Technology Update

Richard Petrie, Anne Bates, Mo Shivji

Wednesday 20 November 2024
LINX123

- Mailman Update
- Service Incidents
- Projects Update
- Review of LON2 Technology
- Problem Management
- Route Server Update

Mailman Update

Mailman Update

- Recap
  - Migration from legacy MLMMJ setup to mailmain v3
  - Bugs found on how MLMMJ deals with mail security
  - Very few people are using MLMMJ, community support and documentation is limited

- What next?
  - Tested successfully in a dev environment, found some minor bugs
  - Deployed a production environment and tested with a test mail group (LPC)
  - Documentation and training completed
  - API and workflows updated for NOC tooling
  - Rollout planned for late November

Review of LON2 Network Technology

Review of LON2 Network Technology

- Recap
  - Last LON2 refresh project was in 2016/2017
  - We selected a fully disaggregated solution with Edgecore providing the hardware and IP Infusion providing the NOS
  - The NOS and hardware are coming to end of life and will need to be phased out in 2025
- Research happening now
  - Review of the next generation Edgecore and IP Infusion offering
  - Look to the market on alternate solutions
  - Assessment of the disaggregated market in general
- Next steps
  - Q4 final review of the project given the outcome of a full RFP phase
  - Q1 decision and procurement as well as implementation planning
  - Q2 phased migration from legacy solution

Service Incidents

Incidents

LAN	Incidents
Systems	5
LON1	1
LON2	1
LINX Manchester	0
LINX Wales	0
LINX Scotland	0
LINX NoVA	1
JEDIX	0
ManxIX	0
Nairobi	2
Transmission PoPs (London)	1

Portal: https://portal.linx.net/maintenance-and-outages

Incidents

Shared on LINX Community

1. Transmission PoPs
    - LINX LON1 & LON2, Virtus Hayes, Dark fibre between VPOP Virtus Hayes and Telehouse North went down. Dark fibre provider confirmed a large-scale fibre break involving estimated 30 fibres across their London metro. [8th May]
2. LINX System Incidents
    - Firewall Issue, resulted in loss of member facing stats. [10th June]
    - Loop on management network, engineer accidentally created a loop, portal stats were affected by this short outage. [1st July]
    - Firewall issue, affected stats on the portal. [4th and 9th September]
    - Self-service automation offline. An issue with TACACS meant that automation couldn't push changes to the switches [23rd September]
3. Collectors
    - LINX NoVA, Instability with BGP to LINX Collector, caused by our arp sponge device malfunctioning. The status list under the sponge had the collector's IP address mapped to a member's mac address. [6th September]

Incidents

4. Nairobi Dark Fibre
    - Fibre cuts between iColo NBO1 (ICN) and African Data Centres (ADN). Simultaneous outages on both dark fibres from ADC resulted in site being isolated. [10th September]
    - Fibre cuts between iColo NBO1 (ICN) and IXAfrica (IXN) on both dark fibres resulted in ICN being isolated. [3rd October]
5. LINX LON1
    - Digital Realty LON1 (TCM) line card 1 on edge2-tcm rebooted unexpectedly. [20th September]
    - FPC4 on core4-thw (THW) rebooted [12th November]

Problem Management

Problem Management

- LINX Internal Network, Firewall Issues
  - Minor, but annoying, issues with staff VPN access has been seen since a firewall software upgrade to patch a critical vulnerability earlier in the year. Mostly causing connections to need multiple retries.
  - Software updated in response to a critical vulnerability contained a memory leak. Resource exhaustion caused processes to be restarted without triggering a failover to the standby firewall. This has caused interrupted access to LINX hosted services such as Portal.
  - We have worked with the vendor to diagnose and resolve the memory leak. Hardware has been replaced and software upgraded.
  - Since the end of October the firewall has been stable.

Problem Management

- NTP2 losing signal
  - Potential hardware fault
  - Work around has been to sync NTP2 to another NTP device, to allow it to validate external clocks (GPS & MSF) when they disagree with each other
  - Project initiated to refresh NTP servers

Problem Management

- LON2 MAC and hardware MAC issue
  - This is a longstanding issue, where topology changes from either provisioning or network incidents can cause software and hardware MAC address tables to get out of sync.
  - A work around is in place, in that, checking hardware MAC address tables when rebooting a switch or when there are ISL flaps.
  - Issue hasn't reoccurred since last year, will likely be mitigated by migrating to new vendor next year.

Problem Management

- LON1 power related longstanding issues
  - After core3-thw power issue, identified and remediated several issues, including audits, maintenances, power cables replaced, routers moved onto 3 phase power. Grounding of edge1-eq4 remains.
  - PEM issue edge1-th2 – after i2c errors on PEM3 resulted in errors on all fabric planes across edge1-th2 causing all FPCs to reboot. Advised by JTAC minimal chance of reoccurrence. Bug fix (21.4R3 onwards, currently on 21.4R1-S3) will be incorporated in next round of software upgrades due Q1/Q2 2025.

Problem Management

- Ongoing issue with **unknown unicast traffic** on LON1 (highlighted by member)
- Some members are responding to unknown unicast traffic incorrectly, forwarding on traffic from their own MAC which results in members receiving incorrect ARP entries.
  - IP addresses bound to the member MAC addresses are responding erroneously to the unknown unicast.
  - For all identified members, this issue has been resolved.
- Last year, based on our partner vendor, Juniper's recommendations, the unknown unicast policer was adjusted to police at the lowest value (8K) rather than drop completely.
- We have an active investigation into dropping unknown unicast traffic over the current practice of rate limiting it.

Problem Management

- Lab testing
  - Performed with a view to drop unknown unicast on the Juniper MX10Ks.
  - Filter was applied dropping unknown unicast, but reachability issues were observed between test CEs when doing this, despite the PE routers mac-ip/proxy arp table correctly propagating.
  - Further investigation ongoing with vendor.
  - We do drop all unknown unicast at SAP ingress on the LON1 Nokias and have not observed issues in doing this with them.

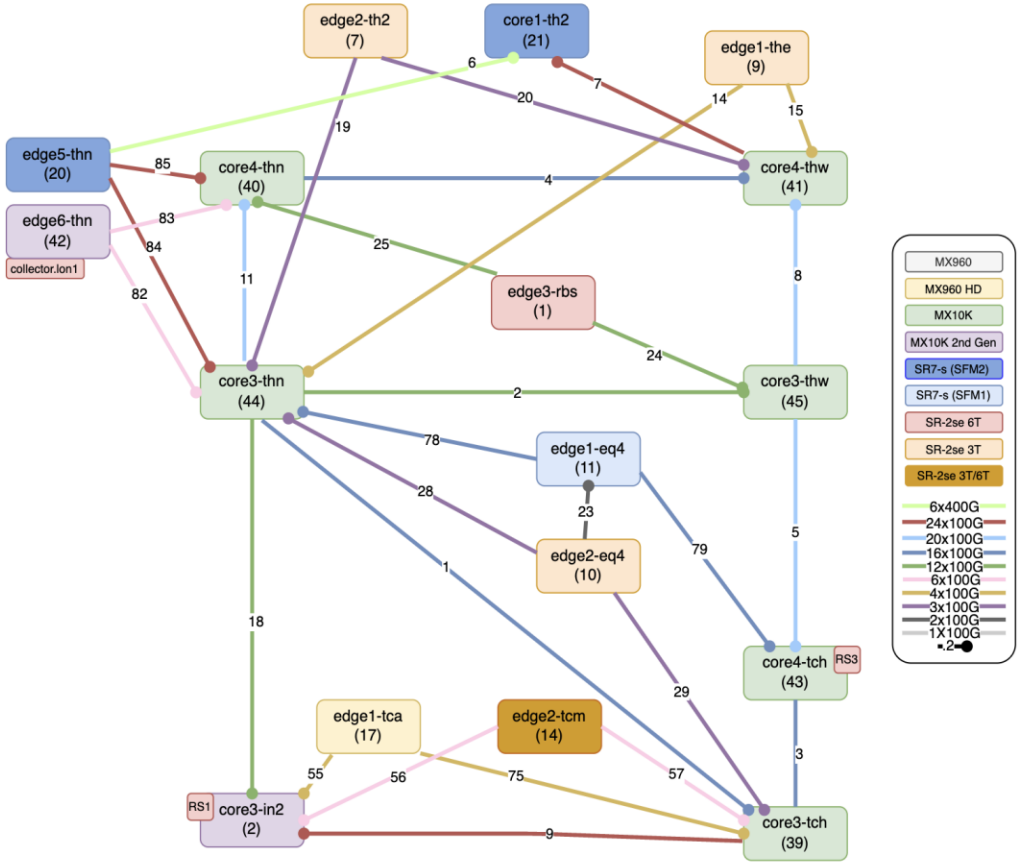
Problem Management

- Lab testing
  - There is also another Junos bug (PR1770350) that impacts unknown unicast behaviour
  - Whereby there is delay in the unknown unicast policer being applied when a MAC is withdrawn.
  - During that delay period, traffic is flooded briefly without being policed.
  - This is fixed in a newer software version, upgrades planned for Q1-2025.

Projects Update

LON1 Update

- Migrated all 10G members in Slough to Nokia SR-2se; decommissioning MX960.
- One final MX960 to remove from network in Equinix LD9 (involves moving to a new rack) – work progressing on that Q4.
- Nokia SR7-s installed in Telehouse North 2
  - 40 100G/400G connections from TH2 that were longlined to THN and THW have been migrated back to TH2
  - SR7 is being used to connect members directly in TH2
- Work also ongoing to increase core capacity to 6*400G/24*100G across Telehouse campus and Harbour Exchange.



Projects

- ISO27001:2022
  - Full ISMS audit completed in August against the new standard
  - Recommended pass of audit in August, certificate awarded in October
- Riyadh
  - New Internet Exchange now live in Riyadh as of May 2024, Saudi Arabia in partnership with Center3
  - Has hit peaks of over 900G
- Accra
  - Expected date Q1
- Mombasa
  - Expected date Q1

Projects

- Nairobi
  - Expansion plans to PAIX data centre in Q1
  - MAPS Nairobi is live from September 2024

Change Freeze

NoVA Thanksgiving Change Freeze

- From 17:00 EST on Wednesday, 27 November to 09:00 EST on Monday 2 December 2024

Christmas and New Year Change Freeze

- From 17:30 UTC on Friday, 13 December through to 09:00 UTC Thursday, 2 January 2025

Route Server Update

BIRD

- Currently running BIRD 2.13.1 on LINX other LINX RS's
  - No issues reported with this in operation.
  - Has been stable
- Will plan to upgrade to newer release in the new year.
  - Possibly BIRD 2.15.1
  - Test in lab first
- Deployment of LINX Collectors completed.

OpenBGPd

- Still running OpenBGPd 8.3 on all RS except LON1
  - Current release is 8.6
- Has been stable in operation since initial deployment
  - Tested successfully for 1000+ peers for LON1 on 8.5
  - Looking to deploy to LON1 in Q1 of 2025
  - Will be 8.6 or newer release
    - Test in lab first
- Expand deployments to:
  - Jeddah IX
  - Riyadh IX

Looking Glass

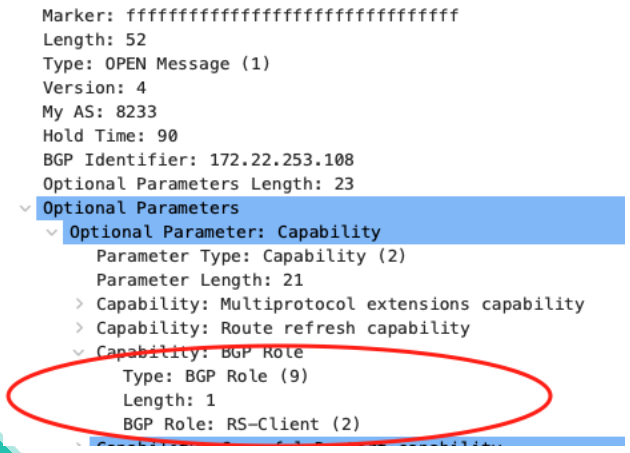
- AliceLG for route-servers crashed for customer Extended communities
  - AliceLG version 6.1.0 has operational bugs with OpenBGPd.
  - Currently using version 6.0.0 for route-server looking-glass.
  - Bug case open with developer for this issue.
  - Workaround to filter non-transitive extended-communities in OpenBGPd.
- AliceLG UO to API connection issue
  - Raised through a member bug report.
  - API disconnects and not collecting.
  - Raised to developer.

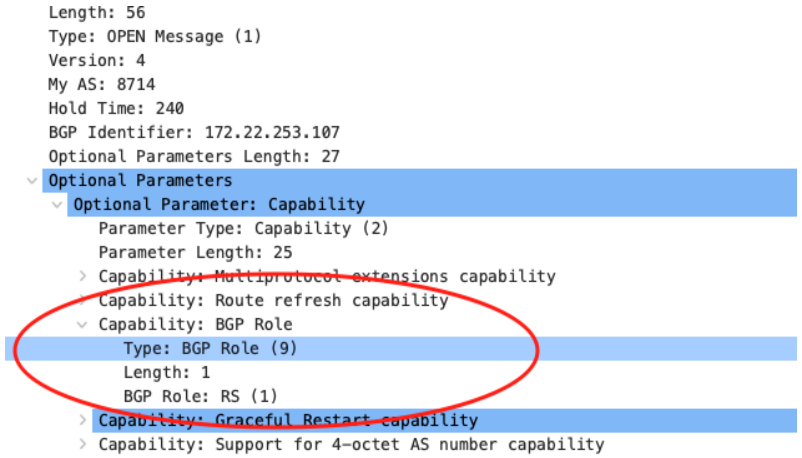
RFC9234 - Route leak detection & prevention

- RFC9234 detects and prevents BGP route leaks commonly caused by errors or misconfigurations.
- Detects and prevents BGP route leaks by enhancing the BGP OPEN message to establish an agreed peering relationship on each eBGP session.
- Peering relationship is agreed on a role.
- RFC9234 does not prevent route-hijacks.
  - Only route leaks.
- Not yet supported by major vendors.

RFC9234 – How it works ?

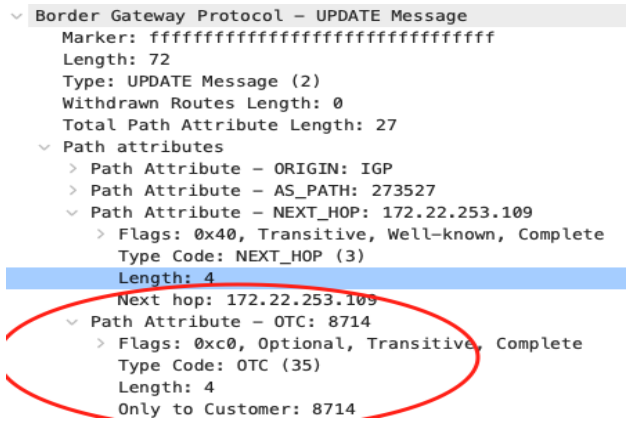
- Each peer has a role type configured
  - Route-server and route-server-client roles
  - Roles have to match for peering sessions to establish,
    - Other roles also configurable for providers and transit customers etc.
      - provider, customer, peer.
    - Role is added to the OPEN message as a parameter

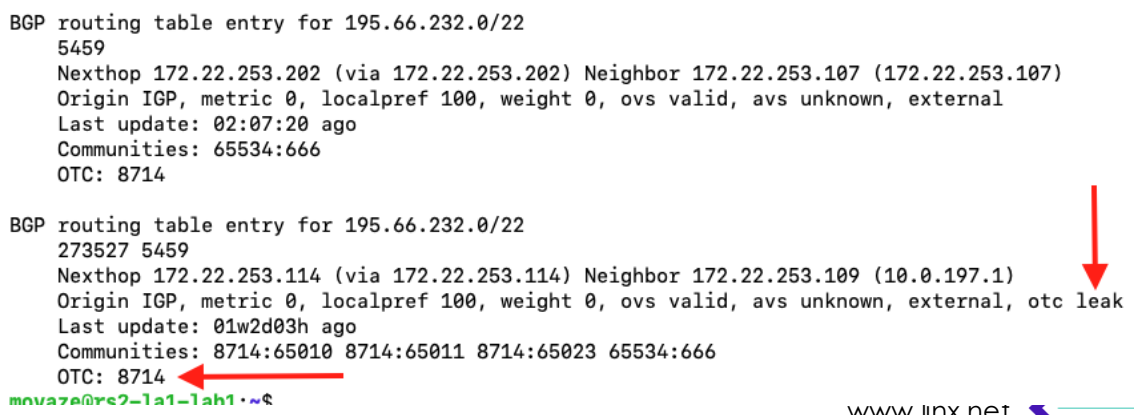




RFC9234 – How it works ?

- When RS-Client sends the prefix to the route-server.
  - OTC attribute is set in UPDATE message the with the value of its own ASN.
  - If the RS-client has no role configured or does not support RFC9234
    - OTC attribute is set in UPDATE message by the route-server.
    - If an update is received from a RS-client with an OTC the route for that update is rejected and considered a leak.
    - It is all in the code and no extra filtering required in configuration.





RFC9234 – Real World

- Announced on NANOG mailing list

Let's zoom in on 1 entry:

```
    $ bgpctl show rib 157.185.154.0/24 detail

    BGP routing table entry for 157.185.154.0/24
        6939 38040 54994
        Nexthop 206.126.225.20 (via 206.126.225.20) Neighbor 206.126.225.20 (216.218.252.194)
        Origin IGP, metric 1911, localpref 100, weight 0, ovs not-found, avs unknown, external, otc leak
        Last update: 11:58:08 ago
        Communities: 0:2906 0:16265 0:16276 0:18638 0:41690 0:48641 0:49029
        Ext. Communities: ovs not-found
        Large Communities: 53339:11:1 53339:11:3
        Aggregator: 54994 [163.171.131.254]
        OTC: 51706
```

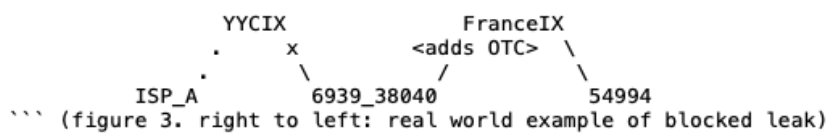
Prefix announced By FranceIX

OTC of FranceIX Route-Server

Marked as a leak

(figure 2. inspecting an leaked route using OpenBGPD's CLI)

In figure 2. one can see the route is marked as 'otc leak', this was made possible because FranceIX's route server's attached the OTC attribute with the ASN value set to their Route Server's ASN (51706).



(figure 3. right to left: real world example of blocked leak)

RTBH Improvements

- Baisc RTBH deployed on all LINX exchanges.
  - Initial deployment was strict on prefix validation.
    - ROA's had to be strict
    - Strict prefix objects required for IRRDB
  - User experience was not great for few members who have used it.
    - Became an obstacle when being attacked and use RTBH.
    - Agreed with PM need to look at making this more user/member friendly.
      - Followed up on some research
      - Will validate prefixes using "loose" filters for RPKI and IRRDB under required condition.
        - Upto /32 for IPv4
        - Upto /128 for IPv6
      - Loose ROA's via RTR on BIRD
      - Loose Static ROA's on OpenBGPd
  - Changes should be deployed in time for change freeze.

Thank you