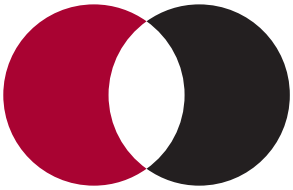
Compute Canada Pre-Budget Submission 2014-08-06

To: The House of Commons
Standing Committee on Finance



compute • calcul
CANADA

"Effective high-performance computing capabilities are going to play an increasingly crucial role in driving Canada's economy and advancements in research and innovation."

Honourable Ed Holder
Minister of State (Science and Technology)
June 24, 2014

Table of Contents

Introduction

Today's most spectacular global science efforts depend not only on our excellent scientists but also on our digital infrastructure. This infrastructure is now a critical ingredient for all science, research and engineering.

Compute Canada encourages the government to commit to predictable, sustainable funding for all the components of Canada's digital infrastructure with a particular focus on "big data" sharing among researchers and exploitation by industry.

Digital infrastructure is directly linked to our ability to innovate in mining, energy, aerospace, drug development and medical technologies, and clean technologies. Computer modeling, simulation, visualization and design power today's industrial research and development.

This requires continued investment in networking, storage, computational resources and highly qualified personnel (HQP) working with this data.

> The research environment has changed considerably over the past decade.
>
> Most notably, massive digitization initiatives, high-throughput devices, sensor networks and computational modeling and simulation drive many of today's scientific endeavours and generate datasets that are unprecedented in size and complexity.
>
> When harnessed to its full potential, this enormous pool of information paves the way for radically new breakthroughs in knowledge and understanding at all levels. Data-intensive research, using these resources, also lays the groundwork for the development of revolutionary approaches to problem solving and decision making, the very drivers of innovation in academia and in the private, public and not-for-profit sectors.
>
> ***Canada Foundation for Innovation***

A commitment to sustained funding will allow Canada to keep pace with its neighbours and trading partners and ensure continued competitiveness of Canadian science and innovation. Compute Canada is positioned to ensure that these investments are made prudently and that the benefits to Canada are maximized.

Increasing the competitiveness of Canadian businesses through research, development, innovation and commercialization.

Global Context

Around the globe nations recognize advanced research computing is central to the economy, innovation and science. The Partnership for Advanced Computing in Europe (PRACE) believes the competitiveness of European science and industry will be jeopardized if sufficiently capable computers and skilled people to exploit them are not made available.

There is a global race for leadership in advanced research computing to address detection of diseases like cancer, to forecast natural disasters, and to support industry in manufacturing new materials and products.

High-performance computing (HPC) has played a central role in establishing the importance of simulation and modeling as the third pillar of science (theory and experiment being the first two), and the growing importance of data is creating the fourth pillar. HPC must encompass the ability to efficiently manipulate and manage vast quantities of data. It must also simultaneously address innovations in software and algorithms, data analytics, statistical techniques, fundamental operating system research, file systems, and innovative domain-centric applications.

National Science Foundation (US)

The Role of Compute Canada

Compute Canada operates and maintains Canada's advanced research computing infrastructure — combining high performance computers, high efficiency data storage and sophisticated research software services working in partnership with four regional organizations, research-intensive institutions and CANARIE.

Compute Canada's vision is to make Canada a world leader in the use of advanced computing for research, discovery and innovation.

Our mission is to enable excellence in research and innovation for the benefit of Canadians by effectively, efficiently and sustainably deploying a state-of-the-art advanced research-computing platform supported by world-class expertise. Compute Canada uses this platform to support a growing base of excellent researchers, and to serve as a national voice for advanced research computing.

Our unique federated model maximizes resource availability through a shared national platform and maximizes leverage of funding from numerous sources. By a large margin, Compute Canada more than doubles its federal investment through matching funds with provincial programs and initiatives.

Compute Canada in partnership with its regions is helping to train highly qualified people — in both industry and academia — able to exploit computation for the advancement of our economy.

Canada is in the attractive position to be a global leader in data storage and management. With increased capacity for storage, our research community can house the world's most compelling big data turning data libraries into data laboratories.

Considerations for the Committee

Despite the success of our service model, our ability to provide Canada's researchers with the levels of capability needed to retain our global competitiveness is strained. The Canada Foundation for Innovation supports this statement in its recent Cyber-infrastructure initiative announced in June 2014:

> *"The Compute Canada national platform is now reaching the limits of its capacity, and the current computational services being offered are not necessarily designed to meet tomorrow's research challenges."*

Recommendations

In Canada, our government addresses advanced research in DIGITAL CANADA 150. The report recognizes that increasing the potential of Big Data is essential for Canada.

> *"Canada will be one of the global leaders in applying "big data" to change how we think about and carry out health care, research and development, as well as the myriad activities of business and government."*

In order to achieve this leadership role we must recognize that Canada's digital infrastructure is a central component for science, innovation and industry.

Compute Canada recommends the government consider flexibility with its digital infrastructure funding mechanisms to more efficiently support Canada's research and innovation advanced computing services by:

- Recognizing Compute Canada and CANARIE are as essential to our global competitiveness as broadband access is to the prosperity of everyday Canadians.
- Aligning mandates and funding cycles between Compute Canada and CANARIE would allow more cohesive and strategic planning for the foundational components of digital infrastructure

Recommended Priority Investments

The following recommendations will in part be addressed through the Canada Foundation for Innovation's Cyber-infrastructure initiative or similar means. These investments are strategic and assume that additional sustained funding mechanisms will be established that can build on these investments.

In addition are strategic in nature and assume the future predictable

Capital Investment

- $45M Federal capital investment over three years *(a separate consideration from other "domain specific" funding)*. This would be matched by $45M provincial capital investment over 3 years.

Operations Cost Support

- $30M investment over three years ($10M a year) in addition to the baseline Compute Canada operating costs covered by the Canada Foundation of Innovation's Major Science Initiative program through 2017.

Priority Investment Overview

1. **Two large systems (for redundancy) (install early 2016, 2x 1.5 Pflop/s)**
    - Capital cost: $20M + $20M provincial match
    - Replace old systems, increase Canadian capability by a factor of 2
    - Add 2x 20 Petabytes of storage
    - Highly shareable, cloud-like capability aligns with most researchers' needs

2. **Large pilot (~40 Petabytes) of redundant durable storage for data management**

    - Capital cost: $10M + $10M provincial match
    - Multiple systems in multiple locations
    - Starting platform for Canada for international big science/big data projects

3. **Middleware/Digital Services**

    - Development/acquisition in partnership with major international initiatives
    - Operating cost: up to $10M over three years
    - Create common systems architecture for Canadian research community to exploit – a truly shareable advanced research infrastructure

4. **Domain-specific investments**

    - Suggest these be supported through the CFI and/or Granting Agencies. Compute Canada would provide storage and software platforms on which specific domains can build (see #2 and #3)

5. **National HQP development and training** (academia and industry)

    - Operating costs: up to $5M over three years — match possible

6. **Software R&D to optimize use of limited hardware**

    - Operating costs: up to $5M over 3 years — match possible

7. **Hardware R&D to address challenges to Moore's law**

    - Optimization of system engineering, challenges of power provisions to systems, support for international exascale efforts
    - Operating costs: up to $10M over 3 years — match possible

8. **One large "capability" system**

    - Capital cost: $15M + $15M provincial match

Compute Canada has historically relied on funding from the CFI, which manages a world-class competitive process that ensures our most transformative science opportunities have the necessary support.

We encourage the government to consider additional models of funding for essential, shareable infrastructure, where operational excellence and cost efficiency are also important objectives.

We support the CFI's recent commitments to sustained operating funding for Major Science Initiative (MSI) projects. We also support CFI's request for continued funding of this program.

However, Compute Canada recommends that the CFI be given the flexibility to adjust the parameters of that funding to allow centralized expenses such as management, development of architectural standards, technology planning to be exempt from matching requirements.

Similarly, we recommend greater ability to balance operating costs and capital expenditures to ensure the most appropriate expenditures are made and older systems with higher operating costs are not maintained.

Preparing for Canada's Future: Recommendations for Funding beyond 2017

Capital

- $30M/year Federal investment
    - $10M/year additional provincial investment
- Mixed investment in computing, performance storage and long-term storage
- Estimated aggregate national capacity by 2022: 70 PF and 1,000 PB (1 Exabyte)

Operations costs

- $30M/year Federal investment
    - Enabling continued leadership in advanced research computing expertise, systems architecture, data management, outreach, training and R&D
    - Includes 40% contribution to on-campus support personnel
        - 60% covered by institutions and provinces
- Power costs will escalate and could reach $30M/year by 2022. *(Could be recovered through new funding mechanisms.)*

Compute Canada has initiated a comprehensive planning exercise, to forecast national needs for digital infrastructure through 2022 and beyond, based on detailed projections by different research communities, institutions and innovative industries. This forecast will encompass compute, data storage, security and digital service requirements, but already it is clear that by 2020 a single discipline (genomics) will require all of the compute capacity available in Canada today and almost 1 Exabyte of storage. There are also indications that Canada has a real opportunity to establish leadership in international data management and storage through sustained and balanced funding commitments.

Over the next few months Compute Canada will launch a number of new digital services to meet the needs of the science and innovation community. These services will be part of a systems architecture that will reflect best international practices and maximize the opportunities for collaboration with science and innovation partners in the US, Europe and Asia.

This architecture will also harmonize service development here in Canada, allowing research communities to leverage existing digital capabilities, maximize opportunities for sharing and data reuse. As part of our leadership role, Compute Canada is committed to working with other leaders in that ecosystem, notably CANARIE, Canada's research-intensive universities and hospitals, and Canada's research granting agencies, to coordinate the continued renewal of our digital infrastructure.

CANARIE and Compute Canada are working in collaboration to develop and promote a shared vision of digital infrastructure in Canada; and to explore the potential for collaboration and alignment of each organization's technical, marketing and operational programs in support of each other's respective missions in support of world class research, education and innovation in Canada.

Conclusion

Compute Canada encourages the government to commit to predictable, sustainable funding for all the components of Canada's digital infrastructure and increase flexibility to enable efficiencies between operations and capital expenditures.

These investments will ensure that our digital infrastructure systems will become world-leading virtual laboratories.

About Compute Canada

This made-in-Canada federated model manages an enterprise providing 2 Petaflop/s of aggregate computing power, 20 Petabytes of data storage, and more than 140 experts in advanced research computing located at almost institutions across the country. Compute Canada serves 70 academic organizations and almost 10,000 researchers across the country, including many of Canada's most influential scientists. Our model was designed to respect, respond to, and leverage the needs and priorities of institutions, provinces and regions across Canada.

Our structure balances integrated leadership, service management and technology planning, with a federation of regional service organizations with expert staff working in 30 locations across the country, making sure that national services accommodate and integrate the priorities of institutional, provincial and regional stakeholders.

compute · calcul
CANADA

36 York Mills Road, Suite 505, Toronto, Ontario, M2P 2E9

www.computecanada.ca | 1-416-228-1234