Final Project Report

Workshop on Clusters, Clouds, and Data for Scientific Computing
(CCDSC 2016)

October 3–6, 2016

Châteauform'
La Maison des Contes
427 Chemin de Chanzé, France

Sponsors:

National Science Foundation (Award 1606551); AMD; PGI; NVIDIA; Intel; Mellanox; GENCI; CGG-Veritas; University of Tennessee, Knoxville/Innovative Computing Laboratory; Grenoble Alps University

Accomplishments

Context: Two Exceptional Trends

The research areas of cluster, cloud and data analytics computing, which today provide fundamental infrastructure for all areas of advanced computational science, are being radically transformed by the convergence of at least two unprecedented trends.

The first trend is the ongoing emergence of multicore and hybrid microprocessor designs that are ushering in a new era of computing in which system designers must accept energy usage as a first-order constraint, and application designers must be able to exploit parallelism and data locality to an uncommon degree. As the research community is rapidly becoming aware, the components of the traditional HPC software stack are poorly matched to the characteristics of systems based on these new architectures, which have hundreds of thousands of nodes, millions of cores, GPU accelerators, and reduced bandwidth and memory per core.

The second trend is the dramatic escalation in the amount of data that leading-edge scientific applications, and the communities that use them, are either generating or trying to analyze. A key problem in such data-intensive science lies not only in the sheer volume of bits that must be processed and managed but also in the logistical problems associated with making the data of most-current interest available to participants in large national and international collaborations sitting in different administrative domains, spread across the wide-area network and wanting to use the diverse resources—clusters, clouds, and data.

For the cyberinfrastructure research community, these new conditions raise a broad set of challenging design and deployment issues in areas such as scalability, programmability, performance, interoperability, resilience, resource virtualization, data logistics, and system management.

Project Goal

The Workshop on Clusters, Clouds, and Data Analytics for Scientific Computing (CCDASC) is directed at addressing the aforementioned design and deployment issues by providing a forum for creative dialogue between leading-edge researchers in the field to produce the kind of coordinated review, penetrating analysis, and new insights that can help translate the unprecedented trends into highly consequential benefits for the scientific community.

This workshop articulated a set of concrete approaches and proposals to develop software infrastructure for data analytics used in scientific computing across a broad spectrum of research domains.

We were able to accomplish the following:

- Survey and analyze the key deployment, operational, and usage issues for clusters, clouds, and grids, with special focus on the discontinuities produced by multicore and hybrid architectures, data-intensive science, and the increasing need for wide- or local-area interaction
- Discuss the current state of the art in clusters, clouds, and data analytics and identify interesting questions and limitations
- Document experiences with clusters, clouds, and grids with respect to the science research communities and science domains that are benefiting from the technologies
- Discuss interoperability among disparate clouds, as well as between various clouds and grids, and the impact of this interoperability on the domain sciences
- Discuss directions for future research and development in the context of disruptive trends and technologies and the recognized gaps in the current state of the art.

Specific Objectives

The workshop examined new insights concerning the commonalities that exist between clusters, clouds, and data analytics and explored the ways they must interface and interoperate with one another to support today's data- and compute-intensive collaborations.

Significant Results

Workshop participants were able to focus on the specific issues in cluster, cloud, and grid computing delineated below. Thirteen journal articles based on the workshop are now under review for publication.

Development issues in cluster, clouds, and data computing:

- Designing to scale: parallelism, energy, interoperability
- Component and system interoperability
- Reliability and fault tolerance
- Improvements to the storage and I/O paradigm
- Campus-level challenges with data-intensive research

Operational and usability issues for cluster, cloud, and data computing

- Programming abstractions
- System issues (performance monitoring, system software)
- Power management

In terms of the organizers and participants involved, the following results were produced.

- This workshop represented one in a series of extremely productive biennial meetings involving collaborations between US, European, and Asian colleagues.
- The organizers are ideally situated in several respects to recruit outstanding participants from the field and to leverage the resulting ideas, insights, and innovations for the benefit of the broader research community.

Key Outcomes or Other Achievements

This workshop has a history of nearly twenty years. The series represents a collaboration between US and French colleagues who are well situated to recruit outstanding international participants from the field and leverage the resulting ideas, insights, and innovations to benefit the broader research community.

These meetings provide a model of what such international collaborations should be and continues to supply the underpinnings for international partnerships in cyberinfrastructure (hardware and software) that remain essential for rapid progress in scientific fields vital to the long-term welfare of our society.

Opportunities for Training and Professional Development

In making our final selection of invitees for the workshop, we developed a list of promising junior faculty from which to draw. We worked with Shirley Moore of the University of Texas, El Paso, to recruit appropriate individuals from underrepresented groups to this meeting.

Products from the Workshop

Journal Papers

- Abdelbaky, Moustafa; Diaz-Montes, Javier; Parashar, Manish (2017). Software-Defined Environments for Science & Engineering. International Journal of High Performance Computing and Supercomputer Applications.
- Aupy, Guillaume; Benoit, Anne; Dai, Sicheng; Pottier, Loïc; Raghavan, Padma (2017). Co-scheduling Amdahl applications on cache-partitioned systems. International Journal of High Performance Computing and Supercomputer Applications.
- Badia, Rosa; Conejero, Javier; Corella, Sandra; Labarta, Jesus (2017). Task-based programming in COMPSs to converge from HPC to Big Data. International Journal of High Performance Computing and Supercomputer Applications.
- Benoit, Anne; Pottier, Loïc; Robert, Yves (2017). Resilient co-scheduling of malleable applications. International Journal of High Performance Computing and Supercomputer Applications.
- Chalios, Charalampos; Georgakoudis, Giorgis; Tovletoglou, Konstantinos; Karakonstantis, George; Vandierendonck, Hans (2017). DARE: Data-Access Aware

Refresh via Spatial-Temporal Application-Resilience on Commodity Servers. International Journal of High Performance Computing and Supercomputer Applications.
- Deelman, Ewa; Peterka, Tom; Altintas, Ilkay; Carothers, Christopher; Kleese Van Dam, Kerstin (2017). The Future of Scientific Workflows. International Journal of High Performance Computing and Supercomputer Applications.
- Desprez, Frederic; Videau, Brice; Pouget, Kevin; Genovese, Luigi; Méhaut, Jean-François (2017). BOAST: a Metaprogramming Framework to Produce Portable and Efficient Computing Kernels for HPC Applications. International Journal of High Performance Computing and Supercomputer Applications.
- Fox, Geoffrey; Kamburugamuve, Supun; Wickramasinghe, Pulasthi ; Ekanayake, Saliya (2017). Anatomy of machine learning algorithm implementations in MPI, Spark, and Flink. International Journal of High Performance Computing and Supercomputer Applications.
- Ho, Minh Quan; Obrecht, Christian; Tourancheau, Bernard; Dupont De Dinechin, Benoît; Hascoet, Julien (2017). Improving 3D Lattice Boltzmann Method memory bound stencil with asynchronous transfers on many-core processors. International Journal of High Performance Computing and Supercomputer Applications.
- Jeannot, Emmanuel; Georgiou, Yiannis; Mercier, Guillaume; Villiermet, Adèle (2017). Topology-Aware Job Mapping. International Journal of High Performance Computing and Supercomputer Applications.
- Lefevre, Laurent; Benoit, Anne; Orgerie, Anne-Cécile; Rais, Issam (2017). Combined shutdown policies for large scale computing systems: models and simulations. International Journal of High Performance Computing and Supercomputer Applications.
- Lusk, Ewing; Butler, Ralph; Pieper, Steven (2017). Evolution of a Minimal Parallel Programming Model. International Journal of High Performance Computing and Supercomputer Applications.
- Robert, Yves; Bosilca, George; Bouteiller, Aurélien; Dongarra, Jack; Thomas, Herault (2017). A Failure Detector for HPC Platforms. International Journal of High Performance Computing and Supercomputer Applications.

Impacts

Development of the Principle Disciplines

Having the capability to exploit the potential interoperability and convergence of high-speed-computing and data-intensive applications is essential to the future of cluster, cloud, and data computing, especially given the kind of global research collaborations that are increasingly dominant in many important application areas.

Even taken in isolation, these powerful trends are widely viewed as disruptive of conventional models and methods; considered in combination, however, they will clearly demand significant innovation across a range of different related areas. Perhaps the biggest point of divergence between the scientific-computing and big-data ecosystems are programming models and tools. Accordingly, this workshop provided, in the context of these new conditions, the necessary coordinated review and analysis of the major issues of cluster, cloud, and data-driven computing to elicit new insights.

The architectural similarities between clusters, clouds, and data analytics infrastructure—and the fact that high-performance clusters typically make up the major compute nodes of the other two—means that programming, deployment, operational, and usage issues surrounding computational clouds and grids form a superset of the issues that revolve around clusters. This workshop captured ways of leveraging both the similarities and the differences that exist between these basic types of computing cyberinfrastructure. At the same time, the continuing growth of wide-area, data-intensive collaboration in science requires that users be prepared to work with clusters, clouds, and data in combination with one another and strive to develop the kind of coordinated work flow that facilitates maximally productive, multiparty research.

Clearly, as we continue to experience explosive growth in the data resources—particularly in contexts where different disciplines are attempting to combine and synthesize their efforts—new thinking will be required to develop novel protocols and APIs on which such emerging communities can converge. Achieving the kind of seamless interoperability required for workflows was a point of emphasis for the participants of this workshop. Consequently, they provided inputs based on absorbing and understanding the experiences of application communities, working on campuses, and collaborating across the globe. Those types of information will be critical to the development of viable solutions.

Other Disciplines

This meeting produced a set of concrete approaches and proposals to develop software infrastructure for data analytics for scientific computing across a broad spectrum of research domains.

Human Resources

A major part of the success of this workshop and the previous ones in the meeting series is attributable to the outstanding quality of the individuals drawn from national and international communities to attend and present at this meeting. As compared with other gatherings, the people at these gatherings are uniformly of “invited speaker” quality.

Furthermore, we worked with our colleagues to recruit promising junior faculty—especially women, minority and underserved individuals—before invitations were sent. A primary consideration in developing our list was to expand the range of workshop topics to include issues such as economic efficiency for cyberinfrastructure, workflows across all infrastructure tiers down to the campus level, and the experiences of different research fields in using various types of cyberinfrastructure innovatively and productively.