CANADIAN ASTRONOMICAL COMPUTING, DATA AND NETWORK FACILITIES: A WHITE PAPER FOR THE 2010 LONG RANGE PLAN

L. Jonathan Dursi (CITA), David A. Bohlender (HIA/NRC), James Wadsley (McMaster University) and JJ Kavelaars (HIA/NRC)

ABSTRACT

Significant investment in new large, expensive astronomical observing facilities spanning a substantial portion of the electronic spectrum was a dominant theme of LRP2000 and continues to be necessary for Canadian astronomy to maintain its world position. These developments are generating increasingly large volumes of data. Such investments only makes sense if they are balanced by strong infrastructure support to ensure that data acquired with these facilities can be readily accessed and analyzed by observers, and that theoreticians have the tools available to simulate and understand their context. This will require continuing investment in computational facilities to store and analyze the data, networks to ensure useful access to the data and products by Canadian researchers, and personnel to help Canadian researchers make use of these tools.

In addition, large parallel simulations have become an essential tool for astrophysical theory, and Canadian Astronomy has world-leading simulators and developers who rely on world-class High Performance Computing facilities being maintained in Canada to do their research effectively.

We recommend that Compute Canada be funded at $72M/yr to bring HPC funding per capita in line with G8 norms; that part of every Compute Canada technology renewal include a Top-20 class computing facility; NSERC and other funding agencies begin supporting software development as an integral component of scientific research; that the staff funding for consortia be tripled, including local access to technical analyst staff; and that the last mile bottleneck of campus networking less than 10 Gb/s be addressed where it is impacting researchers, with particular urgency for the current 1 Gb/s connection at the CADC.

1. CANADIAN ASTRONOMY USE OF HPC

Canadian astronomers have historically been a driving force for scientific computing within their institutions. Driven by the increasing demands of world-class research, and with techniques learned from using the previous generation of Canadian computers or larger foreign systems, Canadian astronomers doing theoretical astrophysics — and increasingly observational astronomy — have always pushed for, and made immediate use of, Canada's largest computing, network, and storage resources.

1.1. *Theoretical astrophysics*

Large numerical simulations, too large for desktop computing, are now essential to most areas of theoretical astrophysics. This includes solar-system dynamics, planetary physics, star formation, single and binary star evolution, turbulent convection, galactic structure and evolution, large scale structure formation, cosmological evolution, and the very early universe. Canadian astronomers have historically been world-leaders in numerical astrophysics in both development and applications, contributing to the current availability of research codes used throughout Canada and worldwide. Codes developed in Canada include GASOLINE[1], HYDRA, PMFAST, FLASH in self-gravitating hydrodynamics, PARTREE/GOTPM in pure N-body work, and SWIFT in solar system integrations. Canada has a long-standing involvement in Numerical Relativity, with Choptuik's development of the first AMR codes in this field, the discovery of critical collapse, and continuing development of both the code and science (Choptuik 2009). Presence in this area has strongly increased during the last year with the hire of Lehner at Guelph/PI and Pfeiffer at CITA, who both focus on simulations of compact object binaries with applications to astrophysics and gravitational wave astronomy. Pfeiffer uses a multi-domain pseudo-spectral code, named Spectral Einstein Code (SpEC). Additional groups use and modify existing codes to perform increasingly high-impact, world-class simulations.

Some highlights: the first ever successful simulation of the merger of a binary black hole system (Pretorius, 2005); the largest (640^3) cosmological simulations of galaxies with AGN (HYDRA, Thacker et al. 2006) and the highest resolution cosmological dwarf galaxy to date (100 $M_\odot$ per 'star') (Mashchenko et al. 2008) . Such groups typically require millions of CPU hours per year, in the form of several parallel runs with 100-1000 of processors over several weeks or months., e.g. Dubinski et al. will use 5000 cores for 500,000 CPU hours for a simulation of the local group in 2010, and Pen plans to spend 6 million CPU hours to do a trillion-particle cosmological simulation, which would be the largest ever done.

1.2. *Observational Data Analysis*

Simulation has always been a large client of large-scale computing; but with the increasing size of observational datasets, simulators are in the unfamiliar position of competing with observers for time on large computers.

At the CITA McKenzie and then Sunnyvale clusters, recommended by LRP2000, data analysis — whether the primary analysis of CMB experiments (Readhead et al. 2004; MacTavish et al. 2006) or the combining of many data sets to produce interesting new cosmological constrains (Seljak 2006) — have not only been amongst the largest users of the systems, but also among the highest impact. The expertise developed at CITA on these machines made it feasible for the main analysis for the much larger datatset from ACT to be done by CITA researchers, which currently makes the ACT project one of

[1] A list of the many acronyms and abbreviations used in the text is provided in Table 1.

the largest single users on the new SciNet GPC cluster, currently Canada's largest supercomputer. It's no exaggeration to say that the fifteen million CPU hours used on SciNet since last June — an amount of work which would have required over three years of all of Sunnyvale — have been absolutely necessary for the success of the ACT project. "SciNet is essential for the Atacama Cosmology Telescope (ACT) project. The computer has enabled a new frontier in producing maps of the early universe, and is changing the way cosmologists make sense of the cosmos." (L. Page, private communication).

2. THE NATIONAL CONTEXT

2.1. *Computing and Storage*

LRP2000 occurred towards the end of an era of funding significant department-scale computing facilities with significant computational power and local storage. Investments in large-scale computing has been centralized into 7 HPC consortia across Canada (Table 2). This effort permitted significant economies of scale in operating costs and hardware. The consortia maintain facilities that represent together close to a petaFLOPS of computing capability and and several PB of short- (day long) and long-term (year long) storage with access to Canada's national, provincial and territorial high-performance networks. This storage does not currently provide any 'archival' storage and is only intended for active computing projects.

Compute Canada, an umbrella organization, is seeking stable funding to continue developing and maintain a national HPC platform for research, the Compute Canada National Platform, by unifying access to the resources of these consortia; researchers in Canada can apply for accounts at any of these resources, and request a significant amount of time on any of them through a competitive allocation process. Astronomy and Astrophysics accounted for approximately 20% of the large scale computation and storage resources requested in 2009.

2.2. *Networking*

The backbone of the Canadian high-speed research network is the CA*net 4 connection, maintained by CANARIE and the Government of Canada, which provides a dedicated 10 Gb/s connection into each province. From there, ORANs connect this network to institutions serving more than 39,000 researchers at nearly 200 Canadian universities and colleges, and scientists at many research institutes, hospitals, and government laboratories across the country. The ORANs provide links of various speeds up to 10 Gb/s; however, not all researchers have immediate access to connections of this speed, and even when they do, a lower-speed link may intervene between themselves and their desired resource or, commonly, the hardware and software network configuration is not tuned to take full advantage of the available network capacity.

2.3. *Astronomical Data*

The CADC, operated by the HIA with support from the CSA, archives and serves data from the HST, FUSE and MOST space telescopes as well as the national optical, IR, sub-mm and radio facilities Gemini, JCMT, CFHT, and DRAO (i.e., the CGPS). At the end of 2009 the CADC's data collection totalled 411 TB and an additional 220 GB are added each day. These data are archived in a storage facility that receives stable funding from the NRC but whose growth is now limited by conflicting mandates between Compute Canada and NRC.

The CADC also provides community storage to research groups working on projects with significant data loads and access requirements. Data collections and proprietary access portals are maintained for projects such as BLAST, MACHO and CFHT Large Programs such as MiMeS and NGVS. Details of the CADC infrastructure and projects can be found in a separate WP.

For some projects, such as the ACT experiment which generates approximately 2 TB of data per week, there is currently no straightforward way of distributing such data amongst researchers. The CADC is actively involved in the pilot stages of a new project, CANFAR. Under the leadership of the University of Victoria, CANFAR aims to develop an operational system that enables the effective delivery, processing, storage, analysis and distribution of very large datasets produced by astronomical surveys. Details of the CANFAR project are also contained in the CADC WP.

For simulation results and analysis there is no comparable method of sharing data amongst researchers or dividing analysis between facilities. With cutting-edge but otherwise unremarkable research projects routinely generating tens of TB of outputs, it is often infeasible to make these generally available, despite the fact that such results could often be analyzed in many different ways to different purposes by various groups; see for instance the wide range of purposes put to use by the data products from the Millennium Simulation of large-scale structure formation.

A community accessible resource for model/theory result storage and distribution, perhaps operated out of the CADC or one of the consortia, is required for these data-intensive projects. However, such a resource must provide not just physical storage but also the expertise to ensure the collection is properly curated and accessible to other members of the given project or the community at large.

3. CURRENT ISSUES

The development of serious national-scale computing since LRP2000 has meant that most researchers have access to much more computing resources, even in terms relative to the growth of computing power over that time, than in the past. However, serious issues remain.

3.1. *Computation*

The recent addition of top-20 class supercomputing[2] to the Canadian HPC ecosystem is an extremely positive development, and one which the Canadian astronomy and astrophysics community was able to make immediate use of. However, unlike the PRACE initiative in Europe or the NSF or DOE supercomputing programs in the United States, there is no commitment by Canadian funding agencies to maintain such a world-class level of computing. Without such a commitment, this could be a 'one-off' caused by funding of the new and likely last consortium, SciNet, and within five years Canadian

[2] Top twenty relative to the Top-500 computing list, a twice-yearly list of the largest computing facilities in the world; `http://www.top500.org`

astronomers may again be restricted to having reliable access only to mid-range computing at home, effectively shutting them out of many projects and research opportunities.

Absent any other infrastructure, remote large-scale computing and large data stores do not obviate local computing resources; for some projects, users must still take their simulation results or survey data to their local cluster for interactive visualization or analysis to derive knowledge from data. The continuing roll-out of high-speed networks assists with this, making some sorts of remote interactive analysis and visualization more feasible; however, even if the high-speed networks were ubiquitous and effective, which they are not yet, it does not solve every problem.

For those analyses of data where the analysis product is much smaller than the input data for individual/focused analysis, the network bottleneck is largely avoided; the computing resource with most ready access to the relevant portion of a dataset will be given a task, and segments of the resulting data product can be sent over the network to the user. This infrastructure does not, however, help with interactive analysis and visualization of multi-TB outputs from processes whose output datasets are of a size comparable to the inputs; the large data-transfer requirements weaken the advantages of such a cloud-computing approach, where it is explicitly assumed that locality doesn't matter. For these sorts of projects, local signficant computing resources such as visualization clusters will still be required — these might be maintained by the consortia (as SHARCNET does with some visualization machines at different instututions) or by the institutions themselves.

3.2. *Personnel*

When computing resources were more local, generally the computational staff were limited to those keeping the computers running; most institutions had few or no staff available to assist researchers in making use of the resources – by, for instance, porting, parallelizing or developing new computational tools for analysis or simulation.

NSERC's Discovery Grant Program is not geared to support software development. As a result having existing packages retooled to make effective use of HPC facilities cannot be properly funded; hiring students and/or postdocs to develop innovative new scientific tools enabled by access to new HPC facilities is impractical, if not impossible. Making the best use of these resources on this scale requires not just tinkering, but fundamentally rethinking the entire approach to a problem from the algorithm level up, and every implementation detail matters. This requires extensive input from someone well versed in the issues. The advent of 1000 processor machines (including three of the 100-largest computers in the world) within the Canadian research landscape has made the lack of expert personnel an acute problem.

At present, the level of staffing is insufficient to the need. In the United States, at Argonne Labs' Leadership Computing Facility, there are over a dozen staff to work with users with one aspect or another of their code and algorithm development. There are similar numbers at NERSC (16, including staff pushing the envelope on visualization, tools for multicore programming, and science grids), and still greater at ORNL (23, where they like to use a rule-of-thumb of 'one FTE per three supported projects', B. Messer, private communication), SDSC, and projected for the upcoming Blue Waters project. At HECToR in the UK they have 20 FTE, *not* including 'first line' support for things like help with compilers. At SciNet, whose biggest computer is very similar to that at HECToR, the number has recently increased to three *including* 'front line' support, with plans to increase to six. No consortium funds the levels of personnel seen in these foreign centres; current funding for HPC support staff in Canada is at about a quarter of international levels as discussed in the Computing Long Range Plan update (2007), and other consortia have no plans or funding to significantly increase staffing levels. Canada stands to lose out if it cannot support the group development model, including dedicated programmers, that is becoming the international norm.

These staff are HQP with skills analogous to both instrument builders and support astronomers (an analogy due to F. Herwig). Indeed, finding and retaining such well trained people in Canada is not easy, even for the consortia, as due to the aforementioned funding issues there is a lack of people trained in making effective use of large-scale HPC resources. Current HPC consortia staff have a role to play in training scientists in the use of these tools, and almost all the consortia offer such classes but they can't replace direct per project support, and staff rarely have time to gain expertise into the advanced material needed for cutting edge development. The CADC CANFAR project, which consists of seven postdocs and five programmers dedicated to making 'cloud' or 'grid' computing more accessible to observational astronomy projects is a model example of the level of support and activity required to ensure that HPC resources can be fully exploited by the research community. Unfortunately the CANFAR project is only funded for the next year and the road towards continued support is not clear.

Within the theory/modelling community, the lack of such expert HPC personnel not only means that there is less astronomical HPC tool development in Canada than there could usefully be, but also that there is less support tool development occurring that would make astronomers use of HPC easier and less time consuming - there is very little Canadian scientific visualization tool development being done, despite the fact that there is real demand for this from the community both for observational data such as that produced by the SKA pathfinder survey, and for interpreting the results of large scale 3D computations. Data caves can aid this process, but current technology is rather primitive. Neither is there development of better (possibly domain-specific) programming languages or analysis tools. A distributed-memory tool that could do most of IDL's job in parallel on distributed memory machines, for instance, would be of immense use to the community, but there are no groups in Canada that would conceivably work on such a thing. Many observers use the KARMA and Starlink software suites as visualization tools; however Starlink is no longer funded (some critical pieces are supported by other agencies) while KARMA is no longer being developed, despite the fact that there is no equivalent tool to replace it. Groups doing tool development of this sort would require steady, dependable funding of HPC centres of excellence, something which has historically been intermittent even in the US, and certainly does not exist in Canada - we could usefully learn from mistakes made by the US NSF supercomputing program (Smarr 2010).

The dual roles analogous to instrument building and support astronomer suggests another tension – where should these staff be located? The advantage of central locations at the consortia is that one can have pools of broad expertise available to everyone; but for some projects, working together over email may not be sufficient. Supporting astronomers at the leading edge requires local availability of these HQP; depending on the needs of the projects at consortium member institutions, this may mean staff located locally to the institutions with expertise specific to the projects occuring at those locations; for others it may mean improved telepresence facilities or increased travel budgets with support for multi-week stays to work on particular projects full time.

3.3. *Networks and Data*

Basic astronomical questions of today will be addressed with the help of panchromatic data acquired at many observing facilities, possibly over many epochs, on scales many orders of magnitude larger than we have been used to. Instruments such as the ACT experiment are already producing data at rates of 2 TB/week, and MegaPrime at CFHT approaches a TB/day; but new facilities such as ALMA and, in particular, the SKA will generate data volumes of many EB and even YB, terms many of us have never before encountered (see Fig. 1).

Storage concerns aside, the importance of data mining, the exploitation and visualization of these massive data sets, will become unavoidable for astronomers in the current and next decade. This has been reflected in the US by the emergence of a named discipline, "astroinformatics" and is the subject of separate WP's in the US decadal survey and LRP2010. When data sets are as large and high dimensional as next generation surveys will be, it will only be possible to extract meaningful information from them by using highly sophisticated and automated mechanisms. Development of such exploitation capabilities using large computing resources and data mining algorithms will become a full-time specialty.

In a survey of the Canadian research community, many observational astronomers stated that they "rely crucially" on the CADC and urge continued "rock-solid" support for it. Explicit mention is also made about the need for routine access to SCUBA2 and ALMA data in the future, as well as access to the computational resources that need to go along with these large datasets. But it is clear that many also feel that even the current CADC archives already require more powerful and intuitive querying capabilities. This applies not only to advanced data products produced by the CADC itself and their data providers, but also to unprocessed data.

Today's observational astronomers already struggle with the increased image processing needs for their individual research programs. They often, at considerable expense, purchase their own local machine with dozens of GB of RAM and several TB of disk space. When they are not using the machine it sits idle and in a few years will need replacement. This is a clear waste of resources. With adequate network support, both in terms of physical infrastructure and technical support to ensure that networks are being used to their full capacity, the proper use of existing consortia computing resources is a much more efficient use of computing dollars.

The movement of progressively larger number and volume of files through the network will become even more important (and routine), see Fig. 2. We must urge CANARIE to continually expand the advanced networking capabilities available to researchers and support the development of new technologies and approaches that allow researchers to fully exploit this capability in discovery research and development.

4. KEY RECOMMENDATIONS

Canadian researchers have made great advances in computing infrastructure, but we are not yet fully competitive, and there is tremendous need for technical expertise and networking to enable the full exploitation of resources that do exist, and maintaining these resources to at least the current level over the period of LRP2010.

Future research goals in astronomy will be held back without a consistent approach, and long-term, reliable funding for computing, data, and networks. Increasingly powerful computers, which enable and are required by cutting edge research in theory and observation, must be matched by personnel to work with researchers to make use of them; by proper funding for high-capacity research networks which must support the physical network, the technical management of that network; by well-funded, reliable, data stores; and the implementation of research tools that make utilizing the computing, network, and data stores straightforward and scientifically fruitful.

In this era of an growing dependence on Compute Canada infrastructure, a stable and long-term funding solution must be created. Storing the research community's archival data on storage that is funded on 5-year time cycles is a risk we should not take. Building computers without a community of professional staff to build the tools to help researchers use them for all areas of their research is a mistake made by NSF centres (Smarr 2010) which we need not make. However, our current staffing levels are even lower than NSF centers. We need to ramp up rather than just maintain our staffing levels.

Thus we recommend that Compute Canada be funded at $72M/yr, and in particular that a top-20 class computing facility be part of every refresh; that technical analyst staff funding be tripled; and that this funding be stable over timescales much longer than 5 years so that centres and staff can make long-term plans and work on long-term projects. Meeting these requirements would bring Canadian HPC spending to *below or just at* the current G8 average, the *minimum* required to remain competitive.

In addition, we recommend — and require — that national networking be brought up to a level commensurate with the rapid increase in available computing power since the last LRP. There are too may institutions with 'last mile' bottlenecks in their network which will limit their ability to make full use of the Compute Canada Consortia, or of the CADC. These bottlenecks may be in their network connection themselves (e.g., a 1 Gb/s link vs. 10 Gb/s), or in sub-optimal tuning of the network for performance. The 'first mile' bottleneck of a 1 Gb/s link to CADC will become inadequate on a timescale of months, not years, and must be addressed immediately.

REFERENCES

Choptuik, M.W. & Pretorius, F. 2009, astro-Ph, 0908.1780
Computing LRP update 2007, https://computecanada.org/modules/cms/checkFileAccess.php?file=local.cc_nic/LRP_update.pdf
Pretorius, F. 2005, Phys. Rev. Lett., 95, 121101
MacTavish, C.J., *et al.*, 2006, ApJ, 647, 799
Mashchenko,S., Wadsley, J., & Couchman, H.M.P. 208, Science, 319, 174
Readhead, A.C.S. *et al.*, 2004, Science, 306,836
Seljak, U., Slosar, A. & McDonald, P. 2006, J. Cosm. Astro-Part. Phys., 10, 14
Smarr, L. 2010, http://www.calit2.net/newsroom/release.php?id=1632
Thacker, R.J., Scannapieco, E, Couchman, H.M.P. 2006, ApJ, 653, 86

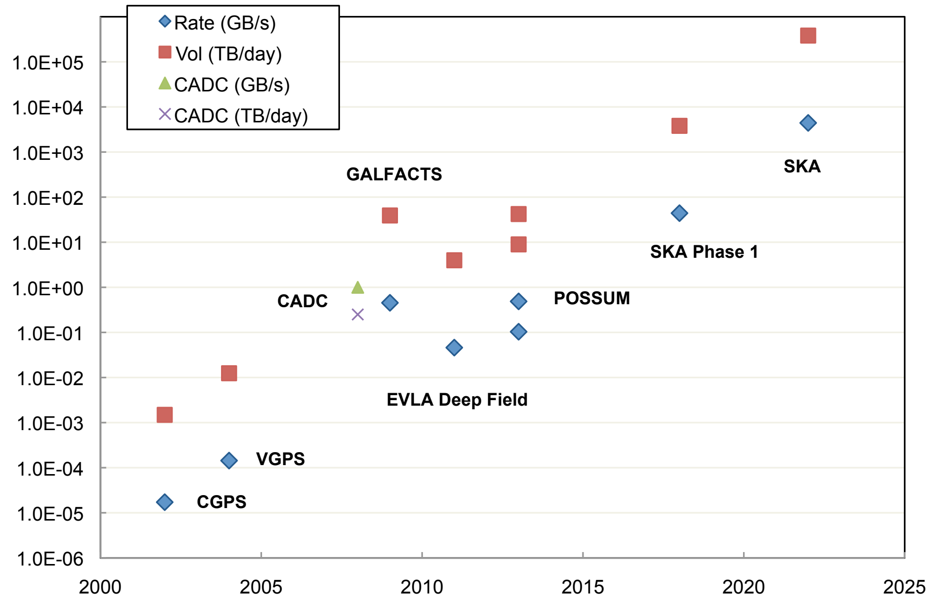

Fig. 1.— Current and anticipated data acquisition rates for existing and planned radio facilities. (Courtesy of Russ Taylor, U. Calgary, and JJ Kavelaars.)

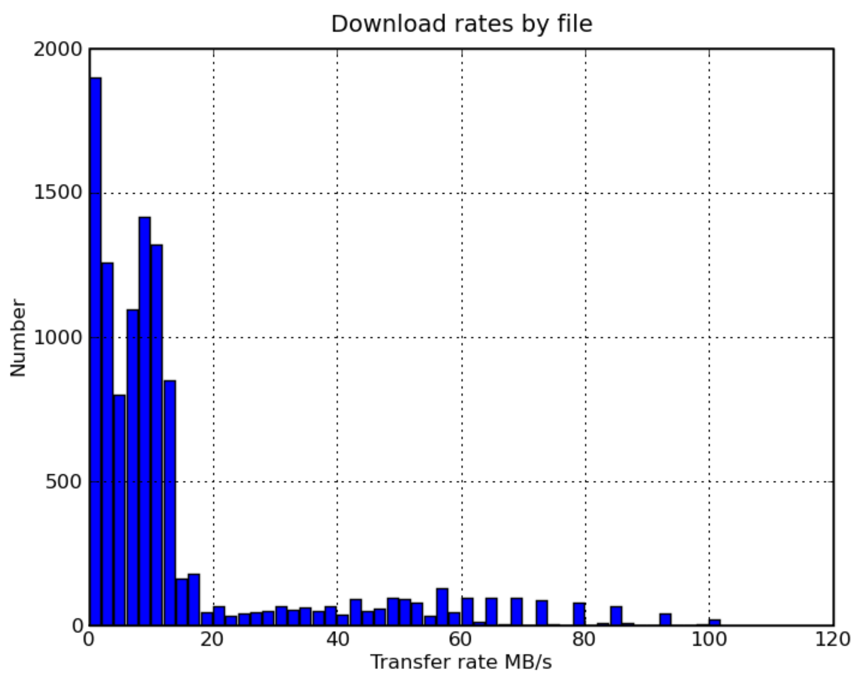

Fig. 2.— Typical download rates binned by rate for access to between Canadian Research Universities and the CADC. Aggregation is over all universities in Canada accessing files from the CADC. A number of Canadian institutions have tuned their networks to allow transfers at a rate close to the theoretical limit of the CADC's connection (120 MB/s) but the majority of institutions experience very low transfer rates due to poor network configurations and 'last mile' issues.

TABLE 1
Acronyms and Abbreviations Used in the Text.

Abbreviation	Description
ACEnet	Atlantic Computational Excellence Network, a CC Consortium
ACT	Atacama Cosmology Telescope
ALMA	Atacama Large Millimeter Array
AMR	Adaptive Mesh Refinement; a technique for using high resolution only where it is needed in a simulation
CADC	Canadian Astronomy Data Centre
BLAST	Balloon-borne Large Aperture Submillimeter Telescope
CANARIE	Canadian Network for the Advancement of Research, Industry and Education
CANFAR	Canadian Advanced Network for Astronomical Research
CFHT	Canada-France-Hawaii Telescope
CGPS	Canadian Galactic Plane Survey
CFI	Canadian Foundation for Innovation
CITA	Canadian Institute for Theoretical Astrophysics
CLUMEQ	Consortium Laval, Université du Québec, McGill and Eastern Quebec
CPU-hour	A unit of computing time, requiring one core or processor for an hour
EB	Exabyte = 1000 PB or 10^{18} bytes
FLASH	A multiphysics astrophysical reacting self-gravitating AMR fluids code, `http://www.flash.uchicago.edu`
FLOPS	FLoating OPerations per Second, a rate of computation; the fastest computer in the world can in principle compute at $\approx 2 \times 10^{12}$ FLOPS.
FTE	Full Time Equivalent
FUSE	Far Ultraviolet Spectroscopic Explorer
GASOLINE	An SPH code developed by James Wadsley, `http://imp.mcmaster.ca/images/`
Gb/s	Gigabit/s, a data rate; 1Gb/s is a modest-speed research internet connection, 10Gb/s is the fastest currently available in Canada.
GB	Gigabyte = 10^9 bytes; 2 GB of RAM is typical for a desktop computer.
HECToR	High End Computing Terascale Resource
HIA	Herzberg Institute of Astrophysics
HPC	High-Performance Computing
HPCVL	High-Performance Computing Virtual Laboratory
HQP	Highly-Qualified Personnel
HST	Hubble Space Telescope
HYDRA	N-body, hydrodynamical simulation code developed by Couchman
IDL	Interactive Data Language
JCMT	James Clerk Maxwell Telescope
LRP	Long Range Plan (2000 and 2010 versions)
MACHO	Massive Astronomical Compact Halo Object
Mb	Megabit
MB	Megabyte = 10^6 bytes
MiMeS	Magnetism in Massive Stars (CFHT Large Program)
MOST	Microvariability and Oscillations of Stars Space Telescope
NCSA	National Center for Supercomputer Applications
NERSC	National Energy Research Scientific Computing Center
NRC	National Research Council of Canada
NSERC	National Science and Engineering Research Council
NSF	National Science Foundation
ORAN	Optical Regional Advanced Networks, responsible for regional connections to the cross-Canada CANARIE network.
ORNL	Oak Ridge National Laboratory
PB	Petabyte = 1000 TB or 10^{15} bytes
petaFLOPS	10^{12} FLOPS, approximately the speed of the worlds fastest supercomputer.
PMFAST	A self-gravitating MHD code developed by Pen, Merz, and Trac.
PRACE	Partnership for Research Advanced Computing in Europe, a European program for providing HPC resources to its researchers.
RQCHP	Réseau Québécois de Calcul de Haute Performance
SciNet	U. of Toronto HPC Consortium
SDSC	San Diego Supercomputer Center
SHARCNET	Shared Hierarchical Academic Research Computing Network
SKA	Square Kilometer Array
TB	Terrabyte = 1000 GB or 10^{12} bytes; 1-2 TB are the sizes of the largest single hard drives currently available.
WestGrid	Western Canada Research Grid
WP	White Paper
YB	Yottabyte = 10^9 PB or 10^{24} bytes

TABLE 2
Compute Canada's HPC consortia.

Consortium	Web Page
ACEnet	http://www.ace-net.ca/
CLUMEQ	https://www.clumeq.ca/
HPCVL	http://www.hpcvl.org/
RQCHP	https://rqchp.ca/
SciNet	http://www.scinet.utoronto.ca
SHARCNET	https://www.sharcnet.ca/
WestGrid	http://www.westgrid.ca/

TABLE 3
LRP2000 Recommendations for Computing and Data.

Recommendation
The LRPP strongly recommends that the CADC host archives of data from upcoming space and ground-based observatories, and develop innovative data mining techniques for their exploration. This should be one of the highest priorities among the computational projects.
The LRPP strongly recommends that funds be allocated toward the support and upgrade of a mid-range parallel computer plus a local user-support person. This should be one of the highest priorities among computational projects. Furthermore, this capability should be located at CITA to provide national high performance computing for modelling and simulations.
The LRPP recommends that the funding towards equipment grants in the country be substantially increased to enable researchers to keep pace with the huge volumes of data and computation that will shortly become standard in astronomy and astrophysics.
The LRPP recommends that a sustainable, nationally funded multidisciplinary HPC network be established through initiatives made possible by the CFI program.

TABLE 4
LRP Midterm Review Recommendations relevant to this WP.

Recommendation
The MTRC recommends that NRC-HIA conduct a review of Canada's role in global data management and the CADC's contributions to this role, particularly in light of the new ground-based and space-based facilities such as those described in the LRP. Meanwhile, LRP support for CADC should be continued to help maintain the strengths of the existing programs.
The MTRC recommends that CASCA, through its subcommittees, conduct a review of the data retrieval and analysis requirements of all LRP facilities, and then consult with NRC-HIA and ACURA to formulate a coherent strategy to address this issue. This should precede and provide input to the review by NRC-HIA of the contributions by the CADC to Canada's role in global data management contained in the previous recommendation.
The MTRC recommends that the HPC community urgently develop and implement a strategy for providing access to a cost-effective Tier 1 computing system for astrophysics, i.e. one which is competitive with the leading systems over all disciplines worldwide. The emphasis should be on covering the need for the next three-year technology cycle. The strategy must ensure access which satisfies the demand of its theoretical astrophysics community, and ensure a national leader- ship role for this community in computing and an international leadership role in the science achieved. Concurrently the community should engage in a cross-disciplinary dialogue with the aim of ensuring long-term sustainability for Tier I HPC.