# Astronomy and Astrophysics Research Computing Needs: Present and Future

**Computing, Data, and Networks Committee of the Canadian Astronomical Society (CASCA)**

## *Executive Summary*

Advanced research computing resources have never been so essential to the Canadian Astronomy and Astrophysics research community. In the past few years, astronomical researchers have benefited greatly from modern large-scale computing systems; a diverse range of resources, which are a good match to the diverse computing needs of our scientists; and good working relationships with existing providers, allowing flexibility and collaboration between these centres and research groups.

However, CASCA has concerns about the near future of advanced research computing available to its researchers. Here the Computers, Data, and Networks Committee of CASCA present, on behalf of the Society, a summary of the current state of the computing needs, successes, and concerns of our researchers taken from previous consultative summaries and their updates. This is the first step of a process that will continue through the first half of 2013, which will include a comprehensive survey of research computing needs of the Canadian Astronomy and Astrophysics community, and will investigate a variety of strategies for meeting those needs.

Early systems funded by the CFI NPF are already showing their age; in many cases they are out of their maintenance contract and are already starting to fail. The lack of any clear signs of new investment on the horizon means that even if existing systems were to continue operating perfectly, as other nations continue to invest in new research computing platforms, our researchers, using stagnant computing hardware, will not only fall behind our international competitors as data volumes continue to increase, but also be unable to make full use of prior investments.

When new funding does become available, the Canadian astronomy community would like to see changes in emphasis taken as lessons learned from the CFI NPF procurement. Previous investment focused largely on computing hardware. While this addressed a real and pressing need resulting from years of underinvestment, the research endeavor requires a more holistic approach. Computing hardware investments must be balanced with similar investments in storage, highly qualified personnel, software development, and networking to maximize results.

In this report, we recommend an urgent search for new and sustainable sources of funding for advanced research computing funding; an increased focus on personnel, software development, and storage; maintaining a diverse range of systems; enabling major longer-term projects by committing resources for longer than the one-year allocation window currently of the RAC process; continuing to enable close working relationships with research groups and computing providers, preferably as close to the researchers as possible. In addition, we recommend that CCI's board, through the proposed Researcher Advisory Committee or otherwise, establish a direct relationship with CASCA (and similar professional groups), with via persons charged with representing the needs of these research communities in planning for Compute Canada.

## Contents

# Overview

## Astronomy and Research Computing

Astronomy and Astrophysics is one of nine research fields in which Canada leads the world in scientific impact[1], as measured by normalized output of highly cited papers. As an essentially observational discipline – one can't experiment on stars, which is probably for the best – Astronomy relies on the interplay of careful observations and detailed theoretical modeling to advance knowledge. Both the analysis of observations, and simulations to explore theoretical understanding, have long depended on research computing; astronomers were heavy users of FERUT, Canada's first "electronic computing machine", as early as 1952.[2] Today, Astronomy and Astrophysics, although just 0.5% of Canadian researchers, typically comprise 10%-20% of Compute Canada resource requests.

Astronomy is defined by the objects it studies, rather than particular processes, and its remit is unusually broad – the study of every object and phenomenon that occurs somewhere in the Universe besides Earth. It involves the study of fluid dynamics of both hundred million light-year sized galaxy clusters and millimetre-sized flames inside of supernovae; the calculation of reaction rates for individual atomic nuclei and the integrated burning rate of all of the stars in a galaxy; the analysis of the change of light over time from a single star to detect a planet and the analysis of the light from millions of galaxies at once to detect dark matter. Because Astronomy and Astrophysics encompass an exceptionally broad range of phenomena, its computing needs are exceptionally diverse.

Astronomy in Canada has thus far kept abreast of – and even pushed forward – the rapid increase in data volumes which provides such exceptional scientific possibilities. However, exponentially increasing data volumes provides challenges both for our scientists' observational data analyses and the theoretical simulations needed to ensure that our understanding keeps pace with our observational discoveries. Canadian astronomical and astrophysical researchers are well posed to take advantage of the opportunities around the corner – with world-class expertise in fields such as observational data analysis, general relativity simulations, galaxy modeling, cosmological simulations, and grid and cloud environments for developing insights – as long as the necessary resources and support are in place.

---

1 See for instance, the Sept 2012 study "The State of Science and Technology in Canada, 2012", commissioned by the Ministry of Industry and performed by the Council of Canadian Academies, http://www.scienceadvice.ca/en/assessments/completed/science-tech.aspx

2 Even then, both observational data analysis and simulation were uses; the Dominion Observatory needed to better analyze stellar positions to improve their time services, and models of stellar atmospheres were calculated. See this 1954 article in the Journal of the Royal Astronomical Society of Canada, http://articles.adsabs.harvard.edu/cgi-bin/nph-iarticle_query?1954JRASC..48..176M .

## CASCA and Canadian Astronomy

The Canadian Astronomical Society / Société Canadienne d'Astronomie (CASCA) is the national learned society representing Canadian astronomy and astrophysics researchers; and within CASCA, the Computing, Data, and Networks Committee (CDNC) is responsible for providing timely information to the CASCA board and executive about computing needs and relevant opportunities. Since 2000, CASCA has played a leading science policy and planning role for the discipline, issuing a decadal Long Range Plan (LRP) after a year-long process supported by NSERC, NRC, CSA, and CFI and involving nationwide consultation, proposals, town halls, and written submissions.

Both LRP2000 (http://www.casca.ca/lrp/) and LRP2010 (http://www.casca.ca/lrp2010/) recognized the key importance of computing to astronomical research. LRP2000 proposed the establishment of a medium-sized astronomy-specific computing resource, along with personnel for programming support, to be available to Astronomy and Astrophysics researchers across Canada; this was soon implemented at the Canadian Institute for Theoretical Astrophysics. LRP2010 lauded the CFI NPF and Compute Canada as being significant steps forward, and recommended a doubling of annual funding to such initiatives to bring Canada closer to the G7 average for research computing funding; increasing the fraction going to research support; ensuring that a top-20 class system was always available; and support cloud-computing environments and more longer term data management.

## This Position Paper

In some respects, LRP2010's optimism now seems somewhat misplaced. At the time, several new large scale computer systems were being installed, and there was increased hiring of support staff to go along with these new resources; but since then, operating funding has stalled and no new hardware funds are anywhere to be seen on the funding horizon. Already, systems bought in the years up to 2009 are falling out of maintenance contracts and parts are failing without prospect of being replaced; Canada's high point of having a very large scale computing system – amongst the top 20 in the world – available to research seems to have been an accident of timing rather than an achievable, sustainable, goal; and overall high-performance research computing funding seems to have permanently fallen back from its high point of being "only" a factor of three below the average of the other G7 countries that Canadian researchers consider their main competitors to a level permanently below even that rather inglorious high-water mark.

In this report, the Computing, Data, and Networks Committee of CASCA lays out the current state of advanced research computing for the Astronomy and Astrophysics research community, and areas of potential improvement for the future, at the beginning of a process which will include in the first half of 2013 a comprehensive and quantitative survey of the community's needs and an assessment of ways to meet those needs. We note strengths, weaknesses, threats, and future opportunities of performing research under the current systems, and propose recommendations for Canada's advanced research computing environment which would enable still more world-leading astronomical research in the years ahead.

## Current Canadian Astronomical Research Computing

We begin by describing the overall the advanced research computing environment needed to support astronomy and astrophysics research, then give an overview of how Canadian Astronomy and Astrophysics use advanced research computing today, and finally investigate issues with the current environment. Although we cannot hope to describe in detail the many accomplishments of and challenges overcome by Canadian computational astronomy and astrophysics research here, we describe four representative highlights in the "Astro Computing Highlights" sidebars which span a range of sub-disciplines and computational requirements and provide a broader overview below.

### Astronomy and Astrophysics Research Computing Needs

Performing world-class computation-powered research requires a variety of inputs, most of which are now solely within Compute Canada, Inc.'s purview to provide. It is important to emphasize that the different aspects of the advanced research computing infrastructure – including software, tools, high performance computing resources, rapid-access large-scale storage, and high performance networks, all play important roles. Overemphasis on software at the expense of sufficient storage, or network services without compute resources, will lead to a fundamental imbalance.

### Computer Resources

Computing hardware is necessarily fundamental to the endeavor, but there a variety of different needs in computational Astronomy and Astrophysics research which must be met by different sorts of systems. Often, all of these systems are lumped together under an umbrella of "High Performance Computing" systems, a term which obscures more than it clarifies. Here we distinguish the terms as we use them below; they are defined in rather more detail in the following section in the contexts in which they are useful. Note that there are no universally-agreed upon definitions for some of these terms.

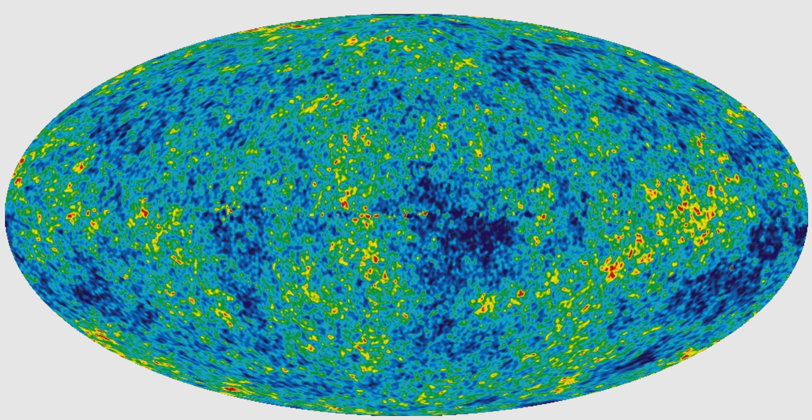
## Astro Computing Highlight: CMB

The Cosmic Microwave Background (CMB) is the sky-filling radio signal from very early in the history of the Universe – before matter started to clump into what would become later become galaxies and clusters of stars. These "echos of the big bang" which were once very energetic light comes to us from what is now very distant source – the hot, dense, Universe of 13 billion years ago – as low frequency radio waves.

A map of this very early phase of the Universe provides enormous information about its structure – the distribution of sizes of dense regions, the magnitude of their peaks, and their correlations – and that information, particularly when compared to what we know about the structure now, 13 billion years later – provides extremely stringent tests of models of the physics of the Universe. It is these measurements, more than any others, which have brought the field recently into the era of "precision cosmology" - so that we now know numbers like the age of the Universe, and the proportion which is composed of dark matter and dark energy, to several significant figures. This sort of accuracy is crucial to understanding the nature of dark matter and energy.

However, making these maps is enormously challenging; the entire history of the Universe and all the objects in that history stand between us today and the signal that we wish to measure. The signal-to-noise ratio for these measurements is extremely small; the noise overpowers the signal by factors of ten-thousand to one or more. As a result, very large amounts of data must be taken, and these huge data sets must be analyzed all at once to find the necessary signal. This requires large scale computing; thousands of processors at once, all acting on a single shared data set. Further, these analysis must be performed many times for each data set, to examine different aspects of the data. Large numbers of processors on tightly integrated supercomputing systems are thus needed, and significant storage as well – modern CMB experiments can take up to two terabytes of data a week, and upcoming experiments will produce much more. In addition, some of this data is essentially irreplaceable, and facilities for reliable archival storage are needed. Once maps are generated, high-throughput style computing – large numbers of independent computations which do not need to be performed on a single, large-scale "supercomputer" - can be performed to test various theoretical models against the observed data.

Canada a leader in the analysis of CMB data. A history of research computing expertise, including the LRP2000-recommended system at CITA and now the LRP2010-endorsed Compute Canada systems, along with close involvement in early balloon-borne projects such as BOOMERANG and SPIDER, and satellite missions such as WMAP and Planck, means that the Canadian Astronomy and Astrophysics research community has developed world-class experience and expertise in this area. This expertise means that Canadian astronomers are taking a leading role in the analysis of data from current and upcoming projects such as the Atacama Cosmology Telescope in the mountains of Chile: "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, Principal Investigator, Atacama Cosmology Telescope).



Continued leadership is endangered by the lack of new supercomputer class systems with appropriate storage, and upgraded network connections to international institutions. If these can be made available, Canada can continue to play a principal role in larger experiments on the horizon.

*Figure left: WMAP map of Cosmic Microwave Background, courtesy NASA*

- **Advanced research computing (ARC)** is the umbrella term we use for any sort of research computing environment where one has to do something specialized to get the answers you need beyond simply using standard software on one's desktop.
- **Supercomputers** are the traditional high performance computing systems; the “big iron” which might feature on the top 25 of the Top 500 list of global supercomputers. These are machines built for enabling very large-scale computations which make use of thousands or more processors at once working very closely together on one single computational task. These machines often include fairly new technologies which then trickle down to smaller-scale systems in future years. This class of systems are also called “capability” systems in that their defining feature is new capability to tackle larger problems.
- **High throughput research computing**, sometimes called “capacity” research computing, focuses rather on the ability to run large *numbers* of computations rather than large computations. In the context of Astronomy and Astrophysics computing, however, these individual computations may still be relatively large, spanning many processors and many nodes, and thus will still require specialized networking, server, and storage hardware. This distinguishes it from the next category,
- **Commodity Cloud computing**, of the sort which might be provided by Amazon, which typically consists of commodity networking and server hardware within the compute cluster.
- **Specialized hardware** refers to typically new computing equipment which may not be generally useful across all research computing needs, but can be invaluable for particular tasks. This category often changes rapidly, with specialized tools morphing into new technologies or becoming mainstream. Current specialized hardware of particular interest is GPU (graphics processing units) for image processing or numerical simulation; FPGAs for signal processing; and more generic co-processing units like the new Intel Phi. Having ready access to these newer technologies can not only greatly advance particular research programs, but give the community experience with up and coming computing technology before they become widespread.

Most current Compute Canada systems might be thought of as small/old supercomputer systems or high throughput systems, with a small number of distinctly similar specialized GPU systems. Astronomy computing will need “big metal” traditional supercomputing, high-throughput computing for medium-scale computations, cloud-style computing for handling large numbers of small computations, specialized systems, and all stages in between. A broad spectrum of resources, balanced across capabilities, will always be required.

### Storage

All astronomical and astrophysical research computing either ingests or produces very large quantities of data. The data must be readily accessible, safe, and secure; there must be a mechanism for archiving data for long term and making data widely available. It is generally understood in our community that Compute Canada has thus far greatly underinvested in all forms of storage. The situation for long-term storage is particularly dire; Canadian astronomy finds itself in the untenable position that important, irreplaceable data which will be useful and needed for decades to come is currently stored on systems with at very best a 5-year funding horizon.

As with computing hardware, storage can be architected very differently to address different sort of uses. Sometimes there is a great need for fast local storage for intermediate results as well as larger global systems for initial and final results. Large scale data analytics requires rapid random read access to data whereas supercomputing-scale simulations generally require only raw bandwidth in writing to disk.

### Software, Personnel

The very best computing and storage hardware is of no use to researchers who cannot yet make effective use of them. High-quality research software which can take advantage of the high-end computing resources are, in themselves, highly specialized (and valuable) tools. As described in the "Codes as Instruments" white paper for LRP2010[1], high-quality, trusted, "community codes" can shape research directions; and likewise, lack of research software which can take advantage of advanced research computing systems for a particular problem can hinder a program.

Such software for research problems is necessarily bespoke, and must be written by personnel very knowledgeable in advanced research computing, and with enough scientific background to be able to understand the scientific problem being solved. Compute Canada provides access to Technical Analysts to collaborate with research groups in generating such software or adapting such software to new advanced research computing systems; this is in many cases an improvement over the earlier departmental or university-scale centres, but the level of staffing for advanced level consulting, support, or programming is vastly less than international norms. As a result, every astronomy and astrophysics research group in Canada is required to develop a significant amount of expertise "in house" independently, which is enormously inefficient (and poorly supported by existing funding mechanisms at any rate.)

1 "Codes as Instruments", http://www.casca.ca/lrp2010/Docs/LRPReports/CAI.pdf

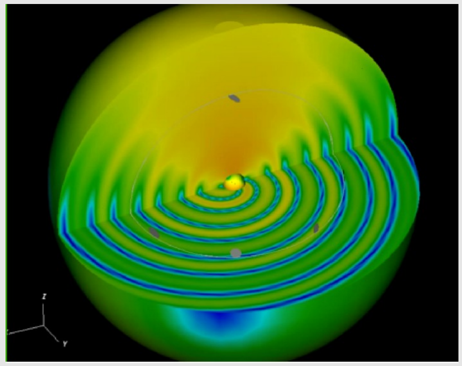
## Astro Computing Highlight: Merging Black Holes

One of the most surprising predictions of general relativity is of gravitational waves – ripples in the fabric of space-time which propagate outward from accelerating massive objects. These waves move out in all directions, affected very little by intervening matter, and can open an entirely new window onto our Universe, just as radio-wave astronomy did starting in the middle of the last century. Observations of gravitational waves will provide insights into black holes, their numbers, and their interactions. As black holes attract and orbit each other, or other dense objects like neutron stars, they will emit gravitational radiation, losing energy to space-time itself, slowly falling closer in to each other and orbiting faster and faster until a spectacular merger occurs.

Gravity is much weaker than light – we can see objects that emit light at much greater distances than we significantly feel their gravitational pull. Detecting gravitational radiation is experimentally extremely challenging; while the effects of gravitational radiation has been seen, the waves themselves have not been detected. This may shortly change. Advanced LIGO, the 2012/2013 upgrade of the international Laser Interferometer Gravity-Wave Observatory (LIGO) project which will increase its observing power by a factor of ten, is widely expected to be begin observing these disturbances. But interpreting the signals from the observatory isn't straightforward. As astronomers started looking at spectroscopic images of our Sun, we could recognize hydrogen emission in the spectrum because we could perform terrestrial experiments with hydrogen and recognize the signals as the same. But there are no such similar Earth-bound benchmarks we can use for black hole mergers.

The benchmarks for understanding the signals coming from Advanced LIGO must then come from theory and simulation. Canada has a long-standing involvement in computational studies of General Relativity, from Choptuik's development of the first adaptive mesh software in this field, making large-scale simulations of general relativistic feasible, his student Preteorius's work developing the excision technique which allows simulations of merging black holes, and through to the current day and the recent hires of Lehner and Pfeiffer. These researchers' work includes building libraries of merging black hole pairs, or black hole/neutron stair pairs, in all combinations and configurations, to be able to identify signals as they are seen. Another important goal is to have a "rapid response" facility to be able to simulate – and thus interpret – unusual signals soon after they are detected. The computational challenge here is not that any one of the simulations are enormous – they require perhaps 64 processors, although they must run for weeks or months at a time, and have only relatively modest storage requirements individually – but that an exceptionally large number of them must be performed. What's more, when LIGO observations start coming in, there will be a desire to run simulations as quickly as possible to match to the incoming data.

Here, then, the use case is of "high throughput research computing"; a large ensemble of independent computations must be performed. However, the individual computations are still quite big – requiring several nodes with substantial amounts of memory, connected to each other with a high speed network. Thus it requires advanced research computing hardware.



Further, the need to cut the length of time these computations require from months down to weeks or days is pushing these researchers to specialized "accelerator" hardware such as GPUs (using graphics hardware as specialized computing equipment) or completely new devices such as the Intel Phi coprocessor. While such new architectures hold a great deal of promise, the requirement of adapting existing, mature scientific software to run on a rapidly changing set of accelerator hardware places a great deal of strain on these research groups who must then be both experts in programming cutting edge computing equipment and in numerical general relativity. The lack of adequate programming support as compared to institutions like the Oak Ridge National Laboratory in the US slows down this research program.

*Figure right: Gravitational waves after black hole merger.*
*Courtesy H. Pfeiffer*

### Networking

Research performed using advanced research computing necessarily involves a great deal of data transfer between the scientists' home institutions and computing resources. If data that must be analyzed cannot be effectively moved to the computing resources, or outputs cannot be returned to the home institutions, the computing facilities are of limited usefulness. In several cases, national networking has not been upgraded at the same rate as the increase in computing resources, which introduces science-limiting bottlenecks. Still worse, there are too may institutions with `last mile' bottlenecks in their network which limit their ability to make full use of advanced research computing for their research. Although the national network is CANARIE's, rather than Compute Canada's, direct responsibility, CANARIE's proposed transition to a user-pay model to provide the resources to even maintain, much less upgrade, existing networks suggest that this may very shortly involve Compute Canada.

## How Astronomy uses Advanced Research Computing: An Overview

### Theoretical Astrophysics – Simulation

For much of the 1990s and early 2000s, the use of advanced research computing in Canadian Astronomy and Astrophysics was driven by **Theoretical Astrophysics** – performing simulations based on well understood physics (such as fluid dynamics, gravity, nuclear and chemical reactions, and radiation transport) to explore the behaviours of very complex astrophysical phenomena whose mechanisms were unclear. (See for instance the Astro Computing Highlights "Merging Black Holes" and "Cosmological Simulations".) The Canadian astrophysical research community has had a long history of both developing simulation software ("codes") for theoretical astrophysics simulation – a necessarily incomplete list would include GASOLINE for galactic simulation, HYDRA for simulations of galaxy clusters, PMFAST, CUBEP$^3$M, and PARTREE/GOTPM for gravitational simulations of wide swaths of the Universe as a whole, and SWIFT for exploring the very long-term evolution of planetary systems, and several codes for the evolution of general relativistic systems such as black holes. In addition, researchers have had enormous success using and extending software developed elsewhere to perform simulations of situations relevant to their research.

Some particular recent successes of this approach include the first ever successful simulation of the merger of a binary black hole system[1]; the then-largest simulation of active supermassive black holes in the centres of galaxies[2]; and the highest resolution simulation of a dwarf galaxy[3].

1 Pretorius, 2005, Phys Rev Let 95 121101

2 Thacker *et al*, 2006, ApJ 653 86

3 Maschenko *et al,* 2008 Science 319 174

World-class simulations of these sorts can be extremely computationally intensive. Often the simulations must capture both the very largest scales in the system and the very smallest scales on which interesting dynamics occur, meaning the simulated volumes must be enormous – encompassing the largest scale at the resolution of the smallest – and thus contain billions or trillions of zones, or particles, or “stars”. This alone means that millions of processor hours can be required to calculate the evolution of such a large system, or more if the system is complicated and requires simulating lots of types of physics. Further, since the problem is broken into many interacting pieces over many processors, the processors must be connected by very high-speed networks; and a fast, large, storage system capable of being written to by all of the processors at once for saving outputs must be available. These **supercomputer**-scale resources – very large numbers of processors connected very “tightly” by a very high-speed network and with a high performance storage system – are needed for the largest scale simulations, and a lack of access to these systems can limit the sorts of science that Canadian astrophysicists can perform. Data management – maintaining the results of these very large simulations – isn't straightforward, but carefully done can be very valuable. These simulation outputs, being very costly to have produced, may well be needed to be accessed for some time and perhaps available widely; the results of one such large computation, the “Millennium Simulation”[1] were made public and were analyzed in many different ways by many different researchers internationally, resulting in dozens of follow-on publications.

On the other hand, not all theoretical simulation work requires the largest scale resources. As an example, often the mechanisms behind a particular phenomenon are fairly well understood, but a large number of computations must be performed to match a particular observation – one might need to perform a large number of stellar atmosphere simulations, for instance, to determine the composition of a particular cluster of stars. In this case, the computing need is for a particular form of **high-throughput research computing**; a large number of simulations need to be performed, but they are independent of each other and they have the flexibility of being performed at once or over length of time, and not necessarily even on the same computing resource; thus supercomputing resources may not be necessary. However, many of these smaller computations may still require many processors and span many nodes; thus high speed interconnects and large fast servers may be required for these tasks. In addition, the storage requirements are different but still significant; the (many, independent) outputs from the large number of jobs will typically need to be stored and accessed as a library for looking up similar sorts of problems in the future.

1 http://www.mpa-garching.mpg.de/galform/virgo/millennium/

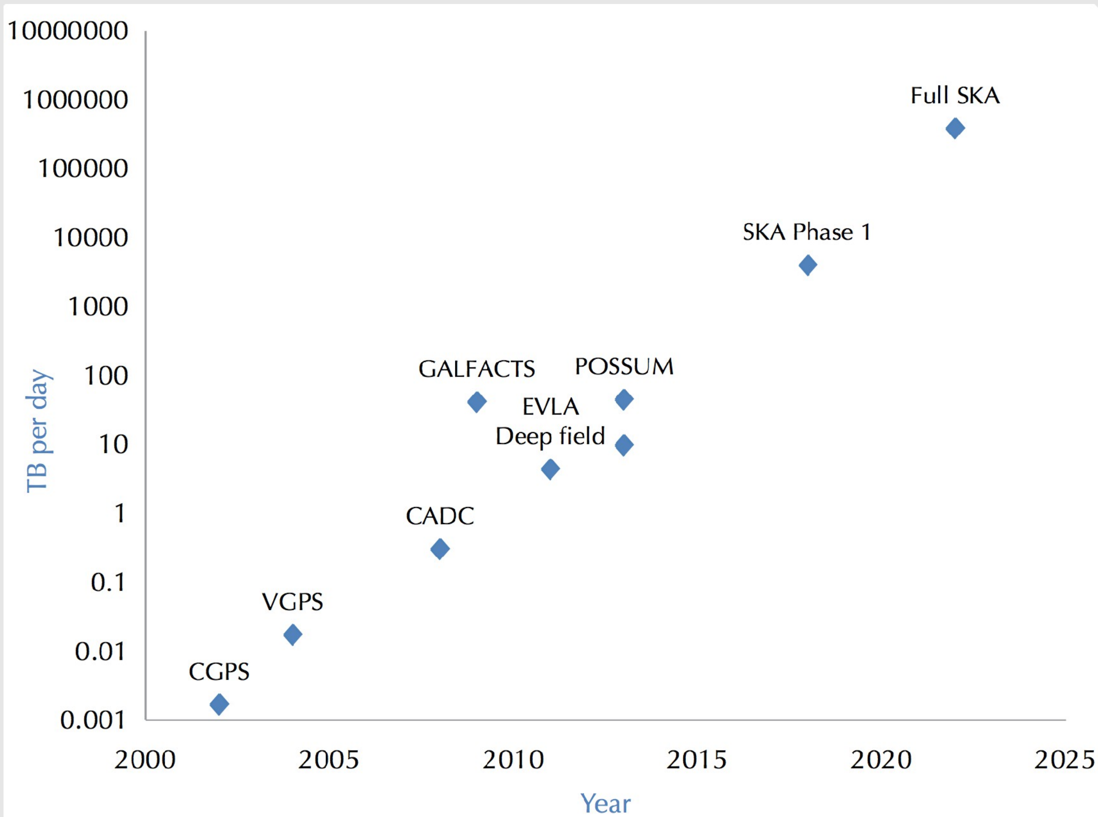



Figure 1: Incoming data rates of upcoming projects will simply exceed the capability of existing resources. A number of planned facilities, such as ALMA, LSST, and SKA, will produce Petabyte to Exabyte data sets. Not only does this challenge observational data analysis capability, but an increase in fidelity of data demands an equivalent advance in theoretical understanding. To meet the challenge posed by upcoming observations, simulations must grow in size, and become more sophisticated in terms of the physics modeled. *Image from CASCA LRP2010*

## Observational Data Analysis

As observational data sets increase rapidly (see Figure 1) theoretical astrophysicists are in the unfamiliar position of competing with their observational colleagues for time on advanced research computing facilities.

One important difference between theoretical simulations and observational analysis is that **data management** is frequently far more complicated for observations. Theoretical simulations may result in substantial data outputs, but where those outputs are written is not particularly important as long as they can be later accessed. In addition, while it would be unfortunate to lose the output of a simulation, it could be regenerated if necessary. On the other hand, observational data may be too large to easily copy yet must come from an authoritative source; access to the data may not yet be public, so that authentication must occur; and finally much of the observational data may be literally irreplaceable. Thus data management plays a much larger role in observational astronomy computing than theoretical computing. Canada is uniquely well positioned in observational data analysis in having the Canadian Astronomical Data Centre (CADC), an NRC operated and CSA supported national central repository for the distribution, archiving, and processing astronomical data from a wide variety of large surveys. In operation since 1986, CADC maintains world-class expertise in astronomical data management throughout its entire life cycle.

As with theoretical astrophysics, the nature of the computational challenges faced in observational data analysis can vary widely. Some observational projects must sift through huge amounts of data to extract observations from large amounts of noise, or – in the case of many radio astronomy projects, for instance, must be able to correlate all the data from a given signal at once to be able to extract any astronomical observations at all. (For an example of this sort of project, read the "CMB" Astro computing highlight). These projects require similar sorts of supercomputing resources as the large-scale simulations described above; many processors tied together by very high speed interconnects with very high-performance storage systems.

Other observational projects – such as many large optical "surveys" which carefully observe large fractions of the entire sky – do not require a global view of the data at once to produce results, so supercomputing resources are not required; however, these do require pipelines of tasks of varying computational intensity to be run on each observation. This can result in an enormous numbers of independent chains of amount of compute jobs, particularly in surveys which repeatedly observe the same portion of the sky to look for new events (eg, supernovae, or to detect planets moving in front of stars).

These requirements again require high-throughput computation to perform the large number of tasks necessary; in some cases the tasks may be again individually still quite complicated, requiring specialized advanced research computing hardware; in other cases, a broader **cloud computing** style of less specialized resources may be enough. The CANFAR project – a unique collaboration between CADC, Compute Canada, and CANARIE – is leading the world in the design and construction of an astronomical observational data analysis cloud computing environment combining both powerful computational analysis tools, access to both cloud and high throughput research computing resources, and access to the underlying astronomical data. For more information, see the "CANFAR/CADC" Astro computing highlight.

### Astroinformatics

The increasing volume of processed data that now exists – either "born digital" observations from new surveys or digitized archival data – is giving rise to a new field of astronomy, astronomical data science, or astroinformatics. This is "Big Data for Astronomy," where statistical and data-mining technologies and approaches are applied to extract meaning not about individual objects but in search of unsuspected connections and trends.

Canada is already very well positioned to be a leader in astroinformatics with CADC and CANFAR. However, taking advantages of the opportunities presented by astroinformatics will require computation resources architected somewhat differently to take advantage of Big Data workflows; the computational needs (and opportunities) of big data analytics weren't clear at the time of the CFI NPF. While such computational workloads can take advantage of supercomputing-class computers, much more emphasis must be given to having fast local storage and large amounts of memory per node than would typically be the case in a more traditional supercomputing environment.

## Strengths, Weaknesses; Opportunities, Threats

A great deal of Astronomy and Astrophysics research computing is being done on the existing Compute Canada systems, which represent in almost all respects significant improvements for scientists over the pre-2006 environment. We describe here the strengths, weaknesses, opportunities, and threats of the current environment for doing computation-powered astronomy research as compared with those of our international competitors.

### Strength: Diversity of Systems

As described above, Astronomy and Astrophysics research computing spans a wide range of different challenges, which collectively impose a variety of sometimes opposing requirements. The current diversity of Compute Canada systems is a much better match to the needs of our scientists than any "one size fits all" approach would be.

### Strength: Good Working Relationship with Providers

Many large-scale astronomical research projects require close cooperation with our computing resource providers – whether in tuning or debugging external network connections; making significant changes to how compute jobs are submitted as for CANFAR; or in providing input into the design of new systems. Similarly, training and support provided by existing providers has been extremely valuable for our scientists. While we would like to build still further on these relationships (see "Staffing" weakness below), the existing collaborations and working relationships are a strength of the existing system.

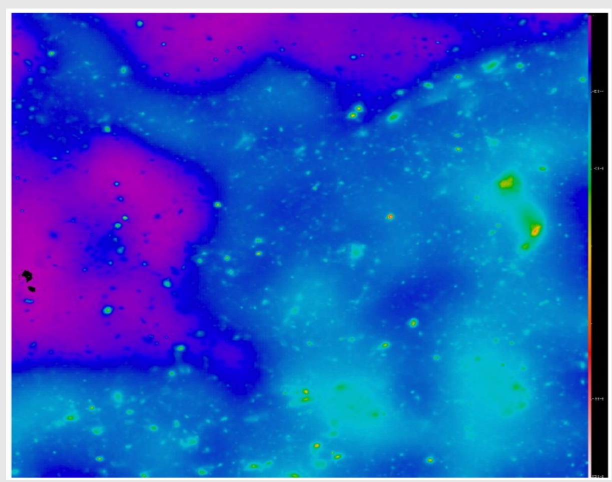
### Astro Computing Highlight: Cosmological Simulation

Simulations of the entire Cosmos – cosmological simulation – allows scientists to compare the outcome of a physical model of the Universe to both the present day and to observations of the much earlier Universe such as observed through the Cosmic Microwave Background or other distant, early objects.

These comparisons must necessarily be statistical – we don't expect a model to generate galaxies in exactly the same places in the simulation that they occur in the actual Universe, for instance – and to perform such statistical comparisons, extremely large numbers of astronomical bodies must be simulated. In addition, simulations must be performed which encompass simulation volumes which are significant fractions of the observable Universe, while also modeling small-scale processes and objects which effect the results the researchers wish to compare – for instance, to directly compare the brightnesses of modeled and observed galaxies, one must be able to simulate the star forming regions of individual galaxies, which might have one trillionth of the volume of the overall simulation domain.

These are then classic supercomputer-style simulations, requiring the largest computers available to perform world-class computations. Once done, the results of these simulations can be later mined for a wide number of results by the simulators or other researchers.

The very largest of these computations cannot currently be performed on Canadian computing resources, but many important simulations can. Ue-Li Pen, of the Canadian Institute for Theoretical Astrophysics, proposed to use the entirety of Canada's largest computer to perform a series of cosmological simulations. These simulations would be used to compare to observations currently being made by telescopes such as the Atacama Cosmology Telescope or the South Pole Telescope of the "kinetic Sunyaev Zel'dovich effect" – the glow given off by moving clusters of galaxies as they are backlit by the cosmic microwave background. These simulations would contain over half a trillion computational zones and would be directly comparable to observations.



*Right: CMB + kinetic Sunayev-Zel'dovich map of the nearby Universe, from simulations. Courtesy Ue-Li Pen*

### Weakness: Stagnant, Decaying Infrastructure

An increasingly urgent weakness is the lack of new hardware investment on the horizon. Compute Canada now has one system barely in the top 100 (#94) of supercomputers internationally, which strongly limits the very largest scale simulations and data analyses our scientists want to perform. Similarly, no new systems which would provide Big Data/astroinformatics capabilities are in the foreseeable future, no increase in storage commensurate with the recent increase in data volumes are available, and local and national networking remains a serious issue.

### Weakness: Staffing well below international norms

Rapid expansion of the scale of simulations and data analyses which need to be performed, combined with an ever-shifting landscape of computing technology makes exploiting the newest technologies extremely challenging – researchers are required to be experts at both the most cutting edge computing technologies as well as their own research areas. Having every research group in Canada independently generate this expertise is catastrophically inefficient.

The Compute Canada national platform is a significant improvement in this regard over previous, departmental- or University-level computing resources, which typically only had enough personnel to keep the systems running. At the consortia, Technical Analysts are, in principle, available to help researchers in their work. But the level of staffing is woefully insufficient to the need. As pointed out in the CASCA CDNC white paper to LRP2010[1], the UK computing facility HECToR which was at the time very similar in size (in terms of compute resources) to one of the Compute Canada consortia, had ten times the Technical Analyst support personnel; and in the case of HECToR, these staff were solely dedicated to very high-level consulting, support, and programming tasks or advanced training, whereas in Compute Canada consortia, these staff typically handle everything all the way down to low-level "help-desk" type questions involving basic login questions or passwords. (HECToR had junior technical staff for these sorts of queries).

HECToR is not alone in this, and in fact has relatively moderate levels of support; at the high end, in one leadership computing facilities in the United States, a rule of thumb for the level of advanced support is `one FTE per three supported projects'. With Compute Canada supporting approximately 200 competitively-allocated RAC proposals, this would require 66+ FTEs focused entirely on high-level consulting and support.

---

1 http://www.casca.ca/lrp2010/Docs/LRPReports/CDandN_WP.pdf

Canadian Astronomy and Astrophysics research stands to lose out if it cannot benefit from this level of support for advanced research computing that is becoming the international norm.

### Weakness: Lack of software development

The low number of dedicated advanced research computing support personnel means that each research group across Canada must develop the advanced research computing expertise to adapt or develop its own ARC-capable software. Not only is this woefully inefficient, it works poorly and is not supported by existing scientific funding mechanisms in Canada. This means that much less scientific software tool development is done in Canada than would otherwise be. Since this software is crucial to performing research on advanced research computing platforms, it puts Canadian astronomers and astrophysicists at a disadvantage. To examine the problems that they would like to study, they must either recuse themselves from doing science for years to develop their own software tools of possibly variable quality to use existing computing resources, or they must abandon the area they would like to study and instead peer under the already-crowded lampposts in fields where there is already well-established scientific software.

It is difficult to innovate when you are reliant on others to provide your tools. Thus a serious weakness of the existing system is, related to under-staffing, is a lack of support for new tool development.

Astronomy has a long tradition of funding both new facilities (large-scale equipment, like telescopes) and instrument-makers who build tools such as cameras, spectrographs, etc to provide tools to enable scientists to make use of the facility for discovery, in addition to funding the “operators” who run the facilities and help with immediate, help-desk like problems. We suggest that this would be a good model for advanced research computing in Canada, where both support staff and “tool-building” teams are funded.

### Weakness: Lack of Support for Long-Term Projects

The current allocation process of computing resources involve annual competitive grant applications and re-applications. This works well for a number of our researchers, but cause significant difficulties for others. Many of our scientists, both observational astronomers and theoretical astrophysicists, work as part of institutional projects based around new telescopes, satellites, or surveys. As a requirement of being part of these international projects, these researchers often have multi-year commitments to these collaborations to perform data analysis or supporting simulations. The tension between the annual process for being granted access to compute resources and the requirement to commit to several years worth of work to be part of an international collaboration puts our researchers in a difficult situation.

An extreme example of the misalignment between the yearly RAC allocation process and scientific commitments is the CANFAR project, described in an Astro Computing Highlight. Storing vitally important data which will be important for decades on storage allocations which are only year-to-year, or attempting to provide a value-added computing platform for observational data analysis on top of compute resources which may not be available next year places the project in an untenable position.

### Opportunity: Engaged, Active Research Community

Because of the good working relationships between scientific communities and existing research computing providers, the new CCI has the opportunity for a constructive engagement with active research communities and societies like CASCA which are responsible for representing the needs of entire scientific disciplines.

As a lead-up to the Mid-Term Review process for the Long Range Plan, CASCA's Compute, Data, and Networks Committee will embark in at the beginning of 2013 on a survey of the Canadian Astronomy and Astrophysics community to produce a report in the first half of the year on the computing, data, and networking needs of the community over the next 3-5 years. This form of domain expertise on the needs of a community could be of use to CCI in its planning process, and a better understanding of CCIs priorities and constraints could aid CASCA in its own planning and assessment of needs.

A formal working relationship could be established between CCI and CASCA, and other such societies who perform regular assessments of the computational needs of their researchers, by having representation from CASCA's CDNC as part of CCI's research advisory bodies. Such representing, charged with representing the needs of an entire community to CCI (and the priorities and constraints of CCI back to the community), would provide a productive way for these two organizations to work together.

### Threat: Lack of support means Canada loses competitiveness in Astronomy research

The most urgent threat our members perceive to research computing for Astronomy and Astrophysics in Canada is the real danger of status quo – that no existing substantial new funding is secured, that existing hardware continues to decay or only minor new upgrades are made, inadequate to handle the upcoming data volumes and requirements for simulation; that no increase in staffing occurs, meaning that many important research projects that could be undertaken simply aren't. In such a situation, Canada will squander its advantages in several areas of astronomy and astrophysics and lose attractiveness as a place to do cutting edge research, eventually losing its position of leadership in Astronomy and Astrophysics.

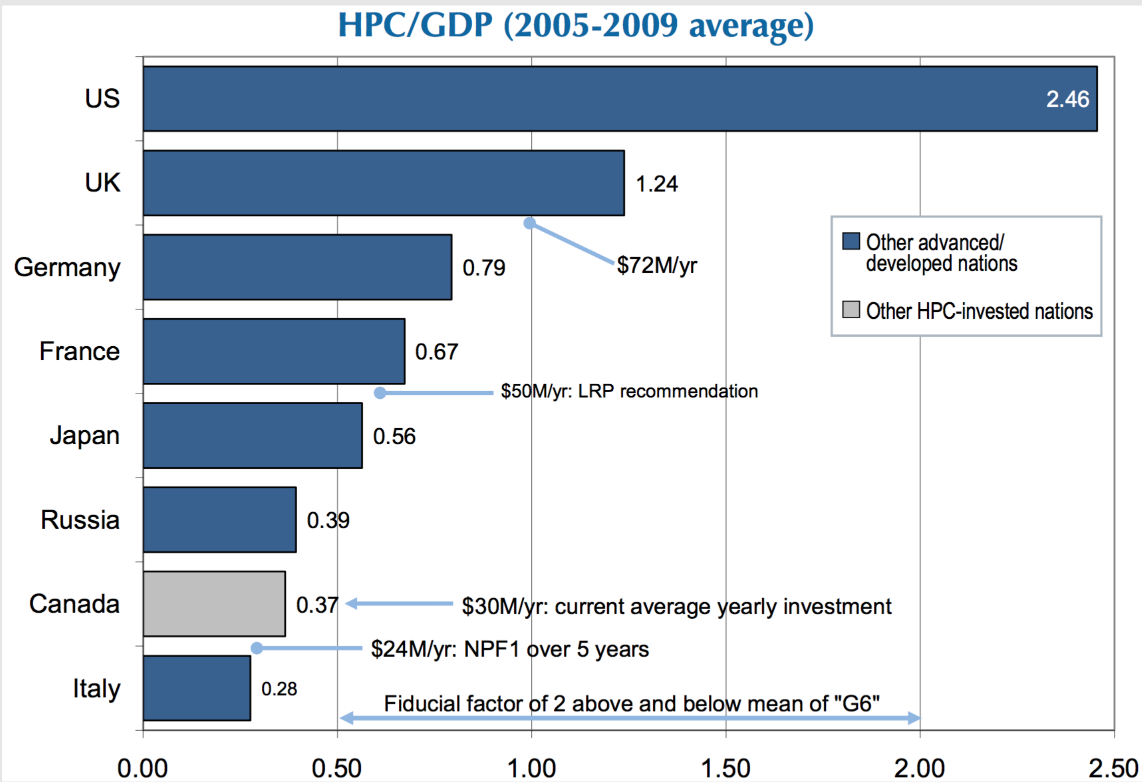



Figure 2: International Investment in advanced research computing relative to GDP, normalized to the mean of the "G6" (G7 less Canada) nations that are our primary international research competitors. The situation has only deteriorated for Canada since 2009. LRP2010 recommended a substantial increase in ARC funding, to $50M/yr, which would have remained significantly below the G6 average. *Image from CASCA LRP2010*

These concerns do not come from nowhere; many of our researchers have been present for several cycles of investment in Canadian advanced research computing, subsequent optimism, and then near-total abandonment of the effort; from the never-followed-up-upon FERUT at UofT; the Cyber 205 installed in Calgary in 1985, only to be taken down in 1991; the "Ontario Centre for Large Scale Computing" at UofT with a computer installed in 1988, only to have the entire centre disbanded in 1992; a Fujitsu installed in Calgary by 1994 with a combined commercial and research mandate, which was gone by 1996[1].

1 e.g., "High Performance Computing in Canada: The Early Chapters", Physics in Canada 64:2 85-88 Apr 2008

Only once in recent memory has Canada had a "top-20" scale machine for academic research; it was commissioned in 2009, and even after a 2011 upgrade, that machine has already slid down to 94th and is rapidly decaying as its maintenance contract has expired. There are no new prospects on the horizon. Only Environment Canada has had a sustained advanced computing presence, and it expressly has little or no research or development component: it largely runs software developed elsewhere.

In the present day, we have had a high-water mark at the time of the NSERC MFA (1996) through the first CFI national platform grant where Canadian advanced research computing investment achieved a factor of three *below* average G7 levels of HPC investment per GDP. But without new investment, as machines purchased with those funds begin to shut down, Canadian researchers increasingly have nowhere else to go.

With this history of abortive attempts at advanced research computing platform development, it has only been though determination and luck that Canadian Astronomy and Astrophysics has managed to maintain the advanced research computing capabilities that it does have. Even still, researchers in other countries are enormously advantaged in access to both computing resources, expertise, and funding for tool-building. If Canadian advanced research computing investment backslides again, Canada will become a significantly less attractive place to perform world-class research.

## Recommendations

Canadian researchers, due to Compute Canada, have experienced 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 for expanding these resources over the coming years.

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.

### Astro Computing Highlight: CANFAR/CADC

CANFAR is a unique project run out of the Canadian Astronomical Data Centre (CADC), which combines open standard cloud software technologies, Compute Canada compute resources, CANARIE networking, and CADCs astronomical data analysis and management expertise to build an astronomical data analysis computing platform for the Canadian astronomical community.

Large surveys of the sky have been a key part of Canadian success in astronomy, but new surveys with high data rates pose both computational and data-management challenges. Not only must enormous amounts of data be processed using complex computational pipelines, but with large amounts of data and large collaborations, even simple data management questions become complex – what is the most recent version of the data, and where is it?

CANFAR provides a working cloud-style computing environment that enables the effective delivery, processing, storage, analysis and distribution of very large data sets produced by astronomical surveys. It provides a survey-agnostic cloud environment that allows astronomers to effectively process the massive amounts of data generated by highly sophisticated and complex observatories using their own custom software. In the past year, it has distributed 805 terabytes of data, and provided over three million processor-hours of data analysis to the Canadian astronomy community.

Having proven itself on existing data sets from CADC projects – all the data coming from Canadian Telescopes as well as Canada's contribution to Hubble Space Telescope processed storage distribution now are now stored on CANFAR resources – CANFAR is now being leveraged to buy Canada into major international projects. Pan-STARRS, a huge new survey being performed with University of Hawaii-owned telescopes, will generate 20 petabytes (over 20,000,000 gigabytes) of data, and is a fascinating project in its own right. But in addition, it provides very important complementary data to a European Space Agency satellite project scheduled for 2015, Euclid. CANFAR is proposing to provide data services for Pan-STARRS, and use that as a way to work with the Euclid project, enabling roughly 20 Canadian astronomers to participate in this expensive satellite collaboration who would not be able to otherwise.

CANFAR's success builds on the availability of Compute Canada resources, but the lack of any long-term storage availability, and having to fit the long-term goals of the project into the year-to-year allocation schedule of Compute Canada, have posed significant challenges.

*Figure right: Image from the Next Generation Virgo Survey. Courtesy CANFAR/CADC.*

Given the above SWOT analysis, it is the position of the CASCA CDNC that research computing in Canada would benefit from:

### Urgent search for new, sustainable funding

Without new, increased, and sustainable funding for advanced research computing investment, all else is moot. Canada has arguably never adequately funded advanced research computing, which has been problematic in the past; but in the current digital era of Big Data, increasing data volumes, larger simulations, and data-driven science, this policy choice verges on the catastrophic. New funding streams must be found.

### Broad focus on entire research computing environment

A broader focus on the advanced research computing environment is necessary than just counting computing hardware – in particular, an increased focus on personnel, software development, and storage is necessary to maximize the scientific results from advanced research computing investment. CASCA CDNC believes that converging international norms provide appropriate guidance as to the balances to strike.

### Maintaining, extending diversity of systems

Canadian astronomical and astrophysical research benefits from the existing diverse range of computing systems available through Compute Canada. Indeed, CASCA scientists would benefit from still more diversity – one or more big data systems, a more explicitly commodity cloud system, and systems with FPGAs and Intel Phi-based acceleration would all be useful additions.

### Continued close working relationship between providers and researchers

The existing close working relationships between research groups and computing providers has been invaluable, both from the point of view of flexibility in use of resources and in advanced training, education, and consulting/collaboration. Having local staff and resources has been an integral part of this working relationship and it is CASCA CDNC's position that these keeping these relationships both figuratively and literally "close" is extremely important to the work of our scientists.

The close relationship extends to diversity of systems, above; having local systems that can be experimented with in collaboration with the providers, such as to modify the submission process to allow cloud-type submission of compute tasks, or to experiment on small scales with accelerator hardware or changes to networking, has been extremely valuable.

## Enabling long-term projects

The current one-year application process for RAC allocations of Compute Canada resources is at odds with the scientific schedules to which many Canadian astronomers and astrophysicists must commit. CASCA CDNC would urge CCI to consider enabling multi-year commitments of resources so that our scientists can fully participate in international collaborations around multi-year projects, and so that projects like CANFAR can build value-added platforms on top of Compute Canada computing resources.

## Focused input from disciplinary societies

CASCA represents the needs of the Canadian astronomical research community. Since 2000 we have held held regular community consultations to lay out long-term plans for research priorities, and we have consistently recognized the crucial role that research computing plays in current and future astronomical research. We believe that we can usefully contribute to ongoing planning and prioritization that CCI will necessarily undertake.

It is the position of the CASCA Computing, Data, and Networks committee that a representative from CASCA, charged with representing the needs and views of the Canadian astronomical community, should be a member of the proposed CCI Research Advisory Committee, and that other disciplinary societies should likewise have representation. We respectfully suggest that this approach is more likely to be productive than the alternative of a committee consisting solely of individual researchers with no broader mandate or responsibility.

CASCA represents the needs of the Canadian astronomical research community. Since 2000 we have held held regular community consultations to lay out long-term plans for research priorities, and we have consistently recognized the crucial role that research computing plays in current and future astronomical research. Our 2013 process for assessing in quantitative detail the needs of our community will be a useful input to CCI, and we believe that we can usefully contribute to ongoing planning and prioritization that CCI will necessarily undertake.

It is the position of the CASCA Computing, Data, and Networks committee that a representative from CASCA, charged with representing the needs and views of the Canadian astronomical community, should be a member of the proposed CCI Research Advisory Committee, and that other disciplinary societies should likewise have representation. We respectfully suggest that this approach is more likely to be productive than the alternative of a committee consisting solely of individual researchers with no broader mandate or responsibility.

We believe that these recommendations, if implemented, would greatly improve the breadth, quantity, and quality of the computational work that could be done by the Canadian Astronomy and Astrophysics community over the coming five years. We look forward to discussions with Compute Canada on how to work together to address shared concerns.