Computation and Data Committee Report to the CASCA Board, Dec 2018

Current Committee membership:

James Wadsley (McMaster) (Chair)	Term ends: 30 June 2021 (OBO)
Pauline Barmby	Term ends: 30 June 2021
Catherine Lovekin	Term ends: 30 June 2021
J. J. Kavelaars (HIA/NRC/CADC):	Term ends: 30 June 2021
Erik Rosolowsky (Alberta)	Term ends: 30 June 2021 (OBO)

The current committee is nominally signed up for three years. However it is suggested that repeat members could rotate of earlier in a staggered way to may new membership smoother (e.g. Wadsley, Rosolowsky, Kavelaars). Note that Kavelaars is an ideally placed member for NRC input. If he chooses to rotate off then it is highly recommended that a replacement from NRC/CADC be found.

CANFAR and CADC

The NRC continues to move forward on the planned 'refresh' of the CADC. The project will complete in 2020 and is currently on schedule. The refresh includes replacing 100% of the CADC storage and computing nodes and upgrading the CADC internal network from 2 Gbit to 40 Gbit [the external network will remain at 10Gbit]. The refresh should enable the CADC to continue to provide archive services for 3-5 years beyond the end of the project. A new 'refresh' is likely to begin when the current project concludes. The current CANFAR resource allocation provides for ~ 4000 TB plus over 1000 cpu years of cloud computing.

The CADC/CANFAR science platform software upgrade is also moving forward. The project is planned to complete in spring 2021. The science platform upgrade includes a new authorization and authentication system, performant network mountable storage, user managed IVOA-TAP enable databases, virtual desktop computing, Jupyter HUB service and small improvements to the OpenStack Cloud VM/Batch system. This project is a major reimplementation of the CANFAR science platform, with the design largely driven by feedback from users and expectations of future requirements.

CADC is continuing to provide archive service enhancements to the ALMA Science Archive, as part of Canada's commitment to ALMA. We will also continue to provide telescope archive services to CFHT. The East Asia Observatory archive agreement, which secures some access for Canadian users of JCMT, is up for renewal and is expected to be renewed but the details are not yet known. The CSA recently extended their funding for the CADC's participation in the development of the JWST archive for 2 additional years. The CADC is not currently participating in the Gemini Data Processing work but has begun a project to mirror the Gemini archive into the CADC systems.

CADC is in discussions with the LSST-Canada consortium to determine if Canada has the capacity to assist Canadian users of LSST achieve the best science goals. The answer is likely no and we expect that a CFI proposal to build new capacity/infrastructure to support LSST users will be needed.

And finally, the CADC is developing a Machine Learning group that will support Canadian astronomers looking to adopt Machine Learning processes into their data analysis. This group will make use of CANFAR infrastructure to provide this support.

CFI Cyberinfrastructure Grants

CIRADA, the Radio Astronomy cyberinfrastructure project, is moving forward with developing prototype systems in coordination with the CADC team.

Compute Canada Governance and Management

Compute Canada (CC) remains the main source of cycles, storage and HPC services for researchers at Canadian universities.   Operational funding for the next few years was recently secured (MSI 2). There was some tension between provincial and national components surrounding this earlier this year but it appears resolved.

The federal government has decided on a major reorganization of computing in Canada based on the Naylor report. This is partly to resolve the fact that computing is an ongoing (operational) support need whereas CFI was designed for one-time infrastructure needs. It is not yet known what the organizational structure will look like going forward. An announcement is expected in Spring alongside a budget announcement for Digital Research infrastructure at the $600M. level.

National Computing Hardware

A fifth new system Beluga, is being installed in Quebec with roughly 30,000 cores.
The other four systems are up and running effectively; SFU's CEDAR (GPU focused ~ 30,000 cores, ~ 3000 GPUs), U-Vic's ARBUTUS (cloud focused ~15000 cores, 1.6 PB storage), Waterloo's GRAHAM (~30,000 cores)and Toronto's NIAGARA system (Large parallel system, 60,000 Intel Skylake cores, 300 TB RAM, 9 PB disk).