Canadian Networks for Particle Physics Research

2012 Report to the Standing Committee on Interregional Connectivity, ICFA Panel
January 2013

This report describes the status and plans of the Canadian network infrastructure used for particle physics research in 2012 HEPnet/Canada (http://hepnetcanada.ca) coordinates the networks[1] and works with the network providers and particle physics laboratories. We describe the status of the CANARIE network infrastructure, the ATLAS Tier 1 and 2 centres in Canada, and highlight other network projects in 2012.

CANARIE Network

The CANARIE Network is among the world's most advanced national research and education networks. CANARIE is a high performance hybrid network that supports data-intensive research and innovation across a range of private and public sector users. This hybrid network infrastructure, equipped with leading-edge optical and routing equipment, enables CANARIE to offer traditional IP network services and Lightpath services (dedicated end-to-end connections) while continuing to develop new network service offerings.

With the hybrid infrastructure, the CANARIE network offers greater flexibility in service offerings in order to meet the changing needs of users including multiple Gigabit connections that are needed by users who require network capacity to support very data-intensive applications, such as High Energy Physics applications.

CANARIE Optical Infrastructure

The CANARIE optical infrastructure is comprised of two optical technologies: SONET and DWDM systems. The lower optical layer is the DWDM layer, which was built using ROADM technology supporting wavelength services up to 100Gbps. In 2006 and 2007, CANARIE completed two ROADM builds: the western-build and the eastern-build. CANARIE continued its dark fibre build in 2011 and 2012. The DWDM system extension from Calgary to Edmonton was completed in March 2012. At the end of 2012, the standalone Winnipeg – Thunder Bay segment was built, paving the way to link the western system to the eastern system.

CANARIE will continue its dark fibre system expansion in the next two or three years. Its vision is to complete coast-to-coast fibre systems from Victoria across Canada all the way to Halifax. The nation-wide fibre system allows CANARIE to deliver high-bandwidth services to data intensive users, high performance computing centers and research facilities. CANARIE started the fibre procurement process in January 2013; responses are expected at the end of March 2013. The first segment will link the western system to the Winnipeg – Thunder Bay system. The second segment will expand the network into New Brunswick, potentially down to Halifax in Nova Scotia.

The second layer of the optical network infrastructure is the SONET layer, which is the foundation of most CANARIE network services. In this layer, the network is composed of individual point-to-point 10G SONET wavelengths with a topology consisting of a series of rings. The ring architecture allows CANARIE to provision redundant network services, as in the case of the CANARIE IP network. Delivering network services over the SONET layer has been a core service delivery method for a number of years. CANARIE engineering continues to evaluate other transport technologies to maximize the infrastructure investment.

[1] Additional information can be obtained by contacting Dr. R. Sobie, Director of HEPNET/Canada (rsobie@uvic.ca)

CANARIE Network Services

CANARIE offers a number of network services to users: Lightpath, IP Network service, and Content Delivery Service (CDS). Recently CANARIE added a VRF (virtual routing and forwarding) based service to its portfolio, the first application to use that service was LHCONE network for the ATLAS project.

CANARIE Lightpath service
The CANARIE Lightpath service is delivered on GbE and 10 GbE client interfaces over SONET or a dedicated wavelength in 10GbE or 100GbE. 10Gps SONET wavelengths are usually partitioned into smaller capacity channels, from 155 Mbps to GbE, and up to a full 10 GbE. With the DWDM systems, lightpaths can also be 10GbE or 100GbE wavelengths that can be dropped directly into researchers' equipment, thus bypassing CANARIE optical switches. TRIUMF has been using lightpaths build the LHCOPN network in Canada since the first years of operation of the LHCOPN.

CANARIE IP Network service
CANARIE IP network service remains the biggest workhorse of CANARIE's service offerings. The core network, consisting of 7 core routing nodes, provides full and equal support for IPv4 and IPv6 unicast and multicast routing. There are eleven internal segments that link the routers in a partial-mesh topology and five external network segments that extend to international R&E layer 2 exchanges: the Pacific Wave in Seattle, StarLight in Chicago, and Manhattan Landing (MANLAN) in New York. Most of these external segments are operated at 10GbE.

International connectivity is important and critical for TRIUMF and the Tier-2 centers. In 2012, CANARIE deployed the LHCONE using VRF technology over the IP network, connecting TRIUMF and the Tier-2 centers to other LHCONE peers in US and Europe.

Content Delivery Service (CDS)
CANARIE's Content Delivery Service (CDS) provides high-speed access to content from approved providers for Research and Education institutions. It has become an important service of CANARIE service offerings. The CDS IP network, which is logically separated from the CANARIE R&E IP Network, links to SIX (Seattle, WA), Pacific Wave (Seattle, WA), TorIX (Toronto, ON) and NYIIX (New York City, NY) to source content. Major content providers like Amazon, Microsoft, Google, Yahoo, Facebook and Box.net are reachable through CDS.

International connectivity

The CANARIE Network is connected to GLIF (Global Lambda Integrated Facility) through three R&E International eXchange Points (IXPs): Pacific Wave, StarLight, and MANLAN (Manhattan Landing Exchange Point). These connections provide CANARIE with the ability and capacity to establish dedicated lightpaths throughout most of the world, enabling researchers to perform international collaborative research. One of the prime examples is the LHCONE deployment linking Tier 1 and 2 sites over common exchange points, most of them are the GLIF nodes in US and Europe.

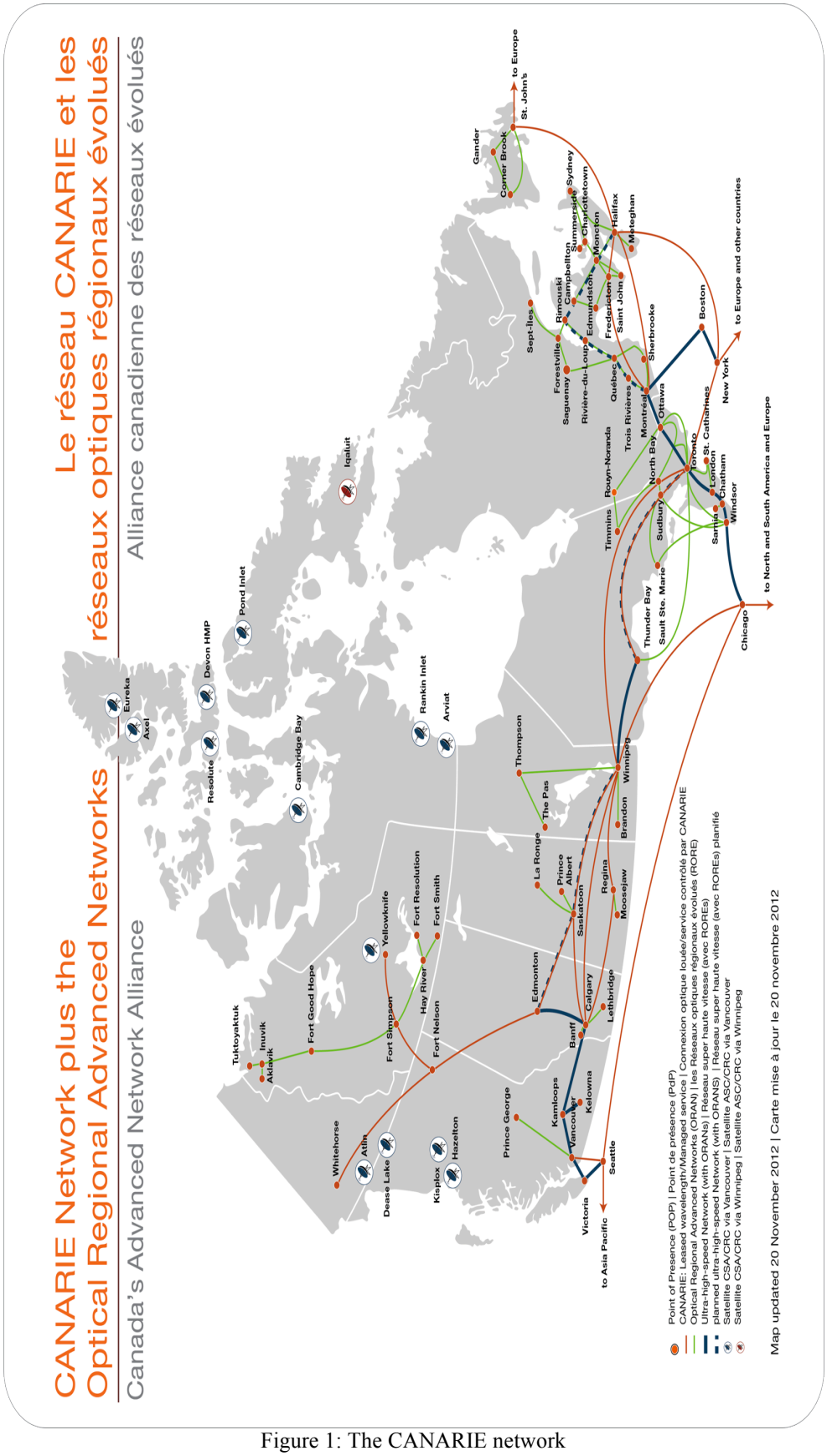

Figure 1: The CANARIE network

ATLAS Tier 1 Computing Centre at the TRIUMF Laboratory

TRIUMF, Canada's National Laboratory for Nuclear and Particle Research operates a Tier-1 (T1) Computing Centre for the ATLAS experiment in Canada. The TRIUMF Centre is linked to the LHC Worldwide Computing Grid (WLCG) and provides an interface to a grid of computing resources at universities across Canada.

In July 2005, CANARIE signed a Memorandum of Understanding (MOU) with HEPnet/Canada, ATLAS Canada and TRIUMF to provide the high-energy physics community with a dedicated 10G circuit across Canada and initial 5G lightpath to the CERN Tier-0 (T0) Centre. This lightpath became active in December 2006.

The TRIUMF T1 to CERN T0 circuit, depicted in Figure 2, runs over the CANARIE 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 Large Hadron Collider Private Optical Network (LHCOPN) address space is allowed to flow over the network. Exchange of routing information is performed using Border Gateway Protocol (BGP) at the T1 and T0 institutions. The 5G lightpath transits Canada west to east. The lightpath travels over the BCNet network from TRIUMF to the CANARIE PoP at UBC, and continues along CANARIEs network, then debarks North America at the MANLAN transit exchange in New York City. The lightpath enters Europe on SURFnet in Amsterdam and then transits Geant network to CERN.

The 5G link has been sufficient for the first few years of LHC operation. However should the demand increase beyond 5G, the lightpath capacity can be increased in 155Mbps increments up to the full 10G. The T0-T1 connection has been adequate and static for a number of years.

The backup link for the TRIUMF T1 passes from the Vancouver CANARIE OME to the Pacific Northwest Gigapop in Seattle, then via Chicago, Toronto and MANLAN to Amsterdam. This link provides an alternate fibre path across most of North America and the Atlantic. Even at a lower capacity of 1G it is expected to be able to temporarily handle loads while the 5G link is restored. The tertiary backup link added travels via Victoria to Pacific Northwest Gigapop and then to the Brookhaven National Laboratory (BNL). In the event of the failure of the primary and secondary links traffic will be carried via the US LHC Network to CERN. The tertiary backup also acts as T1 to T1 link which is particularly advantageous because BNL hosts 25% of the ATLAS Data. In 2008 an additional T1 to T1 link was established with SARA (TRIUMF's Tier-1 partner) following the same path as the primary 5G link.

The TRIUMF Tier 1 supports the Tier 2 centres at the University of Victoria, University of Alberta, University of Toronto, Simon Fraser University (SFU) and McGill University and the University of Melbourne (Australia). From 2007 to 2012 all Canadian T2 centres were connected directly to TRIUMF using point-to-point 1G lightpaths. This was model was extremely successful during this period, however, clear signs of network saturation had developed in 2011 and the 1G lightpaths were clearly inadequate in 2012. In order to leverage the infrastructure developed for the LHCONE, the Canadian T2-TRIUMF networking was completely redesigned in 2012 (see next section).

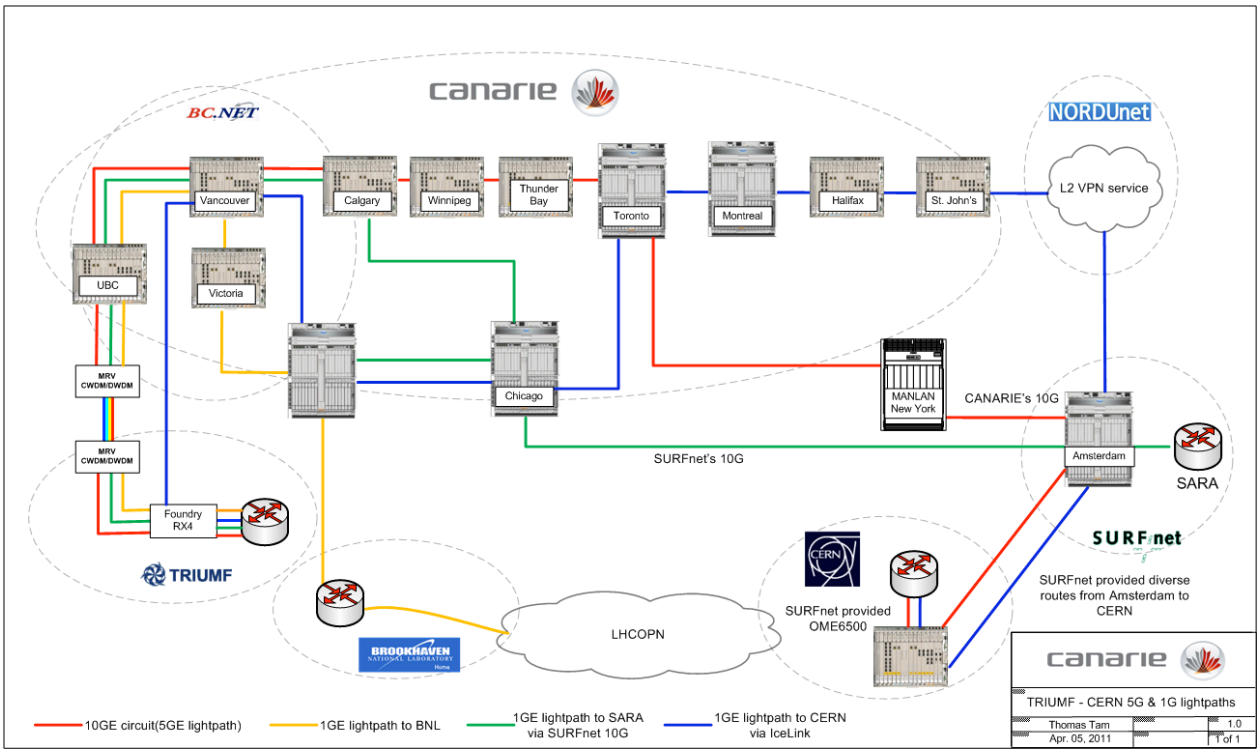

Figure 2: TRIUMF – CERN network

Developments in 2012

LHCONE Network

The LHCONE network was a significant focus of effort in 2012. The shared VLAN LHCONE service was decommissioned in Canada with only the University of Toronto having experimented with the service. We set a goal of having all Canadian sites connected to the LHCONE L3VPN service in 2012. CANARIE developed a purpose built VRF overlay network for use by Canadian T1 and T2 sites spanning all CANARIE Juniper MX series routers. All T2s and the T1 were connected to this service at 10G by December 2012 with the majority of the sites being connected by the late summer. The University of Alberta will be decommissioned as a T2 site in 2013 and hence is only connected at 1G to the LHCONE. In order to achieve the connections, additional 10G circuits were established at McGill and the University of Toronto. At the University of Victoria and SFU, an LHCONE dedicated VLAN was established over lightly used existing 10G circuits. International peerings have been established with Internet2 at ManLan (New York City), and with ESnet at PacificWave (Seattle), with additional peerings planned with GEANT, NorduNet, and ESnet on the East Coast of North America in 2013. Figure 3 shows the Internet 2 peering traffic for the LHCONE between CANARIE and Internet 2.

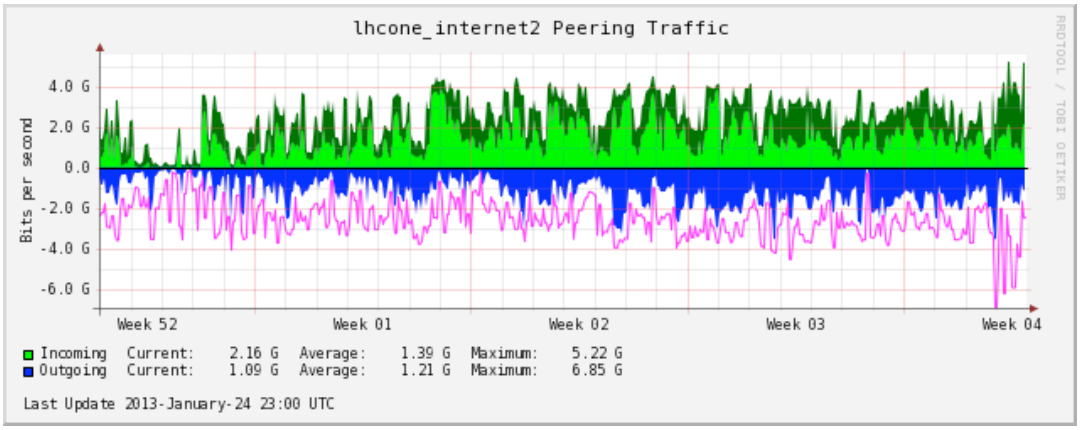

Figure 3: Peering traffic between the CANARIE and Internet 2 VRFs at ManLan in Jan 2013.

One of the principal challenges of the deployment was establishing the routing configuration that would avoid asymmetric routing over the LHCONE. Avoiding asymmetries is absolutely critical to avoid problems with stateful firewalls at the edges of other LHCONE participating sites. Not all Canadian sites had either dedicated equipment or equipment capable of creating VRFs resulting in a diversity of solutions. At the University of Victoria and the University of Toronto, a VRF was established to serve the ATLAS related subnets at each of the site. At SFU no such VRF capable equipment was available resulting in BCNet having to provide a VRF within the Vancouver Transit Exchange over 10 km from SFU. TRIUMF used policy based routing to avoid asymmetry issues (which they plan to remove when they deploy VRF capable Juniper equipment in 2013). Both McGill and the University of Alberta have dedicated equipment.

TRIUMF to Tier 2 Network with Canada

In order to address the aforementioned capacity problems of the existing 1G lighpaths, all Canadian T2s moved to using the LHCONE overlay network for TRIUMF to T2 traffic. This eliminated the need for the point-to-point lightpaths while also increasing the capacity at T2s to 10G. We were also able to provide a single infrastructure to serve both national and international network requirements. However, with TRIUMF using a single 10G connection to serve the needs of Canadian T2s and international LHCONE sites we expect to have to upgrade the LHCONE connection to 2x10G in the near term. Figure 4 shows the already heavy utilization of the TRIUMF LHCONE circuit.

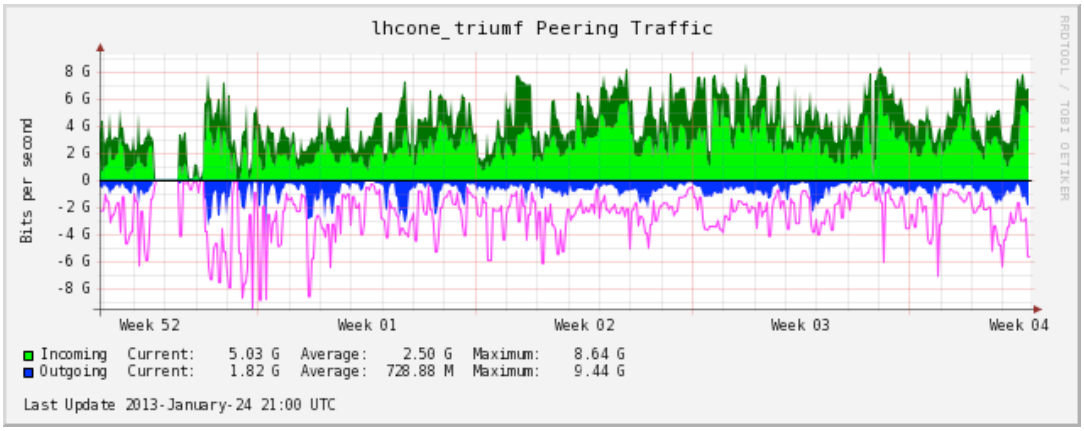

Figure 4: Heavy utilization in Jan 2013 on the TRIUMF LHCONE circuit that is used both for Canadian T2 and other international LHCONE sites.

Network monitoring
HEPnet Canada deployed 10G capable perfSonar boxes to all sites during 2011 and those boxes were used to detect a number of problems in 2012. Further work needs to be done in 2013 to improve the combined display of perfSonar test results and alerting on failed tests. For further details of the hardware configuration please refer to the 2011 report.

Multi 100G Network Demo at SC12
For SuperComputing 2012 in Salt Lake City the University of Victoria in partnership with Caltech, University of Michigan, CANARIE, Internet 2 and BCNet undertook a demonstration of 100G networking technology. With the industrial partners Ciena, Brocade and IBM, a 100GE circuit was established over BCNet, CANARIE and Internet 2 between the University of Victoria data centre and the Caltech exhibition booth of the show floor of the Salt Lake City convention center. The demonstration showed a sustained full duplex throughput of 339 Gbps between all the sites (Victoria, Michigan, Caltech). In addition the team was able to achieve 95 Gbps disk-to-disk throughput from 4 machines at UVic with SSDs to a mix of equipment on the show floor. [2]

[2] For further information visit http://supercomputing.uvic.ca and http://supercomputing.caltech.edu