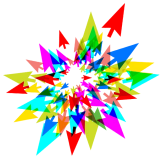
compute canada | calcul canada



Compute Canada/WestGrid Plans and Updates

Patrick Mann, Director of Operations, WestGrid
Lixin Liu, Network Specialist, Simon Fraser University

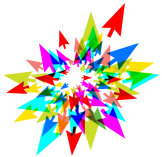
BCNet April 24, 2017

Outline

1. Cyberinfrastructure for Advanced Research Computing (Patrick Mann)

2. High Performance Network for Advanced Research Computing (Lixin Liu)

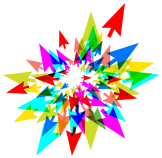
compute canada | calcul canada



Cyberinfrastructure for ARC

Patrick Mann
Director of Operations
WestGrid

compute canada | calcul canada



Advanced Research Computing

- Compute Canada leads Canada's national **Advanced Research Computing** (ARC) platform.
- Provides ~80% of the academic research ARC requirements in Canada.
  - **No other major supplier in Canada.**
- CC is a not-for-profit corporation.
- Membership includes 37 of Canada's major research institutions and hospitals, grouped into 4 regional organizations
  - WestGrid, Compute Ontario, Calcul Quebec, and ACENET
- User Base
  - From "Big Science" to small research groups
  - From Digital Humanities to Black Hole simulations

Funding

Funding from the **Canada Foundation for Innovation**

Matching funds from provincial and institutional partners

- 40% federal / 60% provinces and institutions

Capital: **CFI Cyberinfrastructure Program** + match

- Stage-1 spending in progress ($30M CFI) ← ***We Are Here!***
- Stage-2 proposal being assessed ($20M CFI)
  - Site selection in progress
- Stage-3 planning assumption ($50M CFI in 2018)

Operating: **CFI Major Science Initiatives (MSI)** + match

- 2012-2017, ended March 31, $61M CFI
- 2017-2022, $70M CFI, announced January 9th ← ***We Are Here!***

Planning 2015-2016

SPARC = Sustainable Planning for Advanced Research Computing

In 2016, CC conducted second major SPARC process:

- 18 town hall meetings
- 17 white papers received (disciplinary + institutional)
- 189 survey responses

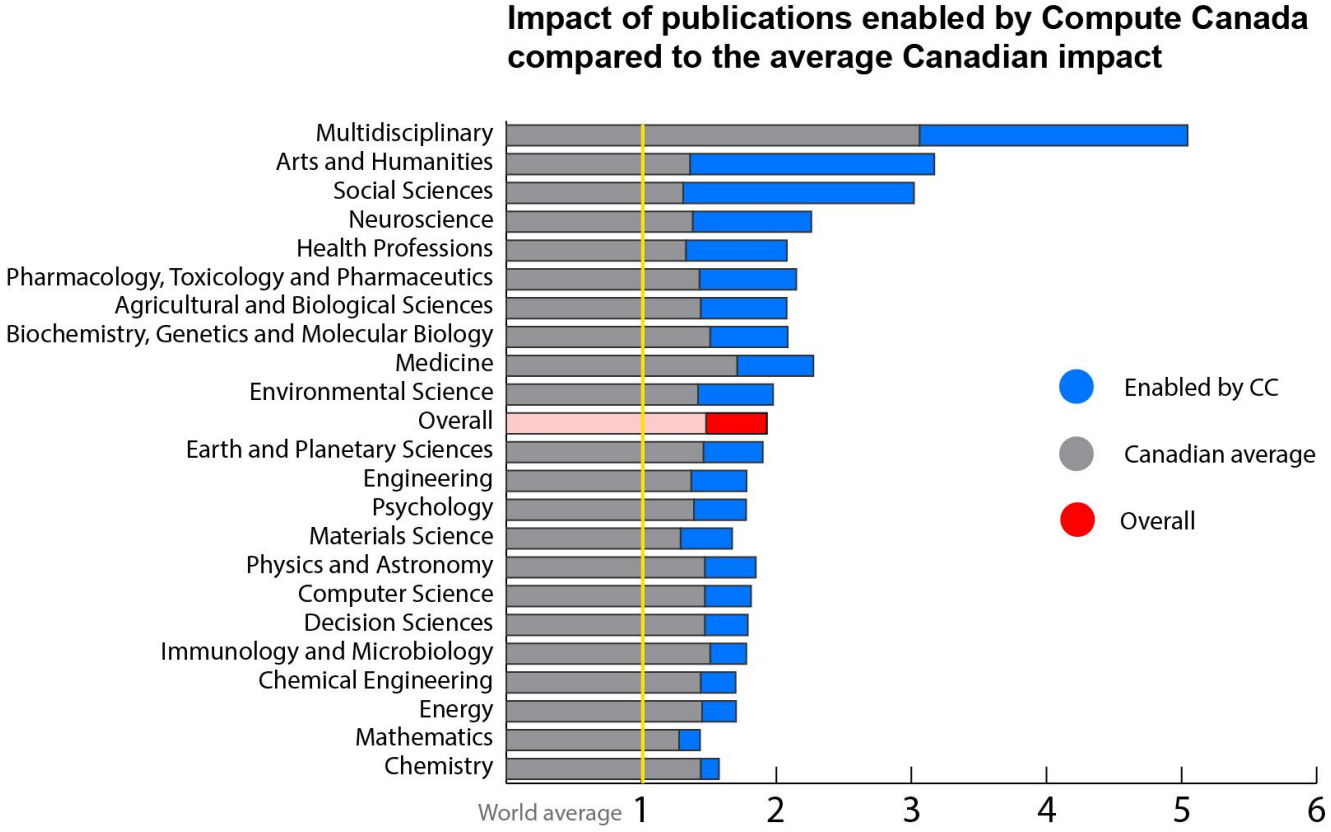
Ongoing consultations on CFI grants:

- Consulted with more than 100 projects in 2015 and 2016.

Several councils of researchers:

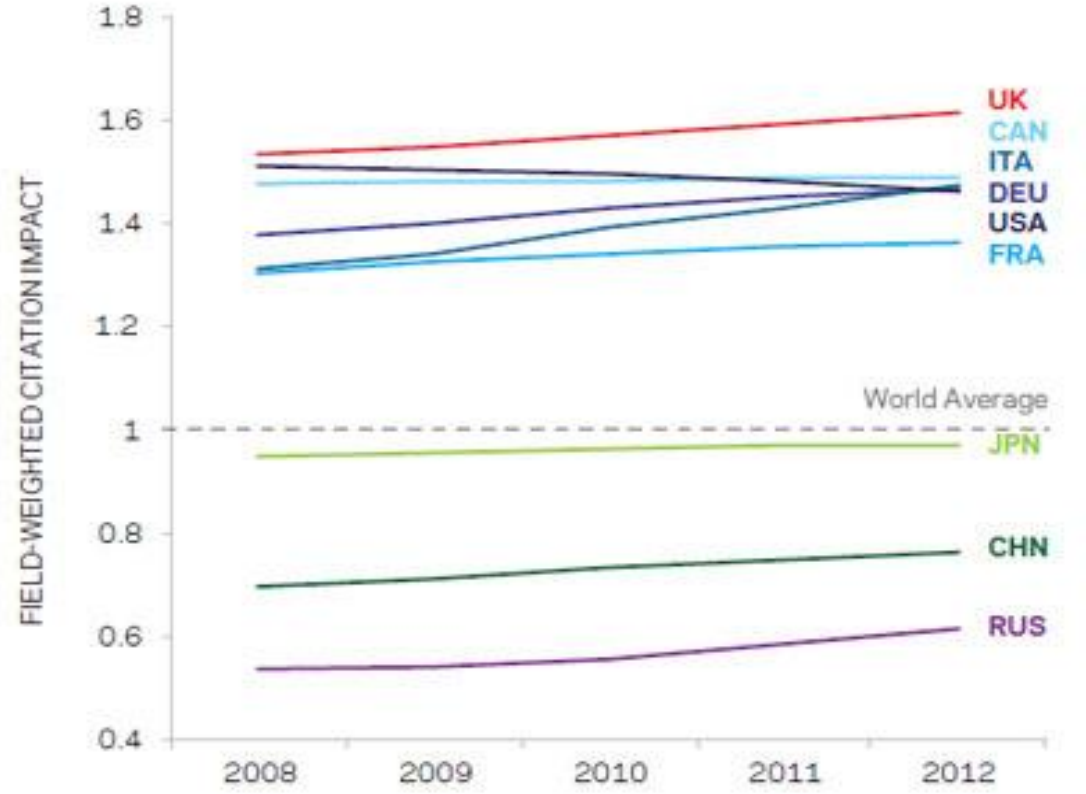
- Advisory Council On Research
- RAC-Chairs
- International Advisory Committee

FWCI



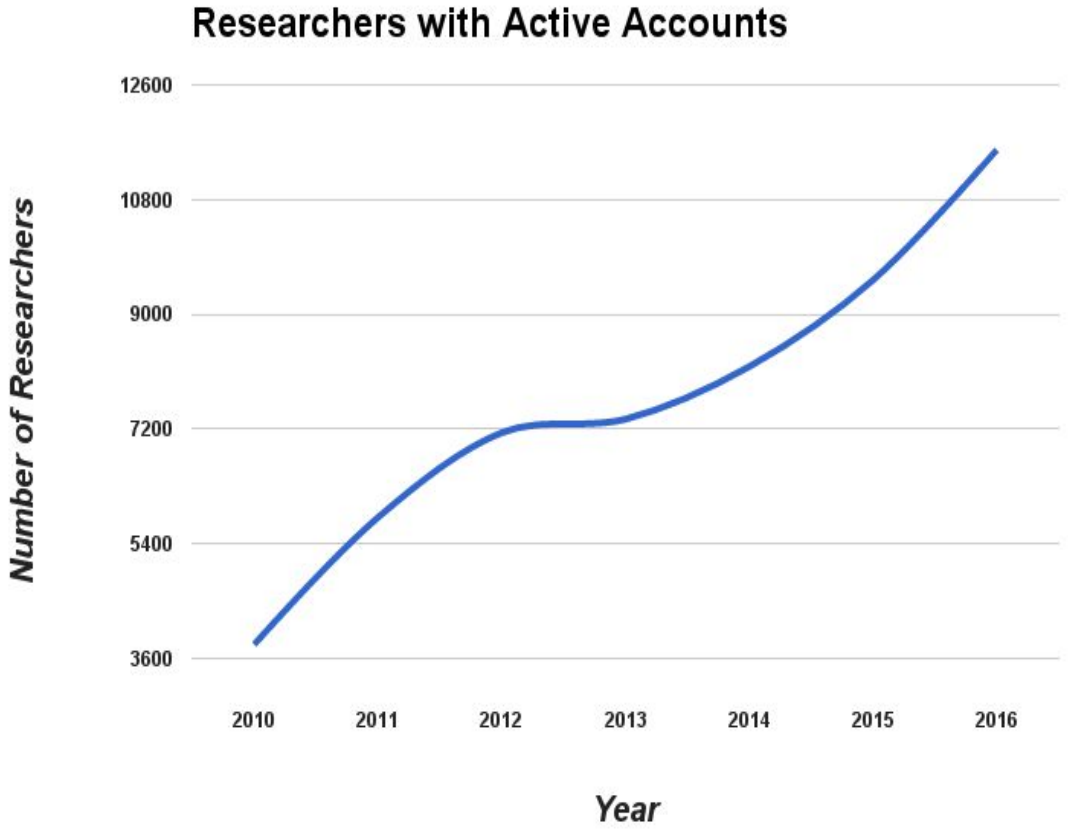
Field-weighted citation impact divides the number of citations received by a publication by the average number of citations received by publications in the same field, of the same type, and published in the same year.

FWCI Canada



From https://www.elsevier.com/connect/report-how-do-the-large-research-nations-compare

User Base Growth



Resource Allocations

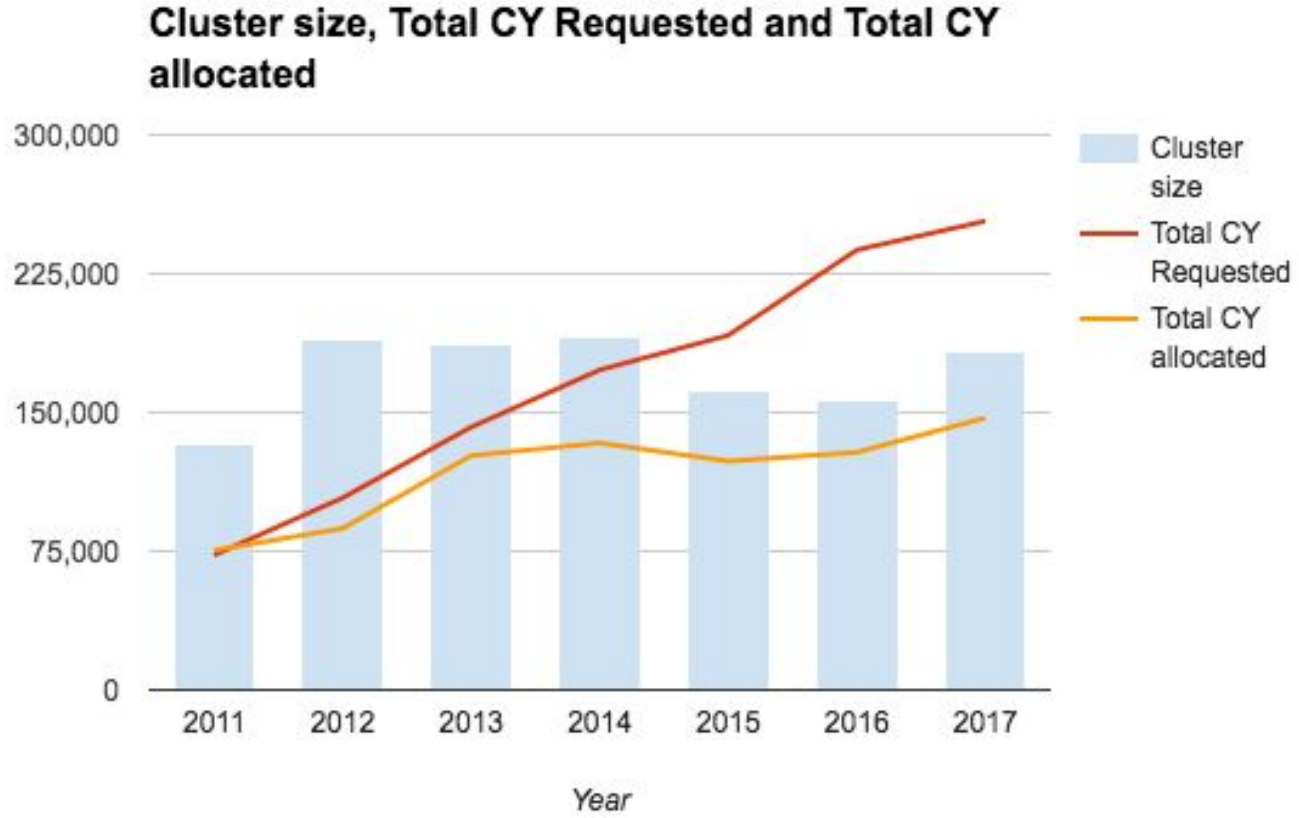
Resource Allocation Competition (RAC)

1. **Resources for Research Groups** (RRG)
   a. Annual allocation. Compute (Cores)+storage
2. **Research Platforms and Portals** (RPP)
   a. Up to 3 years.
3. **Rapid Access Service** (RAS)
   a. 20% for opportunistic use. New users. New faculty. Test/prototype...
4. **Compute Burst** (new systems only)

Competitive:

- Extensive proposals
- Full science review

RAC: More resources, more need



Resource Allocation - 2017

	2017 Requests	2016 Requests	% Change
Compute - CPU-years	256,000	238,000	**+7.5%**
Compute - GPU-years	2,660	1,357	**+96%**
Storage	55 PB	29 PB	**+92%**

	2017 Requested Fraction Available	2016 Requested Fraction Available
Compute - CPU	54%*	54%
Compute - GPU	38%	20%
Storage	90+%	90+%

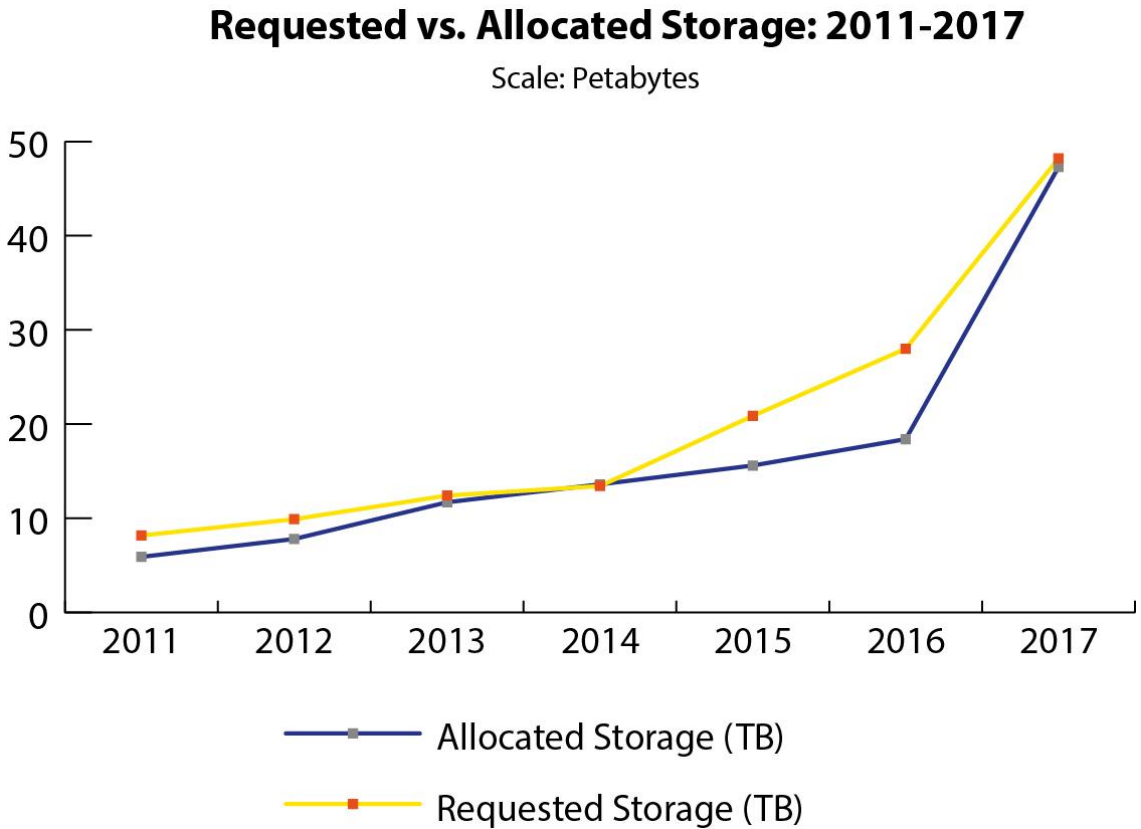
** 54% in 2017 includes 50k+ new cores with better performance*

RAC - GPU

	Total GPU Capacity	Total Requested	Total Allocated	Allocation Request Rate
2017	1,420	2,785	1,042	0.38
2016	373	1,357	269	0.2
2015	482	607	300	0.49

GPU-years

Storage


Requested vs. Allocated Storage: 2011-2017
Scale: Petabytes
50
40
30
20
10
0
2011
2012
2013
2014
2015
2016
2017
Allocated Storage (TB)
Requested Storage (TB)

The Plan

Hardware Consolidation by 2018

- 5-10 Data Centres, 5-10 systems

Systems

- Phase I HPC infrastructure renewal
- National Data Cyberinfrastructure (NDC)
- New Cloud Infrastructure-as-a-Service (IAAS)

2016-2018 Transition years

- Major migration of data and users

Services

- Common software stack across sites
- Common accounts (single sign-on)
- Common documentation
- 200 distributed experts, national teams
- Research Platforms and Portals
  - common middleware services
- Research Data Management
  - Globus: data management and transfer
  - Collaboration with libraries (CARL) and institutions

National Compute Phase I

System	Status	Capacity	Interconnect	Specs	Availability
Arbutus (GP1, UVic)	**In production** East and West Cloud prototypes in service since 2015 Compute and Persistent	7,640 290 Nodes	10 GB Ethernet	OpenStack Local drives Ceph persistent 560 TB (usable)	Sep 8, 2016
Cedar (GP2, SFU)	Datacentre renos complete Racks and servers installed. OS and configuration	27,696 cores 902 nodes 584 GPUs	Intel Omnipath	E5-2683 v4 2.1 GHz NVidia P100 GPU's Lustre scratch ~5PB	May 2017
Graham (GP3, Waterloo)	Datacentre renos complete Racks and servers installed. OS and configuration.	33,472 cores 1,043 nodes 320 GPUs	Infiniband	E5-2683 v4, 2.1 GHz NVidia P100 GPU's Lustre scratch ~3PB	May 2017
Niagara (LP1, Toronto)	RFP issued. RFP closes May 12.	~66,000	??	??	Late 2017

National Data Cyberinfrastructure

System	Status	Availability
NDC-SFU	• 10 PB of SBB's delivered.	May 2017
NDC-Waterloo	• 13 PB of SBB's delivered.	May 2017
NDC - Object Storage All sites	• ~5 PB raw • Object Storage. • Lots of demand but not allocated. • Geo-distributed, S3/Swift interfaces	Summer 2017
NDC - Nearline Waterloo and SFU	• Large Tape systems • NDC file backup • Hierarchical Storage Management	Tape In service HSM in

NDC = "National Data Cyberinfrastructure"
SBB = "Storage Building Blocks"

Silo Interim Storage

Silo: WestGrid Legacy system at USask - 3 PB

Silo to Waterloo completed Jan.11, 2017:

- 85M files, 850TB, 140 Users.

Silo to SFU completed March 9th, 2017:

- 103M files, 560TB, 4,381 Users.

Large RAC Redundant Copies

- Ocean Networks Canada: From ONC to Waterloo
- Remote Sensing Near Earth Environment: UofC to Waterloo
  - ~90M files
- CANFAR (Astronomers): UVic to SFU

Silo was decommissioned Mar.31/2017

Services and Resources

Resources/Services

ownCloud

Globus file transfers

IAAS Cloud

Stable and secure data storage and backup

Object Storage S3

High performance, big data and GPU computing and storage

Videoconferencing

Research Data Management

Expertise

Consultation - Basic and advanced.

Designing, optimizing and troubleshooting computer code

Customizing tools

Installing, operating and maintaining advanced research computing equipment

Dedicated humanities specialist

Visualization, Bioinformatics, CFD, Chemical modelling, ..

Cybersecurity

Training

Group and individual training and ongoing support from novice to advanced

Standard and discipline specific customized training

Livestreaming of national seminar series including VanBUG and Coast to Coast

Quickstart guides, training videos and other upcoming online workshops
www.westgrid.ca

WESTGRID

compute canada | calcul canada

HPCS2017

Kingston Ontario

June 6 -9, 2017

Registration Now Open

(Early Bird $225 - ends April 30)

http://2017.hpcs.ca

WG Training & Summer School

DATE	TOPIC	TARGET AUDIENCE
MAY 04	**Data Visualization Workshop - University of Calgary**	Anyone (in person)
JUNE 05 - 15	**Training Workshops / Seminar Series on using ARC in Bioinformatics, Genomics, etc.**	Researchers in Bioinformatics, Genomics, Life Sciences, etc.
JUNE 19 - 22	**WestGrid Research Computing Summer School - University of British Columbia**	Anyone (in person)
JULY 24 - 27	**WestGrid Research Computing Summer School - University of Saskatchewan**	Anyone (in person

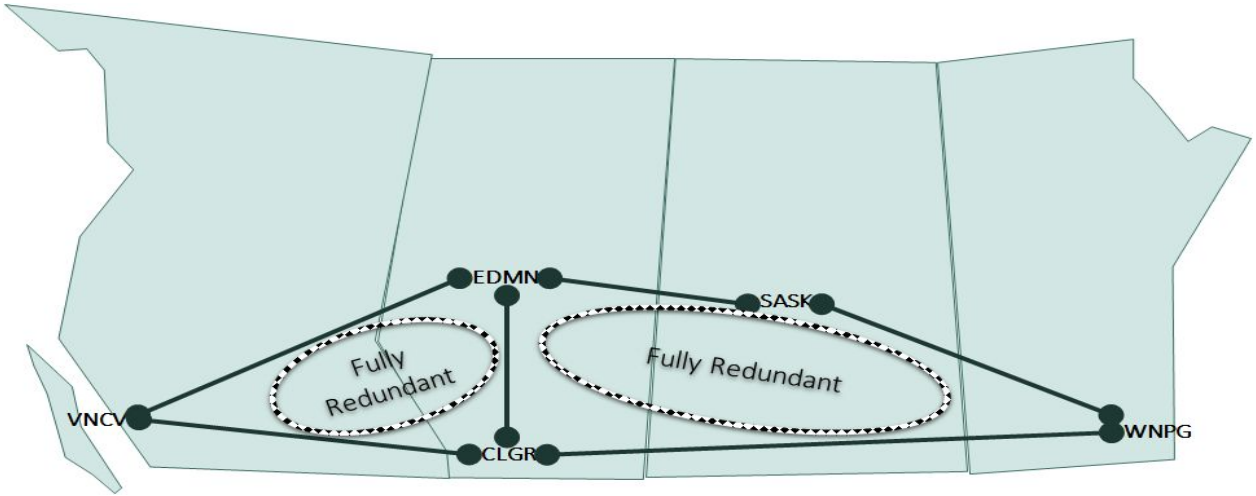
Full details online at www.westgrid.ca/training

High Performance Network for ARC

Lixin Liu
Simon Fraser University
WestGrid

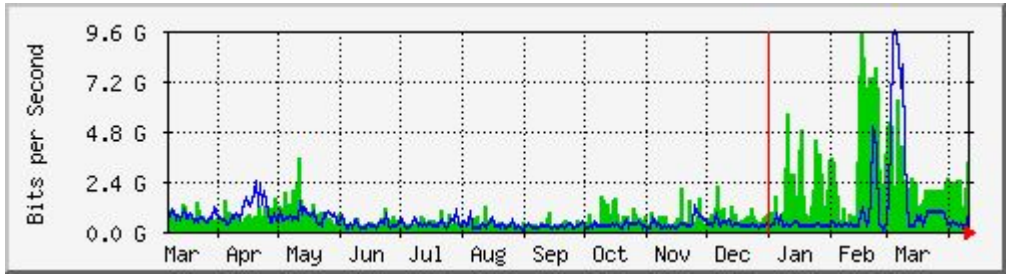
Current WestGrid Network

- WestGrid core uses VPLS Point-to-Multipoint circuits provided by CANARIE
- Endpoints in Vancouver, Calgary, Edmonton, Saskatoon, Winnipeg
- Layer 3 between all data centres, all sites have 10GE connections
- Fully redundant, fast reroute (under 50ms) network



CC Needs Faster Network

- Size of research data grows very fast
- New applications require significant more bandwidth, e.g., WOS
- Fewer data centres means more data to be stored at each site
- 100GE network is very affordable now



Daily network utilization average at SFU WestGrid in 12-month

CANARIE & BCNET 100GE Network

CANARIE

- 100GE IP network available since 2014
- Redundant connections for most major cities

BCNET 100G available in Vancouver & Victoria

- Upgraded Juniper MX960 backplane to support new 100G linecards
- Purchased MPC7e 100GE QSFP28 linecard
- Primary path Vancouver-Victoria
- Alternative path Vancouver-Seattle-Victoria

Network Hardware Procurement

CC Networking RFP

- Issued by SFU in June 2016 for all 4 new stage-1 sites
- To provide 100GE connections for all sites
- Shortlist selected in September
- CC representatives conducted verification on shortlisted vendor products
- Winner (Huawei Technologies) was announced early this year
- Winning Solution: Huawei CloudEngine 12800 Series (CE12800)
- Purchase orders created for SFU, UVic and Waterloo, UofT soon

Hardware for Each CC Site

Hardware Orders

- UVic: CE12804S
- SFU: CE12808 (WTB), CE12804S (Vancouver & Surrey)
- Waterloo: CE12808
- Toronto: CE12804S

CE12804S

CE12808

Huawei CE12800

CE12800 Features

- Switching: 45Tb (12804S), 89Tb (12808)
- Forwarding: 17Gpps (12804S), 35Gpps (12808)
- 100G ports: up to 48 (12804S), up to 288 (12808)
- Linecards: 100G (CFP, CFP2, QSFP28), 40G, 10G and 1G
- Large buffer: up to 24GB
- Virtualization: VS, VRF (vpn-instance), M-LAG, VxLAN, EVPN
- L2: VLAN, STP
- L3: IPv4/v6, BGP, OSPF, ISIS, Multicast, MPLS, etc.
- Management: CLI, SNMP, Netconf, OpenFlow, Puppet, Ansible, etc.
- Availability: ISSU, VRRP, etc.

CC Datacentre Connection
University of Victoria

University of Victoria

- New 100GE network is ready in February after BCNET router upgrade
- CE12804S to replace rented Brocade MLXe as edge router only
- Connect to BCNET new linecard using QSFP28 SR4 module

CC Datacentre Connection
Simon Fraser University

Simon Fraser University

- New 100GE network is ready in February after BCNET router upgrade
- HC CE12804S to connect to BCNET linecard using QSFP28 LR4 module
- WTB CE12808 serves as the core switch for new CC equipments
- HC to Burnaby connection is using 2x40GE (ER4), will upgrade to 100GE
- Surrey CE12804S will be available for failover path
- TSM servers, DTNs and SBB3s to be connected to CE12808 using 40GE

CC Datacentre Connection
University of Waterloo

University of Waterloo

- RFP issued to acquire a 100GE connection from Waterloo to Toronto
- Initially will use existing 10GE provided by SHARCNET
- CE12808 serves as core switch * edge router for CC equipments

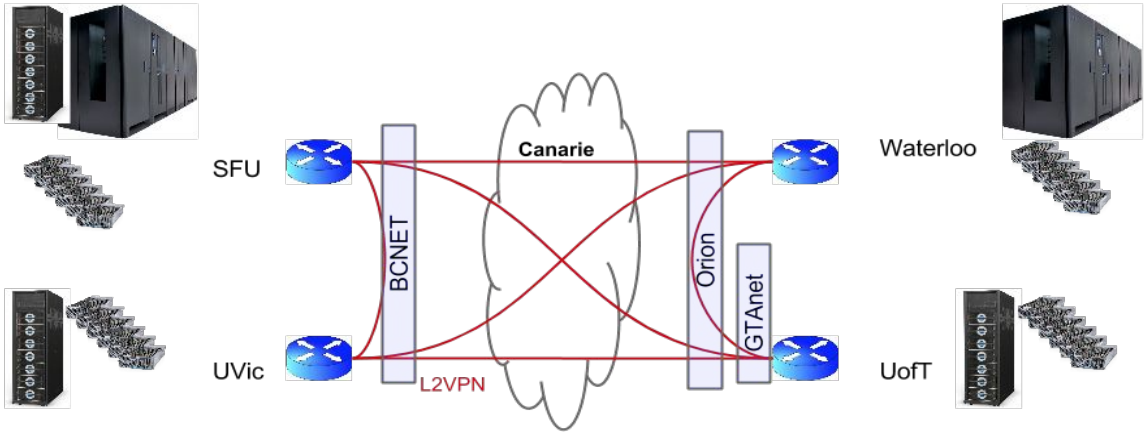
CC Datacentre Connection
University of Toronto

University of Toronto

- TBD (43KM from datacentre to 151 Front Street)

CC Site to Site Network

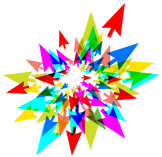
- 4 CC stage-1 sites will be connected “directly”
- CANARIE to provide VPLS circuits in Toronto, Vancouver and Victoria
- Initial plan is to provide L3 services only among 4 sites
- L2 services, if required, will use VxLAN
- IPv4 and IPv6 will be supported



CC Network Applications

- TSM: Replication service between SFU and Waterloo
- WOS: Core nodes traffic may require L2 network
- Globus: DTNs available for each NDC and cluster
- Atlas T1: Use SFU WTB physical connection, but route separately
- Atlas T2: Currently in SFU, UVic, and UofT, will add Waterloo but drop UVic and UofT
- CANFAR: data replication between SFU and UVic, may include UW later

compute canada | calcul canada



Q&A

?

Contact Info

patrick.mann@westgrid.ca
liu@sfu.ca
www.computecanada.ca
www.westgrid.ca
support@westgrid.ca
docs.computecanada.ca