# SC18 Network Research Exhibition: Demonstration Details

## "Supporting Scientific Data-Intensive Collaborations with Distributed Storage and Edge Services"


Jeremy Musser, Ezra Kissel, Grant Skipper, Martin Swany
**Indiana University**
{jemusser, ezkissel, swany}@iu.edu

Joe Breen, Jason Stidd,
**University of Utah**
{joe.breen, jason.stidd}@utah.edu

Kenneth Merz, Andrew Keen, Charlie Miller, Nick Rahme
**Michigan State University**
{merzjrke, keenandr, cdmiller, rahmenic}@msu.edu

Shawn McKee, Benjeman Meekhof, Nicholas Grundler, Amy Liebowitz, Gabriele Carcassi, Bob Killen, Wenjing Wu
**University of Michigan**
{smckee, bmeekhof, grundler, amylieb,carcassi,rkillen,wuwj}@umich.edu

Patrick Gossman, Matthew Lessins, Carlo Musante, Michael Thompson
**Wayne State University**
{pgossman, mjl, carlo, michael}@wayne.edu

Rob Gardner, Lincoln Bryant, Chris Weaver, Iljia Vukotic
**University of Chicago**
{rwg, lincolnb, cweaver}@uchicago.edu



### Abstract

Scientific collaboration with large or diverse data can be a challenging problem which diverts time from core research goals. Excellent network connectivity is central to enabling effective infrastructures supporting science. The University of Michigan and partners at Michigan State, Wayne State, Indiana University, University of Chicago, University of Utah and others will be demonstrating capabilities in this area from the OSiRIS, SLATE and AGLT2 projects.


### I. Overview

We have a broad range of demos we are developing for our various research projects as summarized below. For network details, see the two diagrams (Figures 1 and 2) in Section II below that show the logical and physical network infrastructure that will support our demonstrations.

#### AGLT2

Will participate in state of the art shaped-flow data transfers across 100G paths, showing the impact of flow shaping via OVS/OVN on end-to-end flows between the SC18 exhibition floor and remote data transfer sources and sinks. The dCache infrastructure at AGLT2 is running Open vSwitch version 2.9.1 and we will try to compare the impact of shaping flows

between the SC18 exhibition floor and AGLT2 vs regular data transfers. This demo may be using in conjunction with the SLATE XCache demo when accessing LHC data from AGLT2.

### OSiRIS

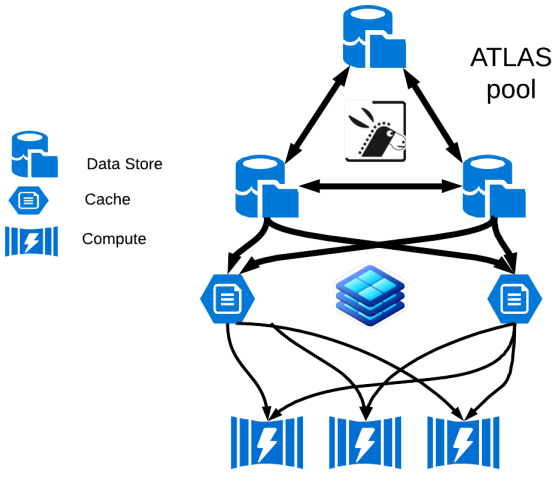
Our demonstration will encompass our use of storage features, advanced networking, identity management/onboarding, and include a collaboration with SLATE infrastructure utilizing OSiRIS storage.

- **Storage Features:** Typical Ceph storage clusters are heavily affected by network latency between components and clients. As such a client accessing OSiRIS storage in Michigan from the SC showroom would not see similar performance to our more proximate users. OSiRIS will deploy dedicated storage hardware for a Ceph cache-tier at SC18 to show that it is possible to achieve reasonable performance even though the backing data pool components are geographically quite far away. A similar demonstration optimized for read-only will be done using a technique to configure pools such that they direct all reads to the storage component of our choice. We will compare/contrast the two as they might apply to various scenarios. The collaboration with SLATE will also leverage these configurations and give us experience with a new use case.
- **Advanced Networking:** The OSiRIS Network Management Abstraction Layer (NMAL) seeks to orchestrate network pathing and quality of service based on real-time feedback. At SC18 we plan to demonstrate the impact of these intelligent network capabilities on the client experience.
- **Identity Management / onboarding:** OSiRIS virtual organizations and the people within those organizations are managed by Internet2 COmanage and a collection of plugins we've written to tie together with Ceph storage. For SC18 we'll spawn one or more virtual organizations for the show and demonstrate self-service onboarding and credential management linked to existing institutional federated identities. Our ideal scenario would be if other demonstrators at SC18 made some active use of OSiRIS storage in scenarios similar to real-world usage. Here again we can somewhat refer to the collaboration with SLATE as they will obtain storage allocation and credentials via this infrastructure.

### SLATE

Demonstrate adding a new SLATE managed resource on the SC18 exhibition floor to an already existing SLATE infrastructure. Show how the new resource performs and how it can be utilized by two or more science domains.

- Distributed caching network using ATLAS data from the Large Hadron Collider:
  - Analysis and simulation processing data from the ATLAS Midwest and Great Lakes Tier2 centers (MWT2, AGLT2) are marked for caching



    - The Rucio data management service[1] is used to identify data in the ATLAS global namespace
    - Xrootd-based caching services are deployed on edge clusters in the University of Michigan and University of Utah SC18 booths, federated using SLATE.
    - Data from the ATLAS global storage pool is accessed from cloud and grid resources via caches at SC18
- Demonstration of a "sandbox" service which emulates deployment of containerized applications on a federated Kubernetes infrastructure using the SLATE API and command line interface.
- Demonstration of MiniSLATE (https://github.com/slateci/minislate) which provides a local copy of a SLATE-enabled Kubernetes federation for local application development
- Demonstration of SLATE-OSIRIS collaboration using tools to manage flows on the I2 SDN testbed and to discover topologies between SLATE instances. (*NOTE: this configuration will only use TCP/IP to access SC18 resources).* The following two figures represent the national topology of the SLATE-OSIRIS collaboration.

Fig: National topology representation

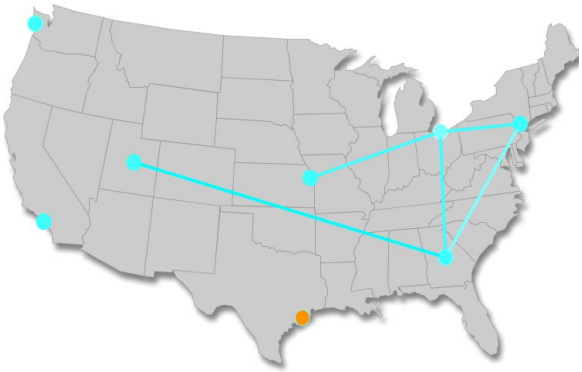
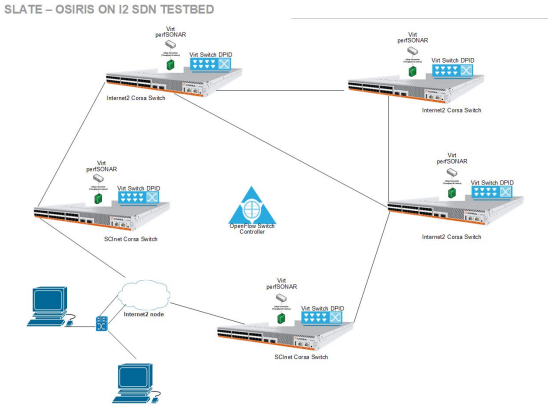
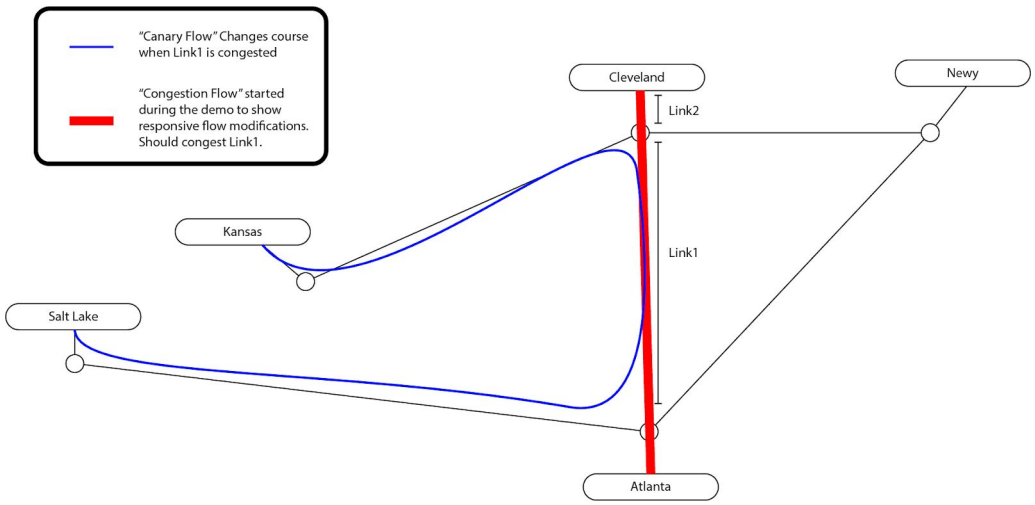


Fig: Physical topology representation




This demonstration is a first phase that allows the OSIRIS-NMAL Flange application running on SLATE to accept specific network characteristics and to orchestrate a path in response to that input. Flange also responds if the network characteristics on the existing path change in order to try to preserve the requested path characteristics. The following Figure describes this action.

[1] http://rucio.cern.ch




The Flange plug-in is able to work with other orchestration techniques beyond OpenFlow. This demo will highlight only the OpenFlow work.

## II. Network Topology

Our demos will use a 100G path from the UM/MSU booth (#1204) directly to SCinet, with VLAN connections to the Starlight booth (#2851) and the Caltech booth (#1413) so we can participate in other LHC-related demonstrations. From SCinet, we will connect to a 100G path back to Starlight, then onto a new SC18 100G MiLR wave back to the University of Michigan where we will have access to the various relevant OSiRIS front-end networks at UM, MSU and WSU as well as the SLATE and AGLT2 resources at UM and MSU.

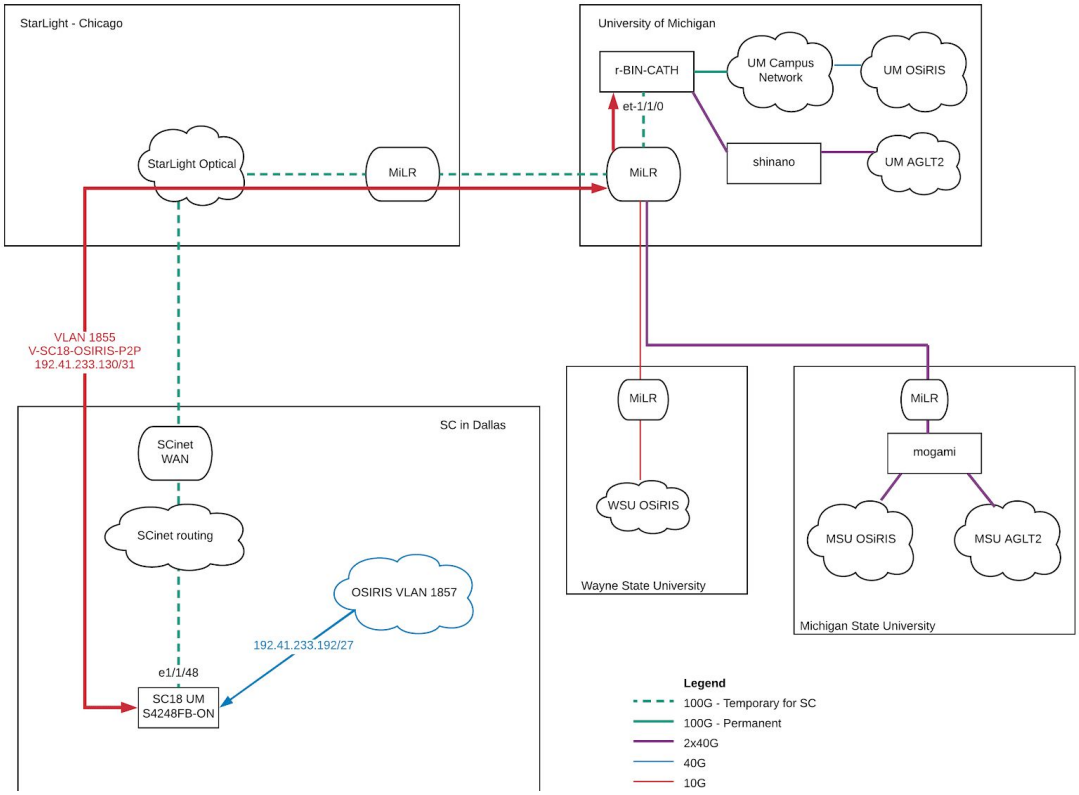



Figure 1: Physical Network Connectivity from UM/MSU Booth

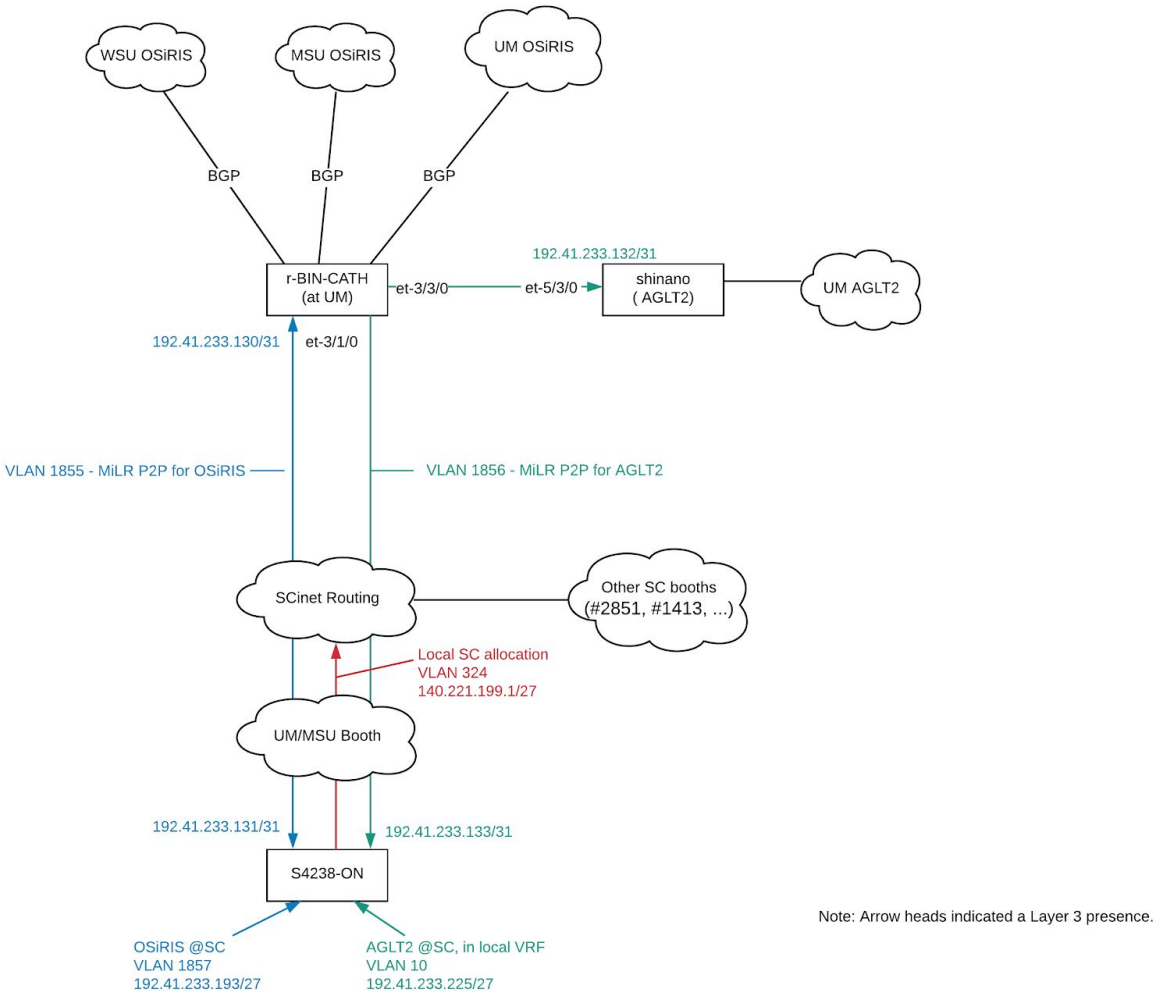



Figure 2: Logical SC18 network connectivity from UM/MSU Booth

## References *(Optional)*

AGLT2: http://www.aglt2.org
OSiRIS: http://www.osris.org
SLATE: http://slateci.io/