Calcul Québec

Abstracts of Calcul Québec's team members communications for the 2019 CANHEIT-ARC conference

Résumés des communications de Calcul Québec dans le cadre de la conférence CANHEIT-ARC 2019 (english)

Big data/Data analytics

Making A Case For Data Visualization. Guys, It's Not Just Pretty Colors.

Julie Faure-Lacroix, Scientific liaison agent

High Performance Computing (HPC) has witnessed a recent shift in its user base. What used to be mainly a select club of physicists and chemists is now opening its doors to groups of non-traditional users. They are often pioneers in their fields because they are the first ones to explore large datasets or extremely complex models. With these new users come new unexpected uses for computing clusters. With these new uses come new demands that used to be considered as trivial or optional, like real-time data visualization. These demands often require access to fast, reliable connections to accommodate the transfer of visual data in real time. At Compute Canada, we still struggle with some aspects of it, mainly because there are many workflows and use cases and some of them are challenging given the limitations and design of our hardware. From medical imagery to hydrology, I present some real-life examples of how our structure answers the needs of researchers and how it struggles to do so. With this, I aim to add to the conversation started by Compute Canada's data visualization team and hopefully stimulate the creation of fresh new solutions to existing problems.

People

How to hire an awesome intern

Félix-Antoine Fortin, Professional services, and Julie Faure-Lacroix, Scientific liaison agent

Interns can inject your team with creative energy, increase productivity and drive new projects…or they can suck all your precious time. How do you hire the elusive intern? You know…the one that comes in with some of the answers, but not all…fits seamlessly into your team culture…and thinks that "initiative" isn't just a buzzword that looks good on their LinkedIn profile, but is rather an essential character trait?

Find out how a clever marketing campaign combined with a hackathon can help you hire your very own awesome intern.

Moving beyond diversity to inclusion

Julie Faure-Lacroix, Scientific liaison agent

Diversity can be a simple numbers game, but inclusion is where the real work begins. Join BOF leader Julie Faure-Lacroix where she shares her real-world insights learnt from her work growing a Women in High Performance Computing (WHPC) Chapter within Québec. And jump in and share your own hard-earned insights into how we can increase the diversity within our teams, our institutions, and our industry…and then how can we make this lasting change by focusing on inclusion.

Teaching HPC to Women: Weirdly, We Haven't Given It Much Thought

Julie Faure-Lacroix, Scientific liaison agent

As Calcul Québec progresses towards the status of Women In High Performance Computing (Women in HPC) chapter, we continue to struggle with recruiting female employees. However, if we shift our focus away from our workforce to our user base, we realize that equally important work needs to be done regarding our teaching approach of HPC to women and minorities. Because courses, summer schools, workshops, and one-on-one tutoring are a major part of Calcul Québec's definition, we must strive to deliver quality services not only to the easy-to-reach cis male audience, which composes the bulk of our users but also to those who historically were not preponderant users of HPC facilities. Women in HPC's main focus is inclusivity, women inclusivity in particular, and this inspired me to create a new type of all-woman introductory class to HPC. This course, taught by a woman to an exclusively female audience, aims at creating a safe learning environment for women and hopefully reach out to those who would not usually dare attend a computer class with a majority of men. Here I discuss the takeaways from that teaching experiment and propose a framework to expand those classes to the province of Québec.

Research and IT

Providing A Unified User Environment for Canada's National Advanced Computing Centers

Maxime Boissonneault, User support team lead

Exploiting an advanced computing platform consisting of several clusters distributed across the second-largest country in the world is challenging. Each cluster may run a different operating system, use a different generation of CPU, GPU, or network fabric, or be managed by a different team of system administrators. Presenting a unified software environment can tremendously facilitate the task of supporting researchers, but is difficult to implement. This is nevertheless what Compute Canada set out to do in 2016, in the midst of deploying a new generation of large clusters. In order to achieve this goal, we had to find software solutions to solve these challenges. Distribution, portability and performance were three important technical criteria for us. We also had to consider the practicality of each approach for our users, and reproducibility of software installations performed by staff located at various sites across Canada.

In this talk, we present the solution that we created, which has allowed Compute Canada to serve the needs of over 10,000 researchers across the country. It is used on over 20 different clusters with heterogeneous configurations, from CPU architectures as old as Nehalem or Opteron all the way up to Skylake, with and without GPUs, with InfiniBand, Ethernet or OmniPath as the network fabric, and with Slurm or Torque/Moab as the scheduler. It presents a unified software environment to the users, providing over 600 different scientific applications that are available in over 4,000 different combinations of version, compiler and CPU architecture.

Bridging Campus Data-Intensive Research to the National ARC Platform

Florent Parent, Systems architecture lead

With the increasing demand for data storage, and the consolidation of the national Advanced Research Computing (ARC) resources, researchers and campus IT are facing an important challenge: bridging the data-intensive research on campus to the national ARC platform.

To make effective use of the Compute Canada Federation resources, data-intensive researchers must be able to move large datasets across the high-speed National Research and Education Network in Canada (NREN).

At the national level, the Compute Canada national sites are providing 100 Gbps connectivity to the CANARIE IP network, including data transfer nodes to make high-speed data transfer possible. The design and implementation of this network connectivity is done in collaboration with the National Research and Education Networks (CANARIE, BCNET, ORION, and RISQ).

For researchers on campus, the environment is more challenging. The data flowing from the campus lab will typically go through multiple network domains, from the research LAN, campus core, firewall, and possibly other devices (middleboxes) before reaching the high-speed R&E

networks. Also, large data transfer flows are competing with a high number of short-lived flows that are characteristic of web and other Internet traffic on campus.
In such context, it is not possible to utilize the full potential of existing high performance networks (NREN) for data-intensive science.
In this presentation, we will describe the implementation of a campus DMZ (based on the Science DMZ framework defined by ESnet) dedicated to research computing. Université Laval has built a new data center specialized in data collection, processing and valorization. Among the applications and tools offered to researchers, a campus DMZ provides an essential service to bridge the researcher on campus to the national ARC resources.
We will describe how collaboration is essential between Campus IT and Research Computing specialists to build and sustain a campus DMZ platform.

Technologies and trends

Terraform for Everybody

Drew Leske, Senior Programmer/Analyst at Compute Canada and Félix-Antoine Fortin, Professional services at Calcul Québec

Terraform is an “infrastructure as code” toolset that provides for the management of servers, devices, and networking via a declarative language. Terraform is open source and well supported by HashiCorp and the community with plugins for OpenStack, cloud providers, network devices, and other infrastructure components and software. In this presentation we will present some of the ways we’re using Terraform in the Compute Canada Federation to solve a variety of problems and hopefully help you find ways to leverage this flexible tool at your institution to promote collaboration, reproducibility, and automation and reduce the need for extensive documentation and operations.

Terraforming the Cloud to Teach HPC

Félix-Antoine Fortin, Professional services

Each year, Compute Canada technical staff train hundreds of new users across Canada on using high performance computing (HPC) clusters through hands-on workshops. High demand for Compute Canada HPC resources lead us to a simple question: do we need an HPC cluster to teach HPC? While there are still a few skeptics among our ranks, we do have a simple answer: yes.
During this talk, we will present the technical challenges that emerge when one wants to teach HPC without access to an HPC cluster and how we were able to overcome them by leveraging cloud infrastructure. We will present how HashiCorp Terraform “infrastructure-as-code” toolset was instrumental in the success of building a scalable and cloud-vendor agnostic solution and why it should be considered for your next cloud project. During the talk, an HPC cluster will be created live in the cloud, and the audience will be invited to experiment with it.

Find more information about CANHEIT-ARC conference on https://canheit-arc-2019.ca/.

Find more information about Calcul Québec on http://www.calculquebec.ca.