AmLight ExP
Americas Lightpaths Express & Protect

12th SIG-NGN Meeting - Oct 25th, 2023

In-band Network Telemetry @ AmLight

Jeronimo Bezerra - FIU/AmLight

Outline

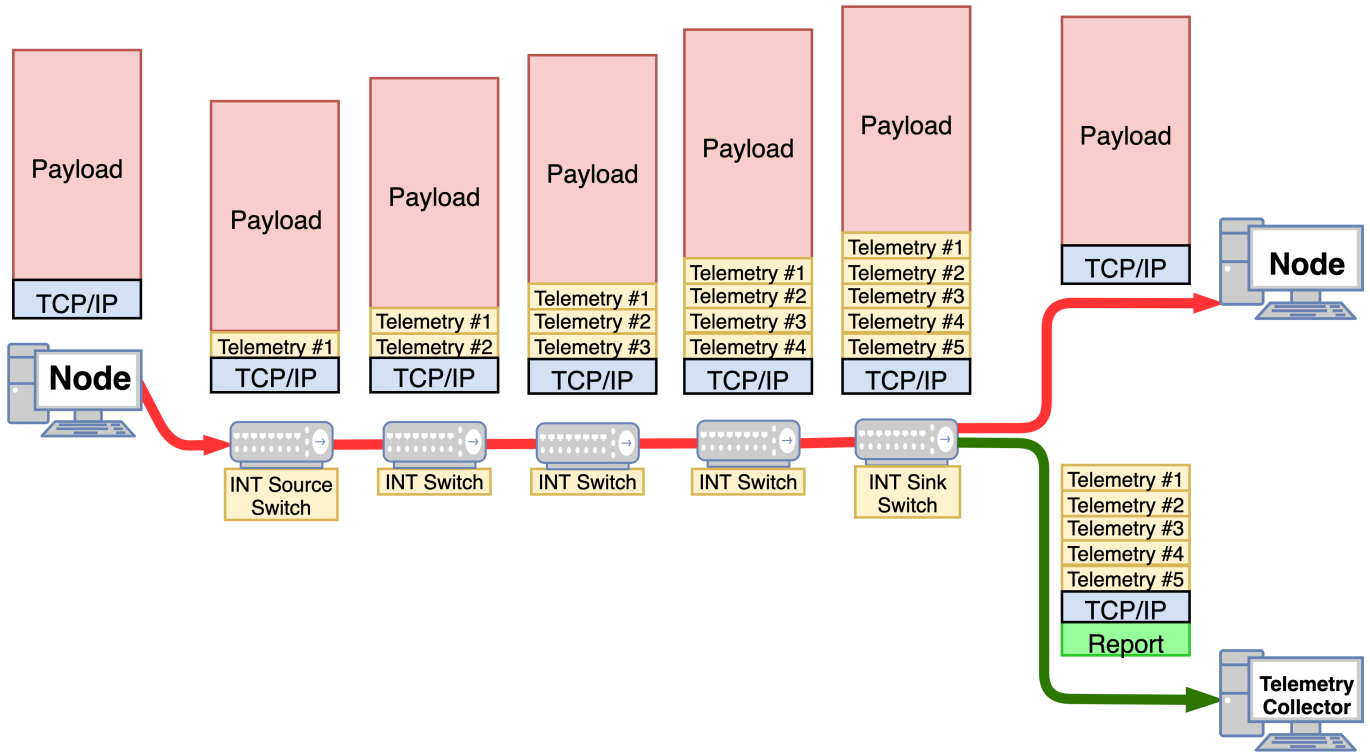
- **What data is available via INT?**
  - And what data is AmLight saving/storing now?
- **How AmLight benefits with the INT data?**
  - Deep visibility and granulkar network monitoring
- **How we use INT data to ”tune our networks”:**
  - *Traffic Engineering*
  - Q-Factor: Sharing INT data with endhosts

In-band Network Telemetry (INT) in a nutshell

INT records network telemetry information (metadata) in the packet while the packet traverses the network

Telemetry data is exported directly from the Data Plane:

- Operator *can monitor EVERY single packet at line rate and real time.*



What data is available in the INT reports?

Per switch:

- Switch ID
- Ingress port
- Egress port
- hop_delay: Egress timestamp - Ingress timestamp
- Egress queue ID
- Egress queue occupancy/buffer utilization

Per report:

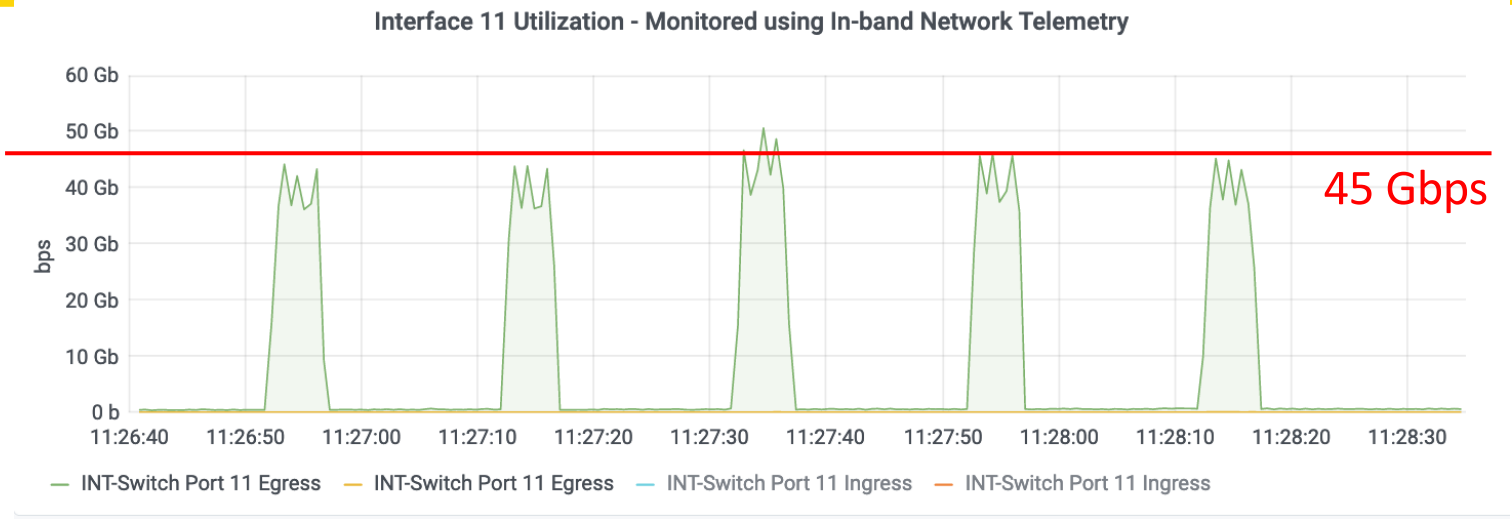
- Report timestamp
- Report sequence number
- Original TCP/IP headers (vlan.id, ip.src, ip.dst, ip.proto, ip.tot_len)

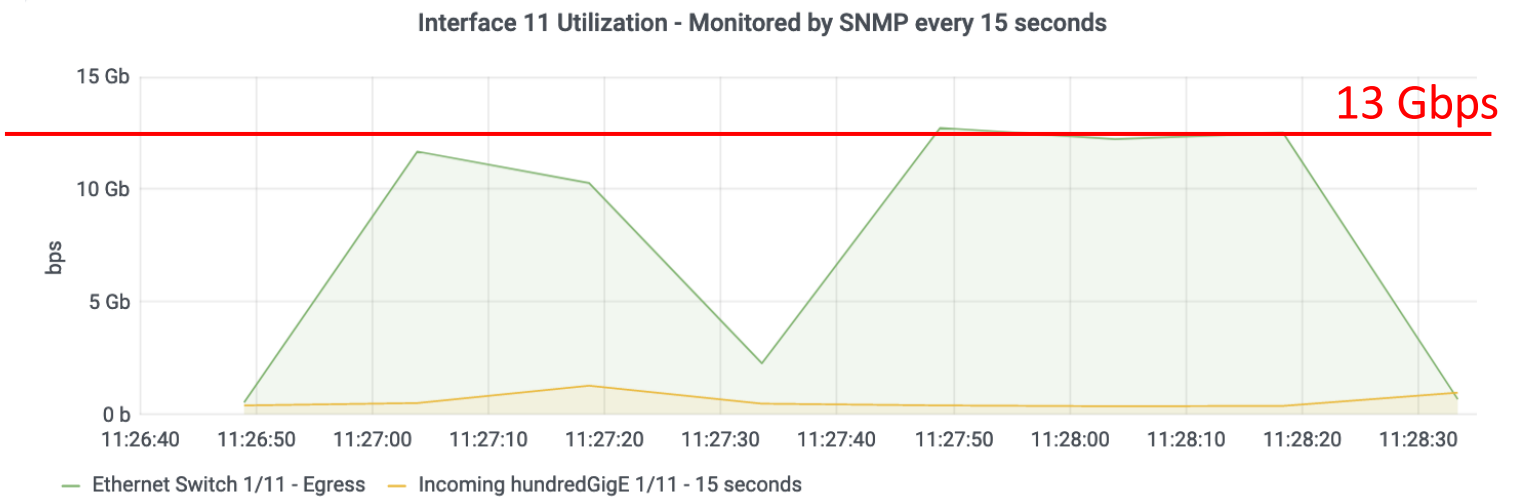
What does AmLight do with the INT data?

- Monitoring & Traffic engineering leveraging INT data
  - Proof of Transit (PoF) or path taken
    - Including LAG member and queue ID
  - Instantaneous Ingress and Egress Interface utilization
    - Accurate bandwidth utilization
    - Microburst detection
  - Instantaneous Egress Interface Queue utilization (or buffer)
    - Source of Packet Drops detection → TCP Retransmissions
  - Instantaneous per-hop delay:
    - Sources of jitter

Examples: Interface utilization

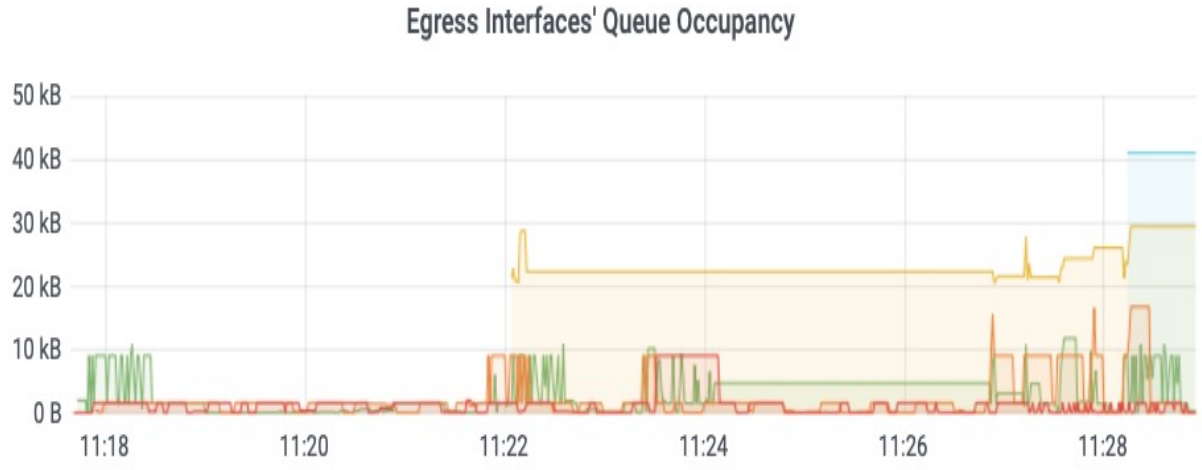
- 5 data transfers/bursts of 40-50Gbps for 5 seconds.
- Top: INT metadata exported in real time, per packet
- Bottom: SNMP get running as fast as supported by the switch: 14 seconds.



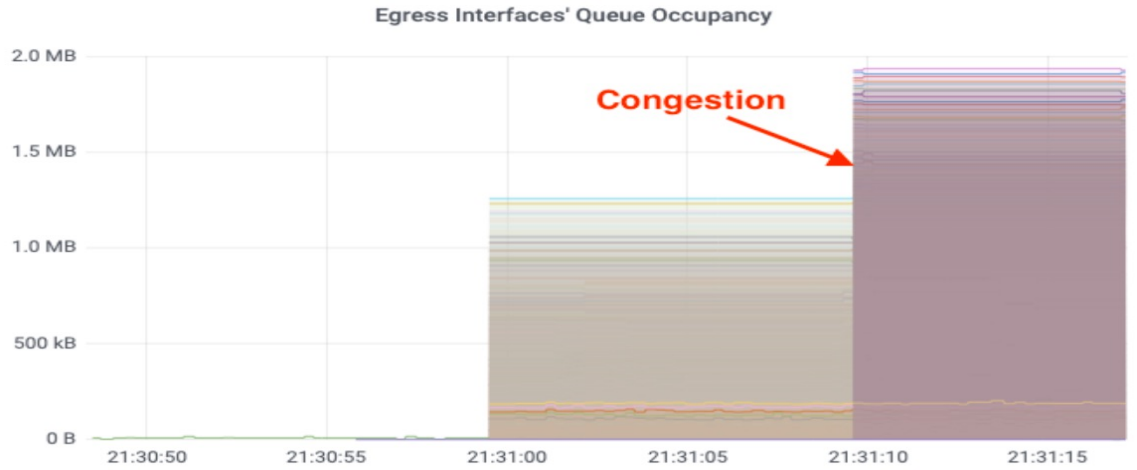


AmLight ExP
Americas Lightpaths Express & Protect

Examples: Instantaneous Egress Queue utilization (or buffer)



Average Buffer Utilization



Under-Congestion Buffers

How can we use INT data to ”tune our networks”

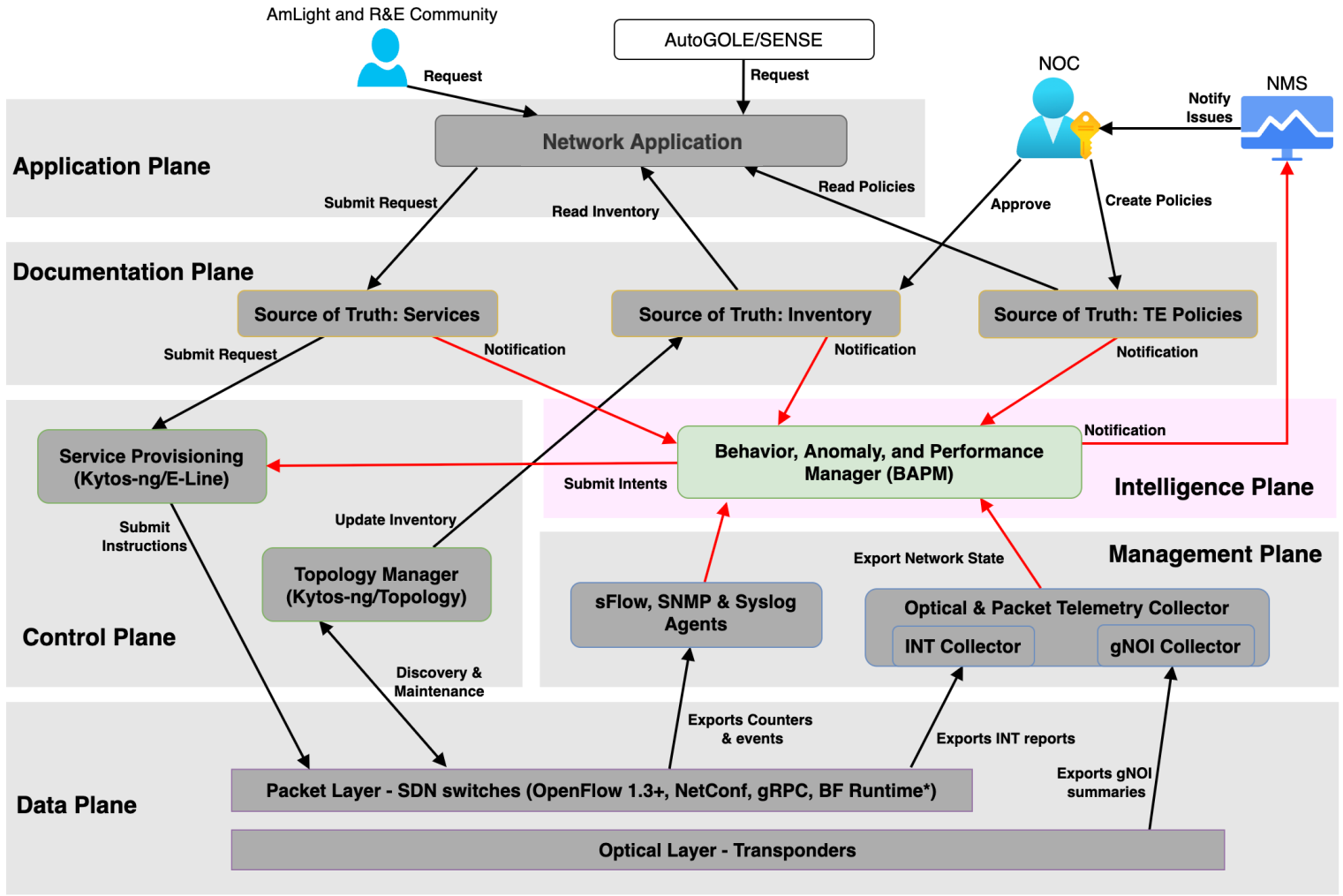
Intelligence Plane:

1. *Gets inventory, policies, and services from the Documentation Plane*
2. *Gets telemetry reports from the Management Plane*
3. *Profiles AmLight’s traffic every 100-500ms*
   - *Discovers performance issues and traffic anomalies*
4. *Makes suggestions to the Control Plane*
   - *Steer traffic, Load balance services, Rate-limit anomalies,*

Self-optimization: be prepared for sub-second reaction and debugging

Example of policies:

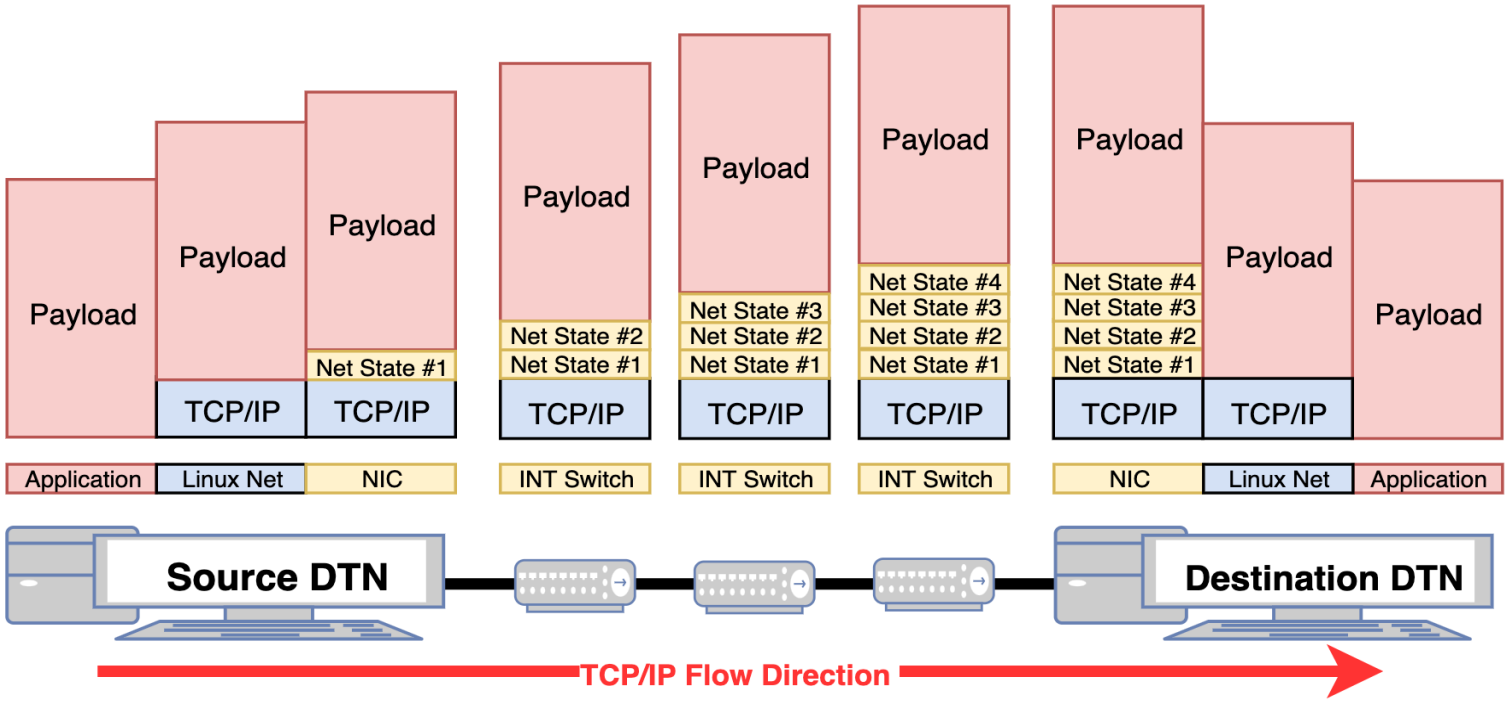
- 80+% BW utilization >= 2s
- 50+% [Queue Occupancy] >= 2s
- Number of path changes >= 5 in 2h



AmLight ExP
Americas Lightpaths Express & Protect

Q-Factor: Sharing INT data with endhosts

- Objective: Improve data transfers over long-haul high-bandwidth programmable networks
- How: Creating an end-to-end framework where endpoints would have network state information to dynamically tune data transfer parameters in real time
  - Bandwidth and resources optimization
  - Tunable TCP pacing based on INT data



- NSF CC*:Collaboration between FIU and ESnet