



Compute Canada's Call for White Papers for Sustainable Planning for Advanced Research Computing Phase II (SPARC2)

Compute Canada is issuing this call for white papers addressing the future needs of specific research disciplines and institutions. The announcement is attached with this message.

More than 20 white papers were submitted in the first SPARC consultation (summer 2014). These had a strong influence on subsequent Compute Canada planning, which resulted in a successful submission for an operations mid-term review and a new \$30M capital award from CFI. This capital award is targeted at new advanced research computing infrastructure, which will come online throughout 2016 and 2017.

Compute Canada will be submitting a new operating proposal (CFI MSI) in the spring of 2016, covering operations from 2017-2022. As part of this submission, CFI expects a plan for capital investments from 2017 onwards. We expect SPARC2 white papers will have a strong influence on capital and operations planning for Compute Canada.

It is important that these proposals reflect a broad cross-section of input from the Canadian research community. Different disciplines have different advanced research computing needs, which may include:

- Analysis of “big data,” running software with a large memory footprint, or other data-intensive workloads;
- Specialized hardware (e.g. GPU accelerators, manycore processors);
- Needs for isolation of data or workloads, for analysis of sensitive private data;
- Access to a very large number of computational “cores” in a single system;
- Dedicated software platforms, scientific gateways, virtual machines in a cloud environment.

For your SPARC2 response, keeping in mind the 2017-2022 timeframe, please address the following basic questions:

- What kinds of problems are you trying to solve?
- What kind of advanced research computing infrastructure is best suited to solve these problems today?
- What quantities and types of infrastructure would be needed to meet the needs of your community?

If you belong to a disciplinary group or institution that submitted a white paper in 2014, and you wish to simply update that paper with new projections and re-submit, this is perfectly acceptable.

Please note that the intent of SPARC2 is for you to inform Compute Canada of what you really need to do world-class research in your field, not to find out what you think you can get in a Canadian context.

For your convenience, we have provided a set of submission guidelines below.



Available Support

Compute Canada can offer support in a number of ways for those writing white papers, for example:

- Translating a scientific need into a list of infrastructure requirements.
- Organizing cyberinfrastructure-focused meetings or workshops among researchers in your discipline.

Deadline

Please submit your white paper by March 1, 2016. Submissions should be sent to sparc@computecanada.ca.

If you need further information or have any other questions please feel free to contact us at sparc@computecanada.ca.

Dugan O'Neil, Chief Science Officer - Compute Canada

White Paper Guidelines

There is no required format for white papers submitted in response to this call. However, we provide an example below to help guide disciplinary groups in providing the most useful information for Compute Canada.

Authors

Identify who wrote the white paper, their contact information, and provide an overview of the scope. Please make it clear whether the needs reflect only those of the authors or represent the broader discipline and how widely the represented community has been consulted.

Science Description

What are you studying and why? This can be a single paragraph or pages of text as you wish. Please note: If you wish to paste-in a page from a grant request or other document, that is also welcome.

Current Use of Advanced Research Computing

There are several relevant questions to answer here:

- Do you make use of existing Compute Canada (CC) systems? If so:
 - Which ones?
 - Do you have a special resource allocation?
- Do you currently make use of non-CC resources? Which ones?
- What is your current level of usage of storage, compute, cloud, portals/gateways, accelerators, etc.? (Note: if you use only CC resources please provide your CCDB IDs - CC will then be able to calculate the current usage of your group for you)
- Do you currently rely on any software packages, middleware platforms, or other sort of shared software infrastructure?

Future Growth

Do you expect your needs to grow? What drives this growth? For example, if improvements in sensors, detectors, imaging and equipment are driving progress in your field, you might expect commensurate growth in advanced research computing needs.



What is the estimated scale of growth in the next five years (2017-2022) compared to today? For example, “we expect datasets to grow by a factor of three and our computing needs to expand by a factor of two.”

Technical Details of Future Needs

Not all groups will be able to translate their scientific needs into technical needs. Compute Canada experts can help. Any information you can already provide on future needs would be welcomed. All questions refer to the 2017-2022 period.

- **Data**

- How much long-term storage do you anticipate needing? Do you have any special requirements around data preservation, curation, access, etc.? What is the time profile of this need?
- Do you need to be able to serve your data to a compute facility (thereby requiring fast disks connected to a significant source of CPU)? How much fast-connected storage do you envision needing and with what time profile?
- Does your data need to be shared by many researchers? If so, will it be served to the international community? If so, how will it be served to the community (e.g. web-based portal/gateway)?
- It would be helpful to have a time profile of your storage needs, like that shown in the table below:

Now	2016	2017	2018 (new sensor comes online)	2019	2020	2021 (new sensor comes online)	2022
100TB	200TB	300TB	1PB	2PB	3PB	8PB	20PB

- **Computation**

- Are your calculations serial or parallel? If you use parallel processing, what is the largest number of cores you use (or would like to use) in a single job?
- Is the nature of your computation constant or does it have peaks and valleys? If there are peaks, please provide both average and peak amounts.
- Based on today’s CPU cores, how many cores do you need and what is the time profile of this need? What is the reason for the increase? Please see the example below:

Now	2016	2017	2018 (new sensor comes online)	2019	2020	2021 (new sensor comes online)	2022
1000 cores (peak) 200 cores (avg)	1500 cores	2000 cores	5000 cores	6000 cores	7000 cores	12000 cores	15000 cores



- What are the memory requirements associated with your calculations? Please specify if the number is a total or is per core. Please specify if you require large shared memory systems and try to quantify the amount of memory needed.
 - Do you benefit from accelerator systems (e.g. GPU, FPGA)? What gain do you expect if CC provides accelerators compared to performing your calculation on regular CPU cores? Please provide details on how many GPUs (or otherwise) you would use and a time profile for that use.
 - Do you expect to need what are being commonly labeled as “Big Data” clusters and services? This may include hadoop clusters, no-sql and distributed databases. Please quantify the level of need (and provide a time profile) if possible.
- **Software**
 - Do you foresee a change in your software/middleware needs?
 - Would you benefit from the work of a software development team to bring new software to your field? What do you need?
- **Networking**
 - Do you envision using multiple sites requiring strong network connectivity between them? What volumes of data do you expect to move from site to site and what is the required transfer rate?

As part of your white paper submission, please indicate whether distribution of the white paper should be limited internally for Compute Canada’s evaluation with its members and partners. If permitted, Compute Canada would like to make white papers available to the broader research community, either by posting the entire paper at www.computecanada.ca or by extracting quotes and summaries for sharing with broader audiences.