What is Scholars Portal

- Shared technology service of the 21 university libraries of the **Ontario Council of University Libraries** founded in 2003
- Provides content aggregation and preservation services for member libraries
- Journals – **16,000** and **38 M** articles
- Books – **610,000** ebooks
- GeoPortal – GIS Data
- ODESI – Numeric Data
- Dataverse – Research Data
What do libraries have to do with clouds?
Ontario Digital Library Research Cloud (ODLRC)

Project Details
3-year project
MTCU-PIF Funding
10 partners
UTL as lead

Technology
OpenStack Swift
1.4 PB (4.6 PB raw)
3x replication
Geographically distributed storage nodes
(5-6 locations initially)
Private network

Goals
Lower cost
Highly scaleable
Replicated
Open technologies and standards
Integrated
Hosted in Canada
Secure

Content
Digital Library resources
Archival resources
Research data

OLRC

Toronto
Ryerson
York
Queens
Ottawa/Carleton
McMaster
Waterloo / Laurier
Guelph
Windsor

orion

Ottawa / Carleton
Queens
Waterloo / Laurer
Guelph
McMaster
Windsor

Toronto
Ryerson
York

gta/Orion
Reassessing Storage Strategies

Level One:
- Server Storage
- Solid State Disks
- $10K / TB

Level Two:
- SAN Storage
- 10K-15K RPM drives
- RAID
- $1.2K-$3.5K / TB

Level Three:
- High Capacity, Low Cost, Online Storage
- $300 / TB

Level Four:
- Tape Storage
- Offline Storage
- 100+ PB

SPEED
COST
Why not go with Amazon/Rackspace/etc.?
Because we secretly hate you.
MTCU PIF Proposal

- Nine partner libraries from OCUL; three year project
- University of Toronto as financial lead
- Develop a 1.2PB object storage service for partners
- Provide subscription storage services to other OCUL libraries
- Develop interfaces with library repository applications
- Create a compute cluster to support text analysis of content in the cloud
Storage RFP

• Storage hardware RFP issued Dec. 20, 2013

• High density disk storage servers (DSS)

• Evaluation and analysis through early March

• Awarded to Dell: 2nd week of March

• All equipment delivered by 31st March, 2014
Data Storage Server

- PowerEdge R720xd
- MD1200 disk drawers
- Each drawer contains 48TB (12 x 4TB NL-SAS drives)
- DSS capacity: 48TB to 432TB:
Infrastructure purchases

- 19 R720xds, 77 MD1200s: 4.6PB raw: 18 server racks
- 26 UPSs, PDUs, 2 x video consoles
- 15 10Gbit network switches + fiber optics
- 4 R620 servers: OpenStack proxy / authentication
- 5 R720xd servers: compute/data processing
GTAnet Pilot

Purpose of the pilot:

To understand how to design and implement an effective network topology to support the operation of the OLRC Storage Cloud
GTAnet Pilot

Execution of the pilot:

Model and record the network traffic generated between four OpenStack Swift storage nodes during routine operation and under various simulated disaster scenarios.
GTAnet UofT HUB - Traffic - UofT subnet1

- Inbound: Current: 5.16 M, Average: 241.27 M, Maximum: 517 G
- Outbound: Current: 5.28 M, Average: 46.77 M, Maximum: 300.46 M

GTAnet UofT HUB - Traffic - RyersonU subnet

- Inbound: Current: 5.03 M, Average: 73.49 M, Maximum: 369.56 M
- Outbound: Current: 5.21 M, Average: 125.71 M, Maximum: 582.86 M

GTAnet YorkU HUB - Traffic - YorkU subnet

- Inbound: Current: 5.13 M, Average: 44.22 M, Maximum: 311.23 M
- Outbound: Current: 5.02 M, Average: 135.11 M, Maximum: 582.85 M

GTAnet UofT HUB - Traffic - UofT subnet2

- Inbound: Current: 5.10 M, Average: 1.13 M, Maximum: 5.45 M
Swift Node Considerations

- How much bandwidth can they provide?
- Will they enable jumbo frames?
- Will they extend VLANs across their network?
- How low are their OTO & ongoing costs?
- Do they have an ORION POP on site?
200 TB data (600 TB RAW) across 5 Zones

- Drive (2TB)
- Drawer (24 TB)
- RAID Card (48 TB)
- Zone (120TB)
Implementation!

- Because what the heck are we going to use all this storage for?

- Or maybe more to the point, how?
Swift Browser

<table>
<thead>
<tr>
<th>Name</th>
<th>Created</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>openstack_swift.pdf</td>
<td>08/05/2013 3:25 a.m.</td>
<td>389.3 KB</td>
</tr>
</tbody>
</table>

Options:
- Temporary URL
- Delete object
ODLRC Hackfest – June 20, 2014
Cloudfuse mount guide

1. Grab a copy of the Cloudfuse source from here.
2. Unzip the zip file, and compile the source

```bash
$ unzip master
$ cd redbo-cloudfuse-21358f1
$ ./configure
$ make
$ sudo make install
```


```bash
$ username=[your username]
$ tenant=[your tenant name]
$ password=[your passport]
$ authurl=http://sample.url.ca:5000/v2.0/
$ region=[your region]
```

4. Create a mount point, and mount. For example:

https://spotdocs.scholarsportal.info/display/ODLRC/Cloudfuse
https://github.com/HackODLRC/docker-wordpress
SwiftFS

A Swift-savvy drop-in replacement for Node filesystem.

https://github.com/HackODLRC/SwiftFS
Actions
- View Details and Locations
- Edit Space
- Create Location here
- Delete Space

Access Protocol
Swift
Path
/
Last Verified
None
Username
toronto_admin
Container
artefactual2
Region
Tenant
toronto
Auth url
http://142.150.191.223:5000/v2.0/
Password

Auth version
2
Actions
- View Details and Locations
- Edit Space
- Create Location here
- Delete Space
Status

- Beta!
- Develop end-user tools
- Repository integration
- Compute cluster and text mining
Acknowledgements

- Our Partner Libraries
- GTA.net – Doug Carson, Lloyd Kwong, Kevin Wong
- ORION – Andy Lam, Mark Grant
- OLRC Admin & Tech Committees
- SP/UTL Systems teams – Steve Baroti, Chris Crebolder, Miki Wong, Harpinder Singh, Bikram Singh
Interested in Learning More or Getting Involved?

cloud@scholarsportal.info

https://spotdocs.scholarsportal.info/display/ODLRC