Ontario Library Research Cloud

future considerations & cost models

Dale Askey - @daskey
McMaster University
Hamilton, Ontario, Canada
Roadmap

---

History & brief technical overview

Current state

Future considerations and issues

Cost model and cost questions
Acronyms & names, decoded

OLRC (oh-lurk) = Ontario Library Research Cloud, built by 11 OCUL partners

OCUL (oh-cool) = Ontario Council of University Libraries - 21 total members

Scholars Portal (SP) = service of OCUL, shared collections and technology infrastructure, hosts/operates OLRC core
History

2012 - idea articulated, original proposal drafted
2013, summer - ready to start, but had a “funding accident”
2014 - purchase, deployment
2015 - final testing, production launch in October
2016 - the fun part starts
Multi-site, widely distributed

~500 km between most distant nodes
Technical overview

Dell PowerEdge R720xd heads & MD 1200 storage shelves
Each carries 12x4TB = 48 TB
ORION (Ontario-based ‘Internet2’) VLAN between nodes
OpenStack Swift object storage
Parent node: SP data centre @ Univ. of Toronto
Current state

1.2 PB (net)

220 TB dedicated for original 11 partners

~1 PB - used by SP, U of Toronto to support digital scholarship / text analysis needs / other projects / other institutions
Work in progress

---

Many partners just getting started with transfers

McMaster: 14 TB, but in testing phase

Bit of trial and error with tools, learning to use object storage
Future plans, considerations, and questions

Tools - we need to develop tools and documentation for existing tools

What to do with (large chunks) of the petabyte?

Repository / Archivematica integration
OCUL Storage Cloud Access Scenarios

Use Cases
- SP Journals / SP Books
- SP News portal
- Backup
- Preservation
- Text Mining
- Projects (Big files, many files, need for sharing)
- Big Research Data
- Dropbox services for individuals

Direct Access
- SP System Administrators
  - Create Swift accounts for
  - Library System Administrators
    - Create Swift accounts for
  - Project Staff (Library and Research)

Indirect Access
- Library Staff
  - Account management via Shibboleth / LDAP / Self registered
  - Indirect access to Storage Cloud via other services
- Researchers
  - To Be Developed
Possible use cases

---

Future:
Text mining / analysis
SP Journals / Books / GeoPortal
Projects - sharing large/many files
Large research data sets
‘Dropbox’ services for individuals

Now:
Backup
Preservation
1. Researcher manages active data in cloud (e.g. via ownCloud → Swift)

2. Researcher sends selected datasets to Dataverse for publication

3. Dataverse uses block storage now but support for S3/Swift is in development

4. Datasets transferred to Archivematica using automation tools and Dataverse APIs

5. Through its Storage Service Archivematica can use Swift as a transfer source and for DIP and AIP storage

6. AIPs stored in cloud and managed in Digital Asset Management System
Questions in search of answers

When do we let others access the storage?
Who, i.e.- how broadly do we throw open the door?
What do we charge them? Ourselves?
Do we sell directly, or via partners?
Will others replicate our model?
Costs: some good news

---

Scholars Portal: established and capable
Dedicated partners; others ready to step in
OCUL governance model and history
Costs: who and how much?

Who:

- original partners
- other OCUL institutions
- other non-Ontario institutions
- researchers
Costs: model needed

Relied extensively on Ratliff and Goldstein (2010):

*DataSpace: A Funding and Operational Model for Long-Term Preservation and Sharing of Research Data*

http://arks.princeton.edu/ark:/88435/dsp01w6634361k
Costs: assume it’s going to rain

---

Calculated in partner costs, even if those are not currently levied

Conservative estimate of annual decrease in storage costs

Five-year hardware refresh cycle
1 CAD ≠ 1 USD

1 CAD = .75 USD, .69 EUR, .53 GBP, 18.6 CZK
Estimated per TB costs

(Include a 20% surcharge to fund expansion on top of cost recovery)

Annual subscription: $416/TB

Pay once, store forever: $2,594/TB
Last thoughts

---

We collaborated to build it; now we need to collaborate on using it and developing tools.

Need to let others use it sooner rather than later.

Need to solidify future governance within OCUL.
Thank you

Much credit due to:

- Alan Darnell, Amaz Taufique, Steve Marks
- partner libraries
- OLRC Admin & Tech Committees
- SP/U of Toronto systems teams: Graham Stewart, Steve Baroti, Chris Crebolder, Miki Wong, Harpinder Singh, Bikram Singh