MetrAN@VA

A Consortium for Advancing Network Observation and Analysis

> Introduction and updates Ed Balas <u>ebalas@es.net</u> July 16th 2024



©2024 MetrANOVA Contributors • Licensed CC BY-SA 4.0 • https://www.metranova.org

What is Measurement and Monitoring?

- Measuring the network over time, and creating useful things with the observations
 - o maps, dashboards, capacity plans, operational alarms, annual reports
 - Whatever helps improve situational awareness, and lets you tell the story
- Typically involve a handful of data types
 - Ports stats, Network Flow Summaries, Optical Performance, Routing Tables, End to End Perf
- Includes systems many of you use today
 - perfSONAR, Nagios, Prometheus, TICK, Stardust, Netsage, Kentik, Arbor Networks, Deepfield

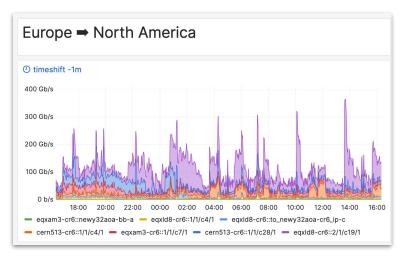






Where we want to be

- 1. Neworks will employ data-driven design and operations
- 2. Diverse Measurements will be combined with metadata to provide a more detailed understanding of the network
- Composite / End to End views of the R&E infrastructure will give stakeholders appropriate awareness



$\leftarrow \rightarrow \triangle$ 0 A	តិដី https://console.internet2.edu/#/?command=show+version&nodes=agg1.da ហ្នំ	${\times}$	\pm	lii\ 🗊	0	2	≫ ≓
📌 🛛 Insight Cons=le	Q. Search organizations and docs		U	sability Ter	ting Provi	de Feedbaci	* 🖬 🤅
Community Interfaces Looking Glass					E	dward Bala	s Sign ou
$Q_{\ }$ Filter by node name or location	[2 modes selected] > show version	Run Comma			Supported	Commands	i i
Aggregation Router	agg1.dall3 > show version	•			nnect inform		
agg1.ashb Ashburn, VA	Cisco IOS XR Software, Version 7.3.3 LNT Copyright (c) 2013-2022 by Cisco Systems, Inc.			show lac LACP infor			
agg1.dall3 Dallas. TX	Build Information: Built By : ingunawa			show lld LLDP neig	neighbor nbors	s	
agg1.egch	Built On : Sat Jan 29 10:33:37 UTC 2022 Build Host : iox-ucs-860 Morkspace : /auto/srcarchive16/prod/7.3.3/8800/ws			show route IP routing table			
Chicago, IL	Version : 7.3.3 Label : 7.3.3			show version Show router firmware version			
Bigg1.losa2 Los Angeles, CA	<pre>cisco 8800 (Intel(R) Xeon(R) CPU D-1530 0 2.40GHz) cisco 8202-SYS (Intel(R) Xeon(R) CPU D-1530 0 2.40GHz) processor with 32GB of memory aqq1dal13 uptime is 7 weeks, 1 day, 10 hours, 6 minutes</pre>			show vrf Show VRF	all		
agg1.newy2 New York, NY	Cíšco 8202 2RU System w/ 12x400GE QSFP56-DD & 60x100GE QSFP28				from router	to supplied	
agg1.sanj San Jose, CA	agg1.dall3 ≻ show bfd		0PY	destination uncheck			
agg1.seat Seattle, WA	IPV4 Sessions Up: 2, Down: 0, Unknown/Retry: 0, Total: 2			Hncheck a	II nodes in th His	tory	Edit
agg2.ashb					w version 1.dall3	core1.d	ienv
Ashburn, VA					w bfd 1.dall3	core1.d	
Dallas, TX				895	1.08113	core1.d	env

NDIANA UNIVERSITY

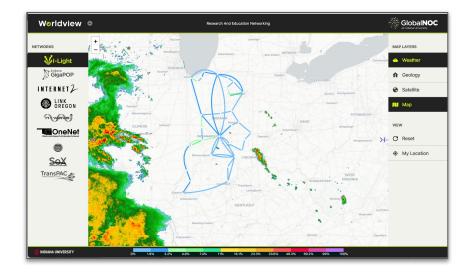
INTERNET





Where we want to be

- 1. Neworks will employ data-driven design and operations
- 2. Diverse measurements will be combined with metadata to provide a more detailed understanding of the network
- Composite / End to End views of the R&E infrastructure will give stakeholders appropriate awareness







Where we want to be

- 1. Neworks will employ data-driven design and operations
- 2. Diverse Measurements will be combined with metadata to provide a more detailed understanding of the network
- 3. Composite / end to end views of the R&E infrastructure will give stakeholders appropriate awareness
 - Recent examples include:
 - International Networks @ Indiana University, NetSAGE, TACC and EPOC
 - GlobalNOC and Global Research Map
 - PerfSONAR

e neusoe		width Dashboard 🕐 La	ist7days × Q C × Q
		n Dashboard	
The map shows	the minimum, maximum, and average bandwidth n	utilization of the circuits and exchange points o eriod.	ver the selected time
	rows below the map show each of the links in more a combined view of the average and maximum ba	e detail, including traffic rate and total volume tr	
Single Link Max A-Z	Single Link Max Z-A	Average Across All Links	Total Transferred
67.5 Gb/s	98.1 Gb/s	6.08 Gb/s	7.36 рв
, + ^{3,0} ∼			(Press
-			
9			
- Hor			
NA / Sec	NORTH		
24	AMERICA		ASIA
200	1	SAN TA	
	f lin		
- /	Pacific	AFRICA	
1			Indian
1 / /	SOUT		
OCEANIA	SOUT		
OCEANIA	Exploring R&I	E Community Connectivity	Ocean
OCEANIA	Exploring R&I	CA TRACT	Ocean
Resources	Exploring R&I	E Community Connectivity	Ocean
	Exploring R&I Find paths between network, compute, sto Path Big find II	E Community Connectivity	Ocean or
	Exploring R&I Port path between network, compute and Path Bg Rei II Network Reinstein Australia	E Community Connectivity	Ocean or
Resources	Exploring R&I Path Big Re II Provide and other methods.	E Community Connectivity	Ocean Oc
Resources Not Frat, saled ratios's tape Lawy p •	Exploring R&I Port path between network, compute and Path Bg Rei II Network Reinstein Australia	E Community Connectivity	Ocean Oc
Resources Net First, salect network layer	Exploring Real Part Balance Andrews Path Manage and the memory are Manage and the memory are Manage and are memory	E Community Connectivity are and other key facilities used in the persuit of acter	Ocean Oc
Resources Net Inst. select halsoft layer Layer p	Exploring R&B Path Bare II Manual Strands returns to the Manual Strands Strandson Manual St	A E Community Connectivity man and other key holitone used in the porsult of ader Manual M	Control of
Resources Net Univ p Source So	<section-header></section-header>	A Subject of the local in the period of and the second of the local in the period of and the second of the second	ocean ate abovey. Graph
Resources Net Ver Prot.steed network type Ver Source Source Ver Compute Ver Compute Ver Ver Compute Ver Ver Ver Ver Ver Ver Ver Ver Ver Ve	Exploration Regeleration Parte Parte Management M	A 1995 E Community Connectivity Research and a second se	Ocean oc
Resources Net Unit stated relatively lage Unit stated relatively lage Unit stated relatively lage Unit state and units Note Source Negen Federa Compute Bagen Bage	Battering	A Surger Community Connectivity A surger A	Anter a series of the series o
Resources Net Text.steed network taget Laver p Source Source Text Text Text Compute Text Text Text Text Text Text Text Te	Experimental Control of Control o	A Community Connectivity apared dive by ballities used in the possil of allow A suggestion of the second of the second of allow A suggestion of the second o	Ocean and documents Graph ASSET ASSET ASSET ASSET ASSET ASSET
Resources Net This set interview Laver p Source Source Gonpda Tepin Hom Anerce Resource Big Red I Destination	<section-header></section-header>	A Summer of the second of the	AGRET
Resources Net Use used releases Source Source Source Use a serve.	<section-header><section-header><section-header><section-header><section-header><section-header></section-header></section-header></section-header></section-header></section-header></section-header>	A Community Connectivity Rear of the hey failthen used in the persuit of team A contract of the hey failthen used in the persuit of team A contract of the hey failthen used in the persuit of team A contract of the hey failthen used in the persuit of team A contract of the hey failthen used in the persuit of team A contract of the hey failthen used in the persuit of team A contract of the hey failthen used in the persuit of the hey failthen used in the persuit of the hey failthen used in the hey failthen used	AGRET
Resources Net User User Source Source Source Source Basere	<section-header><section-header><section-header><section-header><section-header><section-header><section-header><section-header><section-header><section-header></section-header></section-header></section-header></section-header></section-header></section-header></section-header></section-header></section-header></section-header>	A Summer of the second of the	AGRET
Resources Net User Source Source Source Source Source Basers Base	<section-header></section-header>	A Superior State S	ASSES
Resources Net User User Source Source Source Source Basere	<section-header></section-header>	A Superior State S	A A A A A A A A A A A A A A A A A A A

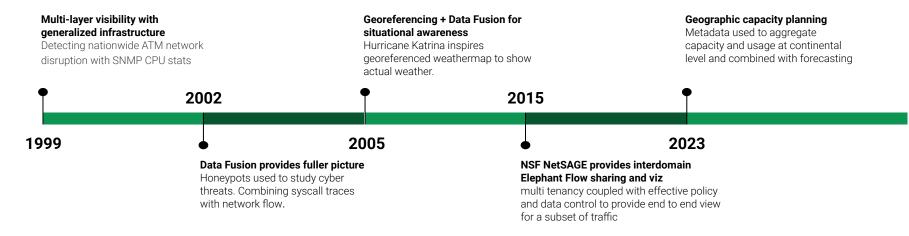
NDIANA UNIVERSITY

INTERNET





Is this a solved problem?



Most of the key ingredients are available however...

- Scaling still difficult at intranetwork level
- Metadata tends to be variable
- Interdomain sharing is limited
- Required domain knowledge a barrier





Imagine the following scenario

- An Engineer gets a report that a customer is seeing poor performance moving data as part of a scientific pipeline
 - o checks their measurement systems and local PerfSONAR results and find all clear
 - Presumes issue is likely at far end of the path close to last mile

• To support the customer effectively, the Engineer needs access to more data





• Best Case:

- External networks have established measurement collections
- federated auth and enough structure and documentation exit to support self service

• More often:

- External network has established measurement collections
- Low fidelity data is externally accessible
- The data needed is private, the engineer however knows his peers and can make a wetware request

• Worst case:

- External network might itself be decentralized with no ubiquitous measurement approach
- Multiple human interactions required to find the right engineer
- The external engineer has to log into the router or other device to debug with you



Where does the engineer do next?

• Best Case:

- External networks have established measurement collections
- federated auth and enough structure and documentation exit to support self service

• More often:

- External network has established measurement collections
- Low fidelity data is externally accessible
- The data needed is private, the engineer however knows his peers and can make a wetware request and has to manually correlate information shared

• Worst case:

- External network might itself be decentralized with no ubiquitous measurement approach
- Multiple human interactions required to find the right engineer
- The external engineer has to log into the router or other device to debug with you



Where does the engineer do next?

• Best Case:

- External networks have established measurement collections
- federated auth and enough structure and documentation exit to support self service

• More often:

- External network has established measurement collections
- Low fidelity data is externally accessible
- The data needed is private, the engineer however knows his peers and can make a wetware request

• Worst case:

- External network might itself be decentralized with no ubiquitous measurement approach
- Multiple human interactions required to find the right engineer
- The external engineer has to log into the router or other device to debug with you





MetrANOVA was created to help

• Advocate for quality ubiquitous collections with appropriate access within all of R&E

- Provide training and policy guidance
- Create knowledge base articles and howtos

• Lower the barriers through technical and policy collaboration

- Reduce need for bespoke solutions
- Amortize software sustainment costs through collaboration.

• Retain Network Measurement as a core competency through the next generation

- $\circ \quad \ \ {\rm Requires \ ongoing \ care \ and \ feeding}$
- Deep domain knowledge in networking, systems, and to an extent stats
- Support next generation of R&E engineers



The Secret Sauce of Research and Education



• Timeless design constructs

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

• Technology != differentiator

- Same software and hardware used in R&E and in Commodity Internet
- \circ $\hfill It's how you use it, not what you use$
- like an artist and a paintbrush

• Combined with community focus

- We are a not for profit community
- Our values differentiate us
- Its how we apply these technologies to address needs and facilitate scientific and educational endeavors.

- Additional Considerations
 - Collaboration and trust are key
 - Ubiquitous access tempered with appropriate access control
 - We need: Design Patterns, Service Definitions, and Policy Guidance









Consortium Details

Goals

- Tools, Tactics and Techniques
- Develop and Share
 - Open Architectures
 - Technical Components
 - Design Patterns
 - Best Practices
 - Policy Recommendations.

Executive Committee

- Provides governance and oversight.
- Decides on new membership organizations.
- Representatives from each member.
 o Inder Monga ESNet
 - Inder Monga
 Ivana Golub
- PSNC/GÉANT
- James Deaton
- Luke Fowler
- Nathaniel Mendoza
- Ed Balas

- Internet2Indiana University GlobalNOC
- TACC
- Consortium Lead

Vision

- <u>https://github.com/MetrANOVA/.github/blob/main/profile/</u><u>vision.md</u>
- 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

Participation Model

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









What have we been up to this year?

- Internal member survey
- community survey launched
- Established near term roadmap
 - Vetted technical stack
 - Policy guidance for appropriate data sharing

• Technical work in progress

- Elasticsearch Time Series Data Stream evaluation
- SNMP vs Streaming assessment
- Data Store evaluations
 - Elasticsearch, Opensearch
 - Clickhouse, mongo, timescaledb
- Science Registry refactor / up keep (TACC)





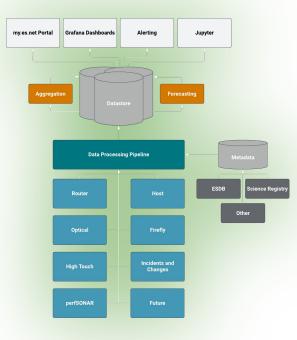
Value Engineering of Elasticsearch Datastore

Broad Elastic Adoption:

- Elasticsearch and OpenSearch used within most member networks and with in PerfSONAR
- Flow, SNMP, Optical, Open Telemetry, Streaming

A few members motivated to explore improved scaling

- scale of > 50 nodes
- new features since adoptions to improve costs / scaling



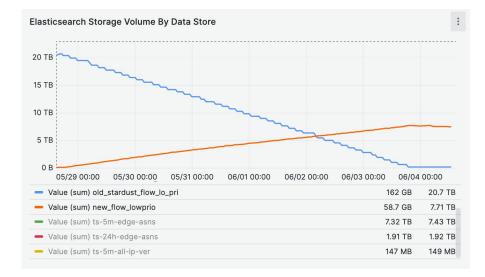
INTERNET



0

Elastic Time Series Data Stream (TSDS)

- <u>https://www.elastic.co/guide/en/elasticsearch/refere</u> <u>nce/current/tsds.html</u>
- As of version 8.9
- Reported savings of up to 70%
 - <u>https://medium.com/squareshift/up-to-70-metr</u> <u>ics-storage-savings-with-tsds-enabled-integrati</u> <u>ons-in-elastic-observability-4cf8b6217c1</u>
- We are evaluating this in particular for both Flow and Port metrics with encouraging results
- ESnet has deployed this at scale a few weeks ago
- We have observed ~63% data reduction for single packet flows

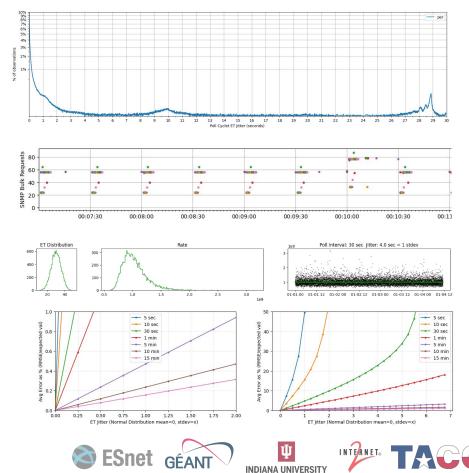






Impact of SNMP timing variance on measurement quality

- Trying to create map of max instantaneous usage we noticed links with impossibly high values
- Exhaustive investigation found:
 - Telegraph SNMP Poller using same timestamp for all values in a getBulk sequence (based on goSNMP)
 - The variance in Elapsed Time is higher that anticipated: O(4) sec
 - We estimate 30 second rate calculations have a 22% average error rate
- Implications:
 - With our ET variance it does not make a great deal of sense to poll at 30 seconds
 - To get < 2% we would need 5 min polling
 - Moving to streaming telemetry we hope will provide qualitative improvements
 - There are ways to through rewriting pollers to improve SNMP
- Data and Scripts
 - Updates shared soon



Data Sharing, Federation, Anonymization, Policy

- Appropriate controls that respect each domains policies and constraints are a must for data sharing
- Having well defined policies is a precursor which today does not always exist
- We will be developing recommendations for the creation of functional policy
 - With technical means to enforce
- Example of constraints you are facing:
 - GPDR, FERPA, HIPPA.
 - NDAs and customers who which to remain low profile
 - Institutional policies, funding bodies, etc.



Supporting multiple levels 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 let consumers and providers helps set expectations
 - What is and what is not shared and at what level of access
- An example:
 - International Networks at IU is an example
 - Supporter of IU's contribution to the consortium.
 - Vital for them to be able to collect/accept data from multiple sources
 - Display in a single pane of glass.
 - Policy constraints what is collected and what is shared







Laws, Rules, Policies, Guidelines, Best Practices

- MetrANOVA is not a substitute for legal advice.
- Just as we do with the technical elements, we aim to give people a set of documents/processes that they can use, or at least start with.
- Covering issues like...
 - How the data is collected and transported to the storage infrastructure?
 - How is it stored, and where? Is it encrypted at rest?
 - Anonymization how, at what stage in the process?
 - Storage, collection, display?
 - Retention how long will the data be kept?
 - Sharing who is the data shared with, on what terms? For what purposes?





- Technical, development, and engineering work (sometimes) have relatively clear-cut ways to define what the "right" solution is
 Performance metrics, etc.
- There's not a direct equivalent there's no "best" set of policies
 We're not trying to set a standard, more give people a starting point
- What we can do is document a set of policies that can work.
 Based on the actual knowledge and experience of participants
- Reduce the amount of effort that people need to put into getting started.





- What does the policy side want to produce?
- A set of documents that define policies re. Collection, transport, etc. etc.
- Align with a MetrANOVA-style architecture
 - But doesn't assume it
- Use as-is or adapt to specific legal/technical/institutional environment





Edward Balas

MetrANOVA Consortium Lead

ebalas@es.net

For more information:

- Github: <u>https://github.com/MetrANOVA</u>
- Web: <u>http://www.metranova.org/</u>





Questions?









"Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Networking and Information Technology Research and Development Program."

The Networking and Information Technology Research and Development (NITRD) Program

Mailing Address: NCO/NITRD, 2415 Eisenhower Avenue, Alexandria, VA 22314

Physical Address: 490 L'Enfant Plaza SW, Suite 8001, Washington, DC 20024, USA Tel: 202-459-9674, Fax: 202-459-9673, Email: <u>nco@nitrd.gov</u>, Website: <u>https://www.nitrd.gov</u>

