The Role of High-Performance Computing in Subatomic Physics

Submitted to the SAP Long-Range Planning Committee
December 2005

1 Introduction

This document is the report of the sub-committee on computing for the 2005 long-range planning exercise for Canadian subatomic physics. The sub-committee surveyed the community to ascertain its computing needs in the medium to long term future (see appendix B). There were relatively few responses to the survey so this report can be short.

The general conclusion of the sub-committee is that most of the computing needs can be met by the high-performance computing (HPC) consortia, which are large-scale multi-disciplinary facilities in the universities, funded mostly by CFI. The notable exception is ATLAS, which will require a dedicated Tier-1 facility [1] at TRIUMF and significant resources in the context of the HPC consortia. This will be extended to future involvement by Canadians in the ILC. A further possibility exists for a dedicated Canadian facility for lattice QCD in the scenario of significantly increased funding to the SAP envelope.

2 Computers in Subatomic Physics

The use of computers for data analysis, simulation of experiments, and sophisticated theory calculations is now essential in subatomic physics. Many projects would not be possible without significant computing resources. In fact, the large-scale storage needed to archive the raw data, as well as the resources, both processors and storage, to do the first-pass analysis of these data and store the resulting data set, should be considered as an extension of the experiment. In some cases these resources are on a scale that makes it impractical to use shared facilities and dedicated computing centres must be built.

Typically, raw data are archived at the experiment site where a preliminary analysis is done using the best knowledge currently available for calibration. The results of the first-pass analysis are usually sent to remote centres for

[1] An ATLAS Tier-1 centre (10 world-wide) is responsible for storing raw data and reprocessing the data 2-3 times per year.

backup and for further analysis. The size of the raw data sets is in the Petabyte range for the larger experiments.

The raw data are reprocessed when better calibration constants and improved analysis algorithms are available. This must be done at the remote centres because the main site is busy coping with the data currently streaming off the detector. The remote centres must therefore also have significant processing power.

The tasks described above require dedicated resources to deal with the data, which come from the experiment essentially continuously. There would be little point in using shared facilities, especially for storage. However, the subsequent steps in the analysis, where specific physics groups or even individual physicists start to access the data are done in a less continuous way. It is therefore possible to share resources with other experiments or with other disciplines at this point.

We are fortunate in Canada to be the beneficiaries of significant investment recently in high-performance computing. These investments, made mostly in the context of CFI, have established so-called HPC consortia across the country. The SAP community has made good use of these resources.

It is important to realize that both the dedicated centres and the shared facilities are crucial to the full analysis chain in particle physics. This division of tasks makes the most efficient use overall of the computing resources.

An example of a smaller experiment where it has been possible to do the whole analysis chain at a shared facility is the TWIST experiment at TRIUMF. TWIST has used the WestGrid facilities at UBC/TRIUMF and Simon Fraser to store and analyze their data, as well as produce the Monte Carlo data sets needed for the analysis.

No mention has been made to date of the computing requirements for theory. While most of the needs of this community can be met by the HPC consortia, there are also theory projects that use dedicated computing facilities. A good example of this is Lattice QCD. Several collaborations have built dedicated compute farms in the US, Europe, and Japan to do their calculations that require extensive CPU time. There is also a relatively modest farm (260 processors) at the University of Regina, for lattice QCD calculations. Canadian lattice practitioners presently rely on dedicated facilities in other countries to sustain their larger-scale computing needs.

3 Requirements for Dedicated Facilities

To set the scale for the dedicated facilities, the numbers for ATLAS-Canada are given below. These are cumulative to 2009. Resource requirements for future years will continue to increase.

- TRIUMF Centre (ATLAS Tier-1):	CPU:	3,425 kSI2k
	Disk:	1,660 TB
	Tape:	1,335 TB
- Tier-2 centres:	CPU:	1,900 kSI2k
	Disk:	1,000 TB

(One 3 GHz cpu corresponds to about 1.1 kSI2k)
(Note that there will be two Tier-2 centres, one in the East and one in the West, and that each of these will be distributed geographically. The total requirements are shown. Note also that the size of the Tier-2 centres is still under debate. These numbers should be considered as lower bounds.)

The estimated cost of the TRIUMF computing centre is $21.6M (using educational pricing), including personnel. The real costs, after vendor discounts, are $13.6M. The estimated cost of both Tier-2 centres after vendor discount is $7.5M. However, this does not include the infrastructure upgrades that would be necessary in the universities to house the hardware. This also includes only a minimum complement of personnel, with little in the way of user support at the Tier-2 centres, so that the actual cost of these centres will be higher.

4 Further Important Points

Two important points should also be made in the LRP report. It is crucial that the networking infrastructure provided in Canada by CANARIE be maintained. The SAP community should write a letter of support when CANARIE is up for funding renewal in 2007. A statement of support should also be made for the HEPNET grant from NSERC. HEPNET coordinates networking issues in Canada and abroad for the Canadian SAP community. In particular, HEPNET negotiates access for Canadians to trans-Atlantic links to CERN and DESY (for example the recent dedicated link to CERN for ATLAS through CANARIE). This service is invaluable.

Furthermore, it is critical that funds for support personnel also be available. Many groups fund such people from their NSERC GSC19 MFA grants. The SAP community also benefits from personnel funded through the MFA grant to C3.ca, the equivalent to IPP for high-performance computing. This category of funding should at least be maintained and is in fact a good candidate to go up in the increased funding scenario.

The sub-committee also notes the growing importance of Grid computing. The scale of the computing requirements for many experiments dictates the use of geographically distributed computing and storage resources. Even if the requirements are such that a project fits into the HPC consortia, these resources will be "gridded" in the near future. It is likely that most large-scale serial type computing will be done on the Grid in the near future.

Finally, note that, with a few exceptions, major funding for large-scale computing for subatomic physics in Canada comes from outside the SAP envelope. However, if we are less successful in the future, GSC-19 should be aware that there could be significant demands on the envelope for computing.

5 Conclusions/Recommendatons

1. Computing is an integral part of many if not most experiments in subatomic physics, and it is playing an increasing role in theory. The LRP report should recognize this and give strong support to funding initiatives for hardware and personnel. Otherwise the community will not have the tools necessary to exploit the considerable previous investments that have been made.

2. Most computing for subatomic physics can be done in the HPC consortia. The LRP report should give strong support to the upcoming CFI request(s) to the National Platforms Fund.

3. ATLAS and the ILC will require dedicated centres. The cost of the ATLAS computing, integrated out to 2009, is estimated at \$13.6M after vendor discounts for the TRIUMF Tier-1 centre, and \$7.5M for both university centres. We can expect continuing costs on the order of \$6-7M/yr after that. All these estimates are out to fiscal 2009 and include personnel and infrastructure costs for the Tier-1 centre.

4. Lattice QCD relies on dedicated facilities. In the scenario of significantly increased funding to the SAP envelope, the community should consider funding for a dedicated facility in Canada for lattice QCD.
5. The LRP report should strongly support continuing funding for CANARIE and HEPNET-Canada.
6. Support should be expanded for computing support personnel in the universities.
7. Deployment of Grid tools in Canada should be encourared and supported.

A Committee membership

- Alain Bellerive, Carleton University
- Randy Lewis, University of Regina
- Randy Sobie, University of Victoria and IPP
- Michel Vetterli, Simon Fraser University and TRIUMF (chair)
- Andreas Warburton, McGill University
- Rob McPherson, University of Victoria and IPP (ex officio)

B Community Survey

Computing Survey for the SAP Long Range Plan - 2005

Dear colleague,

A sub-committee has been formed to prepare a brief to the 2005 Subatomic Physics Long Range Planning Committee concerning computing. As you know, the use of computers for data analysis, simulation of experiments, and sophisticated theory calculations is now essential in our field. Many projects would in fact not be possible without significant computing resources. With this in mind, we are surveying the community to find out what the needs will be for the next 5 years, with an eye out to 10 years.

We are fortunate in Canada to be the beneficiaries of significant investment recently in high-performance computing (HPC). These investments, made mostly in the context of CFI have extablished so-called HPC consortia across the country. The SAP community has made good use of these resources. However, it is clear that some projects require dedicated CPUs and storage, and that this is not compatible with the use of the HPC consortia. The model we are putting forward is that this large-scale dedicated computing, such as first pass analysis of raw data for large experiments, should be done in dedicated facilities, while the HPC consortia can be used for subsequent analysis to extract physics results.

Within this context, we ask that you fill out the following brief survey. We request that the responsible person for computing at each major SAP experiment, major experimental facility, or theory project involving Canadians respond to this request. If you feel you fall outside of this group or if you have general comments about computing for SAP in Canada, please respond as an individual.

With many thanks,

Mike Vetterli; for the SAP Computing Sub-committee

****

Please return by Wednesday, October 19th

- Experiment/Facility/Project:
- Respondant:
- Number of people in the group (faculty/post-docs/students):
- Do you use HPC extensively? If not, you have completed this survey.
- Briefly describe what tasks are performed by computer. In particular, do you see a need for a dedicated (24/7/365) facility for your project?
- What facilities do you currently use?
- How will this evolve over the next 5 years? 10 years?
- Do you currently support computing personnel? From what source (MFA)? How should this evolve with time?

- If possible, give an estimate of the CPU and storage resources you will need over the next 5-10 years.
- How do you see these resources split between the HPC consortia and dedicated facilities?
- Do you have any general comments on computing for SAP in Canada? For example, the community has been asked to comment on what would be possible if there was substantially more money available. We are interested in any ideas you might have.