AAR - MAN LAN Core Node Hardware Upgrade ...

After Action Report

Network: MAN LAN

Issue Name: Core Node Hardware Upgrade

Issue Date(s): 8/21

Trouble Ticket #(s):
CHG0038258
INC0040257
CHG0038386
INC0040258
INC0040272
TASK0038206
INC0040289
INC0040290
INC0040291
INC0040293
TASK0038211
INC0040318

AAR Contributors

Matt Robertson, Chris Robb, Matt Mullins, Mark Brochu, Chris Wilkinson, Camille Davis-Alfs, Ross Shaffer, Vadim Petrov, Andrew Lee, Jason Lomonaco, Mike Robbins

Incident Summary:

A scheduled Manhattan Landing (MAN LAN) Core hardware replacement caused a Layer2/3 network service disruption to the point where the Rollback Plan had to be engaged. The rollback process was kicked off nearly 8 hours after the rollback threshold defined in the MOP, resulting in an accumulated 12 1/2 hours of unscheduled maintenance and service disruption for MAN LAN connectors. Service was restored after the original hardware was re-installed and MAN LAN

connections were moved back to their original ports.

Impact/Affected:

All MAN LAN Connectors were affected in some fashion over the course of the maintenance window and the 12 hours after the window closed. In some instances, networks had redundant international connectivity, but for those networks that didn't have redundant physical links they would have likely utilized Layer 3 backup services of those networks that do have redundant physical connectivity to the United States. For those networks that are multi-homed, it's also possible that they were uniquely routing portions of their entire networks to MAN LAN and would have suffered a complete outage for that portion of their network.

Incident Narrative:

The overall incident can be broken down into several major phases. A detailed Engineering and troubleshooting timeline of the event itself is presented in Appendix B.

Preparation Work Prior to Maintenance Window
Internet2 received the upgraded Juniper QFX hardware several months prior to the installation and resolved to migrate both the hardware and the configuration to an EVPN-based solution. Senior engineers designed and vetted a candidate configuration in a lab environment to ensure traffic passed as expected. This configuration was also tested out against the Internet2 NOC's management systems to ensure that measurement and alarming would perform as expected. The lab configuration was passed off to a Lead Implementation Engineer who took the concepts from that configuration to pre-generate the configuration to be used at MAN LAN during the maintenance window.

Installation Work Covered by the Change Management Window: 0400-1000 GMT
Engineers arrived on-site to perform the maintenance. However, the work was delayed by roughly an hour due to the house electrician arriving late. The maintenance work itself performed without issue, though it took significantly longer than anticipated. By the end of the maintenance window, many of the networks had been migrated to the new QFX switches, but several networks had either not established link or had established link but were not observed to be passing traffic. The decision was made not to back out of the maintenance as because the lingering issues represented a subset of networks and were perceived to be easily resolvable physical issues.

Physical Troubleshooting to bring up remaining peers: 1000-1414 GMT
Engineers worked on site to resolve physical issues with remaining cross-connects. Additional

senior engineers were pulled in to help divide the work up. The issues identified during this time period seemed centered on physical and configuration issues. In one instance, because of confusion between the Cable Run List and the MOP, two Connectors were identified as being swapped. In another case, there was some missing configuration facing a Connector. At one point, an engineer noticed that IS-IS hadn't established between the two new MAN LAN chassis. A portion of the configuration was identified as missing and added to bring up IS-IS at 1414 GMT.

Troubleshooting issues related to traffic replication between switches: 1414-1801 GMT
It's believed that establishing IS-IS adjacency between the two boxes, when combined with a pre-existing configuration mistake on the EVPN configuration, caused traffic replication between the switches and out to MAN LAN connectors. This caused a few MAN LAN Connectors to start shutting their ports down. Early on in this time frame, engineers were unaware that the replication was occurring and were still troubleshooting the remaining physical issues. Once they noticed a high amount of inter-switch traffic and combined that with a few timely reports from MAN LAN's connectors about Layer2 issues, they determined what was happening. Unfortunately the engineers weren't sure what was causing the traffic replication until later in the day. (A full description of the cause of the traffic replication is included in Appendix C.) A few theories were tested until the call was ultimately made to rollback to the original MAN LAN switch.

Rollback to original configuration: 1801-2222 GMT
During this timeframe, MAN LAN engineers ceased troubleshooting the traffic replication issue and focused on restoration of services by migrating networks back to the original Brocade switch. During the original maintenance window, electricians removed the house power feed for the Brocade switch. However, the Internet2 engineers had brought a fuse panel to hang off the new switches' power feed as part of the back-out plan. The Brocade switch was powered up and ready to accept connections at 1948 GMT. Engineers began moving connections one by one, needing to re-run fiber in most cases. In parallel, engineers worked with the OESS software development team to migrate any OESS circuits back to the legacy hardware. The migration of the last connection was complete at 2222 GMT.

Root Cause Analysis

Many of the failures highlighted in this document can be traced back to failures in the Change Management process. The mistakes in the installation materials, for example, can be traced back to not having enough controls built into the CM process that provide checks against those mistakes for a change of this size. Further, the decision to rollback the configuration early on in the process represents a major failure in the work planning and reaction procedures that need explicit guidance in the Change Management process going forward.

Likely Questions and Answers:

Why was the upgrade scheduled as a single window that covered both the hardware change along with the configuration change?

The roots of that decision have been difficult to ascertain, but it appears to have been an assumption from very early on in the project. There was a desire to remove a power circuit from the legacy rack as soon as possible and there wasn't a proper discussion about weighing that goal against operational stability during the hardware transition. It's clear that, while some individuals had misgivings or questions about that much effort, any concerns weren't voiced at the appropriate levels in the organization. It's also clear that the Change Management Change Approval Board wasn't rigorous enough in their examination of the proposed work. This appears to be largely based on some false assumptions regarding the amount of information and comfort the engineering staff had with the level of changes. The CAB also assumed that much of the earlier discussions would have covered that territory, when in fact it did not.

Why did it take so long to properly diagnose the problem?

There were two primary factors at play here: multiple different issues and lack of complete telemetry. Engineers were observing several different types of issues, with many of the early ones looking like mistakes in the installation methods of procedure (MOP) document, its execution, or the configuration of the switch itself. However, engineers were still working through some of those mistakes by mid-morning and the contributed to masking the underlying traffic replication problem. This effectively made the bigger issue look like a combination of smaller and more easily resolvable issues. Engineers were also basing a lot of their understanding on the Internet2 IP network's view of connectivity to MAN LAN than on telemetry coming from the MAN LAN equipment itself. While they were looking at individual portions of the operational state of the MAN LAN switches on the devices themselves, they didn't have an overall view of everything to help identify patterns and potential other issues beyond the ones they were reacting to.

Why was there a mistake in the EVPN configuration?

While Internet2 did have a functional EVPN configuration working within a lab environment that configuration didn't transfer fully into the production environment. The lab configuration was a proof-of-concept with a topology that didn't mirror the MAN LAN environment directly, so the final configuration of the MAN LAN switch needed to be generated as a separate step. At this point, the configuration was handed off to the project's lead implementation engineer who had less direct experience with EVPN than the designer. During the transition of the lab configuration to a configuration suitable for MAN LAN the lead implementation engineer pre-configured the MAN LAN switches with the candidate configuration and observed that traffic wasn't passing. He believed it was due to a missing set of configuration directives on the inter-switch connections. In

reality it was due to the lack of IGP configuration. Unfortunately since both were added at the same time, it appeared that the inter-switch configuration was needed in the production configuration. This would have lead to the same replication issues in the lab, but it wasn't detected at the time. While the exact cause of why it wasn't detected isn't fully understood, it's likely due to the smaller simulated environment and less obvious impact on production services that lead to it being overlooked.

Why did the maintenance extend past its scheduled window?
There were a number of factors at play here. There was a very initial setback with the electricians arriving nearly an hour later than expected, which pushed the migration start time out. But the core issue is that the timeline for the work was vastly underestimated. Further, the timeline for the back-out was also improperly estimated, so it wasn't properly considered in the formation of the overall maintenance window timing. Finally, staff made a judgement call to proceed with the maintenance past its maintenance window, despite having multiple issues persisting past the window. While there were points at which it looked like the remaining issues were being mitigated, it was also true that senior staff didn't have a strong appreciation for the entire breadth of the issue. The lack of telemetry further clouded the ability for both engineers and senior staff to see the wider picture. Finally, the actual scope of the issue itself wasn't fully understood by the engineering team until about six and a half hours after the maintenance window ended.

Relevant Areas for Improvement:

Technical:

1. Time estimates to complete the installation and back-out plans were too optimistic, leading to a window that was too small for the work to be realistically accomplished within.
2. The change itself was risky and was monolithic in nature.
3. Telemetry and remote access wasn't taken into account as a critical part of the installation process.
4. There was some preparation work on the physical installation that could have been handled ahead of the change window, which would have shortened the work within the window.
5. The implementation team didn't have a full technical grasp of the change and wasn't fully aware of the implications of their changes to the tested configuration.
6. There were mistakes in the standard portion of the configuration that contributed to the masking of the root cause.

Process:

1. A change of this size didn't get enough review attention from the Change Management Change Advisory Board

2. A change of this size that involves a new configuration or technology needs to have the direct involvement of the designer of that configuration (or a senior level engineer) or technology plan.
3. There were changes in the cable run list that weren't communicated widely, leading to the CRL and the installation configuration being out of sync with each other.
4. Customers were reacting to outages by disabling interfaces on their end, causing confusion over the source of the issue.

Human/Communications:

1. Many people along the work structure and approval path made assumptions about early decisions in the project and didn't voice their misgivings or questions at the appropriate levels of the organization.
2. Senior technical and management staff didn't have a good overview of the breadth of the problem until late in the morning. This lead to less pressure on the technical staff to rollback their changes.
3. The CM request didn't properly reflect the potential risk and the steps that were taken to mitigate that risk.

Key Recommendations:

1. Add a new technical role into the work planning phase that attempts to play devil's advocate for major installation decisions.
2. Create a tiered CM process that directs the CAB to make a more rigorous review of the proposed change.
3. Develop standards and cultural understandings for proposed change management work:
   a. Smaller, less risky and disruptive maintenances are better than monolithic, riskier maintenances
   b. When determining the estimated timeline for a change, trend closer to worst case estimates or allow some additional headroom beyond the ideal.
   c. Timelines need to explicitly include a go/no-go gate with what is expected to have completed by that point. Telemetry and remote access need to be considered as part of that timeline and must be completed early in the installation process to allow for a complete picture and additional resources to help, should a later portion of the installation MOP encounter unexpected issues. Any decision to push through a final go/no-go gate needs discussion with a member of the CAB.
4. Automate the portions of the configuration that are standard.
5. During an outage that goes beyond its maintenance window, designate someone to determine overall status and communicate frequently with senior staff so they can make informed decisions based on impact data.

6. Develop a set of questions for the CAB and the implementation lead to answer during the CM process:
   a. Is there any way to break this work up and make it less risky?
   b. Have we developed a timeline that accurately reflects what we can expect during the actual work?
   c. When is the go/no-go timeframe and what needs to be accomplished to cross that gate?
7. When possible, create real-time communication mechanisms for engineers and customers to report status. This is especially recommended for situations where Internet2 telemetry paints only a portion of the picture.

Additional Notable Recommendations:

Appendix A: Outage Timeline/Detail -

MAN LAN Core Node Replacement – Event Duration 04:00 – 22:38 GMT

- 04:00 – 13:00 GMT Maintenance 1 of 2 Completed - MAN LAN Core Node (sw.net.manlan)
  - CHG0038258
    - 04:00 Maintenance ticket began until 10:00
    - 05:51 Port swaps and card install/decoms completed (slack)
    - 09:51 Engineering requests extension (slack)
    - 09:55 Maintenance notes extend until 11:00
    - 10:49 Engineering confirms another extension (slack)
    - 10:56 Engineering confirms port links are up (slack)
    - 10:57 Maintenance notes extend until 13:00
    - 11:54 Engineering observes some peers connecting while others aren't (slack)
    - 12:34 Customer Canarie dropped (slack)
    - 13:00 Maintenance ticket & notes closed
- 13:01 – 19:59 GMT Outage Resolved - MAN LAN & IP Various Participants
  - INC0040257
    - 13:01 Outage ticket began
    - 13:02 Engineering confirms Install team leaves (slack)
    - 13:40 Outage Notification sent
    - 13:59 Engineering confirms direct communication with connectors (slack)
    - 16:14 Engineering requests SNMP alarms put into a separate ticket (slack)
    - 16:15 Engineering request SD reach-out to ESNET for verification of disabled ports (slack)
    - 17:02 Major incident declared (slack)
    - 17:36 Engineering requests SD engage NORDUnet (slack)
    - 18:08 Engineering confirms rollback initiated (slack)

        - 19:17 SD requests Emergency change to perform rollback (slack)
        - 22:31 Outage Resolved note sent
- 13:26 -13:27 GMT Brief Outage Resolved – MAN LAN to WIX Interconnect
    - INC0040258
        - 14:40 Engineer confirms via slack outage was a result of troubleshooting
- 15:33 GMT Incident Resolved – I2 Optical Participant NORDUnet
    - INC0040272
        - 15:33 Customer reports flapping
        - 16:50 Customer disables interface to regain stability
- 16:31 GMT Alarms – MAN LAN SNMP
    - TASK0038206
- 16:50 – 21:33 GMT Availability – I2 PX Participant NORDUnet INC0040289
- 16:54 GMT Outage – MAN LAN Participant Ixia INC0040290
- 16:55 – 22:24 GMT Availability – MAN LAN Optical Participant ESNet INC0040291
        - 17:06 Availability notification sent
        - 17:30 Customer disables interface to regain stability
        - 17:31 Availability notification sent
- 16:58 – 20:25 GMT Outage Resolved – I2 Optical Participant ESNet INC0040293
        - 18:01 Resolved notification sent
- 17:24 GMT Major incident – MAN LAN Core Node NEWY32AOA TASK0038211
        - 17:32 Initial Major Incident MAN LAN community notification sent
        - 18:30 Second Major Incident MAN LAN community notification sent
        - 20:49 Third Major Incident MAN LAN community notification sent
        - 22:25 Final Major Incident MAN LAN community notification sent
        - 20:10 of the following day, Major Incident MAN LAN community notification sent
- 19:35 GMT
        - Alarm – MAN LAN CPU
        - INC0040318
- 20:00 – 23:59 GMT Emergency Maintenance Completed - MAN LAN Core Node (sw.net.manlan)
    - CHG0038386
        - 20:00 Maintenance ticket began until 22:00
        - 20:08 Engineering requests SD reach out to customers to re-enable interfaces (slack)
        - 20:44 ESNet re-enables interface (slack)
        - 20:48 NORDUnet re-enabled interface (slack)
        - 21:05 ESNet requests to hold off until after maintenance before re-enabling interfaces due to instability (slack)
        - 21:27 SD inquires about maintenance extension (slack)

- 21:33 Canarie services confirmed to be up (slack)
- 21:54 Engineering confirms maintenance extension (slack)
- 21:56 Engineering confirms MOXY up (slack)
- 21:57 Maintenance notes extend until 23:59
- 22:29 Engineering confirms ANKUBUT will clear with no-mon tag applied (slack)
- 22:38 Maintenance ticket & notes closed

Appendix B: Detailed Engineering and Troubleshooting Timeline (times are GMT)

Roles:

Director of Operations
Director of Architecture
Implementation Lead
Tier 3 Tech #1
Tier 3 Tech #2
Engineering Team Lead
RR Engineer
Network Design Lead
Installation Engineer #1
Installation Engineer #2
Installation Engineer #3
OESS Lead Developer

- 0300 - Installation team arrives onsite and does prep work. NYSERNet staff on-site.
- 0350 - Electricians arrive roughly an hour late. Begin preparing
- 0400 - Installation window begins. Team needs to wait for electricians are ready to begin.
- 0410 - Installation team turns down the MAN LAN switch and backbone traffic is drained away from MAN LAN
- 0445 (est) - House power disconnected. Installation team begins removing connections from Brocade
- 0515 - Fiber panel moved to where the previous power controller was installed. Begin installation of QFX 10002
- 0545 - Both QFX 100002s are installed. Technicians begin the process of powering up the boxes and moving fibers over.
- 0815-0830 - discussion about possible backout. Decision was that it would take longer to rollback than to move forward.
- 0845 - Both QFX 10002 are powered up and have fibers connected. Remotely accessible via

console, but not SSH. Engineers began contending with individual interfaces that were being blocked by BPDU

- 0950 - Installation Tech #2 reaches out to Canarie to move their optic. Unable to get them on the phone.
- 1030 - BPDU troubleshooting on the link to rtsw.newy. Brings up some connector interfaces
- 1203 - Canarie returns our call to work the optic swap on their interface. Interface comes up, but is not passing traffic.
- 1304 - Implementation Lead reports to #ns-manlanwix-working that we have link on all interfaces, but we are not receiving packets on the CANARIE and NEAAR VLANs. At this point, to observers it appears that the issues are restricted to two physical interfaces on these two ports. The engineers troubleshooting also had the additional understanding that logins and snmp weren't functional to either switch.
- 1313 - Director of Operations, upon hearing that GÈANT is still not up, added two additional Tier3 engineers to the #ns-manlan-working chat room to get more eyes on the situation. Tier 3 Tech #1 begins focusing on NEAAR interface. The Implementation Lead focuses on NORDUNet's connection and Tier 3 Tech #2 is working on the MAN LAN-WIX interconnect.
- 1320 - Director of Operations initiates first communication to Internet2 senior staff in the #ns-lead chat room. At the time, it was understood and characterized to be just an issue with a handful of peers that didn't come back up on the Internet2 network.
- 1323 - Network Design lead is able to log into the switch for the first time using a hard-coded username/password to bypass the AAA issues.
- 1329 - WIX interface issue is resolved due to the optical system not recognizing a new optic type. It begins passing traffic right away, indicating success to the team.
- 1332 - Tier 3 Tech #1 is able to get NEAAR interface up and passing traffic by adding LACP configuration. Nordunet still remains down.
- 1345 - Tier 3 Tech #2 mentions TWAREN's interface isn't receiving light. This is the first mention of TWAREN having issues.
- 1346 - Implementation Lead reports that one of GEANT's peerings via Canarie is pingable
- 1346 - Engineering Team Lead reports that he has isolated the login issues to a configuration change needed within the Systems team and that it would take 15-60 minutes to go active.
- 1348 - Implementation Lead takes on the TWAREN issue and tasks Tier 3 Tech #2 with looking at the Internet2 IP network peers that are down on Alertmon and start working there. The tech begins by focusing on OMAN. This is the first mention of OMAN having issues. It's later reported that OMAN rides the Canarie interface that's still down.
- 1352 - Implementation Lead shifts focus to CANARIE
- 1400 - Director of Operations disengages from the discussion to complete a pre-scheduled 1:1 meeting until 1500.
- 1404 - Engineering Team Lead observes that IS-IS isn't functional between the switches

because the family wasn't configured on the interfaces. The Implementation lead reports that he believed IS-IS was up and running. Engineering Team Lead starts working on getting IS-IS up.

- 1414 - Engineering Team lead enables IS-IS peering. It's believed at this point that, since communication is established between the switches, traffic begins looping and MAN LAN customers start seeing Layer2 issues on their connection.
- 1427 - Implementation lead reports that ESNet called about "traps and circuits". It's later stated that they were calling about connectivity to GÈANT over the Canarie interface that is down.
- 1442 - Canarie bounces their report and Nordunet's connection goes down. The team concludes that the Nordunet and Canarie connections are physically swapped at the fiber panels.
- 1457 - While preparing the configuration change to swap the ports, the Installation Engineer #1 reports he is on his way back to the POP to make the physical swap. It's believed this will be faster and more optimal than making the configuration change.
- 1503 - Network Design Lead reports that the "interswitch LAG seems to already be running hot". This is the first mention of excessive traffic between the switches. It's initially believed this is because of the suboptimal state of peers and that traffic will rebalance once things are normalized.
- 1510 - Director of Operations returns to get caught up on the discussion and sends update to #ns-lead channel. At the time the understanding was that the NEAAR issues had been addressed and GEANT was back up with just Canarie and Nordunet being the two lingering issues. In retrospect the impact was improperly understood and communicated a better picture of what was happening on the ground.
- 1522 - Nordunet and Canarie interfaces are physically swapped by Installation Engineer #1. Installation Lead reports that Nordunet is still having problems.
- 1525 - Network Design Lead observes that several VLANs to Canarie aren't showing any traffic. It's believed they might not have shown traffic before.
- 1529 - Tier 3 Engineer #1 points out that Nordunet is still not working. Focus shifts there and whether they're cabled correctly.
- 1548 - Nordunet and SINET swapped. Engineers see an incoming MAC address from Nordunet, but their BGP doesn't re-establish.
- 1552 - Director of Operations asks if a phone bridge would improve communications. Tier 3 Engineer #1 suggests it wouldn't improve communications, so conversation remains in Slack
- 1557 - Director of Operations and Director of Architecture discuss the possibility of initiating a Major Incident. Both comment that they don't feel they have a good overall view of what's affected and the breadth of the problem.
- 1607 - Director of Operations points out that a Major Incident is being discussed and assigns

Installation Lead with the job of creating an overall issue list and keeping Director of Operations and Director of Architecture up to date on overall status.

- 1619 - Engineering Team Lead assigns Tier 3 Tech #2 with the task of looking at NEAAR more closely.
- 1630 - SNMP access working on the devices. They had a misconfigured SNMP string and the devices were set to administratively down in the database.
- 1639 - Director of Architecture reports that ESNet shut their interface down due to "some sort of broadcast storm or some other related L2 hilarity" (their words). Tier3 Tech #1 mentions that he did see a lot of multicast traffic on the interconnect but he got sidetracked dealing with Nordunet.
- 1645 - Director of Operations suggests that traffic may be looping. Some team focus shifts to looking at that while others continue troubleshooting what appear to be physical issues with TWAREN
- 1701 - Director of Operations initiates Major Incident procedures. A Communications Manager and Major Incident Manager is assigned. The Service Desk sends notification to the community.
- 1706 - SNAPP link pasted that shows all ports on both switches are blasting a lot of output.
- 1714 - Director of Operations sends initial communication for distribution to the MAN LAN community to the MAN LAN NOC. It's distributed to the community at 1732.
- 1719 - Director of Operations provides the team with 30 minutes to resolve all issues before the rollback is triggered. The primary theory at this point is that there's some sort of broadcast storm caused by something with the switch configuration, but it's not understood what. The team is focused on shutting down interfaces to test theories.
- 1724 - RR Engineer is added to the channel to add perspective on tickets coming in.
- 1725 - Canarie sent a note to Install Lead and the Internet2 NOC and suggested they shut their port down due to unusual behavior on their Halifax router.
- 1726 - Director of Operations starts a new Slack channel, #i2-incident-manlan-upgrade, to communicate status to relevant Network Services and Marcom staff.
- 1730 - Field team understands there is a potential to rollback and begins to prepare for power supporting Brocade.
- 1745 - decides to shut down one of the LAG members on the connection between the two MAN LAN switches. This doesn't resolve the issue.
- 1801 - Director of Operations makes the call to begin the rollback.
- 1804 - Rollback-focused team is formed and engineers get on a phone bridge to determine a course of action. At this point. Install engineer #1 and install engineer #2 are in the POP and ready to begin the work.
- 1818 - Director of Operations sends an update communication to MAN LAN NOC to distribute to the MAN LAN community. It contains an update on the working theory as to

what's happened at MAN LAN and the decision to rollback. At the time the expected timeline for rollback was 1-2 hours. The message was distributed to the MAN LAN community at 1830.

- 1823 - Decision was made to power Brocade off a single power supply so we can have it and the new hardware up at the same time. The team decided that migrating the connections one-by-one would make the most sense.
- 1910 - Tier 3 Tech #2 starts to bulk move circuits in OESS, consolidating the MAN LAN interfaces on rtsw.newy32aoa. The engineer was using the OESS interface.
- 1923 - Tier 3 Tech # 2 states that the pending diffs doesn't look correct in OESS and decides they need engagement from the lead OESS software developer.
- 1927 - Engineering Team Lead engages with the Lead OESS Developer and adds to the working MAN LAN Slack channel.
- 1948 - Brocade power, management and console are connected. Original sw.net.manlan is powering on.
- 1949 - Lead OESS Developer confirms an issue with the circuit move inside OESS and works on correcting.
- 2001 - Engineering Team Lead and Install Tech #2 start migrating physical connections.
- 2046 - Director of Operations sends an update communication to MAN LAN NOC to distribute to the MAN LAN community. It explains that the MAN LAN legacy hardware has been powered back up and engineers are beginning the process of migrating circuits. The message was distributed to the MAN LAN community at 2051.
- 2054 - Lead OESS Developer states all circuit endpoints should now be correct in OESS.
- 2056 - Tier 3 Tech #2 confirms the OESS diff looks correct and approves.
- 2106 - OESS circuits are migrated successfully.
- 2218 - Following confirmation that all customer-facing circuits have migrated completely, Director of Operations sends a completion message to the MAN LAN NOC for distribution. The message is sent to the MAN LAN community at 2224.
- 2222 - Engineering Team Lead verifies last connection, an internal perfSONAR server, was successfully migrated.

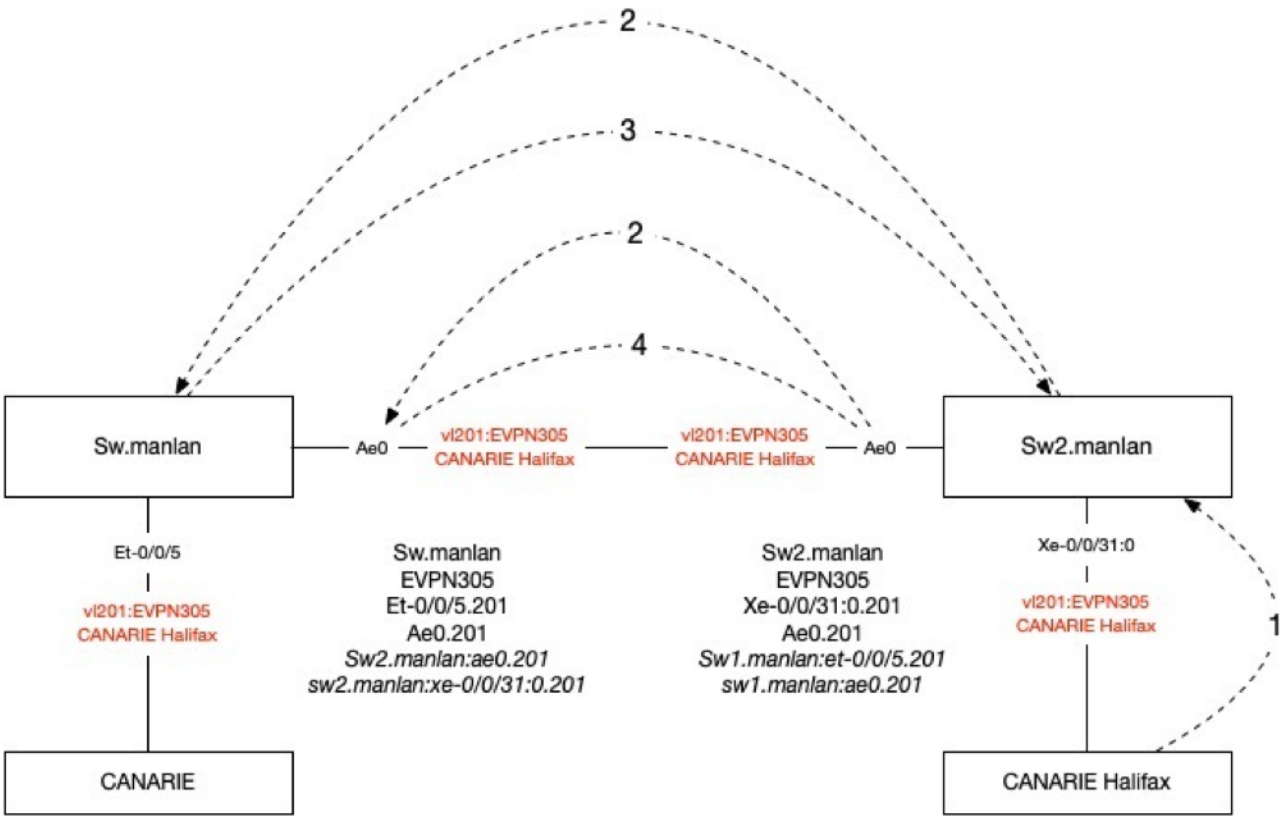
Appendix C: Details of Traffic Replication Behavior and Underlying Cause

An EVPN routing instance is configured in a similar fashion to a Juniper VPLS Service: That is, routing instances are configured on each PE node participating in the EVPN, and interfaces facing the service consumers are added to each routing instance.

In the deployed configuration, the interconnect(s) between the PE nodes were also configured as

endpoints for the EVPN. This is unnecessary, since MAN LAN should have used VXLAN over the backbone interface to encapsulate EVPN traffic. However, due to a misconfiguration, ISIS was not functioning between the nodes, and as a result, VXLAN/Inter-node EVPN was not functioning correctly. Presumably, putting sw/sw2 interconnect VLANs in each VPN solved the issue by creating "manual" VLAN cross connects between the two nodes.

At 1014, engineering re-enabled the ISIS connection between sw.manlan and sw2.manlan, to match the original design. however, this effectively created a loop between the VLAN service interfaces. This is substantiated by time reported by CANARIE indicating performance issues at approximately 10:14AM EDT, 4 minutes after ISIS was enabled. The example below shows what likely occurred in the CANARIE case, however it's understood that this would have been an issue with any network that had ethernet services that spanned the two MAN LAN QFX10002 switches.



1.) Broadcast/Unknown Unicast/Multicast (BUM) traffic sent from CANARIE Halifax to sw2.manlan
2.) sw2.manlan sends traffic out all local ethernet segments as well as remote peers (sw.manlan) via VXLAN
3.) sw.manlan receives traffic via ae0.201 and replicates traffic out all ethernet segments except for ae0.201 (split horizon) but does send it back to the remote peer on sw2.manlan via VXLAN.

sw.manlan also receives traffic via VXLAN and replicates it back via ae0.201 to sw2.manlan.
4.) sw2.manlan recieves traffic via vxlan, as well as ae0.201, loop is complete