The ATLAS-Canada Network

I. Gable$^{3,10}$, M. Bedinelli$^9$, S. Butterworth$^9$, B. Caron$^{7,8}$, R. Chambers$^8$, B. Fitzgerald$^{10}$, L. Groer$^9$, R. Hatem$^2$, V. Kupchinsky$^9$, P. Maddalena$^5$, P. Marshall$^2$, S. McDonald$^7$, P. Mercure$^6$, D. McWilliam$^1$, C. Payne$^7$, D. Pobric$^2$, S. Robertson$^5$, M. Rochefort$^5$, M. Siegert$^6$, R.J. Sobie$^{5,4,10}$, R. Tafirout$^7$, T. Tam$^2$, B. Vachon$^5$, A. Warburton$^5$, and G. Wu$^9$

$^1$BCNet, Vancouver, BC, Canada
$^2$CANARIE Inc., Ottawa, ON, Canada
$^3$HEPnet/Canada, Victoria, BC, Canada
$^4$Institute of Particle Physics, Canada
$^5$McGill University, Montreal, QC, Canada
$^6$Simon Fraser University, Burnaby, BC, Canada
$^7$TRIUMF Laboratory, Vancouver, BC, Canada
$^8$University of Alberta, Edmonton, AB, Canada
$^9$University of Toronto, Toronto, ON, Canada
$^{10}$University of Victoria, Victoria, BC, Canada

E-mail: igable@uvic.ca

Abstract. The ATLAS-Canada computing model consists of a WLCG Tier-1 computing centre located at the TRIUMF Laboratory in Vancouver, Canada, and two distributed Tier-2 computing centres in eastern and western Canadian universities. The TRIUMF Tier-1 is connected to the CERN Tier-0 via a 10G dedicated circuit provided by CANARIE. The Canadian institutions hosting Tier-2 facilities are connected to TRIUMF via 1G lightpaths, and routing between Tier-2s occurs through TRIUMF. This paper discusses the architecture of the ATLAS-Canada network, the challenges of building the network, and the future plans.

1. Introduction
The Large Hadron Colider (LHC) at the CERN Laboratory in Geneva, Switzerland will begin colliding protons in 2008. These collisions will be recorded by four large detectors. Canada, in particular, has focused its efforts on the ATLAS detector. ATLAS, like the other detectors, will produce a volume of data (many 10’s of petabytes) that cannot be stored or analyzed entirely at one institution [1]. Instead the data will be distributed around the world to a hierarchy of computing centres. In order to effectively use these centres, a high performance network is required.

The ATLAS Computing model [1] was developed as a logical mechanism for dividing up the roles of the institutions involved with the detector. The Computing Model employs a multi-tiered approach to dividing the data among the institutions. CERN will host the Tier-0 (T0) facility which will archive the raw data and provide resources for processing the raw data. The raw and processed data will also be copied to Tier-1 (T1) facilities. The T1’s will provide access
Table 1. Distribution of ATLAS T1 computing resources. The last column indicates the fraction of the total resources provided by that site according to the ATLAS computing model.

| Tier-1 Institution | Location                          | share of computing resources (%) |
|--------------------|-----------------------------------|----------------------------------|
| ASGC               | Taipei Taiwan                     | 7.7                              |
| BNL                | Brookhaven NY USA                 | 24.0                             |
| CNAF               | Bologna Italy                     | 7.5                              |
| FZK                | Karlsruhe Germany                 | 10.5                             |
| IN2P3              | Lyon France                       | 13.5                             |
| NDFG               | Scandinavia                       | 5.5                              |
| PIC                | Barcelona Spain                   | 5.5                              |
| RAL                | Didcot United Kingdom             | 7.5                              |
| NIKHEF/SARA        | Amsterdam Netherlands             | 13.0                             |
| TRIUMF             | Vancouver BC Canada               | 5.3                              |

to the data as well as analysis facilities. Tier-2 (T2) facilities provide additional simulation capacity as well as storage for derived data sets.

There are ten T1 computing facilities distributed globally as listed in Table 1. The data flows from the T0 to the T1 facilities via dedicated fibre optic links with a guaranteed high-speed (5-10 Gbps) bandwidth. T2 facilities connect to their regional T1 via a routed or dedicated network connection. In addition, T1 facilities are expected to have a secondary network path to the T0 in case of problems and may have dedicated links to other T1 sites around the globe. The T2 sites normally only transfer data with its regional T1.

The network requirements are outlined in the ATLAS Computing Model Requirements [1]. It is expected that 1600 Mbps of data will move on continuous basis from the CERN T0 to an average-size T1 for the first few years of LHC. In addition, inter-T1 aggregate data transfer rates are expected to be in excess of 1 Gbps.

The network traffic between the T1 and T2 sites is expected to be lower than the T0-T1 and T1-T1 rates. The T2 sites will not host the large samples of reconstructed data but only selected samples. In addition, the T2 sites will send the locally generated simulated data samples to their regional T1 but these rates are expected to be relatively modest. It is expected that a 0.3-1 Gbps dedicated network link between a T1 and T2 site would meet the network requirements.

This paper reviews the network situation in Canada. It describes the current research network infrastructure provided by CANARIE and discusses the dedicated link to the CERN T0 in Geneva from the TRIUMF T1 in Vancouver. The network between the TRIUMF T1 and the Canadian T2 sites is also discussed.

2. The CANARIE Research Network

CANARIE is Canada’s national research and education network organization. The CANARIE network is Canada’s national research and education network infrastructure, designed and operated by CANARIE through funding provided by the Government of Canada.

In 2006 and 2007, CANARIE significantly added to its capacity by deploying a Reconfigurable Optical Add-Drop Multiplexer (ROADM) Dense Wavelength Division Multiplexing (DWDM) network. The Eastern Canadian ROADM network spans Chicago to New York with break-outs in Windsor-Detroit, Toronto, Ottawa, Montreal and Boston. The Windsor-Detroit to Ottawa span was done in partnership with the Optical Regional Advanced Network of Ontario (ORANO). The Eastern Canadian ROADM Network, can support up to 72 x 10 Gbps wavelengths, framed in either 10 GE or OC-192 SONET, and will support 40 Gbps wavelengths
Based on the same technology, the Western ROADM network was done in partnership with BCNET and Netera (British Columbia and Alberta Optical Regional Advanced Networks, respectively). This network spans Seattle to Calgary with break-outs in Victoria, Vancouver and the University of British Columbia/TRIUMF, Kamloops, Kelowna, Calgary and the University of Calgary. Both the Eastern and Western Canadian ROADM networks are depicted in blue in Figure 1.

In addition to the ROADM capacity, CANARIE leases five 10 Gbps wavelengths traversing the country to fill the gap between two ROADM networks. These wavelengths also provide redundancy and connectivity to eastern Canadian regional networks. The 10 Gbps wavelengths are terminating on optical/electrical/optical (OEO) switching gear located at the major service exchange points, as identified in Figure 1. The network infrastructure interconnects with regional networks across the country and with international peer networks at the major access points. The Pacific Northwest GigaPoP, StarLight and MAN LAN facilities in Seattle, Chicago and New York, respectively, serve as international access points.

3. The LHC Optical Private Network
The only technology capable of meeting the requirements in terms of bandwidth for an average Tier-1 institution is the 10 Gbps lightpath circuit. A lightpath is defined as point to point optical circuit based on wavelength division or time division multiplexing that has a deterministic behavior and guaranteed bandwidth.
To meet the demanding network requirements for the LHC, the LHC Optical Private Network (LHCOPN) Project was established. A detailed review of the LHCOPN is available from LHC OPN architecture document [3].

The LHCOPN is a set of 10 Gbps lightpaths connecting each of the T1s to T0. This arrangement results in roughly the star pattern network illustrated in Figure 2. Data must also move between Tier-1s which can be accomplished through either direct lightpath or via routing through CERN.

4. TRIUMF to CERN

The TRIUMF link to CERN is an effort between many organizations that is coordinated by HEPnet/Canada. HEPnet/Canada is responsible for ensuring national and international networks for particle physics in Canada. In 2006 HEPnet/Canada, TRIUMF and CANARIE agreed to collaborate on the link to CERN as well as provide national connections to the Canadian T2 sites.

The TRIUMF T1 to CERN T0 circuit is depicted in Figure 3 runs over the CANARIE CA*net4 infrastructure until it disembarks North America in New York City. Each T1 site must use a small or series of small publicly routable Classless Inter-Domain Routing (CIDR) blocks as only traffic from the Internet Protocol IP address space is allowed to flow over the LHCOPN. Exchange of routing information is performed using Border Gateway Protocol (BGP) at the T1 and T0 institutions. The 10 GE circuit carrying a 5 Gbps lightpath transits Canada east to west, first traveling over BCNet network from TRIUMF to the CANARIE GigaPOP in downtown Vancouver. The circuit continues along CANARIEs network, and debarks North America at the MANLAN transit exchange in New York City. The circuit enters Europe on SURFnet in Amsterdam and then transits Geant2 network before arriving at CERN. The TRIUMF T1 will

**Figure 2.** The logical network diagram of the LHCOPN.
Figure 3. Network diagram of the TRIUMF to CERN 5G Circuit. Note the redundant path within north America for the 1 Gbps backup link (shown in green).

Figure 4. Traffic over the 5 Gbps link between TRIUMF and CERN stating from when the link was first established (February 2006).
Figure 5. Here we see traffic failing over to the backup 1Gbps lightpath (orange). Traffic is reduced to the capacity of the 1G link until the 5G lightpath is restored.

hold 5.3% of the ATLAS data, which is typically half the size of other Tier-1. Hence a 5 Gbps link is expected to be sufficient for the first few years of LHC operation. However should the demand for capacity increase CANARIE is able to provide the complete 10 GE circuit.

In addition, a redundant 1G path to CERN has been established and tested. The route proceeds via Seattle and Chicago before arriving in Amsterdam. The Amsterdam to Geneva connection is the same as the primary link.

A tertiary 1 GE backup path to Brookhaven National Lab (BNL), another T1 centre, is under development. BNL has agreed to carry TRIUMF traffic to CERN in the event of a failure of our primary and backup connections. Correspondingly, in the event of a failure of BNL’s primary backup routes TRIUMF has agreed to carry BNL traffic to CERN. However, for BNL the link to TRIUMF is only a last resort backup because of BNL’s larger demands.

Traffic over the TRIUMF to CERN link since the links inception can be seen in Figure 4. Traffic as of fall 2007 has remained well within the available 5 Gbps capacity. Figure 5 shows a failover event to the backup 1 Gbps link during a planned maintenance period on the 5 Gbps lightpath.

5. The T1-T2 Network in Canada

In addition to the ATLAS T1 facility in TRIUMF, Canada hosts a number of T2 facilities listed in Table 5. The distance between TRIUMF and the site is listed to highlight the long-distance nature of the Canadian network.

T1 to T2 connections in Canada follow the successful LHCOPN model with TRIUMF at the centre of a star pattern. Each of the current 5 Tier-2 institution, the University of Victoria, Simon Fraser University, University of Alberta, University of Toronto, and McGill University
Figure 6. The logical Network Diagram of the ATLAS-Canada Network. The routed networks are marked in green. For T1 to T2 communication failover of lightpaths occurs to the CANARIE research network. T0 to T1 connection fails over to a dedicated 1 Gbps backup connection and if necessary, a BNL lightpath or the regular research network.

| Canadian Tier-2 Institution | Location            | Shortest Distance to TRIUMF (km) |
|-----------------------------|---------------------|----------------------------------|
| Simon Fraser University     | Burnaby, BC         | 23                               |
| University of Victoria      | Victoria, BC        | 87                               |
| University of Alberta       | Edmonton, AB        | 826                              |
| University of Toronto       | Toronto, ON         | 3361                             |
| McGill University           | Montreal, QC        | 3701                             |

Table 2. The T2 institutions in Canada as of fall 2007 and their physical distance to TRIUMF.

have a 1 Gbps lightpath. Currently (fall 2007) Simon Fraser University’s lightpath is under development and the University of Alberta link has not entered a production mode of operation.

In Canada a T2 to T1 connection either transits an Optical Regional Advanced Network (ORAN) before arriving at a CANARIE Gigabit Point of Presence (GigaPOP) to be carried to TRIUMF or connects directly to CANARIE GigaPOP. The largest challenge for the Canadian T2’s has been establishing a connection with either an ORAN GigaPOP or CANARIE GigaPOP. In several cases the T2 must employ Coarse Wavelength Division Multiplexing (CWDM) over existing fibre to arrive at an appropriate POP. This equipment can be expensive and pose a serious financial hurdle for a T2. In addition not all T2’s have routing equipment capable of supporting Border Gateway Protocol (BGP) which is requirement for an ATLAS T2.

For example, the lightpath from the University of Victoria arrives via BCNet to the Vancouver Transit Exchange (VanTX) whereas the University of Toronto and Alberta reach the VanTX
via the CANARIE CA*net4 and the new ROADM POP, respectively. The T2 links from the VanTX to TRIUMF uses MRV CWDM optical equipment.

The Vancouver network infrastructure is particularly important to the ATLAS-Canada network because all incoming lightpaths must transit the city before arriving at TRIUMF. Most important for the network is the CANARIE Western ROADM network POP at the University of British Columbia and the BCnet Transit Exchange (VanTX) in Vancouver downtown.

The logical network diagram can be seen in Figure 6. In the event of a failure of a T2 to T1 traffic will failover to the routed research network connections. Though each T2 site has 1 Gbps routed network connectivity, the routed research network connections do not guarantee bandwidth but are assumed to be sufficient during short term outages.

6. Conclusion
The ATLAS-Canada Network has been designed and commissioned to meet Canada’s commitment to the LHC Computing. The link from the TRIUMF T1 to the CERN T0 has been operational since the beginning of 2007 and has participated in the WLCG and LHCOPN Service Challenges. The T1 to T2 network follows the successful LHCOPN model of using BGP to exchange routing information between sites with a limited CIDR address blocks being used at each T2. It is anticipated that the network demands will grow as the LHC starts recording data in 2008. The current network design is scalable and should meet the needs of ATLAS-Canada over the next 5-10 years.

7. Acknowledgments
The financial support of NSERC is acknowledged. The strong support of CANARIE and the provision of the transatlantic link has been critical to the success of the network infrastructure. The link from Amsterdam to CERN via Geant2 and SURFnet is also due to CANARIE’s key role in the project. The key contributions of the provincial network organization (BCNet, ORANO and Netera) are acknowledged.

8. References
[1] ATLAS Computing Group, 2005 ATLAS Computing Technical Design Review CERN-LHCC-2005-022.
[2] Renshall H, ATLAS Tier 1 Resource Requirements Timetable for 2006/2007, https://twiki.cern.ch/twiki/bin/view/LCG/AtlasPlans
[3] Boe E, Martelli E, ed. Foster D, LHC Tier-0 to Tier-1 High-Level Network Architecture http://lcg.web.cern.ch/lcg/activities/networking/LHC\%20networking\%20v2.dgf.doc
[4] Martelli E, 2007 LHC Optical Private Network Status https://twiki.cern.ch/twiki/pub/LHCOPN/ImplementationDetails/map-lhcopn.png