Computation and Data Committee Report to the CASCA Board, Dec 2017

Current Committee membership:

James Wadsley (McMaster) (Chair)	Term ends: 30 June 2018
Hugo Martel (Laval)	Term ends: 30 June 2018
J. J. Kavelaars (HIA/NRC/CADC):	Term ends: 30 June 2018
Erik Rosolowsky (Alberta)	Term ends: 30 June 2018

Current members are mostly willing to continue. However, if there are potential new members they are also happy to rotate off.   The committee could use some new people so if the CASCA board is aware of anyone with interest please put them in touch with us. The committee would like to identify potential new members and will contact them over the coming months to have nominations ready for the summer meeting.

Computing in Canada

CANFAR, CADC and CANARIE

Organizationally there has been a change within the CADC. JJ Kavelaars has assumed the role of Group Leader within the CADC and is now responsible for the direction of the data centre. David Schade, the former Group Leader, is continuing on at the CADC and focusing on the CADC international partnerships with other data centres, particular via work with the International Virtual Observatory Alliance (which will have its next meeting in Victoria, immediately after CASCA) and Research Data Alliance. The CADC continues to be well supported by NRC, contributing archive support to CFHT, JCMT, DAO, and ALMA while participating in the development of concepts for the SKA Regional Centres.

The CADC continues to work with Compute Canada (CC) to transition the operation of the CANFAR service from a CADC managed service on Compute Canada hardware to being a CC operated service on CC hardware. In the new operational model CADC's effort will be more directly related to expanding the astronomy services that are CANFAR rather than operating the underlying systems. The work of integrating the two organizations in this effort is proving to be a challenge but both sides are committed to a successful outcome.

On the archive storage side, CADC has initiated a project to renew the hardware in their data centre. This work was initiated in spring 2017 and includes new storage hardware, data processing capacity and a higher speed link to the research internet. As many are aware, work with Shared Services Canada (federal government central IT) has been challenging as that single organization was mandated with the largest transformation of government IT ever attempted. Very recently, however, there has been good progress with the CADC renewal project. The new network connection is likely to be activated early in 2018 and the first installment of storage hardware is expected before the end of March 2018.

CFI Cyberinfrastructure Grants

CFI has recently approved funding for the Canadian Initiative for Radio Astronomy Data Analysis (CIRADA), a six-institution initiative to develop national infrastructure needed to support current and next-generation radio telescope surveys. CIRADA is led by U. Toronto (PI: Gaensler) with support from U. Alberta, U. British Columbia, U. Manitoba, McGill and Queens, working closely with CADC.

The Initiative will focus on developing science-ready data products out of survey data from the VLA, CHIME, and ASKAP. The project will use Compute Canada and CANFAR resources wherever possible, and the data requirements of the project will require installing an additional 10 PB of storage on Compute Canada resources. The success of the project requires a constructive working relationship between all these stakeholders, but initial discussions with Compute Canada and regional computing consortia have been promising.

There is an additional CANFAR-based cyberinfrastructure proposal to CFI. Its NOI was successful and a full proposal was submitted Oct 2017, through U-Vic, (PI: Falk Herwig). If successful it will expand the scope of CANFAR to include more extensive processing for upcoming instruments (e.g. SKA, CHIME, ALMA, LSST, CFIS, JWST), production of mock catalogues and the ability to create simulated observations from numerical N-body and hydrodynamics simulations. The goal is develop a flexible framework that can support a broader set of users to organize, access, mine and analyze massive datasets.

Compute Canada Governance and Management

Compute Canada (CC) is the main source of cycles, storage and HPC services for researchers at Canadian universities. CC (http://computecanada.ca) was incorporated under its current bylaws in 2012. The original CC board was unresponsive to the members of Compute Canada (Canada's universities and colleges) and did not initially allow them a say in how the organization was run. This led to a fairly uneven start to the organization. The CC board's initial plan was to reorganize HPC in Canada to be less regional and more national which led to a dispute with the regions/provinces and the pre-existing consortia and a disruption of services and funding. These consortia were made up of universities and colleges then as now. A petition (see below under recommendations) to restore smooth delivery of HPC services and for greater transparency and responsiveness going forward was signed by 238 researchers in 2013 and presented to CFI and VPRs across Canada. Following this, CC was organized in its current form with regional consortia continuing to deliver the bulk of the services and managing the systems.

However, there remained a degree of tension between the CC board and the member institutions. An active subset of the members pushed hard for a reorganization of the board and governance structure which led to the departure of the original board chair and the CEO. A new CC board was established in October 2017. It includes more researcher representation. The current chair is Chris Loomis.

Until recently, CC was led by CEO Mark Dietrich; CSO (Chief Scientific Officer) Dugan O'Neil (Physics, Simon Fraser); and CTO (Chief Technical Officer) Greg Newby. However, the CSO is currently a vacant position (since mid 2017) and the CEO stepped down a few months ago with the CFO (Chief Financial Officer) Terry Lockheart acting as the CEO currently. A process to hire new executive management is underway. Some aspects of the central part of CC are quite active but others are dormant. Established committees such as ACOR (the researcher input committee, including J. Wadsley and R. Thacker as members) have not met in over 2 years.

One might have expected the turnover at Compute Canada central to negatively affect service. However, the technical staff at regional consortia are responsible for managing the HPC systems and have provided normal levels of service over the past year. In addition, several new systems have been installed successfully.

Hardware Refresh

Compute Canada has consolidated all new HPC into just a few major systems in the latest refresh. Four major systems are currently fully funded. Three of the four new systems are up and running; SFU's CEDAR (GPU focused ~ 30,000 cores, ~ 3000 GPUs), U-Vic's ARBUTUS (cloud focused ~15000 cores, 1.6 PB) and Waterloo's GRAHAM (~30,000 cores). Toronto has roughly halved its GPC system to accommodate the new NIAGARA system (Large parallel system ~ 60,000 cores) which should be available for users early in 2018. Other older systems continue to be phased out with large systems in Quebec to be shut down soon. The next planned new system is GP4 for Montreal.

Current storage looks quite modest: ~ 25 PB. CC plans to expand this from 2018 to over 100 PB. There are discussions with CFI over further refresh ("Stage 2" submitted in 2016) that would likely allow for current major sites to be refreshed as well as expanding storage. The biggest issue with this is that Quebec and Eastern Canada may be without new systems for an extended period. The current systems are not well suited to very large parallel (>1000 cores) which may have to wait for Niagara (2018). As a general statement, there is a real cycle crunch occurring now as we wait for Niagara.

Issue of Centralization versus Regional Consortia

For decades Canadian HPC operated very effectively with a minimal central part. CFI has always required matching contributions from the provinces and the regions still operate all the hardware. Thus the regions contribute 60% of the costs. Most HPC staff are located at Canadian universities where they are integrated into local research efforts. From the perspective of many researchers, this model is natural and effective. It is also worth noting that, historically, the regional management has been very close to researchers, with executive committees and management teams made up of researchers in many cases.

A few years ago, CFI pushed for more centralization which led to the new governance structure with an incorporated entity (Compute Canada) in 2012, the creation of central office and the hiring of a full time CEO. There has always been tension between the central and regional parts as noted above.

One benefit to the central part is a more direct relationship with CFI. There are also user level benefits to central coordination such as making machine access (e.g. usernames and passwords) and software stacks more uniform across CC systems. Initially this was done by regional staff acting together in national teams that predate the current Compute Canada. In past few years these teams have been coordinated mainly by the CTO and CSO at CC.

The current teams are shown here: https://www.computecanada.ca/about/governance/national-teams/ With the departure of the CSO (O'Neil), there is a lack of science input. Recently, the CC central office has pushed ahead with executive hires in several areas including security and, most recently, Marc Rousseau as Chief Software Architect (who helped build the CBRAIN project at McGill).

Operational funding for HPC (e.g. staff) has always been tight. The cost of the central office is substantial but exact financial details are difficult to get hold of. Based on salaries, leases, etc… a ballpark estimate would be upwards of $1M per year. This lack of financial transparency makes it difficult to assess the relative value of the central and regional parts of Compute Canada.

The centralization push from CFI is also why the number of national systems has shrunk to four, leaving some regions without new hardware currently and complicating matching fund arrangements. For example, all systems are national (available to all Canadian researchers) but some provinces are not contributing capital funds in the new model.   There are also serious problems with operating costs where universities that host these large systems are now on the hook for large power and building costs.

Canada's computing ecosystem is quite different from other countries. Typically, other countries have substantial mid-sized systems located at many universities (e.g. each having 1000's of cores) with associated local staff to manage them that are not part of a national system. At one time CFI was funding mid-sized systems, often associated with specific universities rather than as part of a network or consortium.   Starting about 20 years ago, CFI began to insist that universities combine into consortia when requesting larger HPC systems. This was instrumental in building the current cohort of experienced HPC staff residing at Canadian universities and colleges. Now CFI insists on a small number of large national systems which serve many thousands of diverse users. This makes simple international comparisons of how national HPC and services are delivered difficult to make. Currently CFI-funded systems are the only option for most Canadian researchers for all computing larger than desktops. Non-Canadian national level systems, on the other hand, focus on large parallel and special purpose users and have a limited user-base (hundreds) mostly focussing on large parallel use.

When it came to operating and services the regions had previously convinced CFI that regional delivery via local staff (the current model) was effective.   However, there is currently a serious dispute between the central part of Compute Canada and the regions. There is now new project management for HPC at CFI.   In the past year, the acting central management for Compute Canada lobbied CFI with the idea that operating funds should be removed from the regions and re-invested in the central part. As noted above, this reflects a difference of vision that was present from 2012 when CC incorporated. This would require some of the current HPC staff at universities and colleges to be fired. The members (Canada's university and colleges) and regional management (e.g. Provinces) do not agree with the new approach.

This dispute is holding up CFI funds for both HPC staff salaries/operating (MSI) and continuing hardware refresh. For example, some universities have not received HPC staff salaries from CC for over 6 months and are paying out of pocket.   This is also causing a hold on releasing funds for already agreed upon hardware installs such as GP4.   The extent of these issues and how they are affecting different institutions could be assessed via direct consultations with CC staff and university contacts for HPC (e.g. CC member representatives, consortium managers and researcher leads).

Recommendations

Discussions with some university (member) representatives to Compute Canada and regional consortia indicate that a letter from CASCA to CFI or similar actions could be valuable for resolving the current dispute and allowing the smooth delivery of HPC in Canada to continue. We suggest that the CASCA board interact with this committee and members of Compute Canada (the institutions) to determine the best course of action.

As a first step to helping resolve this issue, CASCA should undertake with a consultation process with different stakeholders known to members of the CASCA board and CASCA's committees (e.g. such as HPC staff, researcher liaisons, VPR/board members, CC member representatives and CC central management). For example, limited investigation by this committee shows that for at least some institutions the disruption is severe but the extent of the problem needs to be better understood prior to

acting.

In the past, public input from professional organizations and the research community has been useful to help the providers of HPC (Compute Canada, CFI and regional groups) restore their focus on the needs of their user community.

For example, there was a petition in 2013:
https://www.ipetitions.com/petition/restore-confidence-in-compute-canada
when similar issues arose and funding for HPC staff was being held up due to a regions versus central dispute in Compute Canada. One of us (James Wadsley) uploaded the petition letter but did not write the letter itself. At the time CFI indicated that it helped them see what researchers felt was important and it did allow for a continuation of HPC operations at the user level. However, as can be seen from the content of the petition letter, many of the issues that were present then have not been resolved and there are still disconnects among the goals of Compute Canada's central management, the regional providers and the researcher community.

While the best way to respond is not clear, this committee specifically recommends that CASCA provide public support for the following:

1) *CASCA should place top priority on maintaining normal operations for Compute Canada and the resumption of funding flows from CFI immediately.* The current dispute between Compute Canada's central part and the regions is highly disruptive. Canadian Astronomy and Astrophysics needs reliable HPC. Individual researchers and groups require HPC hardware and staff to enable simulations and analysis and also to support projects such as CADC and CANFAR which support instruments.

2) *CASCA should support an open assessment of the current structures and methods for delivery of HPC services to Canadian researchers.* A case has to be made for shift resources based on the benefits to HPC dependent research in Canada. It should also respect the role the provinces and their institutions play in providing 60% of the funding for HPC in Canada. CASCA does not need to take an immediate position on whether more or less centralization is desirable.

2a) *CASCA should support specific transparency requirements on finances and other operations for Compute Canada and for the regions.* There is insufficient information to assess the value of more (or less) centralization for Compute Canada. We lack detailed financial information and also a breakdown of what staff do in the regions and centrally. A proper assessment is required and greater transparency.

2b) *If large reorganizations of staff and funding are to occur, CASCA should support a gradual transition so that staff are not fired and alternative funding mechanisms for local staff can be explored by regions and universities.*

2c) *CASCA should support better and more regular opportunities for researcher representative involvement for each of the regions and Compute Canada.* Compute Canada and the regions have always been uneven in how researchers provide input to and oversight on how systems and staff are managed. For example, the CC researcher input committee (ACOR) has not met in 2 years and there is currently no CSO at Compute Canada (who chaired ACOR and visited member sites).

2d) *CASCA should support the development of mechanisms to provide mid-level HPC that is independent of CFI national platforms.* Canada's HPC ecosystem is unusual and subject to cycles of boom and bust and management instability associated with such a large fraction being dependent on

CFI and using large systems to provide cycles for so many diverse users.   Mid-level HPC could include university and provincially funded hardware and staff or CFI funded systems that are distinct from the so-called National Platforms. CFI currently is very reluctant to fund systems outside the major national systems.