UNIVERSITY OF ALBERTA

OFFICES OF THE PROVOST AND VICE-PRESIDENT (ACADEMIC)
AND THE VICE-PRESIDENT (RESEARCH)

South Academic Building
Edmonton, Alberta, Canada T6G 2G7
www.provost.ualberta.ca
www.research.ualberta.ca

February 26, 2016

Compute Canada
Suite 505
36 York Mills Road
Toronto, ON M2P 2E9

Dear Compute Canada Executive Team:

As one of the top five Canadian universities in annual sponsored research funding, we appreciate the opportunity to advise Compute Canada on the Advanced Research Computing (ARC) needs of our researchers through your Sustainable Planning for ARC phase 2 (SPARC2) exercise.

As you are aware, the University of Alberta has researchers in many disciplines that make intensive use of ARC, from traditional computational fields such as computing science, particle and astrophysics, all fields of engineering, and computational chemistry, to new and expanding ARC fields such as the digital humanities and biomedical fields like bioinformatics, virology, and metabolomics. Tri-Council funding has recognized the excellence of our researchers' efforts in furthering knowledge in all fields of academic pursuit, and ARC resources must be available in plentiful supply lest the funded research be impeded by constrained computational resources. Most fields now make use of technology as a set of research tools to enhance research, either in speed, breadth, or depth. For instance, ARC now complements close reading in the humanities in order to advance knowledge in philosophy and literature. Computer modelling in engineering and science allow examination of questions not possible in physical laboratories.

We understand you are planning the next five years of ARC support in Canada, and we see opportunities to restore and enhance the computational infrastructure necessary for modern research. Many ARC resources are either outdated or in highly constrained supply due to a lack of sustained funding, and we hope that this SPARC2 exercise provides the rationale for such funding. Some of these restorations and enhancements may already be planned, but we include them to underscore their importance in our research efforts.

1. **Research data storage**: The Tri-Council requires research data management plans for each proposal, and the University of Alberta is in the forefront of Canadian efforts with leadership of CARL's PORTAGE project and recent discussions across the U15 around research data management. Without appropriate storage during the research projects and programs, the work cannot be completed. File, block, and object storage are all essential tools in fields as widely varied as bioinformatics, astrophysics, and linguistics.

2. **Training**: As more fields adopt ARC as a research tool, researchers from undergraduates to tenured faculty need training in how to best make use of ARC for their research. Software Carpentry, Data Carpentry, file transfer, visualization, and batch processing are all elements of ARC for which training needs to be available.

../2

3. **Visualization**: From simple two-dimensional plots to more advanced 3- and 4-dimensional plots as well as word clouds and other tools, the examination of collected, open, and purchased data through visual means is crucial to developing understanding, especially as datasets increase in size. An example of how visualization has changed is with the increasing density of points in Computational Fluid Dynamics (CFD) studies of air flows around automotive and aeronautical structures. Support for visualization, through software such as Paraview and hardware for rendering images displayable through local display technologies at all research institutions but also through training for visualization techniques, is crucial for understanding and publication of results.

4. **Fine-grained parallel ARC**: Many fields still make use of fine-grained parallel techniques, and some provision must be made for these algorithms.

5. **Large-memory computations**: As researchers adopt big-data techniques in their fields and as research data sets grow ever larger (either covering more space or finer detail), some provision in the next five years must be made to support shared-memory techniques so that the capabilities of systems like Hungabee hosted at the University of Alberta are not lost to Canadian researchers. Evaluation of the operations on systems like Hungabee can elucidate what volume of memory per core is required for these computations, but the facility must not be lost as big-data techniques become commonplace.

6. **Coarse-grained ARC (clusters)**: The workhorses of ARC will continue to be clusters, with some amount of parallelism supported. These systems are flexible in addressing the widest variety of ARC computations, and should represent a significant portion of CFI investment. Inclusion of GPGPUs (see #7 below) is important.

7. **General-purpose graphics processing units (GPGPUs)**: Many researchers, looking to maximize their computational effectiveness in order to accelerate research outcomes, have adopted GPUs. Compute Canada in the previous generation of systems made a maiden-effort launch of GPUs when GPU computing was a new field. Now that GPU adoption is more commonplace, significantly enhanced quantities of GPU resources are required in order to provide the supply to meet demand as seen in the RAC 2016 requests. Some researchers' techniques have almost wholly adopted GPU programming techniques in order to compute results in reasonable amounts of time, impossible with standard CPU cores. Support of these researchers is crucial to their productivity with Canadian ARC resources.

8. **Cloud resources**: Many researchers require modest computational support, especially those adopting ARC methods from fields not historically represented as ARC-heavy. Databases, websites, and graphical manipulations of data are readily supported on cloud computing infrastructure. While commercial cloud resources are globally available, including those that could host data entirely within Canada, close attention to provision of cloud support at a cost below that commercially available would benefit these researchers. Provision of cloud resources at prices competitive or higher than the commercial providers would be a sign that Compute Canada should exit that provision and to concentrate on optimizing data transfer in and out of such cloud providers.

Thank you for your attention and consideration.

Sincerely yours,

Steven Dew
Provost and Vice-President (Academic)

Lorne A Babiuk
Vice-President (Research)

SD/kb

c David Turpin, President and Vice-Chancellor
Mike MacGregor, Vice-Provost & Associate Vice-President (Information Services & Technology)
John Beamish, Acting Dean, Faculty of Science
Fraser Forbes, Interim Dean, Faculty of Engineering
Lise Gotell, Acting Dean, Faculty of Arts
Richard Fedorak, Interim Dean, Faculty of Medicine & Dentistry
James Kehrer, Dean, Faculty of Pharmacy and Pharmaceutical Sciences
Scott Delinger, Director, Research Computing (Information Services & Technology)