AmLight ExP
Americas Lightpaths Express & Protect

AmLight-ExP (NSF #OAC-2029283)

AtlanticWave-SDX (NSF #OAC-2029278)

Autonomic Network Functions for Science Applications

2023 Internet2 Community Exchange
May 8-11, 2023

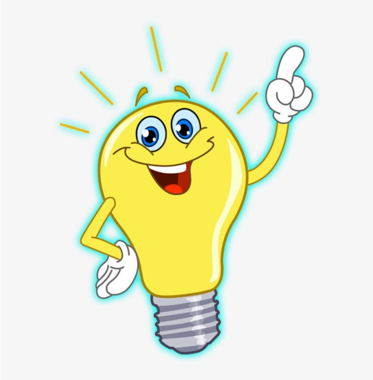
Julio Ibarra, Principal Investigator
Jeronimo Bezerra, Chief Network Architect

Outline

- Introduction
- AmLight Network Infrastructure
- Autonomic Network Architecture (ANA) Introduction
- AmLight SDN Architecture
- Autonomic Network Architecture Use Case: L2VPN
- AtlanticWave-SDX project
- SLA-driven science use case: Vera Rubin Observatory
- Conclusion

Audience and Objective

- This talk is most relevant for network operators
- My objective is to increase your interest about autonomic networking
  - by presenting the work we're doing on AmLight and AtlanticWave-SDX projects



Introduction

Phenomena in the R&E Ecosystem

- Computation, storage and data movement are increasingly distributed
  - Movement of datasets often traverses multiple geographies and multiple R&E network domains
- The number of connected end points continues to increase
  - Sensors and instruments to observe and measure phenomena
  - DTNs, plethora of IoT and network edge devices, end points in the Clouds
    - As a result, management of the R&E network ecosystem is increasingly complex
- Science applications are relying more on network services
  - Dedicated virtual circuit with dedicated network bandwidth between two end points
    - Isolating bulk data transfers from best-effort traffic
  - Dedicated/Scheduled network services are in demand for predictability and reproducibility

Challenges and Expectations

- Science applications expect the network to be available, reliable, and predictable [1]
- Science data management systems are unaware of the network substrate [2]
- Science applications should be able to express time constraints on data delivery
  - *Real-time data transfer, Deadline-Bound data transfer, Background data transfer [3]*
- SLA-driven science applications have explicit requirements of the network
  - Low latency, real-time / on-demand dataset movement [4]

[1] Würthwein, Frank, et al. "Managed Network Services for Exascale Data Movement Across Large Global Scientific Collaborations." arXiv preprint arXiv:2209.13714 (2022).
[2] Barisits, M., et al, "Rucio: Scientific Data Management", Computing and Software for Big Science (2019).
[3] Wu, Wenji, et al. "ROBIN (RuciO/BIgData Express/SENSE) A Next-Generation High-Performance Data Service Platform." 2020 IEEE/ACM Innovating the Network for Data-Intensive Science (INDIS). IEEE, 2020.
[4] https://www.amlight.net/wp-content/uploads/2022/05/Robert_Blum-RubinObservatory.pdf

Example 1: OSG topology and resources

OSG

OSG Compute Federation (OSCF)

UCSD

149 "green dots" listed on this map

Compute Resources at 64 institutions in the USA alone.

(another 16 internationally that support science other than LHC via OSG)

AmLight ExP

as Lightpaths Express & Protect

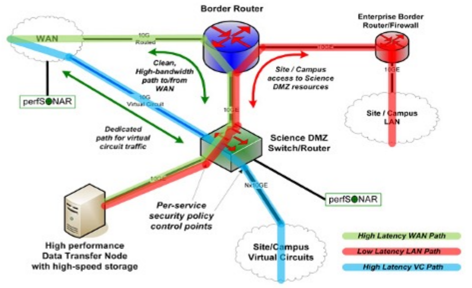
Example 2: CC* Program footprint

The Campus Cyberinfrastructure (CC*) Program

- Networking as a fundamental layer and underpinning of Cyberinfrastructure, driven by scientific R&E needs
- Most awards go to 10/100 Gbps Campus networking upgrades, external connectivity to the national R&E fabric, and campus border redesign prioritizing science traffic.
- ~385 awards totaling >$100M have been made across 50 states and jurisdictions in CC* since 2012
- CC* emphasizes strong campus level partnerships between researchers/teachers and campus IT leadership

Simple Science DMZ Diagram

A simple Science DMZ has several essential components. These include dedicated access to high-performance wide area networks and advanced services infrastructures, high-performance network equipment, and dedicated science resources such as Data Transfer Nodes. A notional diagram of a simple Science DMZ showing these components, along with data paths, is shown below:



The essential components and a simple architecture for a Science DMZ are shown in the Figure above. The Data Transfer Node (DTN) is connected directly to a high-performance Science DMZ switch or router, which is connected directly to the border router. The DTN's job is to efficiently and effectively move science data to and from remote sites and facilities, and everything in the Science DMZ is aimed at this goal. The security policy enforcement for the DTN is done using access control lists on the Science DMZ switch or router, not on a separate firewall.

AmLight ExP
Americas Lightpaths Express & Protect

Motivation for Autonomic Network Functions

- To support science drivers with Service-Level Agreement (SLA) requirements
  - *Example: the Vera Rubin Observatory*
- To increase self-management of the programmable network devices
- To reduce dependency on the human role in network operation
- To reduce costs in network operations by **leveraging** network technologies
  - *AmLight leverages SDN technology, network programmability, network virtualization, In-band Network Telemetry*

AmLight Network Infrastructure…

AmLight Express and Protect (AmLight-ExP) Project

- AmLight-ExP is an international R&E network built to enable collaboration among Latin America, Africa, the Caribbean and the U.S.
- Supported by NSF and the IRNC program under award #OAC-2029283
- Partnerships with R&E networks in the U.S., Latin America, Caribbean and Africa, built upon layers of trust and openness by sharing:
  - Infrastructure resources
  - Human resources

(NSF Award # OAC-2029283)

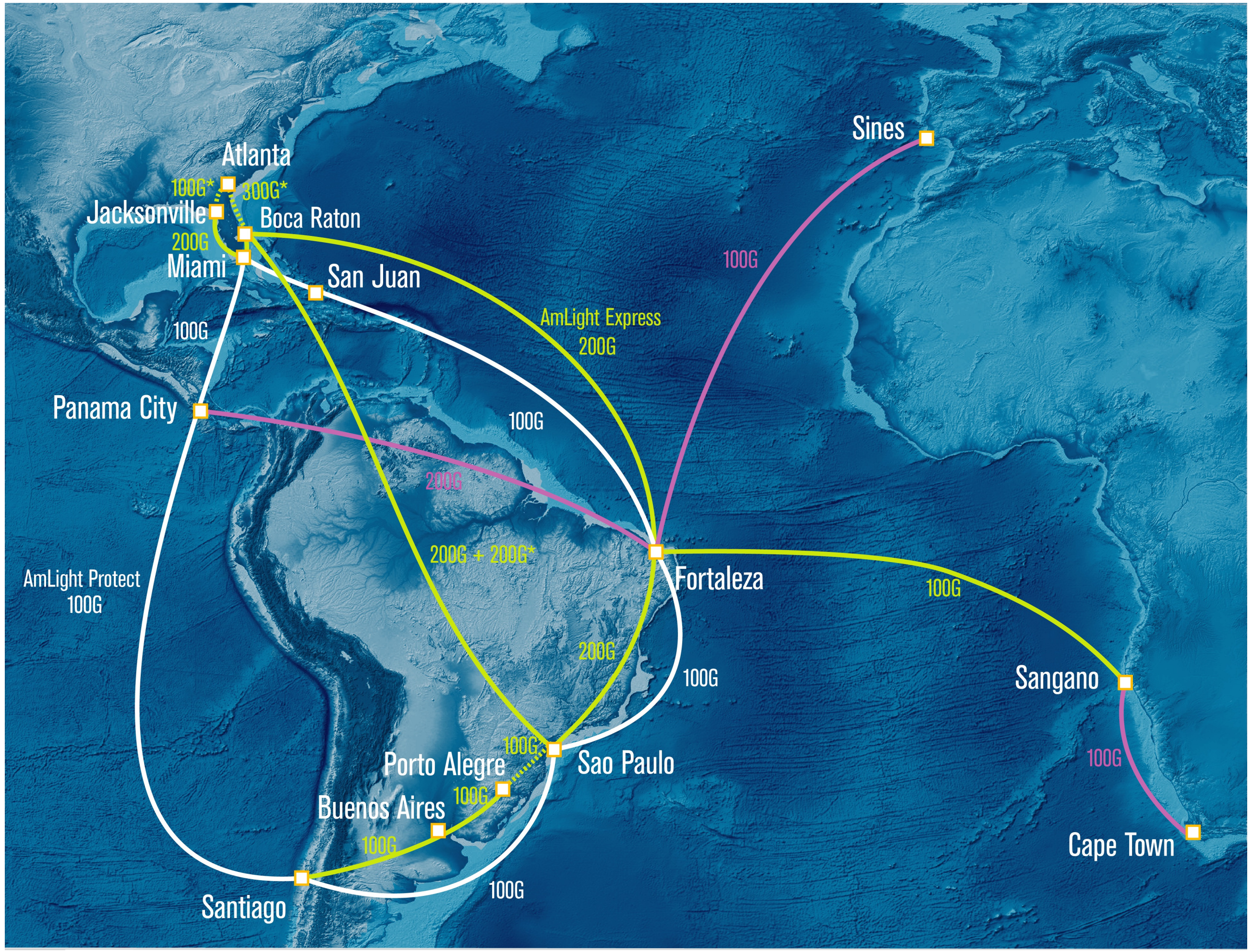
AmLight-ExP Goals

- Vision: To continue enabling collaboration among researchers and network operators in Latin America, the Caribbean, Africa, and the U.S., by providing reliable, sustainable, scalable, high-performance network connectivity and services

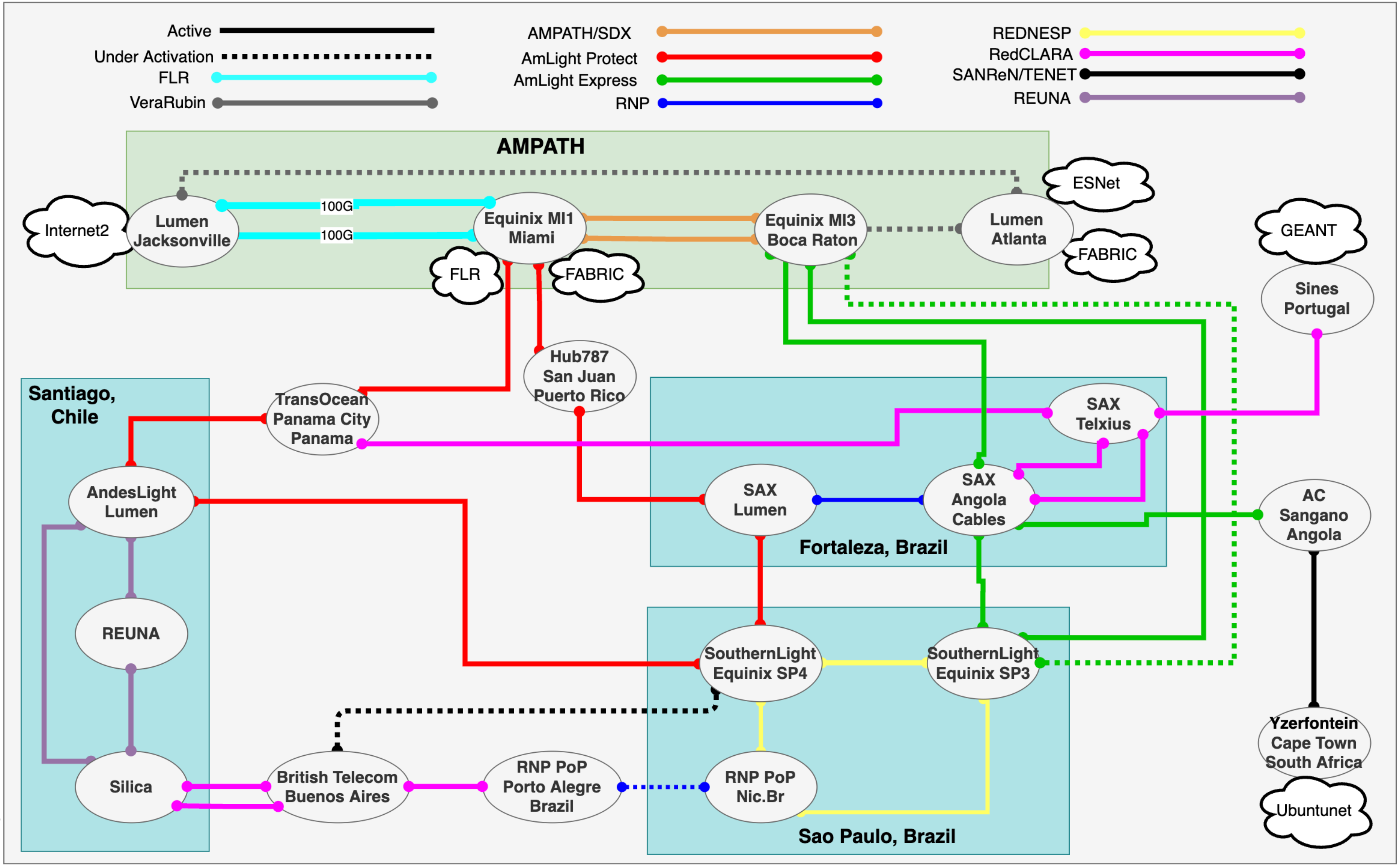
- Focus:
  - Support Service Level Agreement (SLA)-driven science applications
  - Improving network visibility and management
  - Enable integration between AmLight and network-aware science drivers
  - Add new network and cloud services
  - Minimize the human role in network operation

AmLight-ExP Network Infrastructure

- 600G of upstream capacity between the U.S., Latin America, Caribbean and 100G to Africa
  - Blend of Spectrum and Leased Capacity
- OXPs: Florida(3), Brazil(2), Chile, Puerto Rico, Panama, and South Africa, (soon Atlanta, GA)
- Production SDN Infrastructure since 2014:
  - Orchestrators: OESS and Kytos-ng
  - OpenFlow 1.0 and 1.3 Southbound Interfaces
- Programmable Data Plane:
  - P4 -> In-band Network Telemetry (INT)
  - 21 programmable devices in production
- Highly instrumented:
  - PerfSonar, sFlow, Juniper Telemetry Interface (JTI), In-band Network Telemetry (INT)



AmLight ExP
Americas Lightpaths **Express & Protect**



Active
Under Activation
FLR
VeraRubin
AMPATH/SDX
AmLight Protect
AmLight Express
RNP
REDNESP
RedCLARA
SANReN/TENET
REUNA
AMPATH
Internet2
Lumen Jacksonville
100G
100G
Equinix MI1 Miami
FLR
FABRIC
Equinix MI3 Boca Raton
Lumen Atlanta
ESNet
FABRIC
GEANT
Sines Portugal
Hub787 San Juan Puerto Rico
TransOcean Panama City Panama
Santiago, Chile
AndesLight Lumen
REUNA
Silica
SAX Telxius
SAX Lumen
SAX Angola Cables
Fortaleza, Brazil
AC Sangano Angola
SouthernLight Equinix SP4
SouthernLight Equinix SP3
British Telecom Buenos Aires
RNP PoP Porto Alegre Brazil
RNP PoP Nic.Br
Sao Paulo, Brazil
Yzerfontein Cape Town South Africa
Ubuntunet

Autonomic Network Architecture (ANA)…

Autonomic Network Architecture (ANA)

- Autonomic systems were first described in 2001 (Kephart and Chess, 2003)
- Documented in IETF RFC 7575, 7576, 8993, 8994; Internet Protocol Journal v24, no3.
- The fundamental goal is self-management, comprised of several self-* properties
  - Reduces dependencies on human administrators or centralized management systems
  - ANA functions enable a network to adapt to a changing environment
- Closed-loop control
  - Mechanism of self-management functions that include Collect, Analyze, Decide, and Act processes
  - AmLight refers to this closed-loop control mechanism as Closed-Loop Orchestration

Autonomic Network Architecture Definitions

- **Self-configuration**: Functions do not require configuration by either an administrator or a management system. Self-knowledge, discovery and intent to self-configure
- **Self-healing**: Autonomic functions adapt on their own to changes in the environment and heal problems automatically
- **Self-optimizing**: Autonomic functions automatically determine ways to optimize their behavior against a set of well-defined goals
- **Self-protection**: Autonomic functions automatically secure themselves against potential attacks

Automatic to Autonomic Networking Categories

	Automatic	Automation	Closed-Loop Orchestration	Autonomic
Description	Operator runs a script to change a service or configuration	Operator runs a "playbook" to change multiple services and configuration of multiple nodes at the same time	SDX Controller changes multiple services and configuration of multiple nodes. Nodes export state and counters. Application reacts to the new state and performs or not new changes in a forever loop.	Autonomic devices discover assets, policies, and intents. Configure devices from scratch based on policies and intents. Minimal to no-user interaction (plug & play). Resolution of conflicts defined by administrators
User-Input	Scripts, inputs, topology, destination	Scripts, inputs, inventory	Scripts, inputs, inventories, policies/conditions/triggers	Policies and intents

More Human Interaction ⟷ Less Human Interaction

Goal: Closed Loop Orchestration

	Automatic	Automation	Closed-Loop Orchestration	Autonomic
Description	Operator runs a script to change a service or configuration	Operator runs a "playbook" to change multiple services and configuration of multiple nodes at the same time	SDX Controller changes multiple services and configuration of multiple nodes. Nodes export state and counters. Application reacts to the new state and performs or not new changes in a forever loop.	Autonomic devices discover assets, policies, and intents. Configure devices from scratch based on policies and intents. Minimal to no-user interaction (plug & play). Resolution of conflicts defined by administrators
User-Input	Scripts, inputs, topology, destination	Scripts, inputs, inventory	Scripts, inputs, inventories, policies/conditions/triggers	Policies and intents

Goal

More Human Interaction

Less Human Interaction

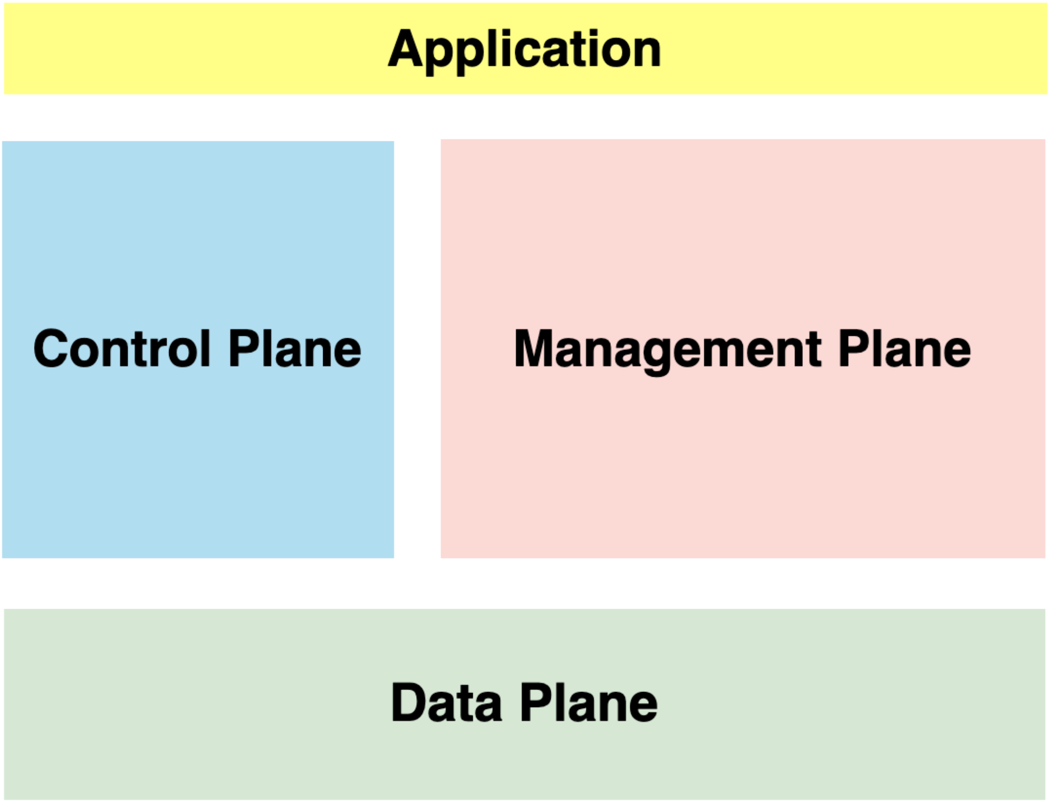
AmLight SDN Architecture…

AmLight SDN Architecture – 2015 - 2020

- From 2014-2020, AmLight followed the “basic” SDN layers, as in [1]
  - Application, Management, and Control Planes were very coupled under the same controller/orchestrator:
    - OESS, ONOS, Kytos, and Ryu
    - Each SDN Plane was operated as modules of the SDN controllers’ software stack
  - Data Plane was a blend of vendors (Dell, Corsa, Brocade) with mixed support
  - Southbound Interface was OpenFlow 1.0 and OpenFlow 1.3

- From the ANA perspective, only *self-healing* was in place
  - Fiber cuts and device outages were handled by finding backup paths.

- Source of Truth was the SDN controllers’ databases.
  - Extra service data had to be documented in an external repository, for instance, contacts in case of issues with the service.

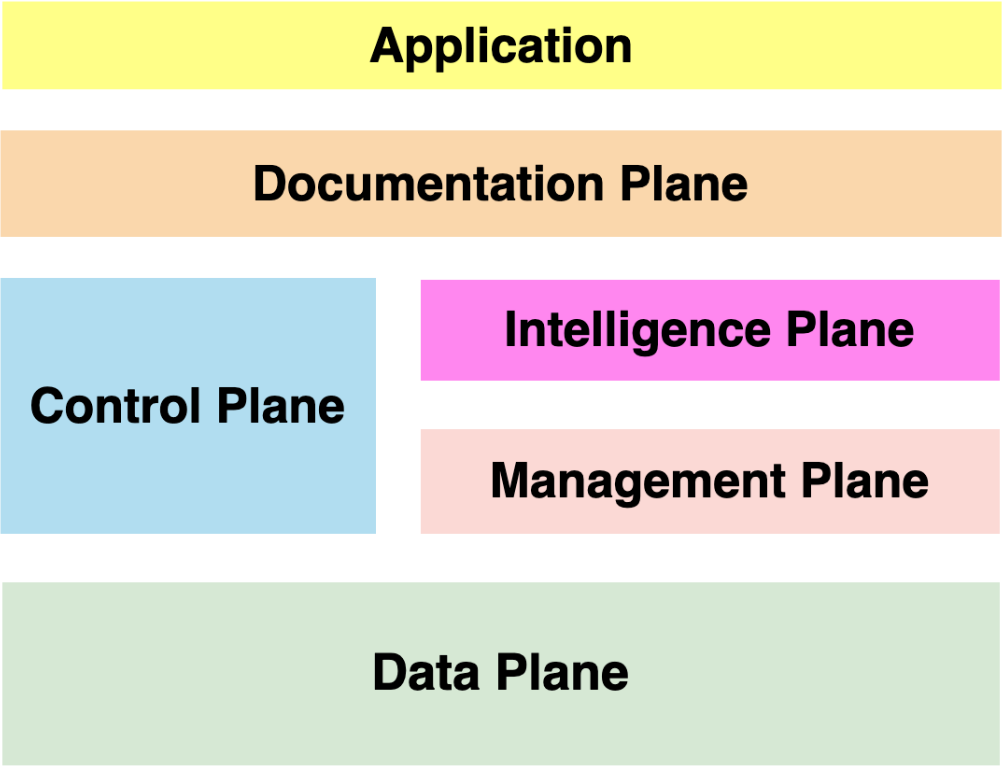


[1] IETF RFC 7626 & RFC 8597

AmLight SDN Architecture – 2021-2025

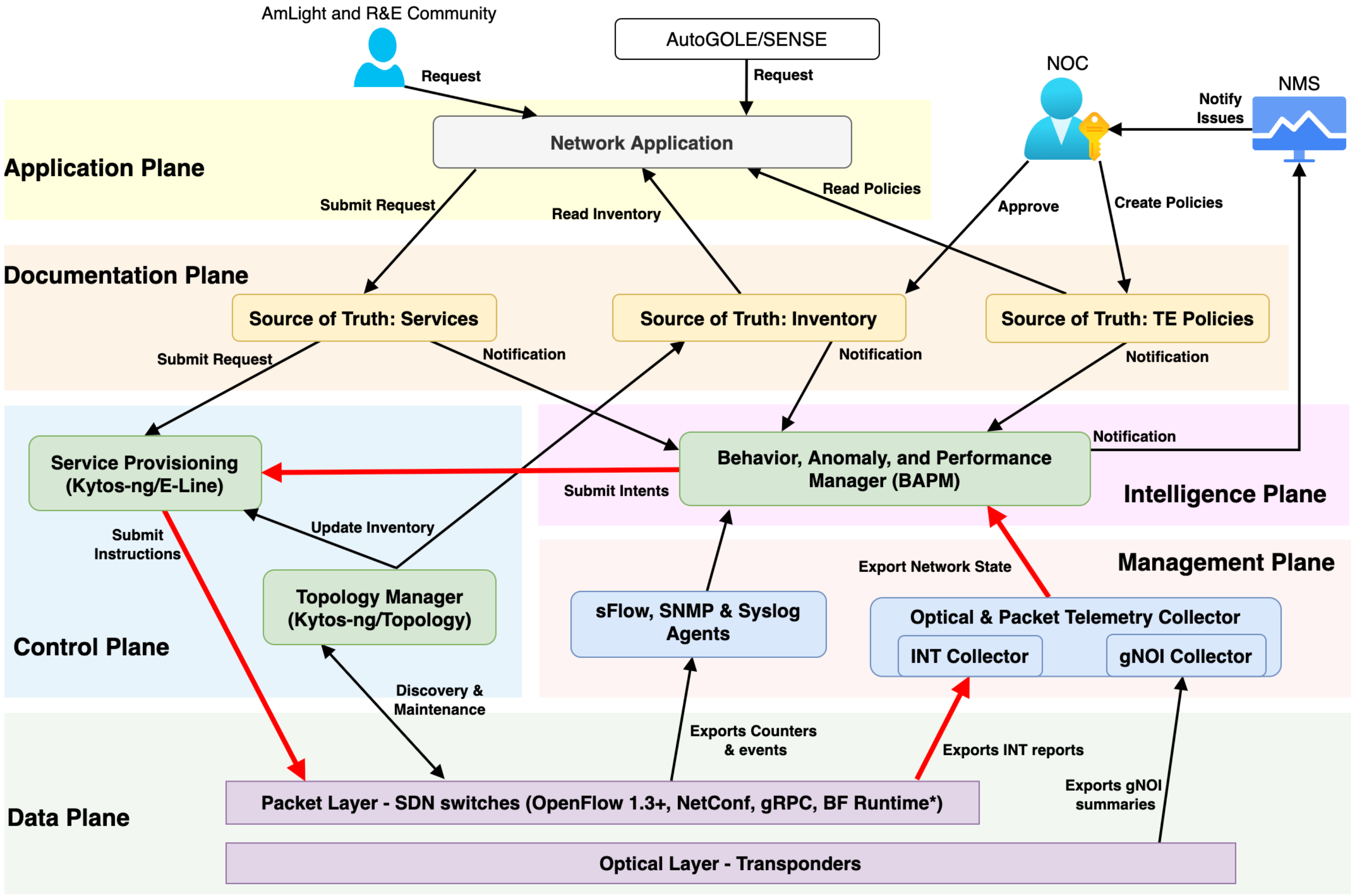
For 2021-2025, specialized components will be added per SDN Plane:

- **Documentation Plane** (Source of Truth): decoupled from the Control Plane
  - NetBox for Source of Truth. Developing plug-ins to use with NetBox.
  - Kytos-ng for the Control Plane
- **Management Plane** (utilization counters): decoupled from the SDN Controller
  - In addition to SNMP and OpenFlow counters, AmLight will use In-band Network Telemetry for counters
- With the Documentation Plane and network telemetry, a new plane was created to add intelligence to AmLight:
  - <u>Intelligence Plane</u> will focus on learning the network state and create a closed-loop control for self-optimization
  - <u>Intelligence Plane</u> will use In-band Network Telemetry to audit how the Control Plane implements the user requests
- **Application Plane** will be decoupled from the SDN Controller
  - To support self-configuration, the Application Plane will interface only the Documentation Plane for services, inventory, and policies. Control Plane will be notified of user requests to act upon



AmLight SDN Architecture – Big Picture

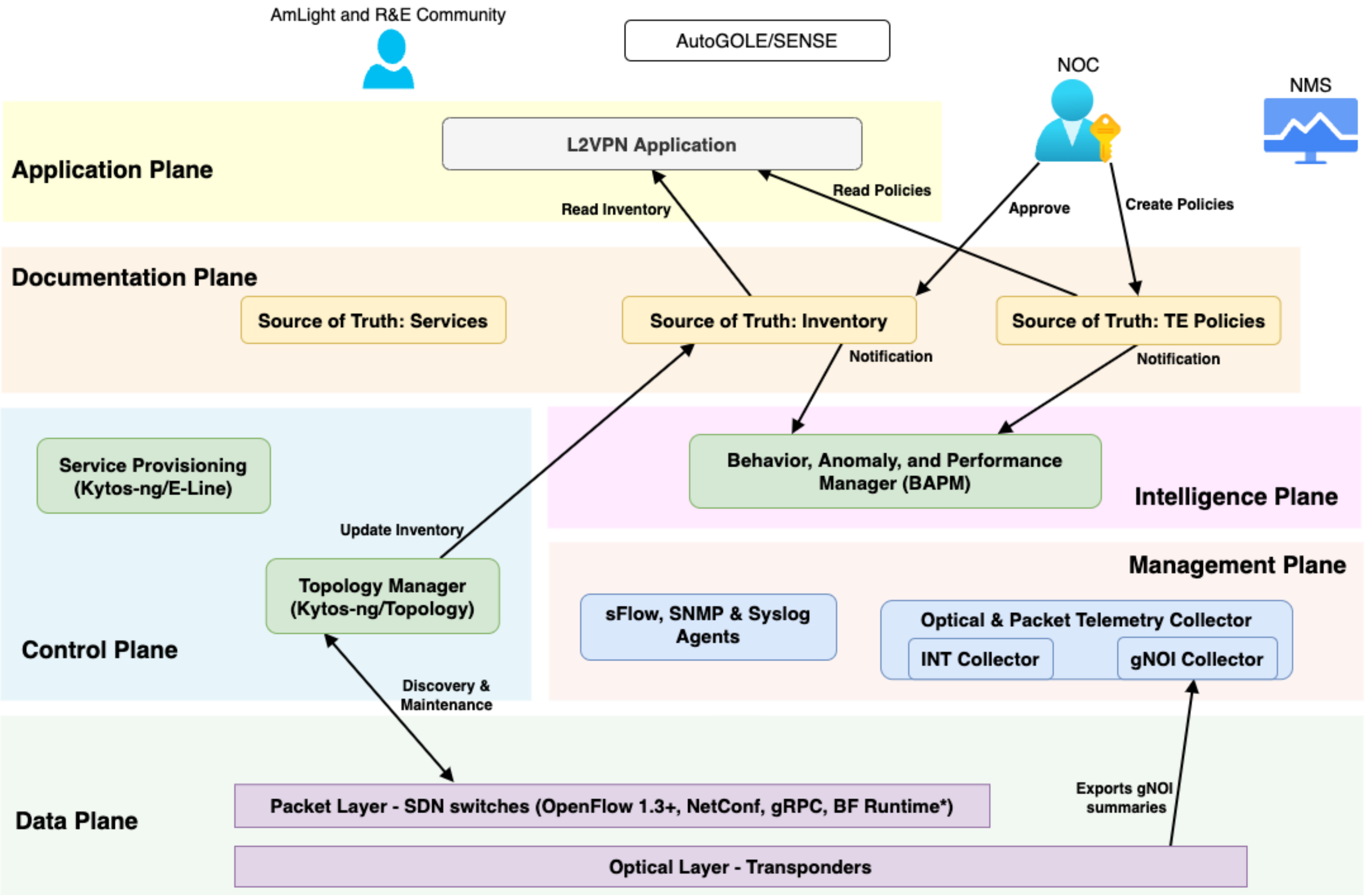
- The Autonomic Network Architecture specification enables partial deployment and support:
  - AmLight won't try to support all ANA functions and requirements
  - Focus will be on a few Autonomic Functions (AF)
- The first Autonomic Function planned is supporting L2VPNs fully managed by this architecture:
  - Administrators will get involved just in case of conflicts and when new policies are needed
- Users will have access to all details related to their services:
  - Topology, counters, telemetry, policies, number of events, and maintenance windows
- A closed-loop control mechanism will be created for self-optimization (red arrows)



ANA use case: L2VPN

ANA use case: L2VPN

- *The first Autonomic Function (AF): L2VPN*
- Topology Manager discovers the substrate and submits it for approval to the NOC
- NOC approves the topology elements
- NOC creates Traffic Engineering policies:
  - For instance, average utilization over 2 minutes should be < 65% of the link capacity
  - L2VPNs with higher priority shouldn't be moved
  - Only move L2VPNs up to 3 times every 10 minutes
  - Don't move L2VPNs that have opted out for no protection

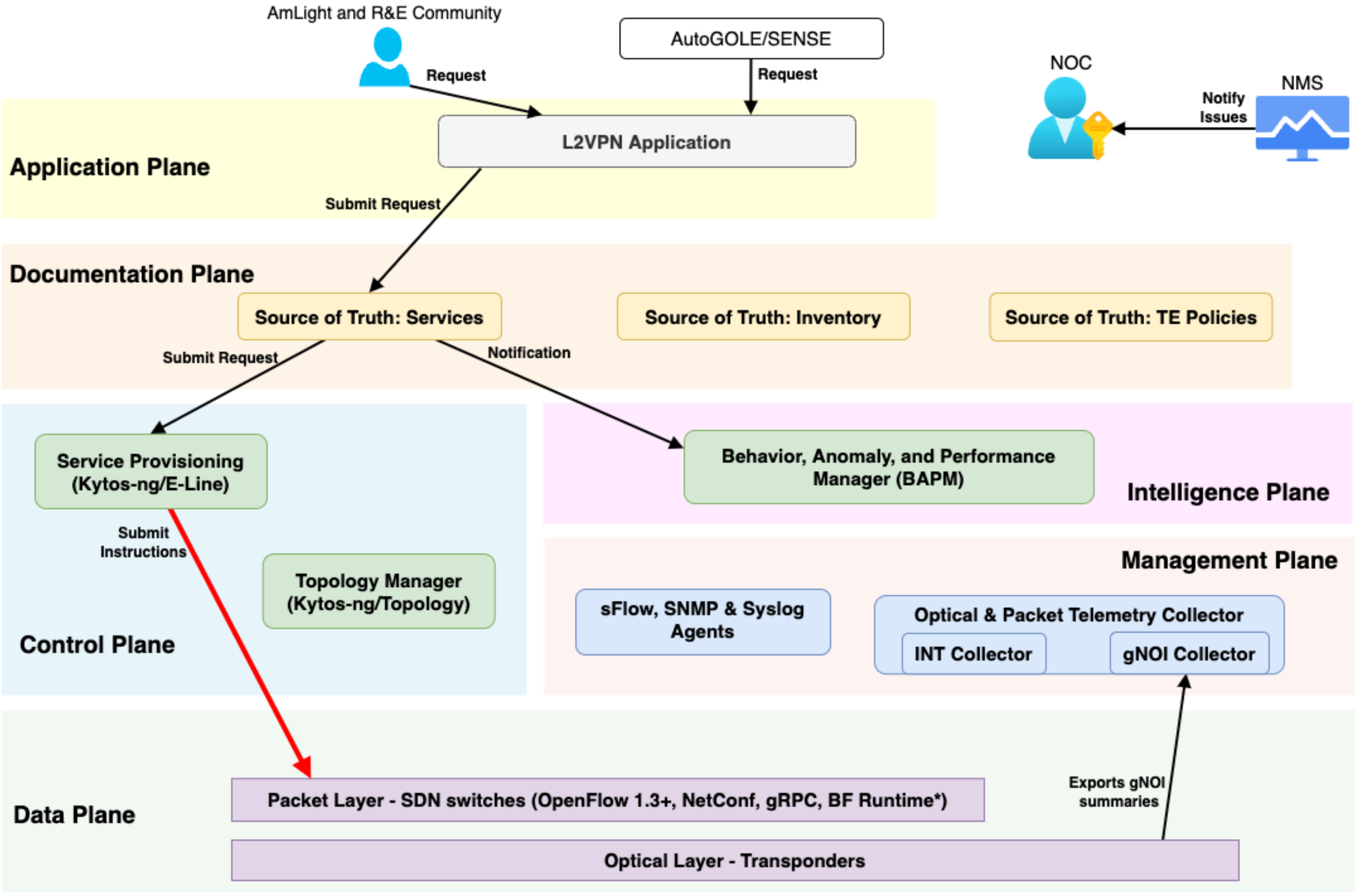


ANA use case: L2VPN [2]

The first Autonomic Function: L2VPN

AmLight user requests a L2VPN via Application Plane

1. Application Plane:
   - submits request to Documentation Plane that stores all metadata provided and triggers an event
2. Control Plane (Kytos E-Line):
   - gets triggered by the event,
   - retrieves the provided metadata,
   - computes a path based on the metrics provided
   - pushes flows to Data Plane
3. Data Plane (SDN switches):
   - creates flow entries
   - forwards user packets as requested



AmLight ExP
Americas Lightpaths Express & Protect

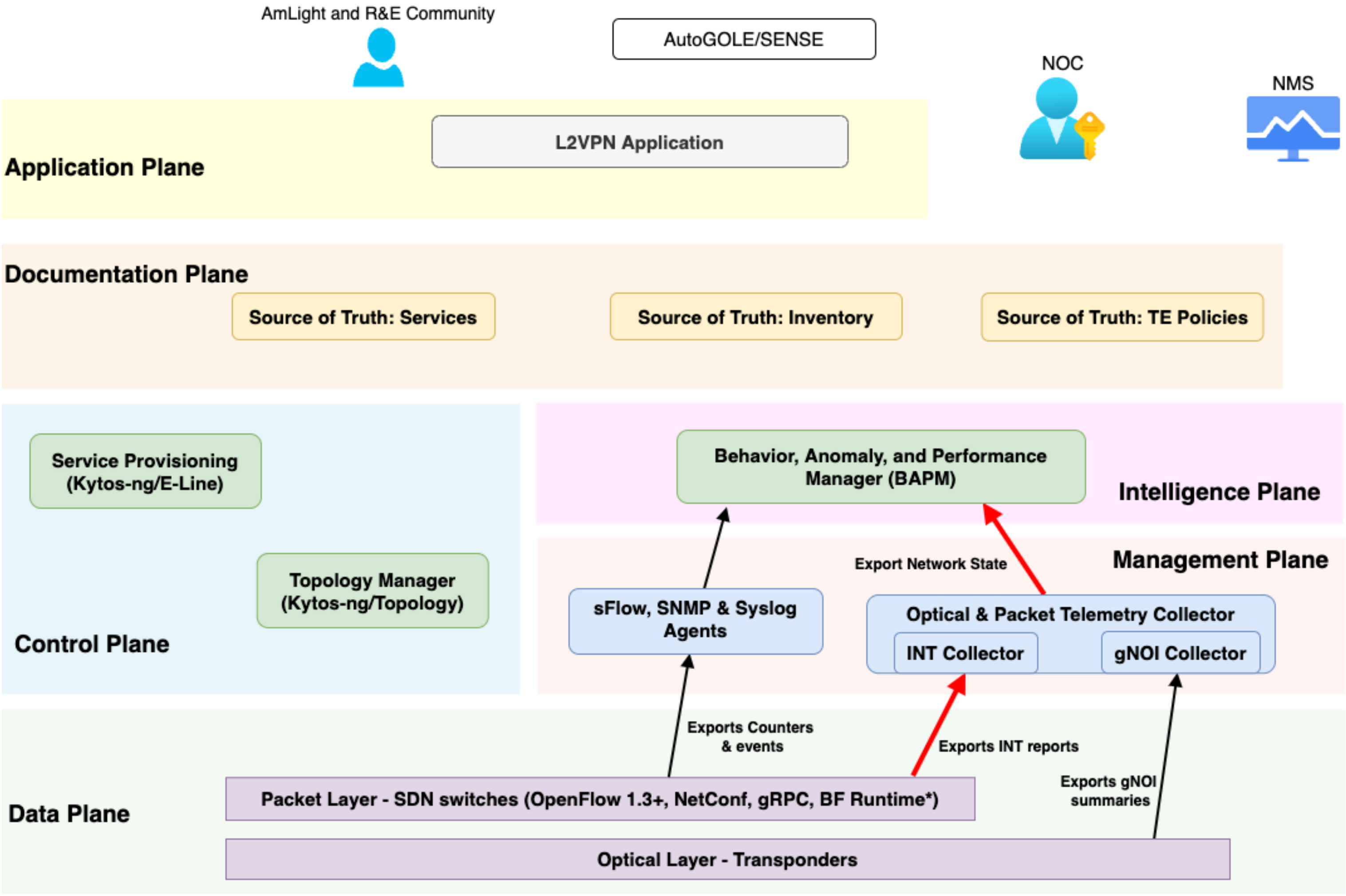
ANA use case: L2VPN [3]

The first Autonomic Function: L2VPN

AmLight user requests a L2VPN via Application Plane

4. Data Plane (SDN switches):
   - Forwards user packets as requested
   - Exports per-packet telemetry data to Management Plane

5. Management Plane (INT Collector):
   - Parses telemetry data looking for thresholds
     - Bandwidth utilization, queue occupancy, hop delay, path taken, …
   - If a threshold is reached, submit a network state report to Intelligence Plane



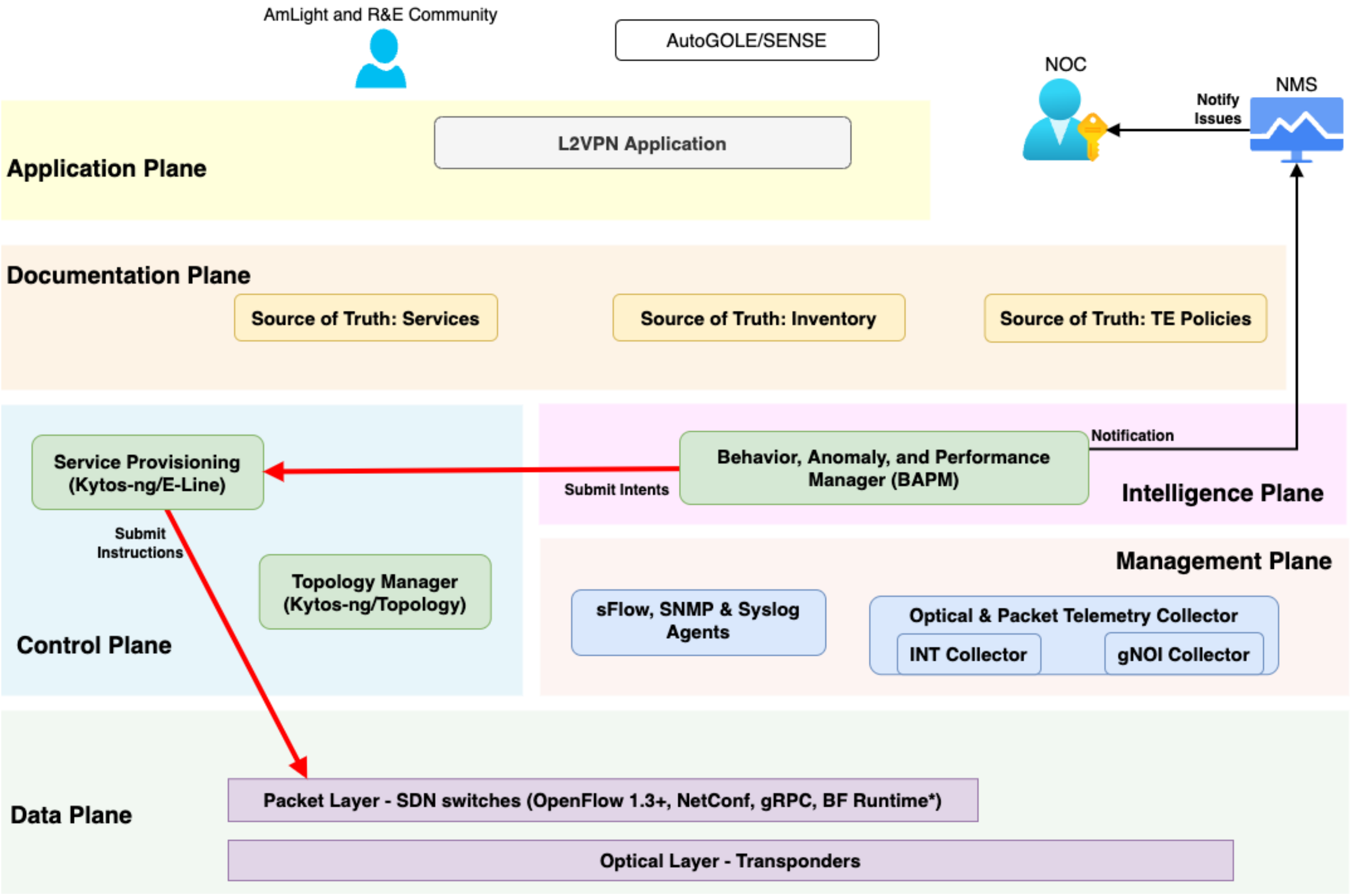
ANA use case: L2VPN [4]

The first Autonomic Function: L2VPN

AmLight user requests a L2VPN via Application Plane

6. Intelligence Plane (BAPM)
   - BAPM analyzes the network state report and policies
   - If a nonconformity is identified, BAPM uses the Services source of truth to identify if/which L2VPNs sharing the same resources (interface, queue, switch) should be relocated.
   - BAPM sends an intent to Kytos-ng E-Line asking to move L2VPNs [A, B, C] out of resource [Z].

7. Control Plane (Kytos-ng E-Line)
   - retrieves the provided metadata,
   - computes a path based on the metrics provided
   - Pushes flows to the Data Plane



ANA use case: L2VPN [5]

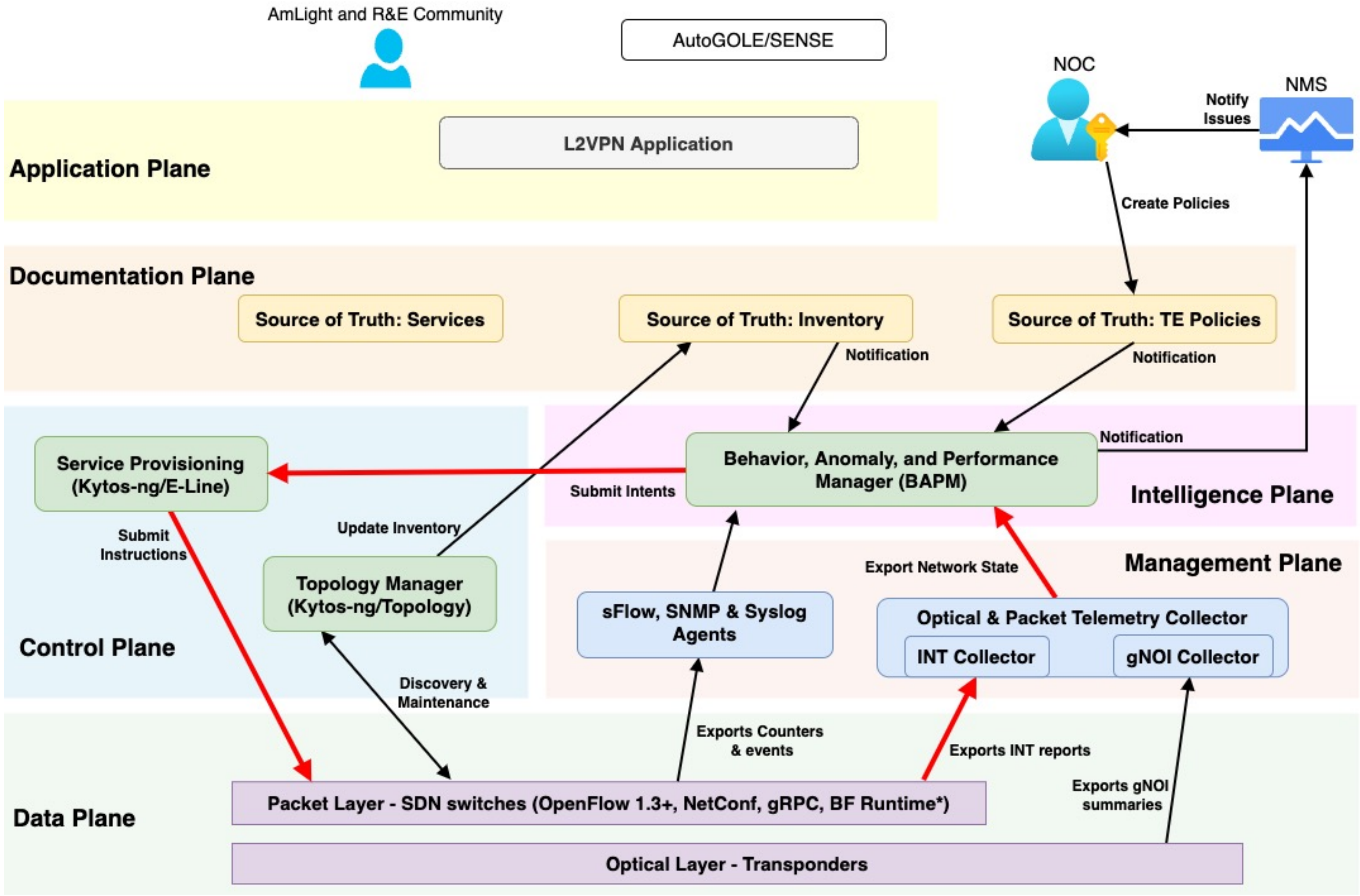
The first Autonomic Function: L2VPN

AmLight user requests a L2VPN via Application Plane

8. Forever loop as long as policies aren't changed or L2VPN decommissioned.

Roadmap: Self-Optimizing the network:

- Year 3: < 2 seconds
- Year 4: < 1 second
- Year 5: < 500 ms



AmLight ExP
Americas Lightpaths Express & Protect

AtlanticWave-SDX project:

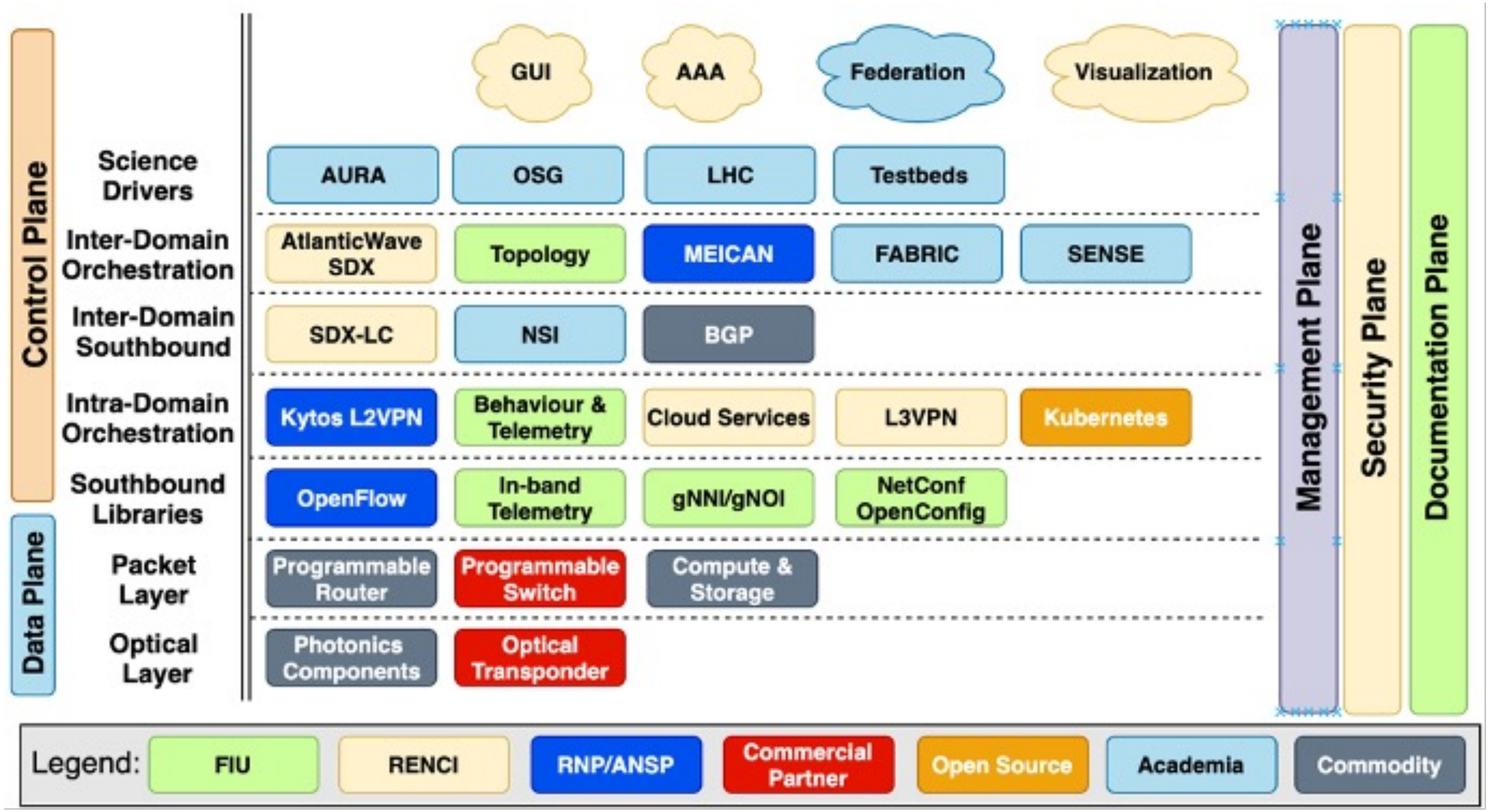
About the AtlanticWave-SDX

- AtlanticWave-SDX is a NSF International Research Network Connections (IRNC) project, award #OAC-2029278
- Vision: To operate a multi-domain distributed Open Exchange Point (OXP) fabric at production quality
  - Having the capability to react to unplanned network conditions and events
- Methodology: Add autonomic network functions to provide self-managing functions
  - for self-adaptability and context-aware behavior changes
  - in response to the emergent properties of the network environment

AtlanticWave-SDX Project Goals

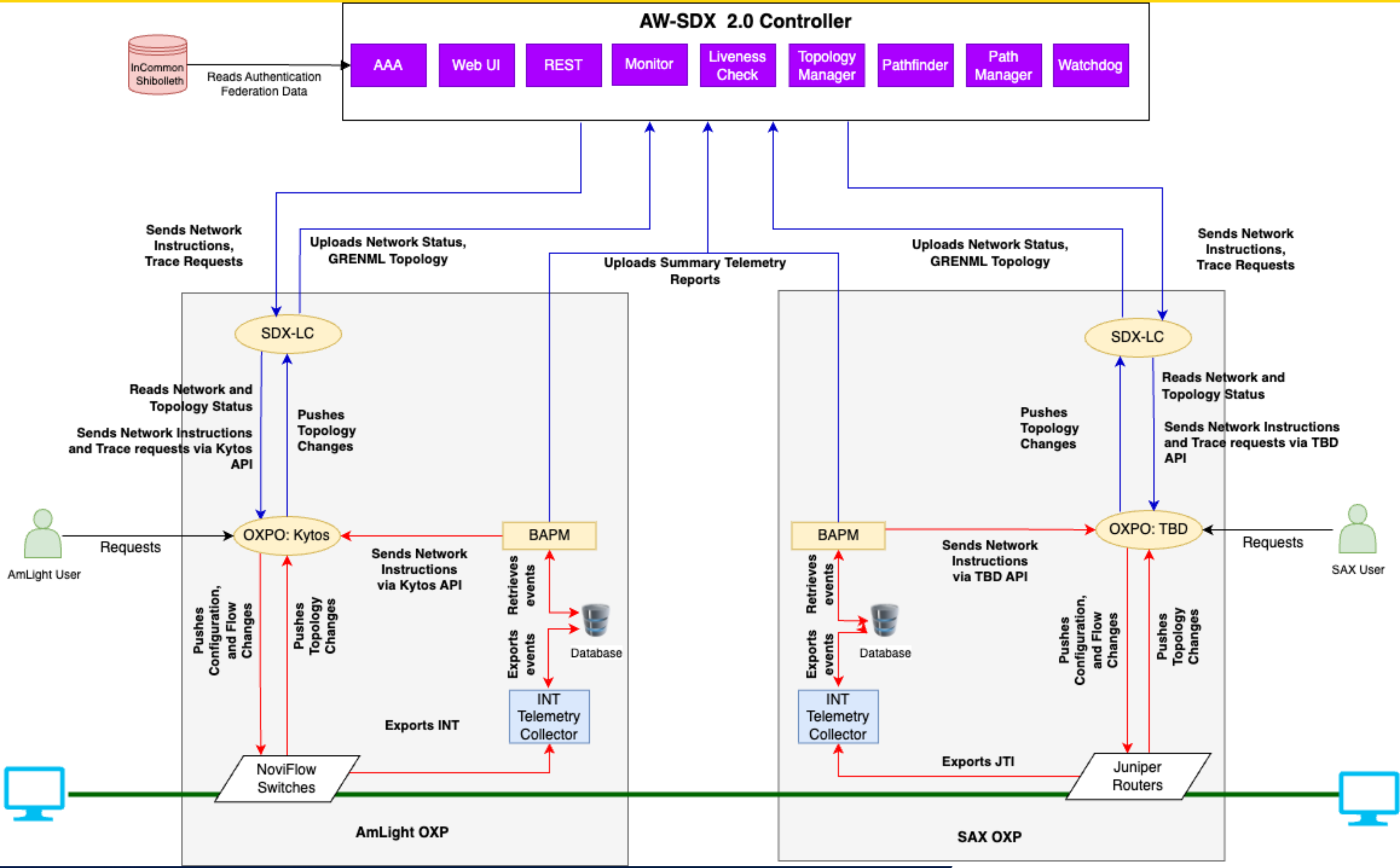
- To improve the multi-domain distributed SDX between the U.S., Latin America and Africa to operate at production quality
  - Improvements are reasoned by use cases of science drivers:
    - SLA requirements of the Vera Rubin Observatory
    - Challenges of the High Luminosity LHC experiments, and
    - To support a FABRIC community in South America
- To evolve the development, integration and deployment of the AtlanticWave-SDX
  - Adds ANA functions to the AtlanticWave-SDX resources to enable closed-loop orchestration capabilities on both the intra-domain (per-OXP), and inter-domain levels
- To coordinate with OXPs and Networks towards adopting ANA functionality
  - To improve support for end-to-end SLA-driven science applications

Architecture and Stakeholders



Closed-loop Orchestration

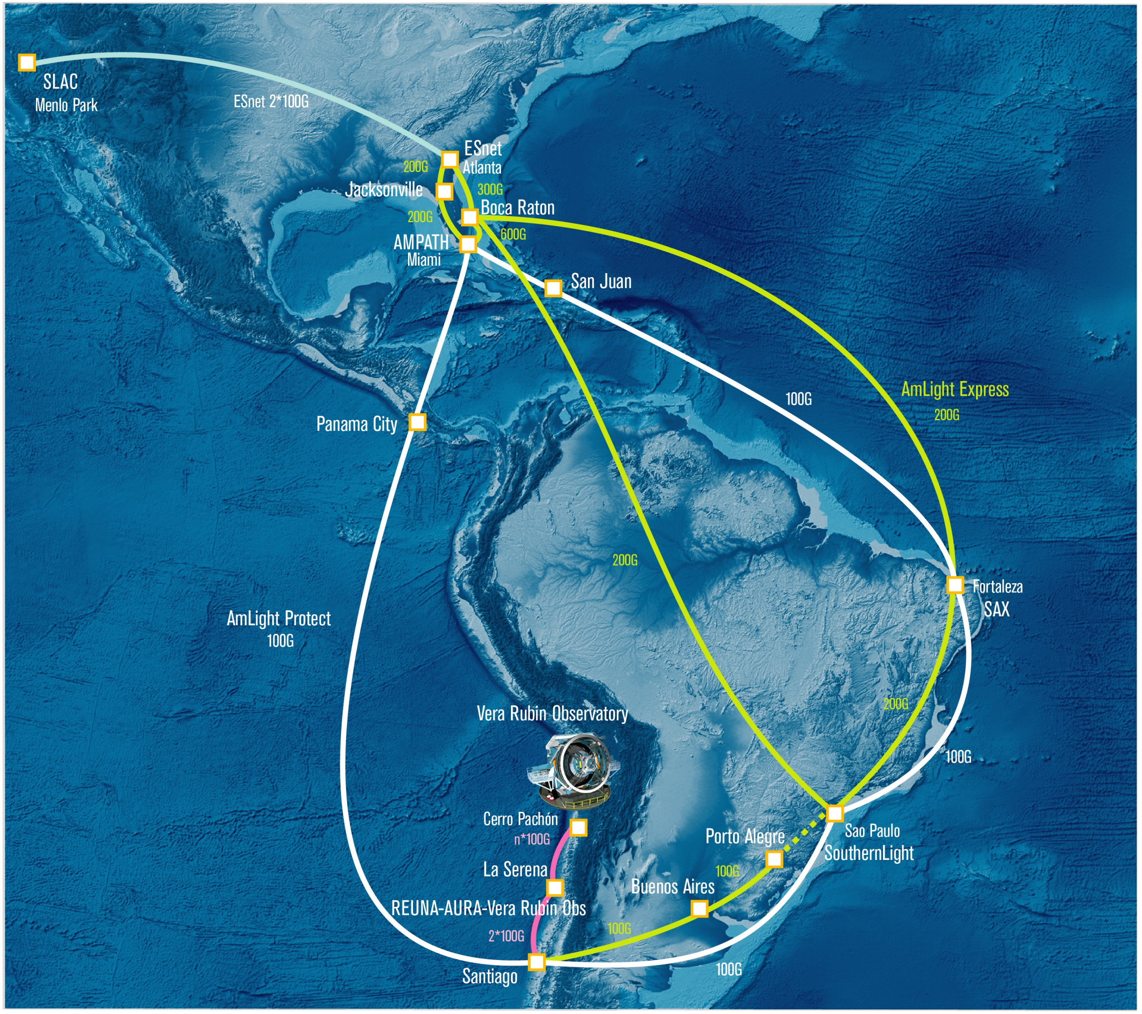
- Per-OXP Orchestration:
  - OXP chooses its Orchestrator to support its operation
  - OXP decides what Autonomic Functions to support
- Inter-Domain Orchestration
  - SDX defines interfaces and data models for OXPs
  - OXPs produce and consume data from the SDX Controller
  - SDX creates a full topology
  - SDX supports all inter-domain network functions



SLA-driven science use case: Vera Rubin Observatory

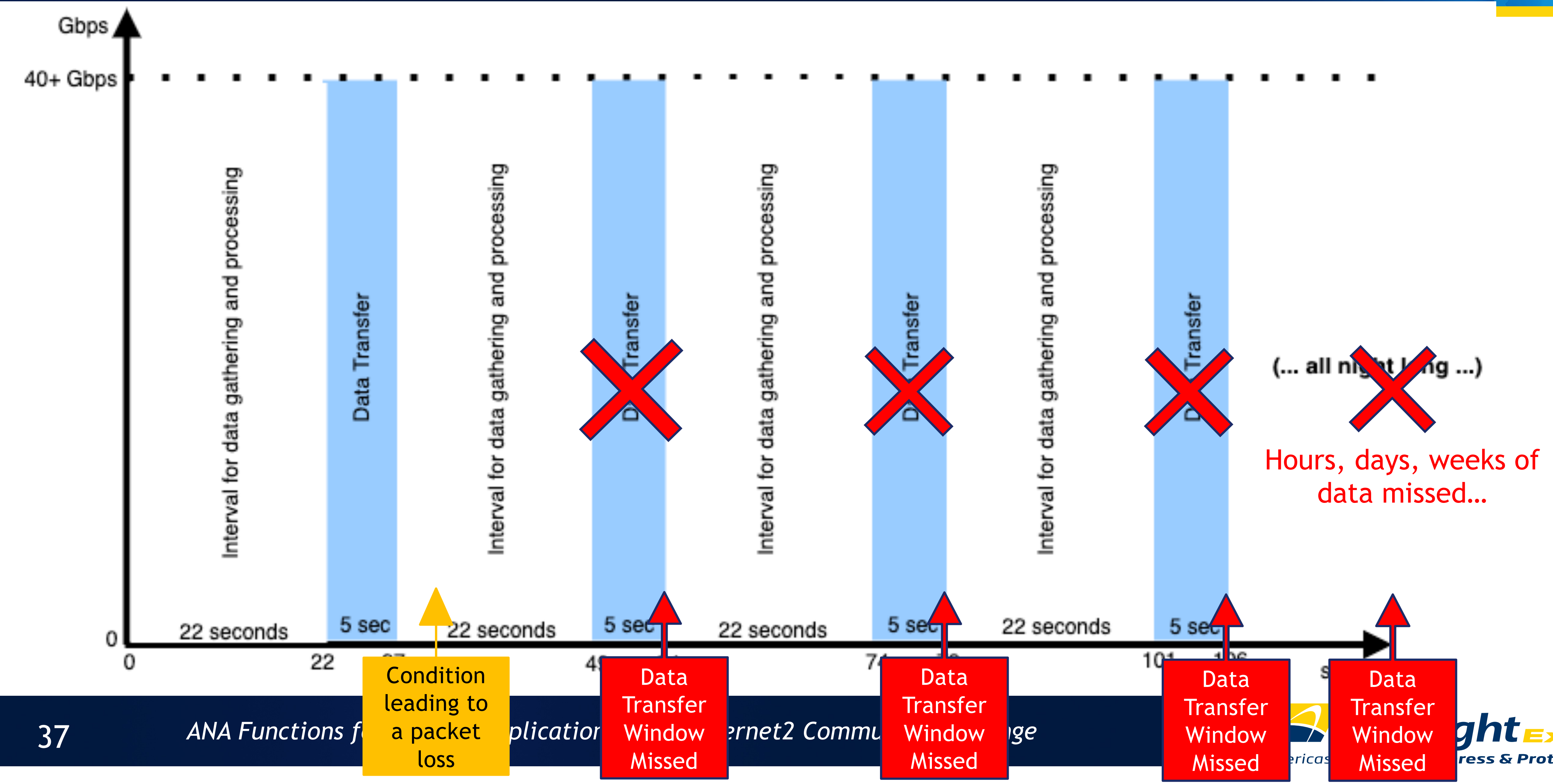
Vera Rubin Observatory operation use case

- Vera Rubin is a large-aperture, wide-field, ground-based optical telescope under construction in northern Chile
- The 8.4 meter telescope will take a picture of the southern sky every 27 seconds, and produce a 13 Gigabyte data set
- Each data set must be transferred to the U.S. Data Facility at SLAC, in Menlo Park, CA, within 5 seconds, inside the 27 second transfer window
- Challenges
  - High propagation delay in the end-to-end path
  - RTT from the Base Station to the USDF is approximately 180+ ms
  - 0.001% of packet loss will compromise the Rubin Observatory application



AmLight ExP
Americas Lightpaths Express & Protect

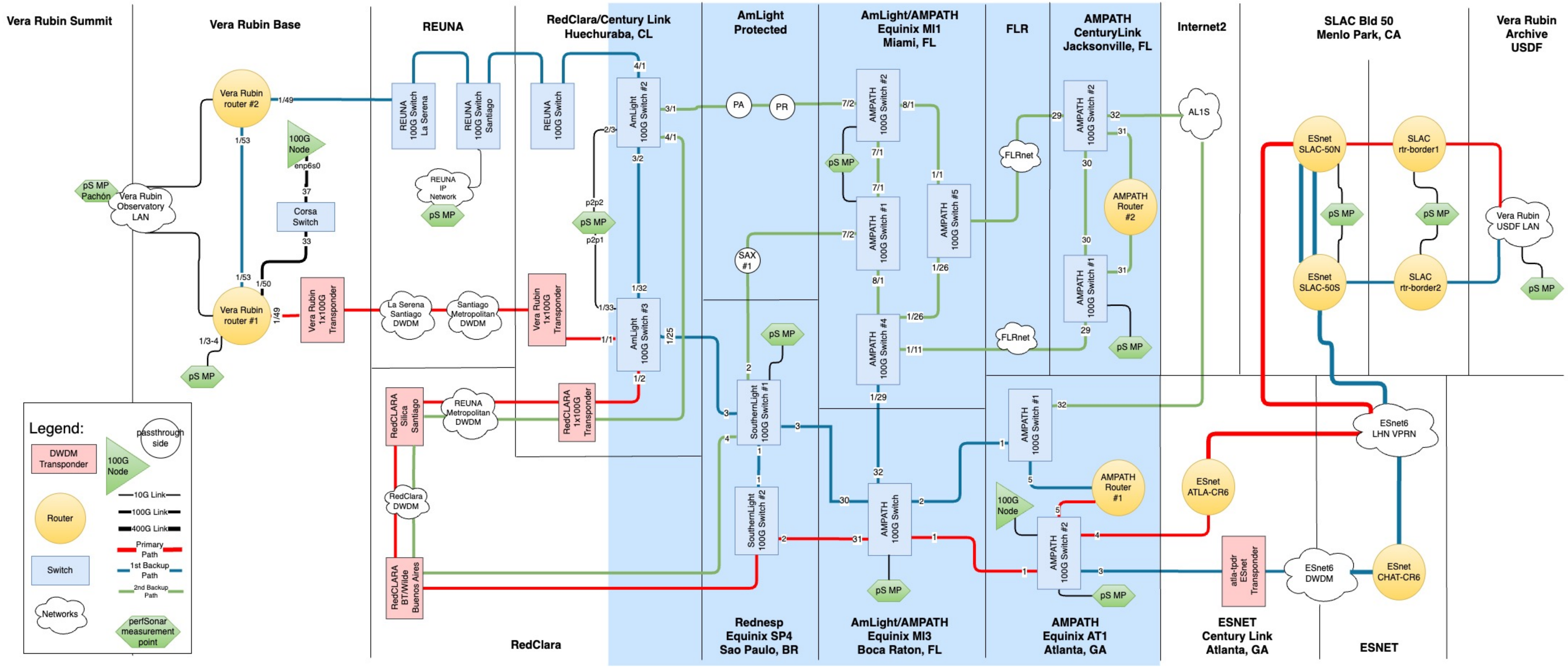
The Use Case: Vera Rubin Obs's operation



Versioning

Author: NET team
Last update: Apr 21th, 2022

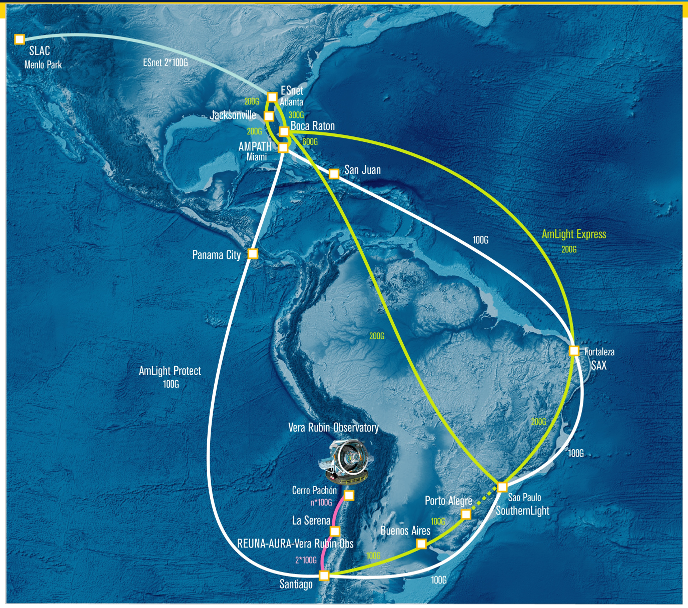
Vera Rubin Observatory LHN - FY2023

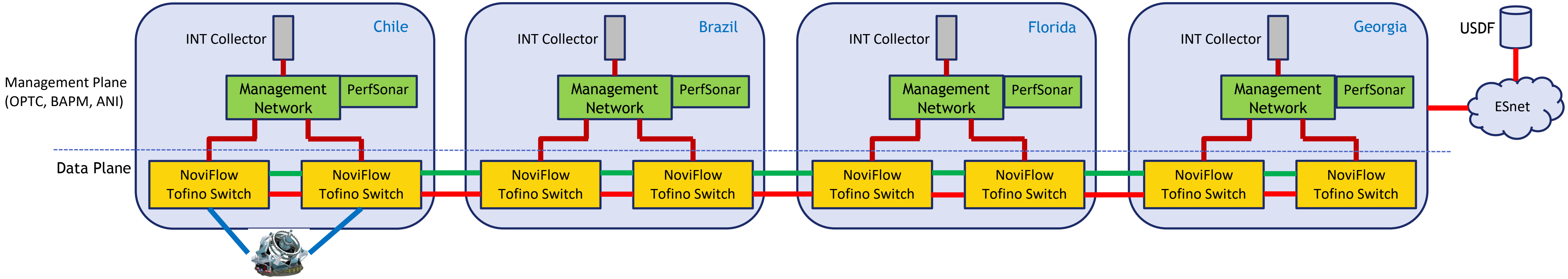


AmLight ExP
Americas Lightpaths Express & Protect

Instrumented for SLA-grade network resilience

- AmLight is Instrumented for SLA-grade network resilience to support Vera Rubin
  - Express and Protect paths are instrumented with INT and PerfSonar
- AmLight's Management Plane
  - Processes telemetry report
  - Isolates and detects traffic anomalies
  - Validates performance thresholds
  - Computes risk profiles of optical and IP layer metrics in a closed loop
  - Reacts to packet loss and packet performance in real-time
- AmLight's metric for success is to not miss a data transfer window





Instrumented for SLA-grade network resilience

Tool/Framework	Accuracy depends on:	Challenges:	Used for:
SNMP	➢ Data Plane counters collection interval. ➢ SNMP collector polling.	➢ Low interval ➔ higher CPU utilization. ➢ High interval ➔ lower accuracy.	➢ General monitoring.
sFlow	➢ Sampling rate.	➢ Low sampling rate ➔ more storage required ➔ higher CPU utilization. ➢ High sampling rate ➔ lower accuracy.	➢ Troubleshooting unusual events. ➢ TOP N reports.
Juniper Telemetry Interface (JTI)	➢ Data sending interval.	➢ Low interval ➔ more storage required. ➢ High interval ➔ lower accuracy.	➢ Environments that require more granular information.
In-band Network Telemetry (INT)	➢ Real time. Complete visibility.	➢ Processing all data collected in real time.	➢ Troubleshooting short-time events.

[2]

[2]- Evaluating INT, JTI, and sFlow @ AmLight – 2022 Internet2 TechEx

AmLight ExP
Americas Lightpaths Express & Protect

Conclusion:

AmLight supports SLA-driven science applications

- AmLight has many links and multiple paths between its sites:
  - From Chile to Jacksonville, there are more than 25 possible paths to take
  - With the new SDN architecture, AmLight expects to effectively load balance network services across links, while respecting user constraints and requirements

- AmLight has a SLA-driven packet-loss-intolerant and sub-minute-response-time-expected science application:
  - With per-packet telemetry and sub-second network profiling capacities, AmLight will be prepared to react to network conditions under 1 second
  - With optical telemetry, AmLight will anticipate issues with the substrate and steer traffic out of the substrate before adverse events happen

- AmLight engineering team prefers to focus on engineering and new services than manual activities:
  - With the closed loop control, some time-consuming operational activities will be performed without human intervention
  - In the end, AmLight engineers will have more time for developing and automating routines


Los Angeles
New York
Miami
Panama
Fortaleza
Rio de Janeiro
Sao Paulo
Santiago
Luanda

THANK YOU