Compute Ontario

Towards a Cloud Strategy for Ontario

Part One

Compute • Calcul Ontario

Towards a Cloud Strategy for Ontario: Part One

Background

Commercial cloud platforms provide robust, agile, highly-reliable and scalable infrastructure that feature the newest hardware available. Cloud providers also innovate continually by rolling-out new services (e.g. for AI, genomics, cybersecurity etc.).

With rapid technology and cost improvements, it is increasingly clear that commercial cloud providers have an important role to play in the academic advanced research computing (ARC) sector. The National Science Foundation in the US, for example, recently funded an initiative (CloudBank) which provides commercial cloud access and support for computer science research and education. The top500 list of supercomputers for June 2019 featured an entry (at #136) which was an ephemeral AWS "cluster" that cost just $5,000 to run the necessary HPL benchmark.

Despite significant investments by the Ontario government, researchers are still able to satisfy only 40% (or less) of their demand for compute resources on publicly-funded platforms. In fact, there are academic groups and collaborations who currently pay commercial providers (using grant funding or other means) to access compute to carry out their research projects. Tapping into commercial cloud resources could be beneficial for Ontario researchers as the scale, capabilities and services of the compute and storage resources available in the commercial cloud are staggering. Combined infrastructure spending by cloud providers worldwide averages more than $200M *per day*!

One of Compute Ontario's strategic priorities is to support the advancement of ARC in Ontario's research community. In order to remain competitive, researchers need enhanced access to the most advanced and innovative ARC hardware, software and services available in a rapidly-evolving environment. As a result, Compute Ontario is actively developing a Cloud Strategy, is collaborating with other Canadian regions on related initiatives, and is currently engaged in a pilot project to benchmark and track the performance and cost of commercial cloud for ARC.

Opportunities

Commercial cloud systems offer several opportunities to advance research within Ontario's ARC ecosystem. Some of these potential benefits include:

- **Flexible hardware:** researchers have access to multiple types of resources and services that can be configured in order to carry out complex workflows
- **Newest hardware:** cloud platforms are frequently upgraded to the newest and most powerful infrastructure available; allowing researchers to access hardware that academia may not be able to purchase, thereby ensuring their research is being conducted on the most up-to-date hardware

- **Skills development:** staff and trainees benefit from experience earned using cloud solutions which are transferable and marketable skills.
- **Industry engagement**: industrial research collaborations are more easily facilitated in the cloud than on typical academic infrastructure.
- **Maximize HQP:** ARC HQP can devote more time to value-add activities directly benefiting researchers such as maximizing research algorithms as opposed to hardware installation, maintenance and monitoring).

Challenges

Cloud computing offers many potential benefits, but there are a number of challenges that need to be addressed:

- **Cost:** most experts agree that commercial clouds are still more expensive than on premises systems within academic institutions (especially when costs for expert user support are included in cost calculations). Unfortunately, comprehensive analyses that accounts for all relevant factors (such as the benefits of flexibility, more services and continually upgraded hardware in the cloud) are lacking.
- **Support:** the current support model for ARC researchers will need to change significantly in order for users to move to commercial clouds.
- **Shifting researchers:** potentially moving thousands of PIs to a new ARC environment can lead to a learning curve that will have to be addressed in order to conduct high level research in the cloud.

Compute Ontario's Approach:

Compute Ontario is committed to strengthening Ontario's ARC ecosystem in order to accelerate research and enhance provincial competitiveness in the global marketplace. The organization's technology strategy focuses on advancing three main components in the region's ARC ecosystem: compute capacity and access, network performance and security, and data storage and integrity.

Access to cloud computing resources was identified as a key opportunity for Compute Ontario in a 2018 study of the provincial ARC ecosystem conducted by *Hyperion Research* - one of the world's leading analysts of supercomputing and technology trends. Their report found that Canadian spending on ARC hardware, as a percentage of GDP, ranked second-last in the G8. This lack of funding has resulted in inadequate access to research infrastructure, and Hyperion recommended providing researchers with access to commercial cloud resources to bridge this need-gap. In other words, exploration of cloud platforms is not being pursued as a replacement for on-premises infrastructure, but a potential complement to existing platforms to better meet the needs of the 60% of researchers not being served today.

Based on Compute Ontario's objectives to advance the Ontario ARC ecosystem, as well as *Hyperion's* recommendations, cloud computing is being explored as a realistic opportunity for Ontario researchers. As a first step towards a cloud computing strategy, Compute Ontario is currently engaged in two pilot projects which examine the feasibility of commercial cloud usage for researchers.

Cloud Pilot Projects

Compute Ontario HPC Cloud Forecast

Compute Ontario is currently conducting a pilot which will assess the ongoing costs and capabilities of commercial clouds for scientific applications. Though there have been numerous published results concerning performance of HPC benchmarks and codes on commercial clouds, there has been no analysis of how performance and cost-performance evolve over time. We aim to run quarterly tests on three commercial cloud providers and one publicly funded platform (Niagara at University of Toronto). This will allow Compute Ontario to provide up-to-date comparisons for the major public clouds in order to provide a concrete understanding of how new capabilities or cost model changes can affect HPC use cases.

CAC Cloud Pilot

CAC, a Compute Ontario partner, is presently carrying out a pilot project with Kingston General Hospital to analyze several years' worth of data on IBM Cloud.