OSiRIS: A Distributed Storage and Networking Project Update

OSiRIS Open Storage Research Infrastructure

Shawn McKee[1] (presenting), Ben Meekhof[1], Martin Swany[2], Ezra Kissel[2], Andrew Keen[3]

for the OSiRIS Collaboration

University of Michigan[1], Indiana University[2], Michigan State University[3]

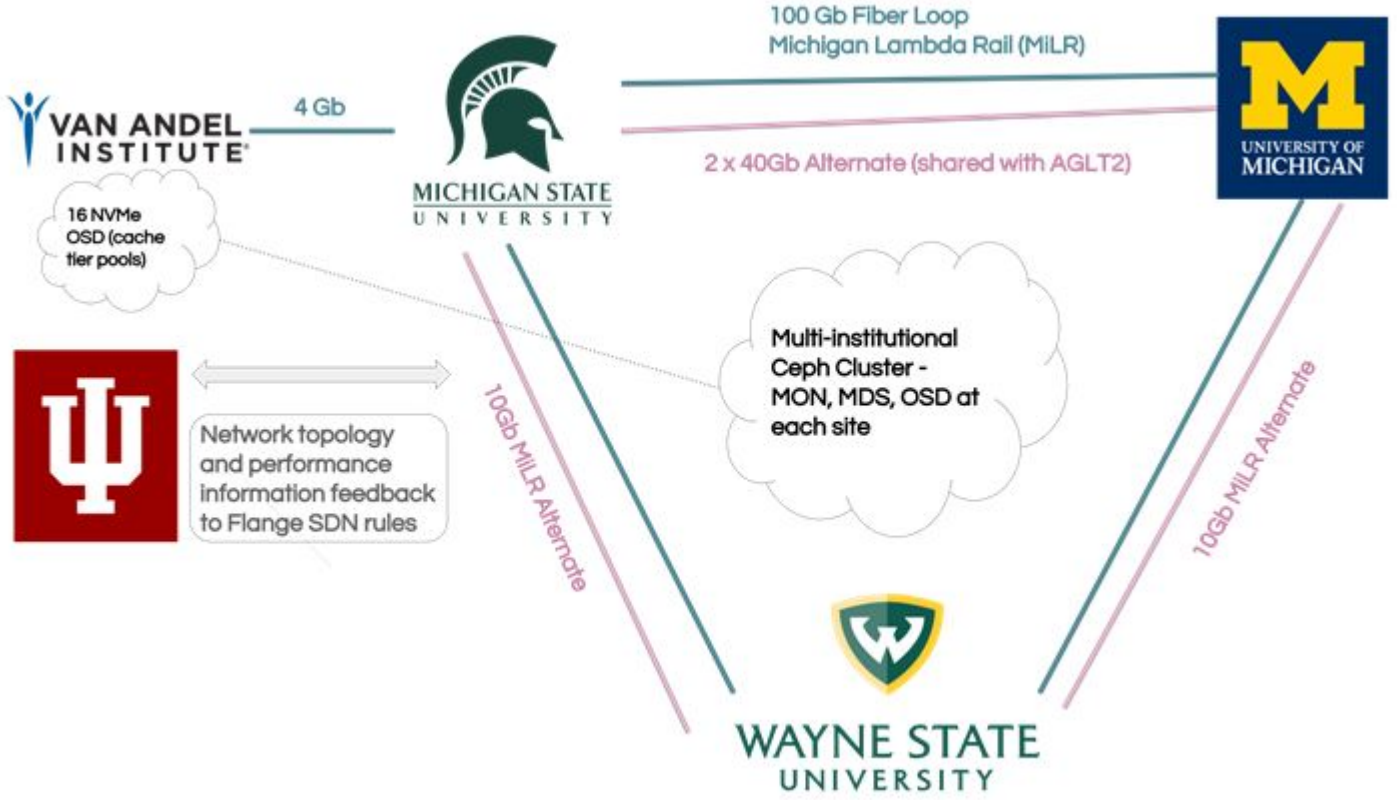
Track #4 at CHEP 2019 in Adelaide Australia

November 4, 2019

OSiRIS Overview

The OSiRIS proposal targeted the creation of ***a distributed storage infrastructure, built with inexpensive commercial off-the-shelf (COTS) hardware, combining the Ceph storage system with software defined networking to deliver a scalable infrastructure to support multi-institutional science.***

Current: Single Ceph cluster (**Nautilus 14.2.4**) spanning **UM**, **WSU**, **MSU** - **840** OSD / **7.4** PiB, **adding 9.6 PiB** in next month)



OSiRIS Science Domains

The primary driver for OSiRIS was a set of science domains with either big data or multi-institutional challenges.

OSiRIS is supporting the following science domains:

- **ATLAS** (high-energy physics), **Bioinformatics**, **Jetscape** (nuclear physics), **Physical Ocean Modeling**, **Social Science** (via the Institute for Social Research), **Molecular Biology**, **Microscopy, Imaging & Cytometry Resources**, **Global Night-time Imaging**
- We are currently "on-boarding" new groups in **Genomics and Evolution** and **Neural Imaging**
- **Primary use-case is sharing working-access to data**

Recent Science Domains

Brainlife.io (Neuroimaging) - Brainlife organizes neuroimaging data and data derivatives using their registered data types. No single computing resources has enough storage capacity to store all datasets, nor reliable enough so that user can access the data when they need them. They will depend on OSiRIS to store datasets and transfer data between computing resources.

OAKLAND UNIVERSITY

Oakland University - Already a user of MSU iCER compute resources, OU will leverage OSiRIS to bring their data closer for analysis and for collaboration with other institutions.

Evolution - Large-scale evolutionary analyses, primarily phylogenetic trees, molecular clocks, and pangenome analyses

Genomics - High volume of human, mammal, environmental, and intermediate analysis data

New and Ongoing Collaborations

Open Storage Network - We will be providing ~1 PB to be included in the Open Storage Network (https://www.openstoragenetwork.org)

- Timeline depends on OSN readiness to engage, some discussions at recent OSN group meeting at TACC

FABRIC - This is a newly funded NSF project to create a network testbed at-scale (1.2 Tbps across the US). OSiRIS will be an early adopter/collaborator, providing ~1 PB to support science use-cases

Library Sciences - OSiRIS roadmap plans for data lifecycle mgmt

- Following detailed analysis of two specific datasets, library scientists at UM are working on automated metadata capture and indexing
- Integration with U-M 'Deep Blue Data' archival system also planned

Recent Upgrades - 100Gb MiLR

MiLR is a high-speed, special purpose, data network built jointly by Michigan State University, the University of Michigan, and Wayne State University, and operated by the Merit Network.

Thanks to combined effort from campus network teams and Merit we were able to deploy direct 100Gb links via MiLR fiber landing directly on our OSiRIS rack switches

- Now we have more options for network management without campus network disruptions

In our first phase of deployment they carry only the Ceph 'cluster network' used for OSD replication data

Normal ceph recovery/backfill operations could easily overwhelm smaller links with this traffic, so removing it was a **huge difference** that let us completely remove throttles on Ceph recovery (see next slide)

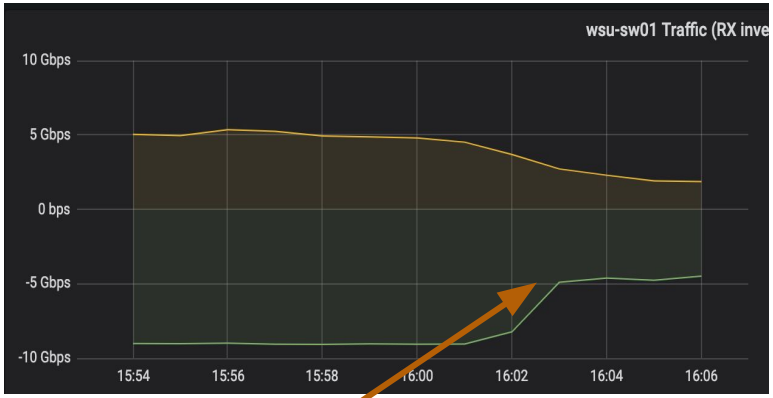
Unbalanced Networks and Ceph

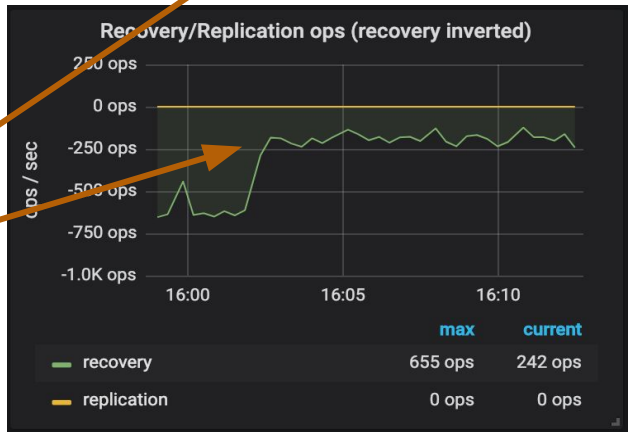
Prior to our installation of 100G links for Ceph cluster backend we had issues with network bandwidth inequality: UM and MSU sites had 80G link to each other but 10G to WSU datacenter

- Adding a new node, or losing enough disks, would completely swamp the 10G link and cause OSD flapping, mon/mds problems, service disruptions

Lowering recovery tunings fixed the issue, at the expense of under-utilizing our faster links. Recovery sleep had the most effect, the others not as clear

```
osd_recovery_max_active:  1       # (default 3)
osd_backfill_scan_min:  8         #(def 64)
osd_backfill_scan_max:  64        #(def 512)
osd_recovery_sleep_hybrid:  0.1   # (def .025)
```

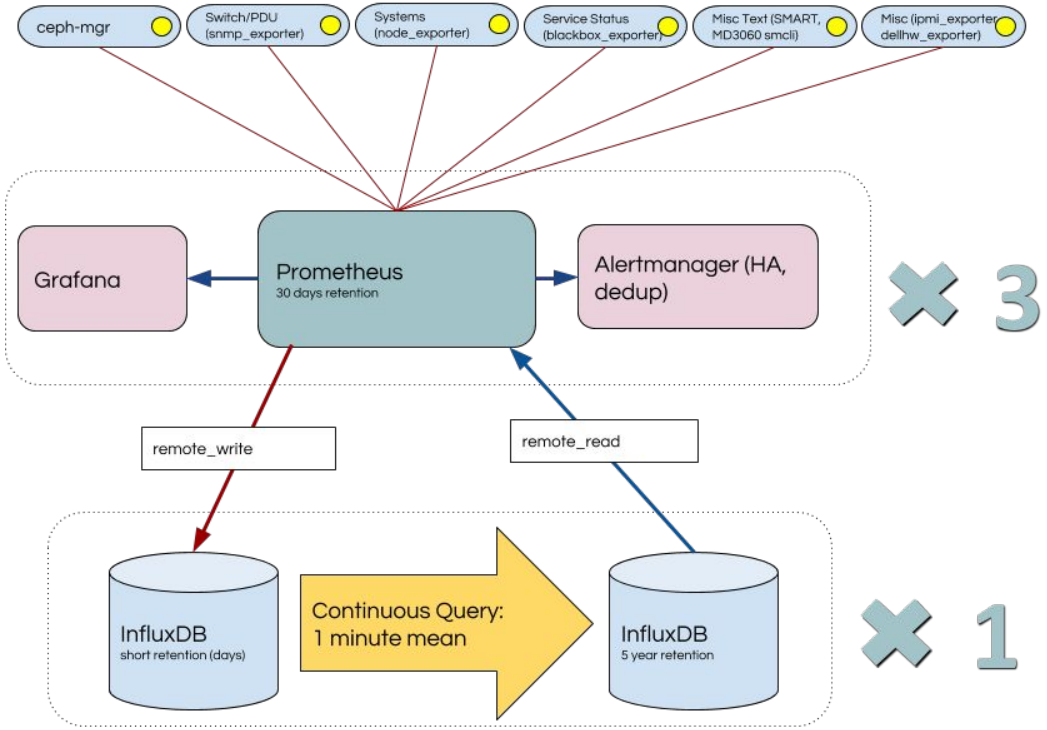




Monitoring and Metrics with Prometheus

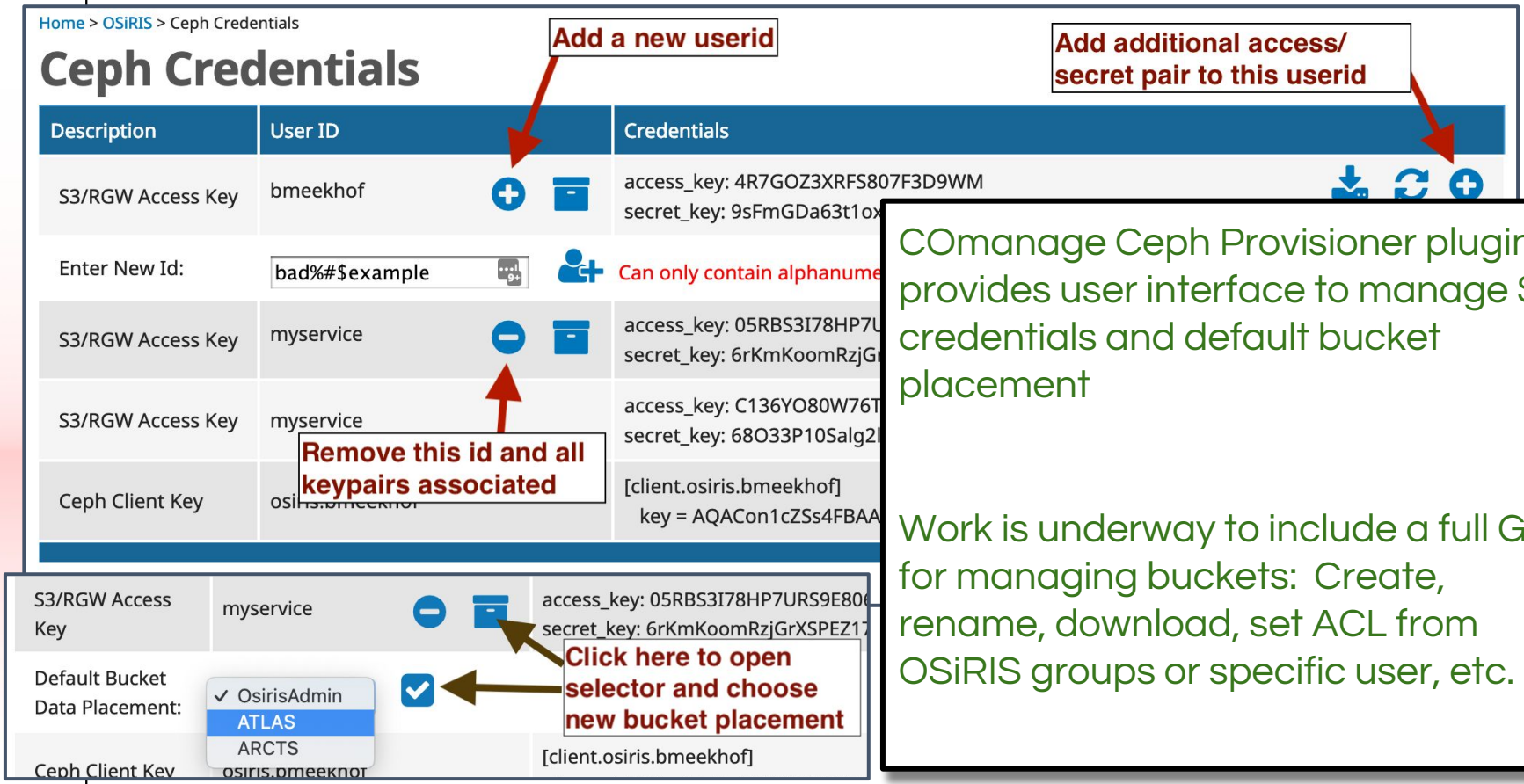
Recently we consolidated all of our metrics, monitoring, alerting to Prometheus

- Migrated from a combination of Check_mk, Influxdb, Collectd
- Continue to use Grafana to visualize, Influxdb for long-term retention
- Consideration was given to standing up more of the influx (TICK) stack, pros and cons each way
- Text collector scripts and alert rules in our git repo (grafana dashboards soon)



https://github.com/MI-OSiRIS/osiris-monitoring

COmanage Credential Management



COmanage Ceph Provisioner plugin provides user interface to manage S3 credentials and default bucket placement

Work is underway to include a full GUI for managing buckets: Create, rename, download, set ACL from OSiRIS groups or specific user, etc.

S3 Fuse Client Bundle

Technically S3 storage makes more sense for most use cases wanting to compute with OSiRIS storage from campus or off-campus locations

- But…not everyone is very familiar with S3
- People often think we are telling them to go use Amazon just by saying S3

We try to make it a little easier by putting together a bundle that automatically FUSE mounts their S3 buckets with s3fs-fuse utility

- Includes setup script, user plugs in credentials
- Auto-detects which OSiRIS S3 endpoint URL is reachable and passes to mount command (our campus cluster users may only be able to reach on-campus endpoint)
- Includes build of s3fs-fuse util made with appimage to be portable to any Linux system.
- https://github.com/MI-OSiRIS/osiris-bundle
- http://www.osris.org/documentation/s3fuse.html

Globus and Gridmap

We provide Globus access to CephFS and S3 storage

- Ceph connector uses radosgw admin API to lookup user credentials and connect to endpoint URL with them

Credentials: CILogon + globus-gridmap

- CILogon DN in LDAP voPerson CoPersonCertificateDN attribute

We wrote a Gridmap plugin to lookup DN directly from LDAP (student project)

- https://github.com/MI-OSiRIS/globus-toolkit/tree/gridmap_ldap_callout_final
- https://groups.google.com/a/globus.org/forum/#!topic/admin-discuss/8D54FzJzS-o

Having the subject DN and lookup entirely in LDAP means it will be easy to add capabilities to COmanage so users can self-manage this information

- Users already self-manage SSH login keys in COmanage (also in LDAP)

Network Management

The OSiRIS Network Management Abstraction Layer is a key part of the project with several important focuses:

- Capturing site topology and routing information from multiple sources: SNMP, LLDP, sflow, SDN controllers, and existing topology and looking glass services
- Converge on common scheduled measurement architecture with existing **perfSONAR** mesh configurations
- Correlate long-term performance measurements with passive metrics collected via other monitoring infrastructure

Recently wrote new Prometheus exporter to collect perfSonar test results from central ESmond store for alerting and visualization

We will demo SDN architecture for traffic routing and traffic shaping / QOS (prioritize client / cluster service traffic over recovery) at SC19

NMAL work is led by the Indiana University CREST team

Summary

OSiRIS continues to improve on our user experience and engage with new collaborators

- ATLAS has been a long time user for Event Service data

Our new hardware purchases this year will increase our node count and make EC pools more efficient

We look forward to participating in more national scale projects such as the Open Storage Network, FABRIC, Eastern Research Network

On our roadmap this year:

- Make our S3 services more highly available with LVS failover endpoints on each campus
- Make S3 services more performant by greatly increasing instance count behind the proxy endpoints
- Improve user GUI for managing storage access
- Build more convenient client bundles, modules, etc to make OSiRIS usage as easy as possible
- Adding **ATLAS dCache storage** to explore using Ceph to manage back-end storage.

Acknowledgements

We would like to thank our **OSiRIS science partners** and our host institutions for their contributions to work described.

In addition we want to explicitly acknowledge the support of the **National Science Foundation** which supported this work via:

- OSIRIS grant, NSF OAC-1541335

Questions or Comments

Questions?

Additional Slides Follow

Backup Slides

Summary of the OSiRIS Deployment

We have **deployed 7.4 pebibytes (PiB) of raw Ceph storage** across our **three research institutions** in the state of Michigan.

- Typical storage node is a 2U headnode and SAS attached 60 disk 5U shelf with either 8 TB or 10 TB disks
- Network connection is 4x25G links on two dual port cards
- Ceph components and services are virtualized
- **Year-4 hardware coming**: **33** new servers (11/site) adding 9.6 PiB (for EC)

The **OSiRIS infrastructure is monitored by Prometheus** and configuration control is provided by Puppet

Institutional identities are used to authenticate users and authorize their access via CoManage and Grouper

Augmented perfSONAR is used to **monitor and discover the networks** interconnecting our main science users.

COmanage - Virtual Org Provisioning

When we create COmanage COU (virtual org):

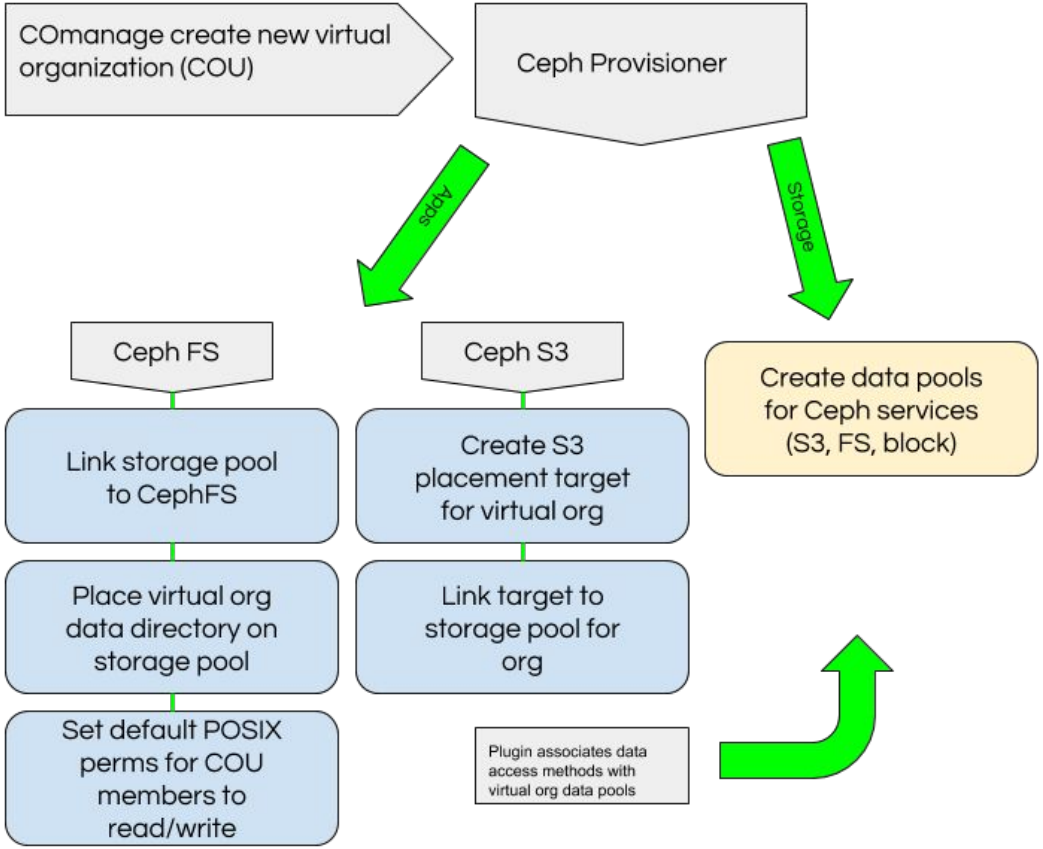
Data pools created

RGW placement target defined to link to pool cou.Name.rgw

CephFS pool create and added to fs

COU directory created and placed on CephFS pool

Default perms/ownership set to COU all members group, write perms for admins group (as a default, can be modified)



Ideal Facilities

If we could have our way, we would have **ideal facilities**:

- CPUs would always be busy running science workflows
- Any data required would always be immediately available to the CPU when needed
- (Oh, and the facilities would be free and self-maintaining and use negligible power!)

As we all know, it is hard to create efficient infrastructures that manage access to large or distributed data effectively

Approaching “ideal” becomes very expensive (in $’s and effort)

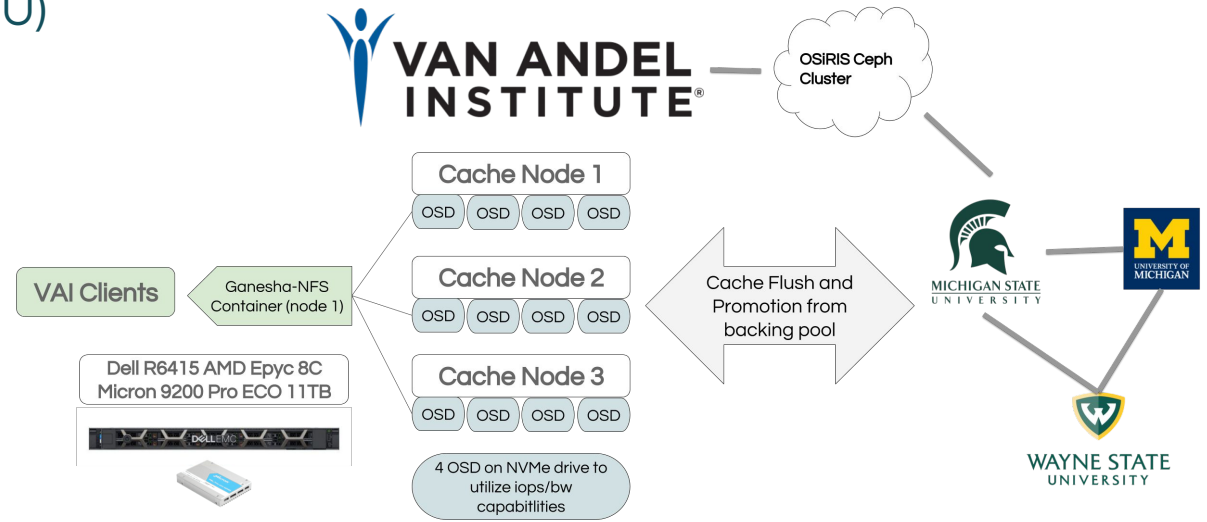
So we need to make progress as best we can.

OSiRIS Ceph Cache Tiering

At Supercomputing conferences (2016/17/18) we've experimented with Ceph cache tiering to work around higher latency to core storage sites

- Deploy smaller edge storage elements which intercept reads/writes and flush or promote from backing storage as needed

Have edge OSiRIS site leveraging this technique at Van Andel Institute (primarily led by MSU)



OSiRIS Topology Discovery and Monitoring

UNIS-Runtime release integrated into ZOF-based discovery app

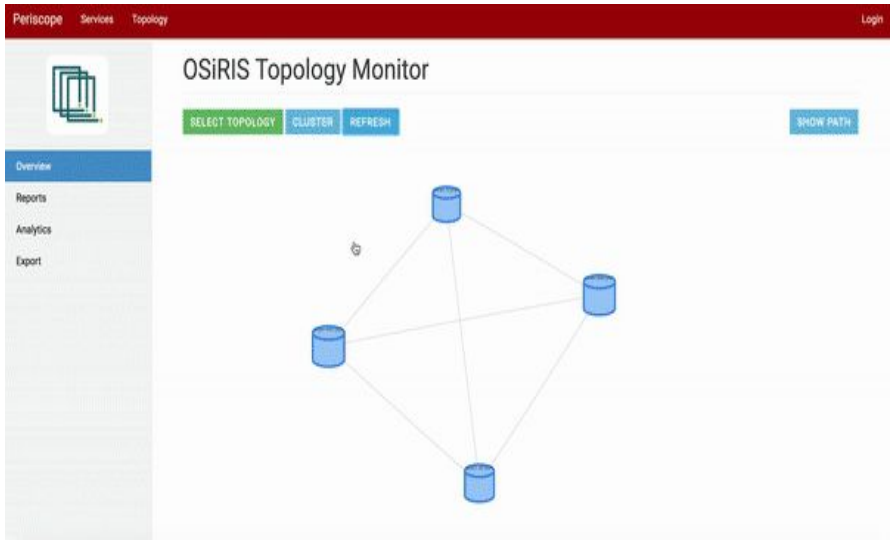
- Increased stability and ease of deployment
- Added extensions for Traceroute and SNMP polling

Web Development has focused on bringing measurements to dashboard

- Link and node highlighting with thresholds determined by link capacities
- Overlay for regular testing results to bring “at-a-glance” diagnostics

Filtering to show layer-2 topology versus layer-3 and virtualized components

- Fault localization, clustering, and zoom are work-in-progress



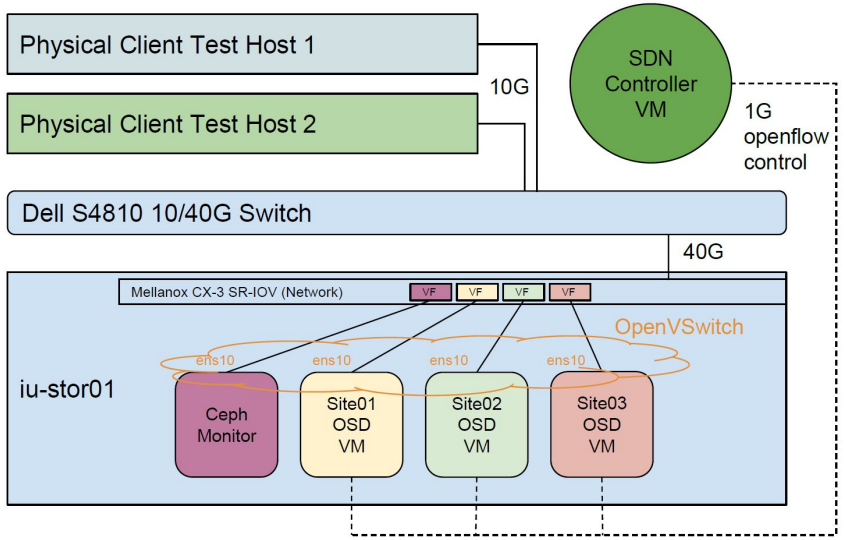
OSIRIS: Quality of Service for Ceph

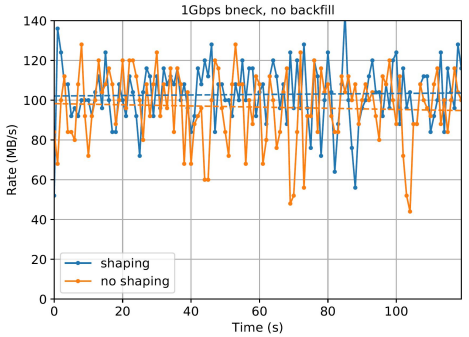
Testbed created to develop QoS functionality
- Explicit control of operations, no noise
- Reduce risk of breaking production

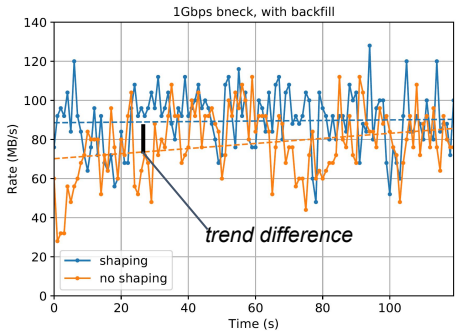
Apply priority queues to ensure that adequate bandwidth exists for Ceph client operations to prevent timeouts and delayed read/write performance

Apply traffic shaping to provide better transport protocol performance between sites with asymmetric link capacities. This is of particular importance when latency between sites is increased

Preliminary results: shaping from sites towards bottleneck can improve client performance, approx 5-10% in early testing.







OSiRIS Lesson's Learned

OSiRIS works very well on a regional scale (networking RTT ~< 10 ms)

We explored scaling for a single Ceph cluster at SC16 where we dynamically added a new site on the exhibition floor 42 ms RTT from the rest of OSiRIS

- The benchmark work-flow data access dropped from **1.2 GB/sec** to **0.45 GB/sec**
- The infrastructure continued to work without problems

Using 'netem' we were able to programmatically add arbitrary delay into the network stack of one of our Ceph servers.

- As we increased the latency we saw the expected impact in throughput
- When we reached **160 ms**, our (untuned) Ceph cluster stopped working
- We needed to decrease the latency back down to **80 ms** to recover

To reach more distributed deployments, OSiRIS would need to start using Ceph Federations (with associated costs) or employ caching to "hide" the latency as much as possible. In DOMA terms, OSiRIS would be appropriate as an element of a data lake.

Further Information

OSIRIS

- http://www.osris.org project website
- Details in various presentations at http://www.osris.org/publications

IRIS-HEP

- https://iris-hep.org/ project website
- Details in various presentations at https://iris-hep.org/presentations/bymonth

DOMA

- https://iris-hep.org/doma.html sub-project website
- DOMA Presentations are available at the above URL

Some Caching studies

https://indico.cern.ch/event/770307/contributions/3301625/attachments/1807559/2952167/Scheduling_with_Virtual_Placement_for_Site_Jamboree.pdf