

WG Town Hall III

Lindsay Sill, Executive Director
Patrick Mann, Director of Operations
Sergiy Stepanenko, Site Lead, USask

Tuesday October 11, 2016

Introduction & Outline

- WG User Survey Results
- WG Incorporation and future plans
- Funding Update
- Tech Refresh - Stage 1
- Migration Plans - Silo (Storage) & Compute
- RAC 2017 Changes
- Upcoming Training Opportunities
- Email questions to info@westgrid.ca or use the Vidyo chat feature, or unmute yourself

Survey: By the Numbers

- 235 respondents
- 36% Faculty
- Over 60% of respondents were from:
 - Engineering, Math or Computer Science
 - Medical, Biological, or Life Sciences
 - Chemistry or Biochemistry
- Thanks! We learned a lot about user needs. For example:
 - 22% of respondents feel they do not have adequate local ARC support
 - Parallel programming is training session most are interested in (49%)

Survey: By the Numbers

- 60% used WestGrid support in past year
 - Over 80% thought support received was GOOD (helpful and quick)
 - 47 respondents felt someone in particular had been extremely helpful

"Excellent support on campus, we just need more like him, [this is a] serious limitation"

"Support services [are not the] problem but cycles are too hard to get"

"WestGrid support is excellent, but with him, it's exceptional"

- **New user support is critical**
 - “It wasn't very clear where to start, first time I wanted (to) start using WestGrid”, “More entry level information in plain language would have been useful when I was getting started.”
- **Training**
 - Parallel programming, scheduling/managing jobs, object store, bioinformatics support (users unaware of resources).
- **Services / Resources**
 - Fast file transfer, Large Shared Memory, secure computing environment, GPUs, long term data storage, VMs.
- **Support**
 - Identifying gaps by domain, institution, etc.
- **Report to stakeholders on findings to follow**

- Governance and management structures are confusing to users, lack clarity between role of the regions and the national organization

Not a legal entity → **Incorporation** planned for next 6-9 months

- define governance model, strategic goals, independence from Compute Canada, partnerships
- no direct effect for research community

Compute Canada received notification of funding for both operations and the next stage of capital for 2017 - 2022

- Very positive
- Confidential until funding announcement
- Allows Compute Canada to proceed with Stage 2 site selection to select national host sites
- Will result in **more** new infrastructure on the floor in 18-24 months.

Next: Technology Refresh/Deployment

Patrick Mann

Director of Operations, WestGrid

Object Storage I

A big new Compute Canada service, but sounds like the community is really not very clear on what it is (survey, informal discussions).

A full exposition would take a while... So one more slide!

Object Storage II

S3 = AWS **Simple Storage Service**. Microsoft: BLOBS
Flat structure: objects inserted (put) and retrieved (get).

- *No POSIX file complications, no directories*

Allows for objects distributed across SBB's.

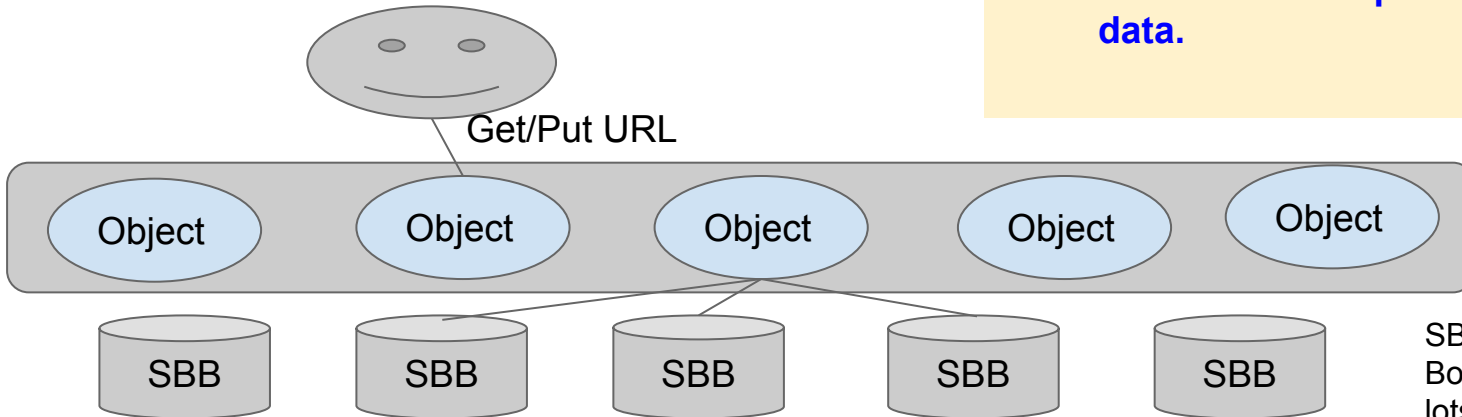
- *Georeplication and striping (erasure coding).*
- *Scale-out is easy*

RESTful interface

- *URL Get/Put from any node*

Example: Youtube - videos are objects.

- Big, flat datasets
- *Metadata/search* provided by application
- *Reliable and Available* (performance not so great)
 - tape backup very poor for such big datasets!
- *Random access patterns, but sparse!*
- **Excellent for experimental/observational data.**



SBB = **S**torage **B**uilding **B**lock
Box with front-end server and
lots of drives!



Current Deployment Status

System	Status	In-production Estimate
GP1 (UVic)	<ul style="list-style-type: none">~6,000 new cores installed at GP1. Total of ~6,600 cores in the cloud.	DONE
GP2 (SFU)	<ul style="list-style-type: none">RFP responses under review.	February, 2017 (Delayed)
GP3 (Waterloo)	<ul style="list-style-type: none">RFP Bids received Oct.6	April, 2017 (ahead of schedule)
NDC - File Storage	<ul style="list-style-type: none">SBB Delivery: SFU (almost) and Waterloo (received)Tape installed.	Oct.31, 2017 (Delayed)
NDC - Object Storage	<ul style="list-style-type: none">Object Storage. DDN WOS. Initial prototype for testing installed on cloud.Various internal training sessions underway.	Early 2017 (estimated)
Network	<ul style="list-style-type: none">Preferred vendor testing. 10 Gbps now at Waterloo	2016/2017
Parallel FS	<ul style="list-style-type: none">RFP Bids received. Evaluation ongoing.	November, 2016
Scheduler	<ul style="list-style-type: none">RFP Bids received. Evaluation ongoing.	November, 2016
LP1 (UofT)	<ul style="list-style-type: none">Early discussion	Late 2017

Cluster Decommissioning

Edmonton	hungabee/jasper	Dec 31, 2016	Storage support expires
Victoria	hermes/nestor	Mar 31, 2017	
Calgary	breezy/lattice/parallel	Mar 31, 2018	Parallel extended to 2018
Vancouver - UBC	orcinus	Mar 31, 2018?	Possible extension
Winnipeg	grex	Mar 31, 2018	Storage extension discussions
Vancouver - SFU	bugaboo	Mar 31, 2018	

Please note these are provisional dates.

For Compute Canada details see

<https://www.computecanada.ca/research-portal/accessing-resources/migration/>

Next: **Silo Migration**

Sergiy Stepanenko

WestGrid Site Lead

University of Saskatchewan

Data Migration: DELAYED

- “Storage Building Block” (SBB) purchase process was very significantly stretched.
- SBB’s have been delivered to Waterloo. SFU expecting any time now.
- Say a couple of weeks for rack and stack.
- So production-ready end of October (estimated).
- Optimistic estimates for Silo user storage migration is **3 months. But we have only 2 months.**
- **So a big problem. Contingency planning now in effect.**

Contingency Plan

- Basic approach: **keep Silo going for a couple of months.**
 - Expensive, but necessary.
- Mitigation plans being implemented
 - Extended support, spare parts, physical space, extended funding, ..
 - *(Many thanks to USask for their continuing support and engagement)*

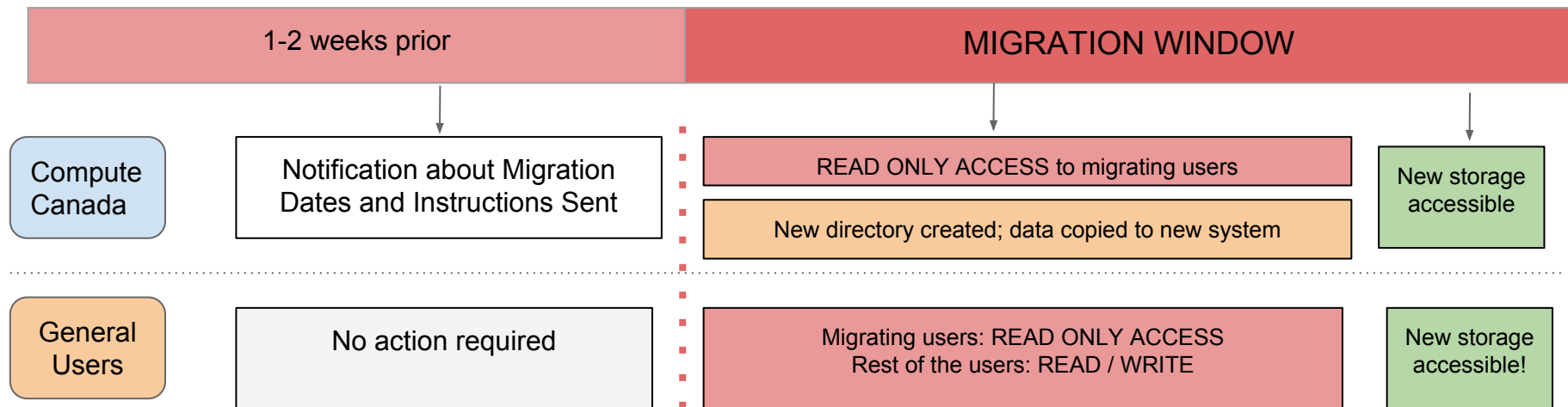
User-Visible

1	Delayed Start	Migration start ~Nov.1 <ul style="list-style-type: none">● Originally planned for Oct.1
2	Read-only from ~ Nov.1	Remaining users on Silo who were actively writing data will have to move their writes to the NSI while we copy their existing data in Jan/Feb.



Silo General User Migration

Silo decommissioning	TBD (early 2017)	Contingency plans for a delay
NSI (SFU) Availability	~Nov. 1, 2016	Significant delay due to SBB delivery!
General User Migration	Nov - Feb	Contingency plans in effect



Note: this period is anticipated to be ~5 days. Silo staff working on detailed plans.

Next: **Compute/RAC**
Back to Patrick

Compute Migration

Some key points from Last Town Hall. Still waiting for the equipment so not much to update here.

- 2016 RAC allocations will extend to **April 2017** (*except jasper/hungabee*)
- GP2 should become available in February 2017. GP3 aiming for March!
 - Overlap period for testing. Hopefully users will be given limited access to start migrating codes. Details still being considered.
- **Comprehensive software/application stack being developed**
 - Expect documentation and details to be available early December.
- For CC details see:
<https://www.computecanada.ca/research-portal/accessing-resources/migration/>

No update - still working on plans

- Hungabee+Jasper scheduled for Dec.31 deactivation.
- GP2 only available in February
 - Working on fallback plans. Under consideration:
 - Provide resources on other legacy systems
 - Other regions may have some new resources available
 - Provide resources in the cloud (HPC/batch partition)
- Hungabee large memory
 - Current plans **do not** include a large, shared-memory system
 - GP2 will have 3 TB (RAM) nodes with $2 \times 12 = 24$ cores
 - WG will contact large memory users. Feel free to contact me:
patrick.mann@westgrid.ca



Resource Allocation (RAC)

Process	Opens	Due
RRG Fast Track (by invitation)	Oct. 11, 2016	Nov. 9, 2016
RRG full application	Oct. 11, 2016	Nov.24, 2016
RPP full application (by invitation)	Oct. 11, 2016	Nov. 24, 2016
2016 RPP Progress Report due	Early December	Jan. 5, 2017
Announcement and implementation of awards	Early March 2017	Mid April, 2017

As usual WG Staff will be available for consultation. We very much recommend that anyone applying work with us. Email rac@computecanada.ca or contact your local WestGrid support staff.

Compute Canada RAC Info Session
Thursday Oct 13, 9am PT

- GP1 Installation - DONE. (Integrated into west.cloud)
- RAC 2017 - Users will be able to request Cloud resources
 - Will be part of the RAC request application form.
- Rapid Access Service (default) allocation will be available
 - Testing/Development: 4 vCPU's, 2 instances, 2 volumes < 1 month
 - Available now. Send request to cloud@computecanada.ca
 - Further details will be available in early 2017.

Fall Training Sessions

DATE	TOPIC	TARGET AUDIENCE
October 26	Compute Canada for Humanists	New Users
November 03	Data Visualization Workshop UBC	Anyone
November 08	Data Visualization Workshop UVic	Anyone
November 16	Best Practices for RDM - Using Dataverse for Research Data	Anyone
November 23	Intro to NextGenSeq Analysis	Anyone

For more information and to view the full schedule, visit www.westgrid.ca/training.

Any questions? Here are a few that have been covered in the presentation but we can amplify or clarify. Feel free to contact your local representatives or myself (patrick.mann@westgrid.ca, support@westgrid.ca).

1. What happens when my primary machine is decommissioned? Where do I go? Can I run tests on the new systems?
2. Do I need to modify my code for the new system?
3. Software stack?
4. Where do we find documentation or ask for support?
5. How safe is my data? Are there any issues with the SBB and Silo delays?
6. Do I need to back it up?
7. What is the timing for the migration and new systems?
8. Do I need new account info / authentication?