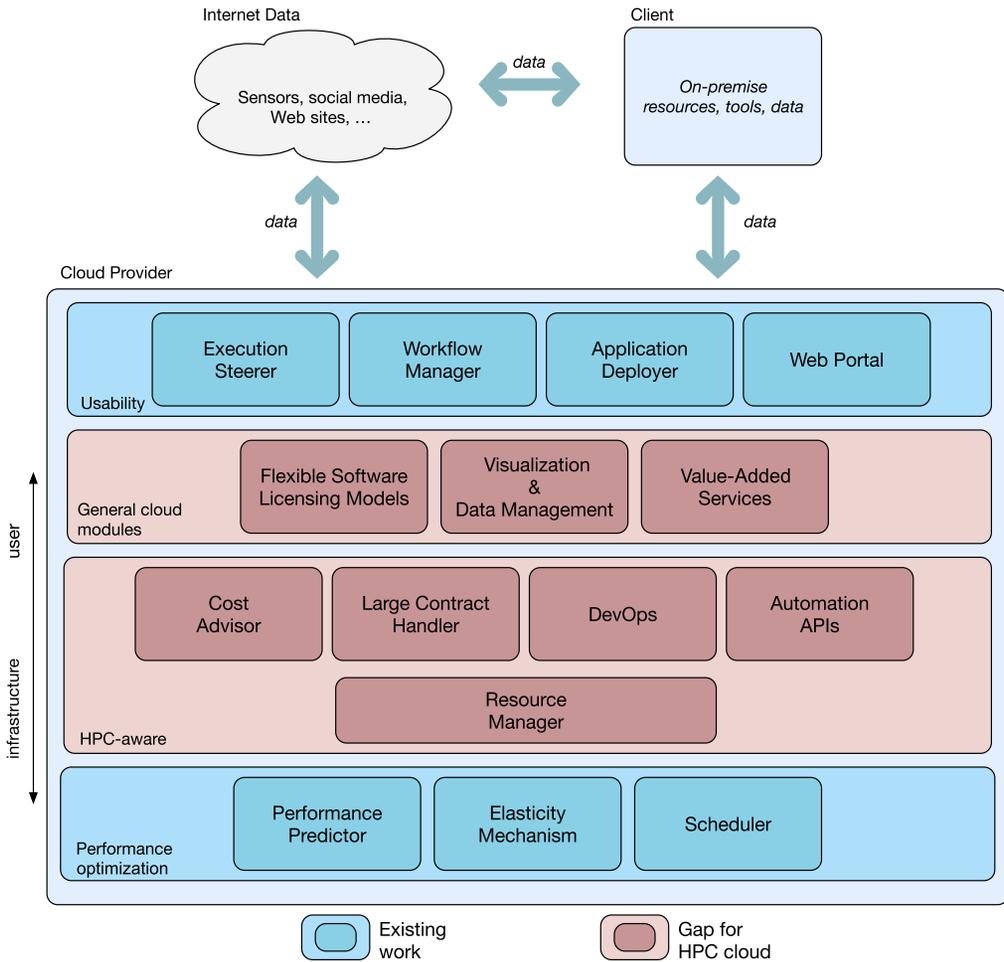


Fig. 5. Vision of an HPC cloud architecture comprising existing modules in the area of usability and performance optimization and modules that require further development. The latter modules have two categories: (i) HPC-aware modules and (ii) general cloud modules that bring benefit to HPC cloud.

3.1 HPC-Aware Modules

In this section, we discuss challenges and research opportunities for five modules that require further development to meet the needs of HPC users. Although most of these modules are already available in various cloud providers, HPC users have different resource requirements and work style that limit the direct utilization of these modules.

3.1.1 Resource Manager. Cloud and HPC environments have distinct ways of managing computing resources. Cloud aims at consolidating applications into the same hardware to achieve economies of scale, which is possible due to virtualization technologies. To extract the best performance of a cloud environment, ongoing efforts have focused on increasing inter-VM isolation and reducing VM overhead. HPC environments, however, aim at bringing the highest performance possible from the infrastructure. User requests to access exclusively portions of a cluster are queued whenever resources are overloaded. One evident benefit of cloud computing is that no queues are

required due to the “unlimited” availability of resources. In addition, HPC hardware, especially network, costs considerably more than that used to build traditional clouds. Hence, placing users in such hardware would be wasteful for cloud providers. Therefore, the challenge on HPC cloud resource management is to have sustainable business for cloud providers via economies of scale and to be able to offer users high performance. Research efforts could find this balance and use the same manager for both cloud and HPC users.

There are a few projects already in this area. For instance, Kocoloski et al. (2012) introduced a system called *dual stack virtualization*, which consists of a VM manager that can host both HPC and commodity VMs depending on the level of isolation required by the users. The configurable level of isolation can relate to different prices that benefit both users and cloud providers.

To advance this area, we envision cloud providers offering queues and having pricing models that consider how long users are willing to wait to access HPC resources. This would allow providers to have more clusters fine tuned for certain types of workloads and for users to have different QoS with respect to the time to access resources. There are some efforts on providing flexible resource rental models for HPC cloud, such as those from Zhao and Li (2012), who were motivated by the different interests from multiple parties when allocating resources. Their models are based on planning strategies that consider on-demand and spot instances. Resource managers can also have user-centric policies (Sherwani et al. 2004) for management of shared HPC cloud resources that offer incentive for users to reveal true QoS requirements, such as deadlines to finalize application executions.

3.1.2 Cost Advisor. Cost advisors have become popular in several cloud providers. It is usually implemented as a simulator for users to specify their resource requirements and obtain cost estimations. Different from traditional cloud users who use this advisor to plan for hosting a service, HPC users need advice on how much their experiments will cost. Therefore, current cost advisors need to be adapted to support HPC users, and this is challenging because HPC user workflows involve tuning their applications and exploring scenarios via execution of several jobs. Such cost comes from software licenses and powerful computing resources that tend to be much more expensive than traditional lightweight VMs.

Researchers have been looking into solutions to handle cost predictions for HPC users. For instance, Aversa et al. (2011) highlighted that cost is a critical issue for HPC cloud because applications are not optimized for an unknown and virtualized environment, and current charging mechanisms do not take into account the peculiarities of HPC users. They also noticed that such an advisory system for cost-related decisions is highly dependent on the user profile, which can vary from a typical HPC user who wonders about the execution time for her job and the configuration of resources to obtain the highest performance/cost ratio to a high-end user who cares more about performance.

Rak et al. (2015) introduced an approach to help users have better predictions of performance and costs when running HPC in the cloud. Using their framework called *mOSAIC*, users can run simulations and benchmarks to give insights about application performance during the application development phase. Another example is the work from Li et al. (2011), who investigated the problem of how to enhance interactivity for HPC cloud services. Their motivation is that HPC users have complex and expensive requirements. Therefore, for them, HPC cloud services are not only about helping users allocate resources but also about their interactions with the computing environment, which include their expectations. Their work allows users to predict how much longer a job will take to complete. Such information can help users reconfigure jobs, and authors showed that users can reduce costs with proper selection of resource configurations as the application reaches 10% to 20% of completion.

The cost advisor module needs not only to be able to predict how long the jobs of a user will take to run but also how long an experiment composed of unknown jobs will take to run. HPC users usually run experiments in batches where they submit a group of jobs, analyze the produced results, and create new jobs based on the intermediate findings. Understanding this workflow and giving feedback to users on estimations of costs is crucial to see if the ongoing strategy of running experiments can meet budget restrictions (Silva et al. 2016). Advisors also need to consider data storage and movement, which is common for HPC users.

3.1.3 Large Contract Handler. Most of the work presented in the survey comes from academic works, reflecting one community of HPC. Enterprises with large HPC demands also look for alternatives to run their compute-intensive applications. For these enterprises, current HPC pricing models may not be sustainable depending on their current HPC infrastructure utilization levels. If the utilization is high, it might be more beneficial for them to maintain their own clusters, but if they use their clusters for sporadic projects, cloud becomes a cost-beneficial alternative. For those enterprises with high cluster utilizations, cloud providers need to come up with sustainable models that are beneficial for them and for their clients. This is a challenging work and a rich opportunity for research projects, as it involves capacity planning, theory for sustainable business models, negotiation protocols for multiple parties, and admission control mechanisms.

For large users, such as enterprises, a contract model for the cloud might be more appropriate. This model works similarly to what is currently seen in other markets, such as energy: instead of buying energy on the spot, large users make contracts with electricity providers, for example. The objective of these contracts is to minimize risks, such as fluctuations in price and discontinuities in supply. In some cases, the contract model for clouds may impose some limitations in elasticity. A contract may specify a minimum and maximum size or amount of resources. For companies that have more predictable workloads, this can still be suitable, as by outsourcing infrastructure management, they can focus on their businesses. Even though there are few publications in the scientific literature in this aspect, there are requests for proposals (RFPs) that show some of these trends (General Accountability Office 2013; NOAA 2016).

HPC environments are different from a traditional cloud infrastructure—clusters tend to have jobs with static number of resources, whereas cloud computing supports execution of elastic jobs on shared resources. It would be expensive, and a waste, to set up clusters with InfiniBand for users who do not run HPC jobs. Therefore, clusters with such high-speed networks need to be well sized because they are not easily expanded/shrunk compared to traditional cloud environments. It is then crucial for the cloud provider to have proper estimates of the demand to use such clusters. In addition, relying on a single client to rent a cluster may have negative impact if the cluster is not rented at full capacity or, even worse, if the client gives up on using the cluster.

One possible sustainable model is a multiparty contract, where multiple parties can have a contract with a cloud provider to create a managed infrastructure that can meet the demand of a group of clients. This helps clients have reduced costs and a cloud provider to set up an HPC environment that is suitable and easier to manage and reduces risks. In this contractual model, cloud providers could offer partitions of large clusters with high-speed networks that they can offer to multiple clients. Whenever clients ask for more resources, depending on the overall demand of all or part of the clients, the cloud provider can increase the cluster size.

3.1.4 DevOps. DevOps aims at integrating development (Dev) and operations (Ops) efforts to enable faster software delivery (Hüttermann 2012). DevOps has become popular with the maturity of cloud computing as an important mechanism for both development and hosting of services. In the HPC world, where most applications are still built to run on on-premise clusters, there are still several opportunities to create tools to support DevOps for HPC workloads and platforms—the

challenge is that HPC workloads are resource intensive and tests can become financially prohibitive. A few projects are pursuing the development of such technologies. For instance, the work from Rak et al. (2015), presented in the previous section, helps developers have insights on application performance, which is an important component of DevOps for the HPC community.

Tests of HPC applications can be much more complex and resource consuming than traditional Web applications being hosted in clouds. With regard to the fact that HPC software developers try their best to develop applications that not only work but are optimized to run in parallel using several resources, the development workflow can become slow. In addition, as cloud resources are usually not exclusive, the performance of the application under development can vary each time the developer runs a test. Therefore, there is a great opportunity for researchers working with DevOps for HPC to create a set of services to facilitate software development. HPC users would benefit from research studies to help understand the balance between different types of tests (those with lighter or heavier resource consumers) considering the various possible performance levels that cloud can offer. DevOps for HPC could also facilitate the use of elasticity (da Rosa Righi et al. 2016), which is a key differentiator of cloud computing compared to traditional cluster environments. As DevOps explores the concept of tests, in an HPC cloud environment, it could also consider different prices for using computing resources depending on estimations of the type of tests necessary to run and the tolerable resource noise from other users.

3.1.5 Automation APIs. Despite many efforts to create GUIs that allow drag-and-drop of components for software development and execution, automation is a cultural aspect that needs to be considered for the HPC community. The more automation, the more comfortable HPC users will be with a cloud platform. Examples of activities that require automation APIs are running a software system with hundreds or thousands of different input parameters, automating when a job should be executed in the cloud or on premises, and defining when new resources should be allocated or existing ones should be released.

Moreover, several HPC users have job submission scripts that were refined over the years. Scientific and business users also have scripts that handle data input and output. Such scripts need to be leveraged so that users do not start in the cloud from scratch. The challenge in the area of automation APIs is to be able to reuse existing scripts from traditional HPC environments and be able to easily extend them to explore the peculiarities and benefits of cloud, such as elasticity and management of resource allocation as a function of the available budget that the users have to run jobs. Researchers with background in software engineering can play a key role on how to achieve a great level of simplicity on this process. Researchers with a background in the human-computer interface can perform studies to understand which mechanisms of automation increase the productivity of users in this platform.

3.2 General Cloud Modules

In this section, we discuss challenges and research opportunities for three general cloud modules that bring benefits to HPC cloud environments.

3.2.1 Visualization and Data Management. Usually data management and visualization are handled separately, and we claim here that they need to be more integrated, especially considering HPC and big data. It is well known that data movement between on-premise and cloud infrastructure is a road blocker for some users (Gantikow et al. 2015). Therefore, it is essential for a cloud provider to offer a service that helps users determine which data needs to be moved from one place to another. It is common for HPC users to process data in batches, with analysis and planning happening between batches. Visualization and data management, when integrated, can minimize data movement, allowing users to analyze intermediate results via remote visualization. Then, when

ultimately needed, data can be transferred. This is a challenging area because it involves detection and modeling of user experience, predictors related to time and costs, and proper visualizations that bring value to users.

To create this module, multiple software components need to be in place. One of them is an estimator to help users determine the amount of time to transfer data between cloud and on-premise environments. Another component needs to be created to determine if a user is inside a session—that is, the user is submitting a stream of jobs with regular think times between new batches of jobs (Zakay and Feitelson 2012). By determining if the user is in a session, it is possible to provide proper suggestions on transferring data versus visualizing it remotely. The determination of the think time can also serve as a clue to understand if most of the time between the submission of new jobs is consumed by the user data analysis or the data transfer itself.

It is also relevant to enable users to verify the progress of their executions with rich visualizations. Users could monitor different metrics at both the system level and application level. These visualizations may help users identify if it is worth downloading data or even continuing their ongoing executions, which would then save precious time and money. Some of these visualizations could be done by using intermediate output files generated by applications. If users are working with popular applications, cloud providers could offer them such visualizations as a service in their platforms.

3.2.2 Flexible Software Licensing Models. For some industries, HPC software licenses are expensive. As pointed out in the UberCloud reports (Gentzsch and Yenier 2013, 2014), users still face unpredictable bills, especially as they lack precise information on how long they need to run their applications and therefore for how long they will need to use software licenses.

There are several types of software licenses, including pay-per-use licenses, shared licenses, floating licenses, site licenses, end-user licenses, and multifeature licenses. Most existing HPC software systems offer licenses to enable their usage on on-premise user environments, which can be single servers or computer clusters. Over time, software companies are enabling cloud pay-per-use licenses.

We claim here that having more flexible software licensing models can help users and cloud and software providers reduce costs and increase usability, respectively. Usage of a software system may not be flat nor peaky for some users. It may happen for periods of a few months and for a few hours a day, and therefore pay-per-use licenses depending on the prices may not be cost effective for the users, which leads them to look for alternative software systems. A cloud provider, with possible partnerships of software companies, could offer flexible software licenses depending on the usage profile of the users, which could be predefined or learned over time. Apart from bringing costs down to users, such flexibility may help users better determine their license expenses in advance, which is critical in HPC settings. In addition, this flexibility allows software companies to have broader usage of their software and cloud providers to attract more users to their environments. Multiparty licenses, involving multiple clients through a marketplace, could also help reduce costs and increase usage of SaaS. Apart from a cultural change required by several companies to start exploring more the usage models of cloud, an interesting research area is to monitor access to software licenses and bring hints to the software owners on the value of different license models per client. This involves analytics and software consumption predictions.

3.2.3 Value-Added Cloud Services. As cloud computing matured, providers started building services on top of infrastructure resources, building higher-value services for users. One example of value-added cloud services is the addition of machine learning APIs to major providers' clouds, such as IBM Bluemix, Google Cloud Platform, and Microsoft Azure. Most of the services currently available in these clouds are meant for users to consume already-existing APIs of pretrained

Table 5. Summary of Research Challenges to Enhance HPC Cloud Capabilities

Module	Importance	Research
<i>HPC-Aware Modules</i>		
Resource manager	Improve cloud performance for HPC workloads, sustainable HPC cloud business	Handle HPC and cloud workloads under same management, new pricing models based on queue, support for queues in cloud
Cost advisor	Avoid unexpected costs for users	Understand user workflow, predict future jobs and their performance
Large contract handler	Sustainable HPC cloud business for cloud providers and clients	Technology to allow simplified contracts for large users, facilitation of multiparty contracts
DevOps	Bring DevOps benefits to HPC users	Handle variable performance, minimize costs with resources, predict performance for various environments
Automation APIs	Keep HPC user tooling in cloud environment	Simplify migration of HPC scripts to cloud, and easily add cloud functionalities to them
<i>General Cloud Modules</i>		
Visualization and data management	Improve user experience and reduce storage and data transfer costs	Identify user workflow, predict data transfer costs
Flexible software licensing models	Reduce costs for providers and clients	Software licensing models based on user workflows, multiparty licenses
Value-added cloud services	Easy-to-use services for application development	Encapsulate and optimize complex services to accelerate development

models. Some providers allow users to upload user-trained models for making predictions using, such as TensorFlow (Abadi et al. 2016) trained models. Current trends suggest that the need to execute machine learning algorithms will increase in the coming years, and although deep learning methods have achieved very good performance in various domains, training procedures still lack in scalability when using stochastic gradient descent (SGD) (Keuper and Preundt 2016; Bhardwaj and Cong 2016), the default optimization method for neural networks. To improve the scalability of SGD-based learning algorithms, work has to be done in areas such as reducing communication overhead, parallelization of matrix operations, and effective data organization and distribution—problems that the HPC community has experience solving. Improving the scalability of such algorithms would allow for better resource usage and larger HPC cloud clusters for machine learning (Awan et al. 2017).

As we observed, there are several research directions in the area of HPC cloud. We have highlighted ongoing research in this article, and in this section we described what we believe needs further work from the research community. Table 5 summarizes these challenges.

4 CONCLUDING REMARKS

This article introduced a taxonomy and survey for the existing efforts in HPC cloud and a vision architecture to expand HPC cloud adoption with its respective research opportunities and

challenges. In the past years, cloud technologies became more mature, being able to support not only those initial e-commerce applications but also more complex traditional HPC, big data, and artificial intelligence applications.

The attempts of moving applications with heavy CPU and memory requirements to the cloud started by verifying the cost benefit of running those applications in the cloud against running them on already owned on-premise clusters. Various researchers used well-known HPC benchmarks and a few applications also common in the area. The goal was to understand not only performance but also monetary costs and how sustainable it would be to decommission their own clusters and move everything to the cloud. The main conclusion was that applications that were compute intensive and with high interprocessor communication could not scale well in the cloud, especially due to the lack of low-latency networks such as InfiniBand. However, a strong support seemed to be present when talking about embarrassingly parallel applications, which showed good performance with current cloud resources. There was also a visible concern about the difference of performance for multiple executions using the same group of allocated resources, which comes due to the resource-sharing aspect of cloud computing.

Meanwhile, several other efforts started to emerge. Researchers started to question the time to provision new machines in the cloud, how cloud could host services to help researchers with no IT background, and how to properly allocate resources in the cloud, with a great focus on hybrid clouds. From a business point of view, hybrid clouds seem to be the current model that brings sustainability for several companies with HPC workloads. With this model, it is possible to leverage existing computing infrastructure, and depending on the peak demands, part of workloads can be moved temporarily to the cloud. The amount of workload that should be moved to the cloud is highly dependent on the actual usage of existing resources—if the utilization level is low, a strategy to reduce fixed capacity and use cloud for peak demands may become a cost-effective alternative.

With the increase of microservices, and technologies for DevOps, transforming existing HPC applications into SaaS that are able to abstract the infrastructure layers can become a trend to make HPC more popular. In addition, research efforts can drive a better understanding of sustainable resource allocation pricing models for both cloud providers and HPC users. From a resource management perspective, it is well known that in some environments, users have the feeling of having access to free resources, making them allocate resources without assessing their actual needs. With cloud bringing the monetary aspect, user behavior may change in HPC settings, which also calls for research arms to have a better understanding about this shift.

As the in-house and cloud environments evolve, new technologies will appear. In this article, we focused on what has been done in the HPC cloud area. However, much research is devoted to technologies that are not essentially HPC cloud, but that will eventually make it to this environment. For example, new virtualization and containers technologies (Pahl et al. 2017; Zhang et al. 2016; Xavier et al. 2013) have been evolving and will play a role to reduce the performance gap between on-premise clusters and public clouds. Moreover, fast and low-latency networks will become more common^{4,5} (Zahid et al. 2016; Zhang et al. 2016, 2017) and may reshape the current offerings found in the cloud. New accelerators, such as GPUs (Giunta et al. 2010; Jermain et al. 2016; Li et al. 2017), FPGAs (Iordache et al. 2016; Kachris and Soudris 2016), TPUs (Jouppi et al. 2017), and frameworks, will become more pervasive and may make viable applications that currently are not present in the cloud environment.

Our main goal was to introduce a much broader perspective of interesting challenges and opportunities that HPC cloud can bring to researchers and practitioners. This is particularly relevant,

⁴ProfitBricks Network: <https://www.profitbricks.com/cloud-networks>.

⁵Azure Network: <https://azure.microsoft.com/en-us/pricing/details/cloud-services/>.

as HPC cloud platforms can become essential for new engineering and scientific discoveries, especially as the HPC community starts to embrace new workloads coming from big data and artificial intelligence.

ACKNOWLEDGMENTS

We thank the anonymous reviewers for their helpful comments in the preparation of this article.

REFERENCES

- Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, et al. 2016. TensorFlow: A system for large-scale machine learning. In *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI'16)*. 265–283.
- Moustafa AbdelBaky, Javier Diaz-Montes, Michael Johnston, Vipin Sachdeva, Richard L. Anderson, Kirk E. Jordan, and Manish Parashar. 2014. Exploring HPC-based scientific software as a service using CometCloud. In *Proceedings of the International Conference on Collaborative Computing: Networking, Applications, and Worksharing (CollaborateCom'14)*. IEEE, Los Alamitos, CA, 35–44.
- Moustafa AbdelBaky, Manish Parashar, Kirk Jordan, Hyunjoon Kim, Hani Jamjoom, Zon-Yin Shae, Gergina Pencheva, et al. 2012. Enabling high-performance computing as a service. *Computer* 45, 10, 72–80.
- Amazon. 2017a. Amazon Web Services. Retrieved December 5, 2017, from <https://aws.amazon.com>.
- Amazon. 2017b. Amazon Web Services (HPC). Retrieved December 5, 2017, from <https://aws.amazon.com/hpc/>.
- Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy Katz, Andy Konwinski, Gunho Lee, et al. 2010. A view of cloud computing. *Communications of the ACM* 53, 4, 50–58.
- J. P. Ashwini, C. Divya, and H. A. Sanjay. 2013. Efficient resource selection framework to enable cloud for HPC applications. In *Proceedings of the 4th International Conference on Computer and Communication Technology (ICCCCT'13)*. IEEE, Los Alamitos, CA, 34–38.
- Marcos D. Assunção, Rodrigo N. Calheiros, Silvia Bianchi, Marco A. S. Netto, and Rajkumar Buyya. 2015. Big data computing and clouds: Trends and future directions. *Journal of Parallel and Distributed Computing* 79, 3–15.
- Rocco Aversa, Beniamino Di Martino, Massimiliano Rak, Salvatore Venticinque, and Umberto Villano. 2011. Performance prediction for HPC on clouds. In *Cloud Computing: Principles and Paradigms*. Wiley, 437–456.
- Ammar Ahmad Awan, Khaled Hamidouche, Jahanzeb Maqbool Hashmi, and Dhableswar K. Panda. 2017. S-Caffe: Co-designing MPI runtimes and Caffe for scalable deep learning on modern GPU clusters. In *Proceedings of the 22nd ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*. ACM, New York, NY, 193–205.
- Mehdi Bahrami and Mukesh Singhal. 2015. *The Role of Cloud Computing Architecture in Big Data*. Springer International Publishing, 275–295.
- David H. Bailey, Eric Barszcz, John T. Barton, David S. Browning, Robert L. Carter, Leonardo Dagum, Rod A. Fatoohi, et al. 1991. The NAS parallel benchmarks. *International Journal of Supercomputing Applications* 5, 3, 63–73.
- Bartosz Balis, Kamil Figiela, Konrad Jopek, Maciej Malawski, and Maciej Pawlik. 2017. Porting HPC applications to the cloud: A multi-frontal solver case study. *Journal of Computational Science* 18, 106–116.
- Denilson Barbosa, Ioana Manolescu, and Jeffrey Xu Yu. 2009. Microbenchmark. In *Encyclopedia of Database Systems*, L. Liu and T. Ozsu (Eds.). Springer, Boston, MA, 1737.
- Mohamed Ben Belgacem and Bastien Chopard. 2015. A hybrid HPC/cloud distributed infrastructure: Coupling EC2 cloud resources with HPC clusters to run large tightly coupled multiscale applications. *Future Generation Computer Systems* 42, 11–21.
- G. Bruce Berriman, Gideon Juve, Ewa Deelman, Moira Regelson, and Peter Plavchan. 2010. The application of cloud computing to astronomy: A study of cost and performance. In *Proceedings of the IEEE International Conference on e-Science*. IEEE, Los Alamitos, CA.
- Onkar Bhardwaj and Guojing Cong. 2016. Practical efficiency of asynchronous stochastic gradient descent. In *Proceedings of the Workshop on Machine Learning in High Performance Computing Environments*. IEEE, Los Alamitos, CA, 56–62.
- N. J. Boden, D. Cohen, R. E. Felderman, A. E. Kulawik, C. L. Seitz, J. N. Seizovic, and Wen-King Su. 1995. Myrinet: A gigabit-per-second local area network. *IEEE Micro* 15, 1, 29–36.
- Julian Borrill, Jonathan Carter, Leonid Oliker, David Skinner, and Rupak Biswas. 2005. Integrated performance monitoring of a cosmology application on leading HEC platforms. In *Proceedings of the International Conference on Parallel Processing*. IEEE, Los Alamitos, CA, 119–128.
- Chris Bunch, Navraj Chohan, Chandra Krintz, and Khawaja Shams. 2011. Neptune: A domain specific language for deploying HPC software on cloud platforms. In *Proceedings of the 2nd International Workshop on Scientific Cloud Computing*. ACM, New York, NY, 59–68.
- Rajkumar Buyya. 1999. *High Performance Cluster Computing*. Prentice Hall, Upper Saddle River, NJ.

- Rajkumar Buyya and Diana Barreto. 2015. Multi-cloud resource provisioning with Aneka: A unified and integrated utilisation of Microsoft Azure and Amazon EC2 instances. In *Proceedings of the International Conference on Computing and Network Communications (CoCoNet'15)*. IEEE, Los Alamitos, CA.
- Rajkumar Buyya, Chee Shin Yeo, Srikumar Venugopal, James Broberg, and Ivona Brandic. 2009. Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. *Future Generation Computer Systems* 25, 6, 599–616.
- Adam G. Carlyle, Stephen L. Harrell, and Preston M. Smith. 2010. Cost-effective HPC: The community or the cloud? In *Proceedings of the International Conference on Cloud Computing Technology and Science (CloudCom'10)*. IEEE, Los Alamitos, CA, 169–176.
- Henri Casanova, Arnaud Legrand, Dmitrii Zagorodnov, and Francine Berman. 2000. Heuristics for scheduling parameter sweep applications in grid environments. In *Proceedings of the 9th Heterogeneous Computing Workshop*. IEEE, Los Alamitos, CA, 349–363.
- Brent Chun, David Culler, Timothy Roscoe, Andy Bavier, Larry Peterson, Mike Wawrzoniak, and Mic Bowman. 2003. PlanetLab: An overlay testbed for broad-coverage services. *SIGCOMM Computer Communication Review* 33, 3, 3–12.
- Phillip Church and Andrzej Goscinski. 2011. IaaS clouds vs. clusters for HPC: A performance study. In *Proceedings of the International Conference on Cloud Computing, GRIDs, and Virtualization*.
- Philip Church, Andrzej Goscinski, and Christophe Lefvre. 2015. Exposing HPC and sequential applications as services through the development and deployment of a SaaS cloud. *Future Generation Computer Systems* 4344, 0, 24–37.
- Philip Church, Adam Wong, Michael Brock, and Andrzej Goscinski. 2012. Toward exposing and accessing HPC applications in a SaaS cloud. In *Proceedings of the International Conference on Web Services (ICWS'12)*. IEEE, Los Alamitos, CA, 692–699.
- Philip C. Church and Andrzej M. Goscinski. 2014. A survey of cloud-based service computing solutions for mammalian genomics. *IEEE Transactions on Services Computing* 7, 4, 726–740.
- Adam Coates, Brody Huval, Tao Wang, David J. Wu, Bryan Catanzaro, and Andrew Y. Ng. 2013. Deep learning with COTS HPC systems. In *Proceedings of the 30th International Conference on Machine Learning*.
- Renato L. F. Cunha, Eduardo R. Rodrigues, Leonardo P. Tizzei, and Marco A. S. Netto. 2017. Job placement advisor based on turnaround predictions for HPC hybrid clouds. *Future Generation Computer Systems* 67, 35–46.
- Rodrigo da Rosa Righi, Vinicius Facco Rodrigues, Cristiano André da Costa, Guilherme Galante, Luis Carlos Erpen De Bona, and Tiago Ferreto. 2016. Autoelastic: Automatic resource elasticity for high performance applications in the cloud. *IEEE Transactions on Cloud Computing* 4, 1, 6–19.
- Marcos Dias de Assuncao, Alexandre di Costanzo, and Rajkumar Buyya. 2009. Evaluating the cost-benefit of using cloud computing to extend the capacity of clusters. In *Proceedings of the 18th ACM International Symposium on High Performance Distributed Computing*.
- Daniel de Oliveira, Eduardo Ogasawara, Fernanda Baião, and Marta Mattoso. 2010. SciCumulus: A lightweight cloud middleware to explore many task computing paradigm in scientific workflows. In *Proceedings of the 2010 IEEE 3rd International Conference on Cloud Computing (CLOUD'10)*. IEEE, Los Alamitos, CA, 378–385.
- Jeffrey Dean and Sanjay Ghemawat. 2008. MapReduce: Simplified data processing on large clusters. *Communications of the ACM* 51, 1, 107–113.
- Jack J. Dongarra, Piotr Luszczek, and Antoine Petit. 2003. The LINPACK benchmark: Past, present and future. *Concurrency and Computation: Practice and Experience* 15, 9, 803–820.
- Ifeanyi P. Egwutuoha, Shiping Chen, David Levy, and Rafael Calvo. 2013. Cost-effective cloud services for HPC in the cloud: The IaaS or the HaaS? In *Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA'13)*. 217.
- Jaliya Ekanayake and Geoffrey Fox. 2009. High performance parallel computing with clouds and cloud technologies. In *Proceedings of the International Conference on Cloud Computing*. 20–38.
- Huston Eubank. 2003. *Design Recommendations for High Performance Data Centers*. Rocky Mountain Institute.
- C. Evangelinos and C. Hill. 2008. Cloud computing for parallel scientific HPC applications: Feasibility of running coupled atmosphere-ocean climate models on Amazon's EC2. In *Proceedings of the Workshop on Cloud Computing and Its Applications (CCA'08)*.
- Roberto R. Expósito, Guillermo L. Taboada, Sabela Ramos, Juan Touriño, and Ramón Doallo. 2013. Performance analysis of HPC applications in the cloud. *Future Generation Computer Systems* 29, 1, 218–229.
- Pei Fan, Zhenbang Chen, Ji Wang, Zibin Zheng, and Michael R. Lyu. 2012. Topology-aware deployment of scientific applications in cloud computing. In *Proceedings of the International Conference on Cloud Computing (CLOUD'12)*.
- Pei Fan, Zhenbang Chen, Ji Wang, Zibin Zheng, and Michael R. Lyu. 2014. A topology-aware method for scientific application deployment on cloud. *International Journal of Web and Grid Services* 10, 4, 338–370.
- Dror G. Feitelson, Larry Rudolph, Uwe Schwiegelshohn, Kenneth C. Sevcik, and Parkson Wong. 1997. Theory and practice in parallel job scheduling. In *Proceedings of the International Workshop on Job Scheduling Strategies for Parallel Processing*.

- Wes Felter, Alexandre Ferreira, Ram Rajamony, and Juan Rubio. 2015. An updated performance comparison of virtual machines and Linux containers. In *Proceedings of the 2015 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS'15)*. IEEE, Los Alamitos, CA, 171–172.
- Ian Foster and Carl Kesselman. 2003. *The Grid 2: Blueprint for a New Computing Infrastructure*. Elsevier.
- Ian Foster, Carl Kesselman, and Steven Tuecke. 2001. The anatomy of the grid: Enabling scalable virtual organizations. *International Journal of High Performance Computing Applications* 15, 3, 200–222.
- Guilherme Galante, Luis Carlos Erpen De Bona, Antonio Roberto Muray, Bruno Schulze, and Rodrigo da Rosa Righi. 2016. An analysis of public clouds elasticity in the execution of scientific applications: A survey. *Journal of Grid Computing* 14, 2, 193–216.
- Holger Gantikow, Christoph Reich, Martin Knahl, and Nathan Clarke. 2015. A taxonomy for HPC-aware cloud computing. In *Proceedings of the BW-CAR Symposium on Information and Communication Systems (SinCom'15)*.
- General Accountability Office. 2013. GAO Protest Decision B-407073.3. Retrieved December 5, 2017, from <http://www.gao.gov/assets/660/655241.pdf>.
- Wolfgang Gentzsch and Burak Yenier. 2013. *The UberCloud HPC Experiment: Compendium of Case Studies*. Technical Report. Tabor Communications Inc. https://www.theubercloud.com/wp-content/uploads/2013/11/The_UberCloud_Compendium_2013_rnd2t3.pdf.
- Wolfgang Gentzsch and Burak Yenier. 2014. *The UberCloud Experiment: Technical Computing in the Cloud—2nd Compendium of Case Studies*. Technical Report. Tabor Communications Inc. http://www.theubercloud.com/wp-content/uploads/2014/06/The_UberCloud_Compendium_2014_rnd1j6.pdf.
- Giulio Giunta, Raffaele Montella, Giuseppe Agrillo, and Giuseppe Coviello. 2010. A GPGPU transparent virtualization component for high performance computing clouds. In *Proceedings of the European Conference on Parallel Processing*.
- William Gropp, Ewing Lusk, Nathan Doss, and Anthony Skjellum. 1996. A high-performance, portable implementation of the MPI message passing interface standard. *Parallel Computing* 22, 6, 789–828.
- William Gropp, Ewing Lusk, and Anthony Skjellum. 1999. *Using MPI: Portable Parallel Programming With the Message-Passing Interface*. Vol. 1. MIT Press, Cambridge, MA.
- Abhishek Gupta, Paolo Faraboschi, Filippo Gioachin, Laxmikant V. Kale, Richard Kaufmann, B.-S. Lee, Verdi March, Dejan Milojicic, and Chun Hui Suen. 2014. Evaluating and improving the performance and scheduling of HPC applications in cloud. *IEEE Transactions on Cloud Computing* 4, 3, 308–320.
- Abhishek Gupta, Laxmikant V. Kale, Filippo Gioachin, Verdi March, Chun Hui Suen, Bu-Sung Lee, Paolo Faraboschi, Richard Kaufmann, and Dejan Milojicic. 2013a. The who, what, why and how of high performance computing applications in the cloud. In *Proceedings of the 5th IEEE International Conference on Cloud Computing Technology and Science (CloudCom'13)*.
- Abhishek Gupta, Laxmikant V. Kale, Dejan Milojicic, Paolo Faraboschi, and Susanne M. Balle. 2013b. HPC-aware VM placement in infrastructure clouds. In *Proceedings of the International Conference on Cloud Engineering (IC2E'13)*. IEEE, Los Alamitos, CA, 11–20.
- Abhishek Gupta, Laxmikant V. Kalé, Dejan S. Milojicic, Paolo Faraboschi, Richard Kaufmann, Verdi March, Filippo Gioachin, Chun Hui Suen, and Bu-Sung Lee. 2012. Exploring the performance and mapping of HPC applications to platforms in the cloud. In *Proceedings of the 21st International Symposium on High-Performance Parallel and Distributed Computing*. ACM, New York, NY.
- Abhishek Gupta and Dejan Milojicic. 2011. Evaluation of HPC applications on cloud. In *Proceedings of the 6th Open Cirrus Summit (OCS'11)*.
- Hanan A. Hassan, Shima A. Mohamed, and Walaa M. Sheta. 2015. Scalability and communication performance of HPC on Azure Cloud. *Egyptian Informatics Journal* 17, 2, 175–182.
- Rashid Hassani, Md Aiatullah, and Peter Luksch. 2014. Improving HPC application performance in public cloud. *IERI Procedia* 10, 169–176.
- Qiming He, Shujia Zhou, Ben Kobler, Dan Duffy, and Tom McGlynn. 2010. Case study for running HPC applications in public clouds. In *Proceedings of the 19th ACM International Symposium on High Performance Distributed Computing*. ACM, New York, NY, 395–401.
- Zach Hill and Marty Humphrey. 2009. A quantitative analysis of high performance computing with Amazon's EC2 infrastructure: The death of the local cluster? In *Proceedings of the 10th IEEE/ACM International Conference on Grid Computing*. IEEE, Los Alamitos, CA.
- Qian Huang. 2014. Development of a SaaS application probe to the physical properties of the Earth's interior: An attempt at moving HPC to the cloud. *Computers and Geosciences* 70, 147–153.
- Michael Hüttermann. 2012. *DevOps for Developers*. Apress.
- InfiniBand Trade Association. 2000. *InfiniBand Architecture Specification: Release 1.0*. InfiniBand Trade Association.
- Intel. 2017. Intel MPI Benchmarks User Guide. Retrieved December 5, 2017, from <https://software.intel.com/en-us/imb-user-guide-pdf>.

- Anca Iordache, Guillaume Pierre, Peter Sanders, Jose Gabriel de F. Coutinho, and Mark Stillwell. 2016. High performance in the cloud with FPGA groups. In *Proceedings of the 9th International Conference on Utility and Cloud Computing*. ACM, New York, NY.
- Keith R. Jackson, Lavanya Ramakrishnan, Krishna Muriki, Shane Canon, Shreyas Cholia, John Shalf, Harvey J. Wasserman, and Nicholas J. Wright. 2010. Performance analysis of high performance computing applications on the Amazon Web Services cloud. In *Proceedings of the International Conference on Cloud Computing Technology and Science (CloudCom'10)*. IEEE, Los Alamitos, CA, 159–168.
- C. L. Jermain, G. E. Rowlands, R. A. Buhrman, and D. C. Ralph. 2016. GPU-accelerated micromagnetic simulations using cloud computing. *Journal of Magnetism and Magnetic Materials* 401, 320–322.
- Norman P. Jouppi, Cliff Young, Nishant Patil, David Patterson, Gaurav Agrawal, Raminder Bajwa, Sarah Bates, et al. 2017. In-datacenter performance analysis of a tensor processing unit. In *Proceedings of the International Symposium on Computer Architecture (ISCA'17)*.
- Christoforos Kachris and Dimitrios Soudris. 2016. A survey on reconfigurable accelerators for cloud computing. In *Proceedings of the International Conference on Field Programmable Logic and Applications (FPL'16)*. IEEE, Los Alamitos, CA.
- Mohammad Mahdi Kashaf and Jörn Altmann. 2011. A cost model for hybrid clouds. In *Proceedings of the International Workshop on Grid Economics and Business Models*. 46–60.
- Janis Keuper and Franz-Josef Preundt. 2016. Distributed training of deep neural networks: Theoretical and practical limits of parallel scalability. In *Proceedings of the Workshop on Machine Learning in High Performance Computing Environments*. IEEE, Los Alamitos, CA, 19–26.
- Brian Kocoloski, Jiannan Ouyang, and John Lange. 2012. A case for dual stack virtualization: Consolidating HPC and commodity applications in the cloud. In *Proceedings of the 3rd ACM Symposium on Cloud Computing*. ACM, New York, NY, 23.
- Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2012. ImageNet classification with deep convolutional neural networks. In *Proceedings of the Advances in Neural Information Processing Systems*. 1097–1105.
- Yu-Kwong Kwok and Ishaq Ahmad. 1999. Static scheduling algorithms for allocating directed task graphs to multiprocessors. *ACM Computing Surveys* 3, 4, 406–471.
- He Li, Kaoru Ota, Mianxiong Dong, Athanasios Vasilakos, and Koji Nagano. 2017. Multimedia processing pricing strategy in GPU-accelerated cloud computing. *IEEE Transactions on Cloud Computing* (2017).
- Xiaorong Li, Henry Palit, Yong Siang Foo, and Terence Hung. 2011. Building an HPC-as-a-service toolkit for user-interactive HPC services in the cloud. In *Proceedings of the IEEE Workshops of the International Conference on Advanced Information Networking and Applications (WAINA'11)*. IEEE, Los Alamitos, CA, 369–374.
- Robert Love. 2003. Kernel kornet: CPU affinity. *Linux Journal* 2003, 111, 8.
- Aniruddha Marathe, Rachel Harris, David Lowenthal, Bronis R. De Supinski, Barry Rountree, and Martin Schulz. 2014. Exploiting redundancy for cost-effective, time-constrained execution of HPC applications on Amazon EC2. In *Proceedings of the 23rd International Symposium on High-Performance Parallel and Distributed Computing*. ACM, New York, NY.
- Aniruddha Marathe, Rachel Harris, David K. Lowenthal, Bronis R. De Supinski, Barry Rountree, Martin Schulz, and Xin Yuan. 2013. A comparative study of high-performance computing on the cloud. In *Proceedings of the 22nd International Symposium on High-Performance Parallel and Distributed Computing (HPDC'13)*.
- Paul Marshall, Henry Tufo, and Kate Keahey. 2013. High-performance computing and the cloud: A match made in heaven or hell? *XRDS: Crossroads, the ACM Magazine for Students* 19, 3, 52–57.
- Gabriel Mateescu, Wolfgang Gentzsch, and Calvin J. Ribbens. 2011. Hybrid computing—where HPC meets grid and cloud computing. *Future Generation Computer Systems* 27, 5, 440–453.
- Marta Mattoso, Jonas Dias, Kary A. C. S. Ocaña, Eduardo Ogasawara, Flavio Costa, Felipe Horta, Vítor Silva, and Daniel de Oliveira. 2015. Dynamic steering of HPC scientific workflows: A survey. *Future Generation Computer Systems* 46, 100–113.
- Viktor Mauch, Marcel Kunze, and Marius Hillenbrand. 2013. High performance cloud computing. *Future Generation Computer Systems* 29, 6, 1408–1416.
- Peter Mell and Tim Grance. 2011. *The NIST Definition of Cloud Computing*. Technical Report. National Institute of Standards and Technology (NIST), Gaithersburg, MD.
- Jeffrey Napper and Paolo Bientinesi. 2009. Can cloud computing reach the TOP500? In *Proceedings of the Combined Workshops on UnConventional High Performance Computing Workshop Plus Memory Access Workshop (UCHPC-MAW'09)*.
- Marco A. S. Netto, Renato L. F. Cunha, and Nicole Sultanum. 2015. Deciding when and how to move HPC jobs to the cloud. *IEEE Computer* 48, 11, 86–89.
- NOAA. 2016. NOAA Completes Weather and Climate Supercomputer Upgrades. Retrieved December 5, 2017, from <http://www.noaanews.noaa.gov/stories2016/011116-noaa-completes-weather-and-climate-supercomputer-upgrades.html>.
- Daniel Nurmi, Rich Wolski, Chris Grzegorzczak, Graziano Obertelli, Sunil Soman, Lamia Youseff, and Dmitrii Zagorodnov. 2008. The eucalyptus open-source cloud-computing system. In *Proceedings of the 1st Workshop on Cloud Computing and Its Applications*. IEEE, Los Alamitos, CA.

- Kary A. C. S. Ocaña, Daniel de Oliveira, Eduardo Ogasawara, Alberto M. R. Dávila, Alexandre A. B. Lima, and Marta Mattoso. 2011. SciPhy: A cloud-based workflow for phylogenetic analysis of drug targets in protozoan genomes. In *Proceedings of the Brazilian Symposium on Bioinformatics*. 66–70.
- Simon Ostermann, Alexandria Iosup, Nezih Yigitbasi, Radu Prodan, Thomas Fahringer, and Dick Epema. 2009. A performance analysis of EC2 cloud computing services for scientific computing. In *Proceedings of the International Conference on Cloud Computing*. 115–131.
- Claus Pahl, Antonio Brogi, Jacopo Soldani, and Pooyan Jamshidi. 2017. Cloud container technologies: A state-of-the-art review. *IEEE Transactions on Cloud Computing* PP, 99, 1.
- Manish Parashar, Moustafa AbdelBaky, Ivan Rodero, and Aditya Devarakonda. 2013. Cloud paradigms and practices for computational and data-enabled science and engineering. *Computing in Science and Engineering* 15, 4, 10–18.
- Dana Petcu, Horacio González-Vélez, Bogdan Nicolae, Juan Miguel García-Gómez, Elies Fuster-Garcia, and Craig Sheridan. 2014. Next generation HPC clouds: A view for large-scale scientific and data-intensive applications. In *Proceedings of the European Conference on Parallel Processing*. 26–37.
- Massimiliano Rak, Mauro Turtur, and Umberto Villano. 2015. Early prediction of the cost of cloud usage for HPC applications. *Scalable Computing: Practice and Experience* 16, 3, 303–320.
- Daniel A. Reed and Jack Dongarra. 2015. Exascale computing and big data. *Communications of the ACM* 58, 7, 56–68.
- Harald Richter. 2016. About the suitability of clouds in high-performance computing. arXiv:1601.01910.
- Eduardo Roloff, Matthias Diener, Alexandre Carissimi, and Philippe O. A. Navaux. 2012. High performance computing in the cloud: Deployment, performance and cost efficiency. In *Proceedings of the 2012 IEEE 4th International Conference on Cloud Computing Technology and Science (CloudCom'12)*. IEEE, Los Alamitos, CA, 371–378.
- Tiago Pais Pitta De Lacerda Ruivo, Gerard Bernabeu Altayo, Gabriele Garzoglio, Steven Timm, Hyun Woo Kim, Seo-Young Noh, and Ioan Raicu. 2014. Exploring InfiniBand hardware virtualization in OpenNebula towards efficient high-performance computing. In *Proceeding of the 14th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing (CCGrid'14)*. IEEE, Los Alamitos, CA, 943–948.
- Iman Sadooghi and Ioan Raicu. 2013. Understanding the cost of the cloud for scientific applications. In *Proceedings of the 2nd Greater Chicago Area System Research Workshop (GCASR'13)*.
- Jahanzeb Sherwani, Nosheen Ali, Nausheen Lotia, Zahra Hayat, and Rajkumar Buyya. 2004. Libra: A computational economy-based job scheduling system for clusters. *Software: Practice and Experience* 34, 6, 573–590.
- Justin Y. Shi, Moussa Taifi, Aakash Pradeep, Abdallah Khreishah, and Vivek Antony. 2012. Program scalability analysis for HPC cloud: Applying Amdahl's law to NAS benchmarks. In *Proceedings of the High Performance Computing, Networking, Storage, and Analysis (SCC'12), 2012 SC Companion*. IEEE, Los Alamitos, CA, 1215–1225.
- B. Silva, M. A. S. Netto, and R. L. F. Cunha. 2016. SLA-aware interactive workflow assistant for HPC parameter sweeping experiments. In *Proceedings of the International Workshop on Workflows in Support of Large-Scale Science in conjunction with the International Conference for High Performance Computing, Networking, Storage, and Analysis (WORKS@SC'16)*. IEEE, Los Alamitos, CA.
- Stephen Soltész, Herbert Pözl, Marc E. Fluczynski, Andy Bavier, and Larry Peterson. 2007. Container-based operating system virtualization: A scalable, high-performance alternative to hypervisors. *SIGOPS Operating Systems Review* 41, 3, 275–287.
- Tamarai Selvi Somasundaram and Kannan Govindarajan. 2014. CLOUDRB: A framework for scheduling and managing high-performance computing (HPC) applications in science cloud. *Future Generation Computer Systems* 34, 47–65.
- Thomas Sterling and Dylan Stark. 2009. A high-performance computing forecast: Partly cloudy. *Computing in Science and Engineering* 11, 4, 42–49.
- Thomas Lawrence Sterling. 2002. *Beowulf Cluster Computing with Linux*. MIT Press, Cambridge, MA.
- TOP500. 2017. TOP500 Supercomputing Sites. Retrieved December 5, 2017, from <https://www.top500.org>.
- U.S. Department of Energy Office of Science. 2011. *The Magellan Report on Cloud Computing for Science*. Technical Report. Office of Advanced Scientific Computing Research (ASCR), Washington, DC.
- Blesson Varghese and Rajkumar Buyya. 2017. Next generation cloud computing: New trends and research directions. *Future Generation Computer Systems* 79, 3, 849–861.
- Christian Vecchiola, Suraj Pandey, and Rajkumar Buyya. 2009. High-performance cloud computing: A view of scientific applications. In *Proceedings of the 10th International Symposium on Pervasive Systems, Algorithms, and Networks*. IEEE, Los Alamitos, CA.
- Jerome Vienne, Jitong Chen, Md Wasi-Ur-Rahman, Nusrat S. Islam, Hari Subramoni, and Dhableswar K. Panda. 2012. Performance analysis and evaluation of InfiniBand FDR and 40GigE RoCE on HPC and cloud computing systems. In *Proceedings of the IEEE 20th Annual Symposium on High-Performance Interconnects (HOTI'12)*. IEEE, Los Alamitos, CA, 48–55.
- Adam K. L. Wong and Andrzej M. Goscinski. 2013. A unified framework for the deployment, exposure and access of HPC applications as services in clouds. *Future Generation Computer Systems* 29, 6, 1333–1344.

- Miguel G. Xavier, Marcelo V. Neves, Fabio D. Rossi, Tiago C. Ferreto, Timoteo Lange, and Cesar A. F. De Rose. 2013. Performance evaluation of container-based virtualization for high performance computing environments. In *Proceedings of the Euromicro International Conference on Parallel, Distributed, and Network-Based Processing (PDP'13)*. IEEE, Los Alamitos, CA.
- Xiaoyu Yang, David Wallom, Simon Waddington, Jianwu Wang, Arif Shaon, Brian Matthews, Michael Wilson, et al. 2014. Cloud computing in e-Science: Research challenges and opportunities. *Journal of Supercomputing* 70, 1, 408–464.
- Feroz Zahid, Ernst Gunnar Gran, and Tor Skeie. 2016. Realizing a self-adaptive network architecture for HPC clouds. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage, and Analysis (SC'16) Doctoral Showcase*.
- Netanel Zakay and Dror G. Feitelson. 2012. On identifying user session boundaries in parallel workload logs. In *Proceedings of the Workshop on Job Scheduling Strategies for Parallel Processing*. 216–234.
- Peter Zaspel and Michael Griebel. 2011. Massively parallel fluid simulations on Amazon's HPC cloud. In *Proceedings of the 1st International Symposium on Network Cloud Computing and Applications (NCCA'11)*. IEEE, Los Alamitos, CA, 73–78.
- Yan Zhai, Mingliang Liu, Jidong Zhai, Xiaosong Ma, and Wenguang Chen. 2011. Cloud versus in-house cluster: Evaluating Amazon cluster compute instances for running MPI applications. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage, and Analysis (SC'11): State of the Practice Reports*. ACM, New York, NY, 11.
- Jie Zhang, Xiaoyi Lu, and Dhabaleswar K. Panda. 2016. High performance MPI library for container-based HPC cloud on InfiniBand clusters. In *Proceedings of the International Conference on Parallel Processing (ICPP'16)*. IEEE, Los Alamitos, CA.
- Jie Zhang, Xiaoyi Lu, and Dhabaleswar K. Panda. 2017. Designing locality and NUMA aware MPI runtime for nested virtualization based HPC cloud with SR-IOV enabled InfiniBand. In *Proceedings of the 13th ACM SIGPLAN/SIGOPS International Conference on Virtual Execution Environments*. ACM, New York, NY.
- Han Zhao and Xiaolin Li. 2012. Designing flexible resource rental models for implementing HPC-as-a-service in cloud. In *Proceedings of the 2012 IEEE 26th International Parallel and Distributed Processing Symposium and Workshops (IPDPSW'12) and PhD Forum*. IEEE, Los Alamitos, CA, 2550–2553.

Received February 2017; revised October 2017; accepted October 2017