

# MetrANOVA

## Data Sharing and Technical Developments

SIG-NOC 23 - October 14th, 2025



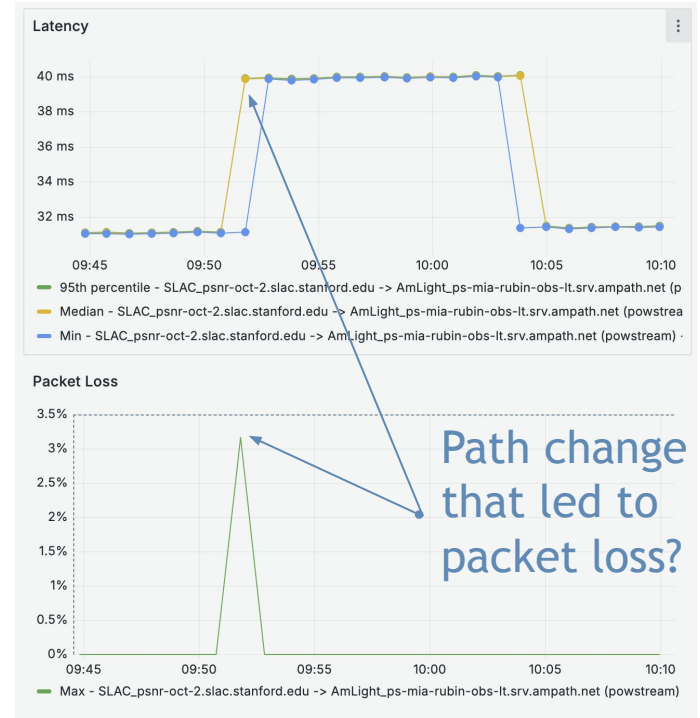


# What is MetrANOVA?



# Motivation: Troubleshooting is hard!

- Something interesting clearly happened
- I can only see my network and no obvious path changes
- What do I do now?
  - Do the other networks have public portals?
  - Do I create a ticket somewhere...even though I am not sure where the problem is?
  - Do I know someone that can help?
  - Do I just ignore it and hope users didn't notice?



## Can we do better?



# MetrANOVA's Mission

1. **Advocate for end to end Network Measurements with appropriate access within all of R&E**
  - a. Provide training and policy guidance
  - b. Create knowledge base articles and howtos
  - c. Train the next generation of engineers to maintain Network Measurement as a core competency
  
2. **Lower the barriers through technical and policy collaboration to create a federated technical capability**
  - a. Reduce need for bespoke solutions
  - b. Amortize software sustainment costs through collaboration.
  - c. Develop federated architectures that preserve local control and autonomy



# Consortium Details

## Goals

- Tools, Tactics and Techniques
- Align with perfSONAR where possible
- Develop and Share
  - Open Architectures
  - Technical Components
  - Design Patterns
  - Best Practices
  - Policy Recommendations.

## Vision

- <https://github.com/MetrANOVA/.github/blob/main/profile/vision.md>
- A collaboratively developed ecosystem exists
- Open Source, loosely coupled, without cloud service dependence
- Solid foundation for production services and innovation
- Facilitate data driven design in engineering and operations

## Executive Committee

- Provides governance and oversight.
- Decides on new membership organizations.
- Representatives from each member.
  - Inder Monga - ESNet
  - Ivana Golub - PCSS/GÉANT
  - James Deaton - Internet2
  - Luke Fowler - Indiana University GlobalNOC
  - Nathaniel Mendoza - TACC
  - Ed Balas - Consortium Lead

## Participation Model

- Member Organizations
  - Requires  $\geq 1$  Full Time Staff Equivalent
  - Participates in governance process
- Affiliates
  - Any organization or individual able to contribute.
  - Lower bar to participate, more flexibility



# The Secret Sauce of Research and Education

- **Timeless design constructs**

- Ubiquitous Access
- Loose Coupling
- Vendor Neutrality
- Open Standards
- Rough consensus and working code

- **Combined with community focus**

- Our values differentiate us
  - not for profit
- technologies applied to facilitate scientific and educational endeavors.





# Data Sharing



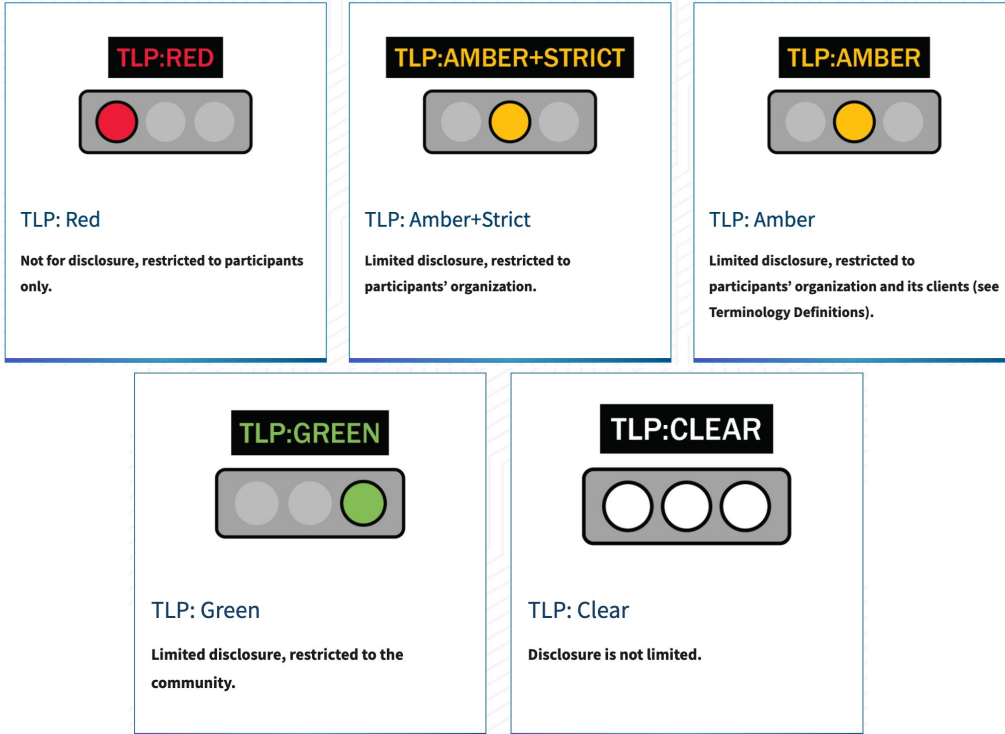
# Different kinds of Sharing

- Different data products have different levels of sensitivity
  - Raw measurements
  - API access to measurement repository with query language
  - Access to online dashboards and reports
- Clear policies helps set expectations
  - What is and what is not shared and at what level of access
- An example:
  - perfSONAR measurement dashboards
    - Performance monitoring data
    - Geopositioning data
    - Display in a single pane of glass.
    - Less policy constraints what is collected and what is shared



# Signalling Sensitivity

<https://www.cisa.gov/news-events/news/traffic-light-protocol-tlp-definitions-and-usage>



- TLP has some drawbacks/areas for clarification
  - What is a “participant”?
- Lots of advantages!!
  - Pre-defined.
  - Easy representation.



# Data Sharing Considerations Document

- Currently working on draft of document that covers a number of considerations when it comes to sharing network data like flow, interface data and more.
- Topics covered such as labeling, classification and enforcements in a number of contexts
  - Example: If you are a provider storing network data on behalf of another organization? What types of things do we have to consider?
- Goal is to not be overly prescriptive while still providing useful guidance on building data sharing federations
- Hope to publish document by end of year

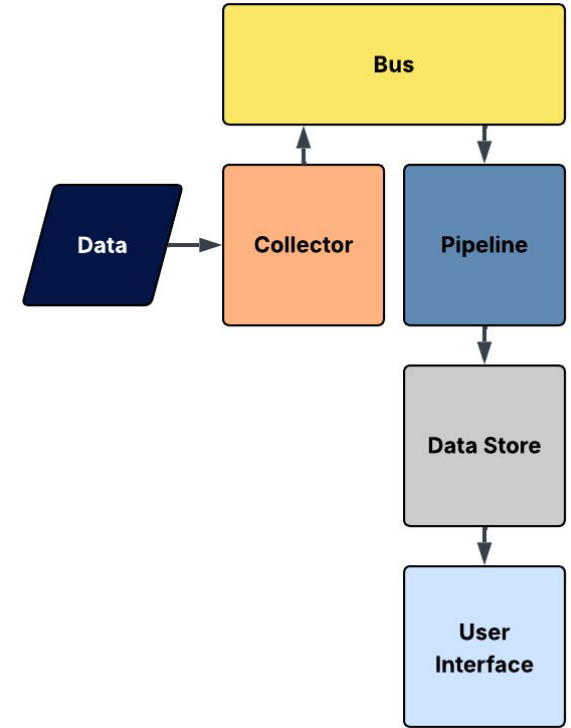


# Technical Recommendations & Software Stack



# Building a Software Stack

- Initial demo between ESnet, Internet2, and I-Light
- Focus on flow data collection and Interface (SNMP/Streaming Telemetry) collection at first
- Implement concepts defined in policy workstream
- Use off-the-shelf open source components as much possible
- Provide a relatively easy to use deployment for small/medium scale
- Provide technical guidance for large scale
  - Also define small,medium and large scale





# Data Store Evaluation

- Performed data store evaluation using representative data sets of snmp interface and flow data
- Compared databases suitable for high-cardinality time series data:
  - ClickHouse
  - ElasticSearch
  - MongoDB
  - OpenSearch
  - TimescaleDB
- Initial findings publick in linked document
- Identified two final contenders: **ClickHouse** and **Elasticsearch**



[Link to results document](#)



# Next Steps

- Finalize data store decision
- Build initial pipelines and dashboards
- SC25 Deployment
- Make a beta deployment available as container images and k8s helm chart
- Deploy on multiple of the partner networks...**because we eat our own dog food!**



ClickHouse



elasticsearch





# Summary

- MetrANOVA is a relatively new consortium looking to expand the scope of network data shared between R&E networks
- We are attacking the problem from both the policy definition and technical deployment aspects
- Expect a document soon on data sharing considerations
- We've already published results of datastore evaluation.
- Working toward a software stack that we can share with community



# Questions?

