Subject: Computing Report
From: Rene Plume <plume@ras.ucalgary.ca>
Date: Mon, 3 Dec 2012 05:59:13 -0800
To: "Ferrarese, Laura" <Laura.Ferrarese@nrc-cnrc.gc.ca>
CC: "manset@cfht.hawaii.edu" <manset@cfht.hawaii.edu>, Gilles Joncas <joncas@phy.ulaval.ca>, Gregg Wade <Gregg.Wade@rmc.ca>, "Parker, Laura (SMTP)" <lparker@physics.mcmaster.ca>, Samar Safi-Harb <samar@physics.umanitoba.ca>, "sawicki@ap.smu.ca" <sawicki@ap.smu.ca>, Christine Wilson <wilson@physics.mcmaster.ca>

Dec 2012

The important recent news is that CFI just agreed (nov 2012) to provide Compute Canada (CC) with MSI funding (operating costs) but not LEF (limited new hardware). Provinces are also willing to provide required matching funds (with some kinks to be worked out). The detailed distribution of the MSI funds has not been settled. However, its seems quite likely that the consortia will be able to continue to operate with current staffing levels for a few more years. Compute Canada will continue to function so that any Canadian researcher can get access to use any system funded through CFI in any province using a single Compute Canada account system.

This means that we will limp onwards with what hardware we have in the hope that a second National Platforms Fund will be announced soon (e.g. 2013 -- the first was announced in 2006!) allowing for new systems in general usage around 2015 or later. At that time Canada will be in dismal shape from an international perspective but we will still be able to run things for a while as long as they don't delay too much longer. From a user perspective it means a lot of contention for resources if you want to do anything substantial. It also means that there are no new resources to assist with increasingly large datasets. Most existing installations are not storage oriented and are not well-suited for analysis of the large data-streams coming from newer instruments.

Over the long term, it also has to kept in mind that CFI has been an unreliable partner for Canadian Computing. Compute Canada could seek funding more directly from Industry Canada, as a budget line item or in partnership with industry, such as the Ontario specific SOSCIP project with IBM which currently hosts Canada's biggest computer http://socip.org . SOSCIP is not directly affiliated with Compute Canada). These approaches may provide much more consistent funding. It isn't yet clear how things like SOSCIP will affect computing in Canada over the longer term. It does show that resources for

Computing can be found outside the problematic CFI path.

Looking outside Canada, several members of the Canadian community are participating in a bid led by Cambridge University to complete pre-construction work on the SKA Science Data Processor. These efforts build on CADC and community work at data management, processing and modelling. The consortia will bid for pre-construction work packages in early 2013. As of yet, a clear channel to secure funding for pre-construction activities has not been identified.

Compute Canada has incorporated and also has a board now and is in the process of seeking a CEO, CTO (Tech) and CSO (Science). This was a requirement for getting MSI funds. The structure of the organization is unclear. Both the old board (VPR) and the new board have elected to defer to the new CEO to let that person design the organization in detail. See here for the new board:

https://computecanada.ca/index.php/en/about-us/news/general-news/12-news-english/245-compute-canada-calcul-canada-inaugural-board-of-directors-will-advance-canada-s-research-and-innovation-agenda

The current consortia provide extremely cost-effective hardware maintenance, access and services (user help) compared to any other HPC framework. CFI continues to suggest (without concrete arguments) that they are inefficient in some vague way and so there may be tension in future if CFI elects to ignore the reality of how HPC is delivered in Canada in favour of some abstract idea of how it should be done based on advice from people who run supercomputer centres. In particular, CFI insists that Compute Canada cater to all levels of use - from slightly larger than desktops, to mid-range use (100's of processors) and serial farming to process data, grid computing and of course supercomputing on very large systems. In all other countries, their equivalent of Compute Canada focuses solely on supercomputing and deals with small numbers of fairly sophisticated users who typically only do large parallel use. Universities, departments and sometimes research groups have clusters of computers for purposes such as analysing data and running mid-range simulations, teaching and learning the ropes by students and developing new codes. However, in Canada, CFI and NSERC together have decided to stop funding this outside Compute Canada. Astronomical data processing, in particular, needs abundant computing power and storage but doesn't particularly need big parallel systems. In this case having local staff and systems that can be flexible to meet the needs of specific users and groups may be as important as just having raw cycles somewhere in Canada. A challenge for the new Compute Canada will be remaining

responsive to local users in the event that hardware is centralized as CFI seems to want it to be.

The biggest issue for the new CC organization that has been left open is what happens to the NIC (regional representatives) and the current consortia -- there is literally no role proposed for them in the new organization. Currently, beyond managing and running all the hardware, the consortia also act to regionally distill researcher/user concerns and needs and get them addressed. It isn't clear than the new organization will have a way to do that. We've been told that all this is not an oversight -- just something to be determined later. However, the provinces pay a similar share to CFI through matching requirements. Compute Canada needs to be mindful of provincial goals and concerns and not let CFI dictate bad choices.

As mentioned previously by this committee, professional, discipline-based societies such as CASCA could play a bigger role in the new CC organization. An interesting prospect is that they could fill the gap left by the marginalization of the regions/provinces by providing detailed planning submissions and guidance on the needs of groups such as astronomers and astrophysicists. In particular, the consortium/region model has not been effective in identifying grand challenge projects for Canadian science that might motivate the installation and use of high-end, world-class systems.

CASCA's submission to the old Compute Canada VPR board was accepted and considered but no specific role for CASCA-like organizations in the new CC governance structure was prescribed. There are proposed roles for researcher committees but it isn't clear that a collection of individual researchers can adequately distil or convey the needs of their research areas. Worse still, they might be asked not to do so and simply provide advice as general users. We should therefore advocate for a discipline based advisory mechanism. A more direct role for CASCA could prove very useful for astronomers. CASCA could lend considerable support to the case for machines to support world-class science generally and also advocate for more specialized systems than the current somewhat generic ones. As noted above, most Compute Canada systems are optimized for mid-level computing with only moderate storage capabilities and many are managed with fairly low-level users in mind because that is the assumed typical user. Using such systems can be frustrating for high-end users and users with special storage/batch processing needs such as astonomers. There is now a new board and a CEO will soon be appointed (and a Chief Science Officer). This could be a good opportunity to press for a bigger role for discpline-specific organizations. Compute Canada will

also hopefully be gearing up to lobby for resources and more reliable funding and it should realize the value of partnering with organizations like CASCA for this.

The new Compute Canada has individual universities and research institutions (e.g. hospitals) as members. The criteria for membership will be reviewed when the organizational structure is firmed up. For example, businesses that do a lot of computing/modelling (e.g. Bombardier) and vendors (e.g. IBM) have been suggested as potential members (perhaps in a different category). One possibility CASCA might advocate for is that professional organizations could be members. We would then be able to vote at the AGM for the board and CEO and so forth. This would give CASCA a voice that is distinct from the universities starting at the base level of the organization.

In summary, CASCA needs to work to engage with the emerging structure that is Compute Canada to ensure that the National Infrastructure needs of Canadian astronomy, for both high performance computing and massive data storage, are addressed at a national level. In particular, the CADC, is now heavily dependent on the success of Compute Canada for delivery of storage and image processing capacity.

James Wadsley
for the CASCA Computing and Data Committee
JJ Kavelaars
Jonathan Dursi
Erik Rosolowsky