

**Internet2 AL2s
6 Month Report
10-1-12 through 3-15-13**

Introduction

This is the first of a series of periodic reviews of the Internet2 AL2s infrastructure. The goal is to examine what is working and what isn't with an eye toward improvements and a focus on information important to the community for their decision-making process.

Report Overview:

Upgrade Details

Maintenance History

Availability Information

Future Plans for Layer 2

Commentary from User: OARnet

Commentary from User: Indiana University

Commentary from the Test Lab

Commentary from the Support Engineers

Timeline of Installs/Upgrades

Node & Circuit Availability

Upgrade Details

A number of upgrades have been done to the core node software as well as the OE-SS application. Feature additions and bug fixes are the primary reason, although the DR Failover test is indicative of the sorts of proactive work that the system should be moved towards.

2012-10-20 – Brocades Upgraded 50 5.4.0.a(Private)

This code drop resolved two bugs. FlowStats, used to gather data on a per flow rule basis, was reporting incorrect data. In addition the 'default drop' rule was being inserted in the wrong order in the CAM table. This code drop resolved these bugs.

2012-11-08 – Controller DR Failover Test

This was a test to ensure that the resilient controller infrastructure would be failed over to correctly if the primary controllers ever became unreachable. It was a success.

2012-12-08 – OESS Upgraded to 1.0.5

This releases added support for hairpin VLANs and fixed a bug affecting some inter-domain provisioning requests.

2013-01-06 – Brocades Upgraded to 5.4.0a.a

This code drop resolved bugs in drop rule processing.

2013-02-04 - Brocades upgraded to 5.4.0b

The 5.4.0b code allowed I2 to deploy vlans in addition to SDN on the same card. It also greatly improved the flow mod processing speed.

2013-02-23 – OESS upgraded to 1.0.6

This code upgrade contained support for the Junipers as well as a feature to allow for 'insertions' of nodes between two other nodes during a new node install. An IDC bug fix was also included.

Maintenance Data

A report was run in the I2 NOC's ticketing system asking for all Scheduled and Unscheduled outages of the layer 2 infrastructure. It was anticipated that a large number of outages would be L1 circuit issues, either Scheduled or Unscheduled, with a smaller number of scheduled L2 outages related to software updates. It was not anticipated that there would be any unscheduled 'crash' outages of the Layer 2 hardware proper.

The report actually indicated that there were 127 outages of the Layer 2 infrastructure. 83 of these were Scheduled events and 44 Unscheduled. This is about a 65%/35% split.

Scheduled Maintenance

Of the 83 Scheduled events 27 (33%) were maintenance on the underlying L1 infrastructure while 45 (54%) were Software Upgrades on the L2 switches. 5 involved customer maintenance, 2 were upgrades on the OESS software. There was one scheduled hardware replacement of a Pluggable Optic and a single DR failover test.

Unscheduled Maintenance

Of the 44 unscheduled maintenances 34 (77%) were related to the underlying L1 infrastructure. Of the remaining 10 events 1 was related to customer maintenance, 9 were related to failing 100g pluggables, 1 was related to a process degrade on the controller, and one was related to a real-life SDN bug.

Failing Pluggables

These outages are clustered mostly around several days in March 2013 and deal with a flapping circuit. Firmware updates on the Brocade 100g pluggables appear to have resolved the issue.

Controller Failing to re-synch rules to node

A bug was detected in OESS where it was failing to provision a new VLAN due to an inability to synch the rules with a switch. It did not impact existing VLANs. Issue was resolve with software reset.

SDN Bug

Customer traffic was disrupted as a result of Brocade bug that gave the wrong priority to a drop rule, resulting in a traffic blackhole. DEFECT000423657 is the issue number with brocade. Expected fix is 5.4.00c.

OESS Circuit States

Some edited circuits would remain in a 'Pending' state in the OESS database rather than be moved to the 'Deployed' status. This would cause them to not failover during maintenance.

Availability

The availability figures are very coarse. The tools currently only allow for a mixed view of the data, combining Scheduled & Unscheduled outages in to one figure. This makes comparisons to commercial providers difficult since they do not include Scheduled outages in their figures. Additionally the vast majority of circuit outages are due to scheduled maintenance on the underlying L1 infrastructure. This tends to skew the circuits figures in to a territory where any other issues are hidden. Similarly, the code upgrades and scheduled maintenance to swap out L3 nodes tends to hide any actual interesting node data.

The community groups are working on a more comprehensive plan for measuring the availability of the infrastructure. It is probably that entire new categories will be necessary because of the separation of the control & data plane. For example, if the existing circuits are available but a new circuit can not be provisioned then is that an outage? An analogous situation would be the inability to create a new L1 circuit because of a failure in the provisioning application that did not impact already provisioned circuits.

Future Plans (April-September 2013)

- The reporting tools clearly need the ability to differentiate between Scheduled & Unscheduled outages to bring I2 in to line with best practices.
- Availability reporting should shift from an individual elements/nod/circuit model to a Services based model.
- Another tracking category is needed. "Impairments", representing something other than a binary "Up/Down" status.
- More granularity is needed in the reporting system to differentiate the layer of the service impact. For example, L2 outages caused by L1 circuit outages.
- Further support training is needed for the engineering and systems support staff.
- Better tools are needed to identify and localize problems. The Layer 2 environment is sufficiently different to require reexamination of the troubleshooting process.
- A better tracking system for major project events (upgrades, major bugs, etc) is under way.
- The community groups are working on a set of metrics to better measure & compare the Layer 2 system.
- The backbone upgrade policy is vague. The only current policy is based on a Layer 3 IP network running at 10gigabit speeds and handling generalized R&E traffic.
- The Layer 3 backbone will begin using Layer 2 (via vlans and SDN-signalled circuits) for 100g transport.
- Openflow 1.3 support is heavily examined on OE-SS, Flowvisor, and the backbone switching nodes, with a desire to move to it quickly to support several gaps in the 1.0 standard.

Raw Commentary from the Users:

OARNET (Paul S)

OARnet was one of the first regional network's to turn up a 100G service on the new Advanced Layer 2 Service (AL2S) service operated by Internet2. While the 100G and AL2S services are operational, it has not been without its challenges. Most of the technical issues have been understandably exacerbated by the nascent nature of AL2S and the immaturity of the operational model. During our turn up of the service there was an outage lasting approximately 10 days and after about a month of operation there was a three day outage.

During our initial turn up, as MERIT was being provisioned, there was an issue with MAC addresses being rewritten to the multicast MAC address, thus rendering them useless. Initially, it appeared there was packet loss, but a packet capture revealed the multicast issue. The problem subsequently appeared on the OARnet connection. Since MAC learning is disabled in the OpenFlow environment, it was very difficult for the NOC staff to run the problem down. The NOC staff performed very well, but the lack of tools and experience hampered their ability to quickly determine the root cause. The root cause was an issue on the interconnect between AL2S and the IP service. The 10 Gbps card on the Juniper is one of the original legacy cards, which are pre-standard.

The second issue was after about a month of operation. The 100G service started seeing packet drops severe enough that BGP was unable to maintain state. OARnet ultimately had to turn down the 100G service. Due to the newness of the service, there were issues in opening a ticket with the NOC and getting the correct information communicated. The NOC was able to roll the service from the circuit it was on (which turned out to be the backup circuit) to a different circuit and service was restored. As of the writing of this, the root cause is still undetermined. It is currently suspected that it is either an optical card issue or a AL2S card issue. Apparently, there have been micro bursts of packet drops but in the this case it was severe enough to render BGP inoperable. The NOC continues to work on the root cause issue.

Indiana University (Tom J)

As the lead Internet2 Network Engineer for AL2S since December of 2012, my experience of the network has improved in depth and confidence. Realistically, the first serious problem I encounter was the OE-SS update which resulted in a bug in not properly comparing the flowmod rules on the switch, the bug inserted many flowmods with the same egress ports causing a lot of packet duplication and in general network outages. After Software Development provided a hot fix, the final step for resolution was to reboot a FPC Type 4 in a Juniper T1600 to restore customer connections

egressing AL2S. Software successfully solved the AL2S problem much earlier but it wasn't until later in the day that we isolated and solved the end-to-end customer "experience" / connectivity.

Since that bug, AL2S has been very stable in my opinion, excluding the 100GE link flapping problem, which is most likely a manufacturing defect. The upgrade to Brocade IronWare 5.4.0b was smooth and has resulted in a huge improvement in flowmod processing and implementation in hardware. My trust in the brocade openflow enabled hardware has improved to the point where I rely on the AL2S nodes not to be at fault and distrust non AL2S elements. Thus far that's been proven true, examples;

1. OARnet - multicast bit being set in Ethernet Frame's, turned out to be vlan-steering on the T1600, a non AL2S device but does affect overall customer service
2. Raleigh - Juniper not supporting 100GE interfaces
3. XSEDE/TACC - problem with Layer1 transport, not Internet2's. Port on AL2S performs correctly
4. BlueWaters - input errors on 100GE customer facing port, found to be failing 100G transponder in customer's optical system

AL2S has been very stable and trustworthy, I have complete confidence that the network performs as expected and anticipated. I think the greatest danger we're facing is control plane software and that has improved greatly over the past month. The newest version of OE-SS should solve the "edited" circuits bug and provide greater robustness by properly utilizing the working/protect path feature of AL2S.

I think the greatest threat will come from network slices, not only risky applications from developers but also the internal process of virtualizing the network, meeting the requirements of the unknown future.

Indiana University (Jeff A)

My impression is mostly positive so far. I only cut over to the AL2S about a month ago so the outage I'm going to describe happened on our backup link to AL2S (we were using our primary 10G to Atlanta). From what I remember (I wasn't brought into the loop til the troubleshooting was over and the issue resolved), software pushed out a new version of OESS and this caused an issue with the Brocade (I2) in Chicago effectively taking down the GigaPOP vlans to I2 (R&E, V6 and TRCPS). The engineer spent all night working the issue and could not pin down the network where the problem was occurring so I never got a call nor did this get escalated until the next morning when other staff (more familiar with how things connect in Chicago) arrived to work and could take over. It was then resolved within 3-4 hours. I guess my first impression is

That it's so complicated that only a select few of us can troubleshoot it or that we need to train our engineers more extensively (so that they aren't so confused that they don't know who to call for help). Or maybe everyone is up to speed now? Either way, since that incident I have had no problems whatsoever and am completely happy with the service.

Raw Commentary from the Test Lab (A):

1. The Testlab has seen major enhancements over the last 6 months. This includes improvement both in terms of deployed hardware and in terms of enhancements made to the automated test suite that is used to catch many of the possible issues that can be introduced with new versions of vendor hardware/software and of the SDN control software used on the network (OESS, FlowVisor). The testlab has grown from just a few NEC and Brocade switches to include several additional vendors (Cisco, Juniper, Dell) and now allows for dynamic, automated network topology changes. This infrastructure allows us to test new and existing OESS code on new revisions of the vendors software, helping to minimize problems when upgrading, and allows for debugging and troubleshooting of issues seen in the production environment.

2. The AL2S testlab is of vital importance to the successful continued development and operation of the AL2S network. The results of this testing allow us to have confidence in our production network. It also allows us to know where trouble spots exist in some part of the current network deployment, allowing us to develop and test workarounds in other components before the trouble impacts production service. This is especially true of a SDN environment, where the network system is composed by integrating control software we've written, community developed virtualization software, and commercial network equipment. This environment is the single place where all of these components are tested together as a system before being deployed into the production AL2S network. The Testlab has show significant improvement over the past 6 months. As we begin to deploy multiple slices on AL2S, leading to additional system complexity, it will be important to continue to invest in test infrastructure improvements. We have already seen benefit in the existing work to automate more of the testing, both in increased speed and accuracy of tests. Many portions of the test suite are still manual and attention should be given to increasing automation over the coming months.

Raw Commentary from the Support Engineers (Tom)

The most common outage is the 100GE link flapping issued.

The second most common outage is OE-SS related.

1. Problems in January when code revision didn't properly diff the flowmod rules
2. Problems in February when ATLA-HOUH was rapidly decommissioned, "edited" circuits wouldn't failover
3. Same problem in March when inserting Raleigh, "edited" circuits wouldn't fail over

Brocade Hardware would be the third;

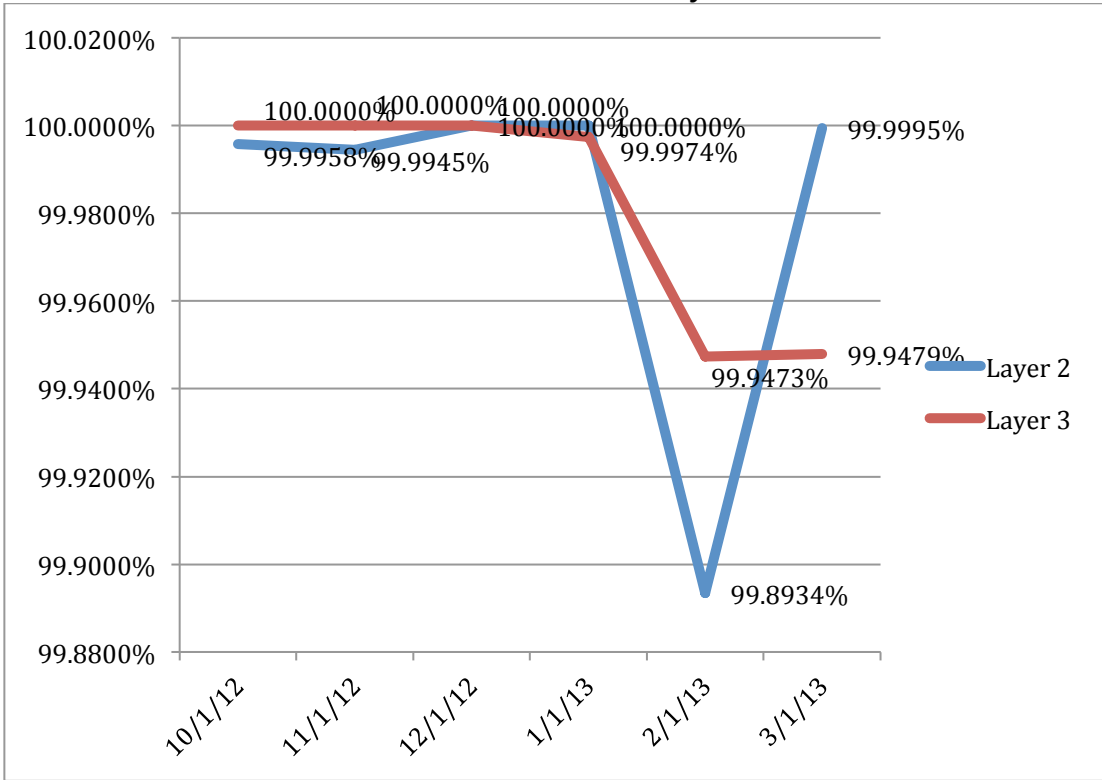
1. 1 faulty 10GE optic in HOUH
2. 1 faulty HSF (high speed fabric) in LOSA
3. 1 faulty MR (management module) in ATLA

Timeline

The Switch Software Upgrades were generally clustered in three batches: 10/16, 11/6, and 1/30. There was another event involving the insertion of the new Raleigh node in to the backbone on 2/12 (physically) and 3/13 (logically.) (need to check with neteng as the tickets don't list the software rev installed)

2012-08-18 – Chicago turned up
2012-08-19 – Washington DC turned up
2012-08-14 – Cleveland turned up
2012-08-16 – Atlanta turned up
2012-08-20 – Denver turned up
2012-08-20 – Houston turned up
2012-08-23 – Kansas City turned up
2012-08-31 – Los Angeles turned up
2012-09-06 – New York/32AoA turned up
2012-09-12 – Starlight turned up
2012-09-19 – El Paso turned up
2012-09-20 – Tulsa turned up
2012-09-28 – Salt Lake City turned up
2012-10-01 – Production Status
2012-10-01 – ONEnet turned up
2012-10-13 – IU turned up
2012-10-20 – Brocades Upgraded to 5.4.0a
2012-11-08 – Controller DR Failover Test
2012-12-08 – OESS Upgraded to 1.0.5
2012-12-12 – OARnet turned up
2013-01-06 – Brocades Upgraded to 5.4.0aa
2013-01-15 – Jacksonville FL In Service
2013-02-01 – FLR turned up
2013-02-04 - Brocades upgraded to 5.4.0ab
2013-02-23 – OESS upgraded to 1.0.6
2013-02-06 – Raleigh turned up
2013-03-07 – Seattle turned up
2013-03-12 – XSEDE turned up

Node Availability



Circuit Availability

