

OSiRIS

Distributed Ceph and SDN for Multi-Institutional Research



**ADVANCED
RESEARCH COMPUTING
TECHNOLOGY SERVICES**
UNIVERSITY OF MICHIGAN

Ben Meekhof
University of Michigan / ARC-TS
Supercomputing - November 2016

What is OSiRIS?

OSiRIS combines a [multi-site Ceph cluster](#) with [SDN](#) and [AAA infrastructure](#) enabling scientific researchers to efficiently access data with federated institution credentials.

The current OSiRIS deployment spans Michigan State University, University of Michigan, and Wayne State University. Indiana University is also a part of OSiRIS working on SDN network management tools.

OSiRIS Goals

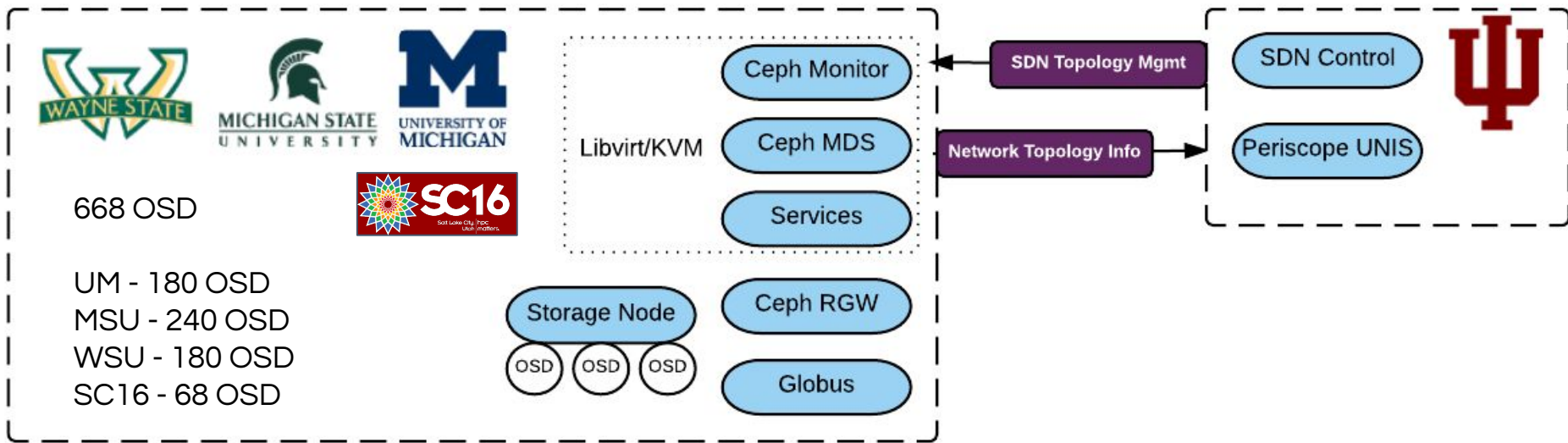


The OSiRIS project goal is enable scientists to collaborate on data easily and without building their own infrastructure.

We have a wide-range of science stakeholders who have data collaboration and data analysis challenges to address within, between and beyond our campuses.

High-energy physics, High-Resolution Ocean Modeling, Degenerative Diseases, Biostatics and Bioinformatics, Population Studies, Genomics, Statistical Genetics and Aquatic Bio-Geochemistry

Our Deployment

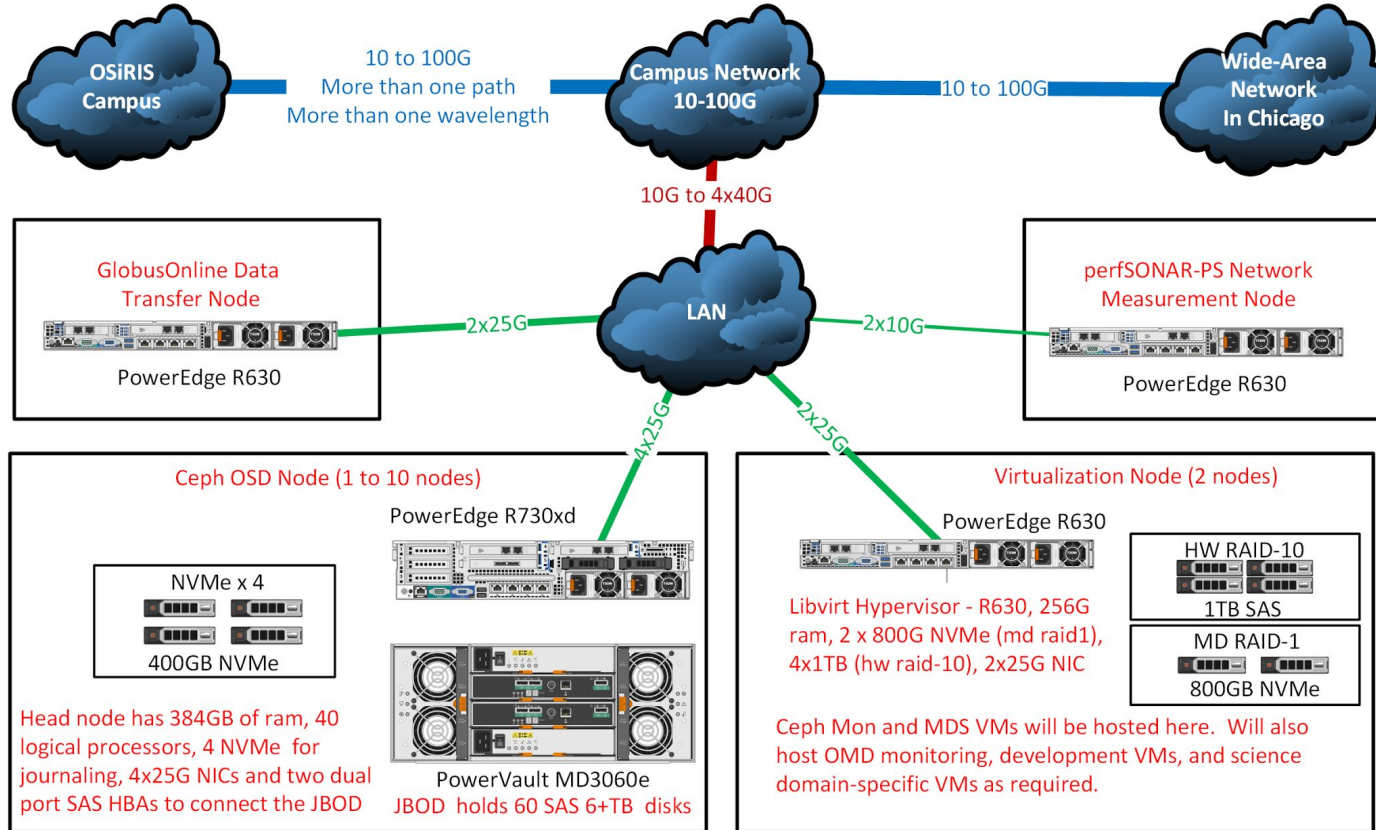


Each site has a monitor, mds, one or more RGW nodes, multiple storage nodes (more every year)

MDS in active-standby

Our Deployment

OSiRIS Data Infrastructure Building Block

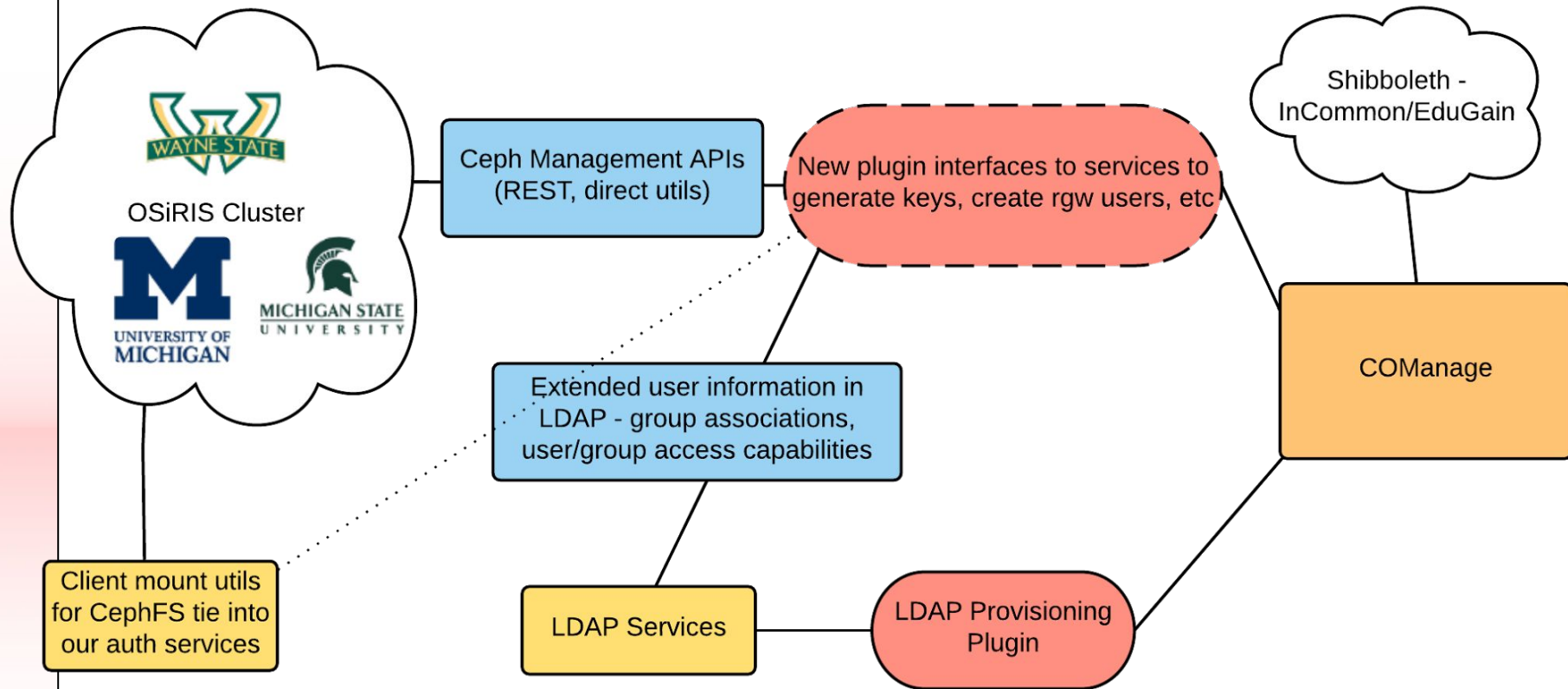


Reality today...

UM-MSU: 2x40GB

WSU: 10GB

Direct links bypassing campus networks leveraging Michigan Lambda Rail (MiLR)



Science Engagement



We are currently working with ATLAS to enable OSiRIS as an 'event service' via S3 gateway.

ATLAS data will be mapped to pools at UM and MSU and accessed via S3 gateways at those sites (as many as it takes!). Allocating 500TB usable, hope to use EC pool.

We are also working with an ocean modeling group at UM which seeks to make data collected by US Navy available to wider scientific collaboration. About 5TB usable, CephFS access.

End

Questions?