

NetSage - A tool to understand data transfers



netSAGE

Doug Southworth

International Networks @ Indiana University
(IN@IU)

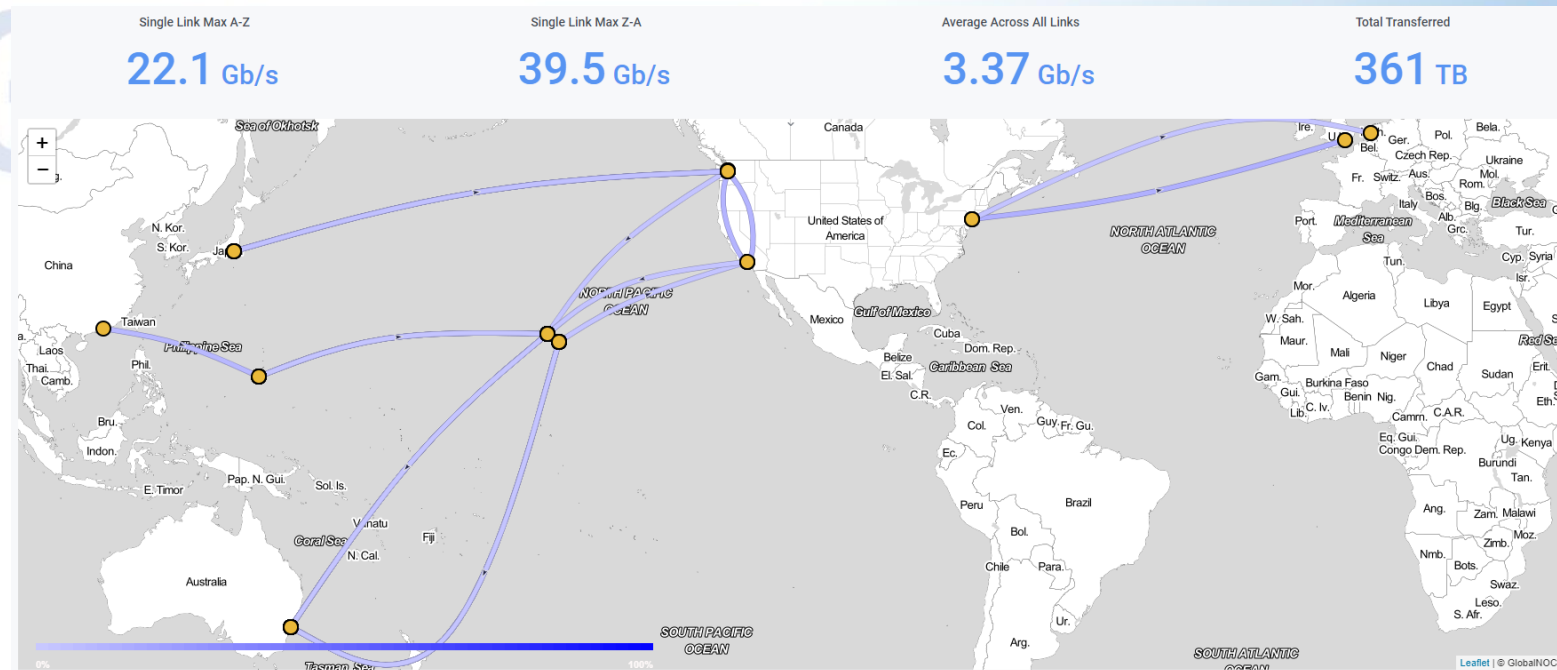
Monitoring using NetSage

- NetSage advanced measurement services for R&E data traffic
 - Better understanding of current traffic patterns across instrumented circuits
 - Better understanding of large flow sources/sinks
 - Performance information for data transfers
- Collaboration between Indiana University, LBNL, and University Hawaii Mānoa
- Originally funded by the NSF international program, software is now being deployed domestically as well

NetSage International:

<http://portal.netsage.global>

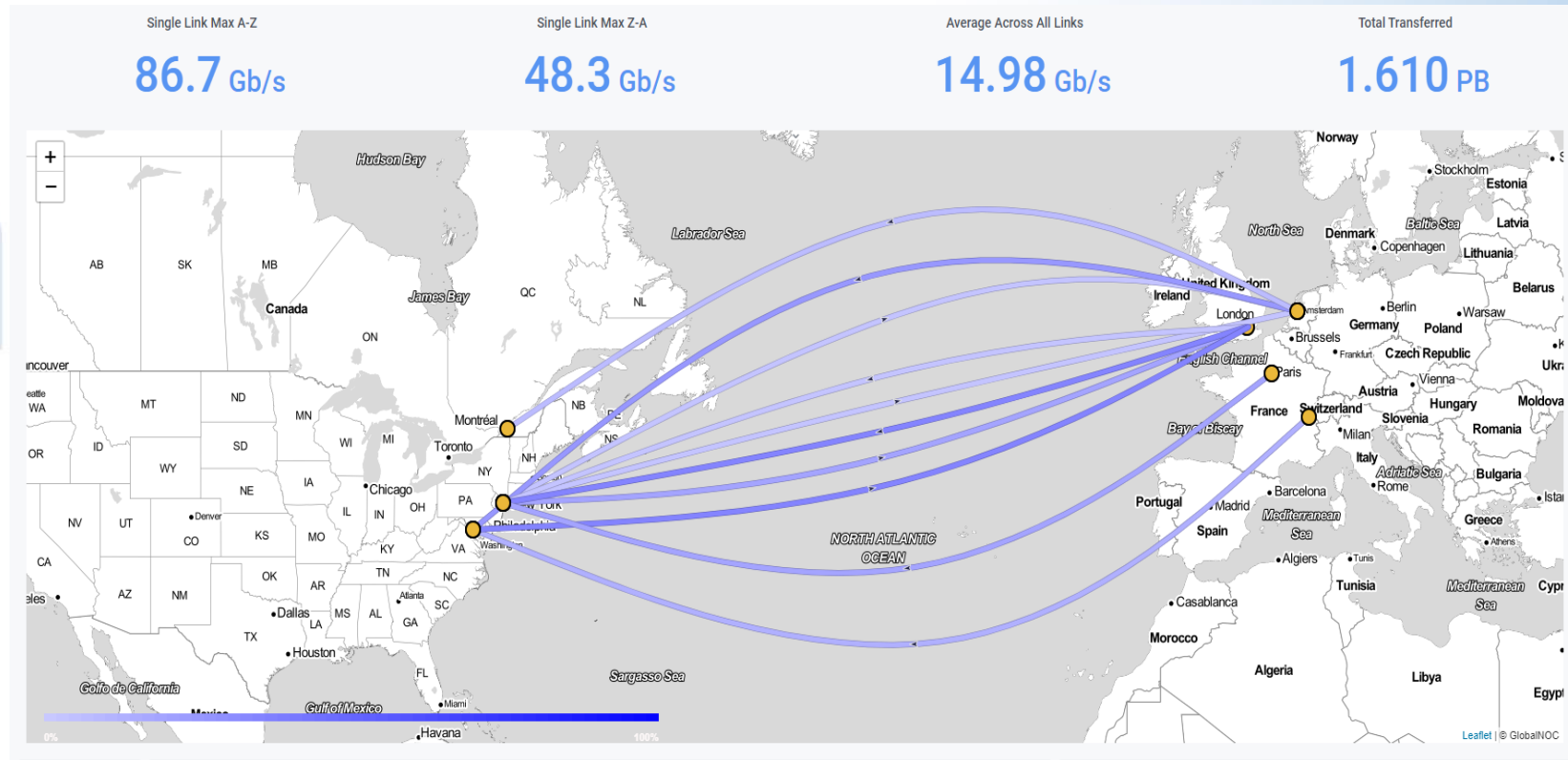
- Treating NSF International Networks like a facility
- Collection points in
 - New York
 - Seattle
 - Honolulu
 - Guam



NetSage ANA:

<http://ana.netsage.global>

- Current and historical SNMP data
- Bandwidth patterns across all ANA links



What NetSage Does Best

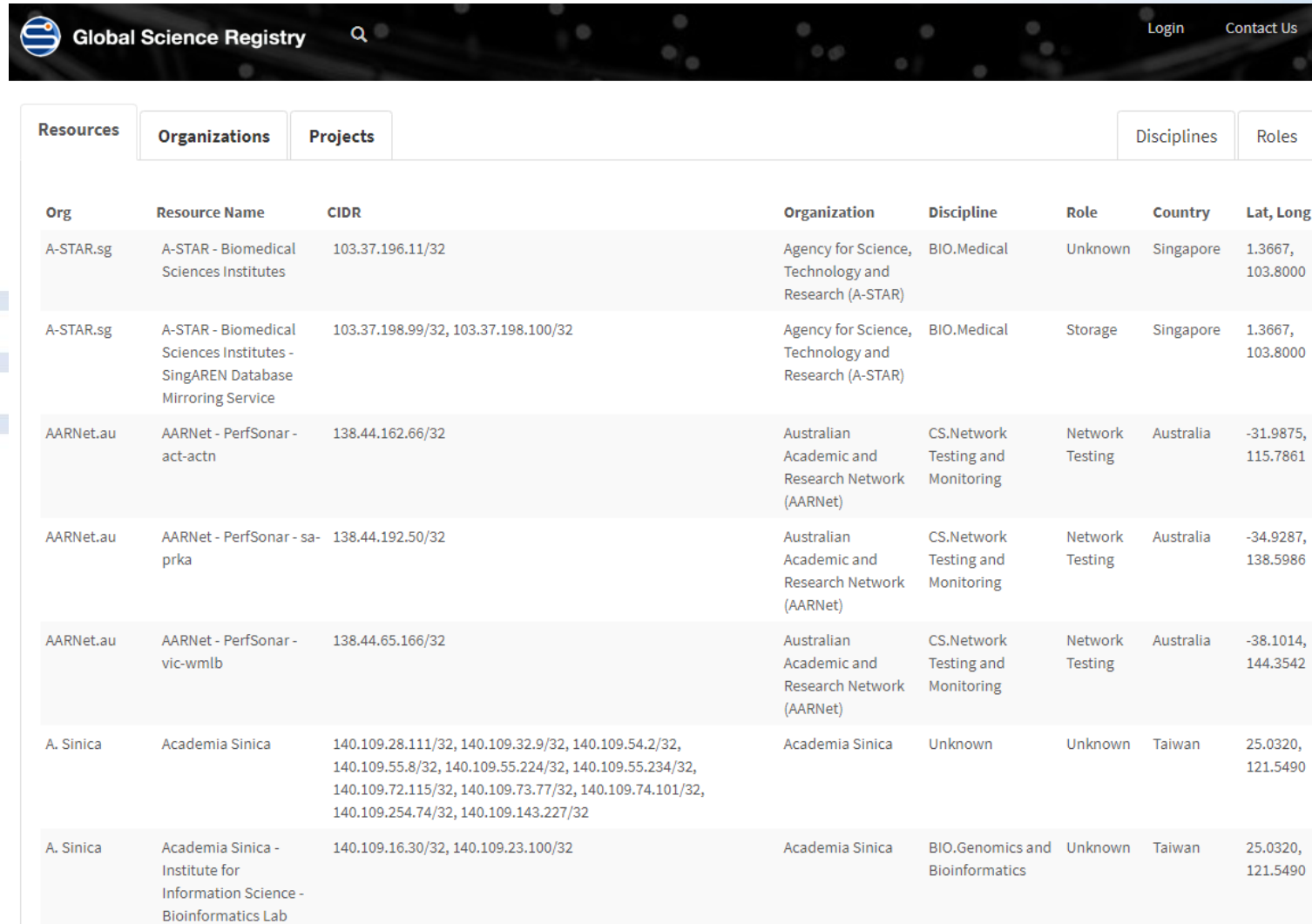
- Answers questions asked by network engineers and network owners
- Human-readable summaries and patterns
- Gives people the higher-level pattern so they can narrow down a time frame and then use local tools that have more detail
- Simplifies and makes accessible basic data

NetSage Data Sources

- SNMP data (Passive) - Basic bandwidth data
- perfSONAR (Active)
 - Active tests between sites
- Flow data from routers (Passive)
 - Only de-identified data collected by NetSage
- Tstat-based traffic analysis for archives (Passive)
 - TCP flow statistics: congestion window size, number of packets retransmitted, etc
 - Also de-identified before stored

NetSage Science Registry

- “DNS” for R&E resources
- Voluntary information sharing
- Allows for more positive flow identification while maintaining de-identification within NetSage



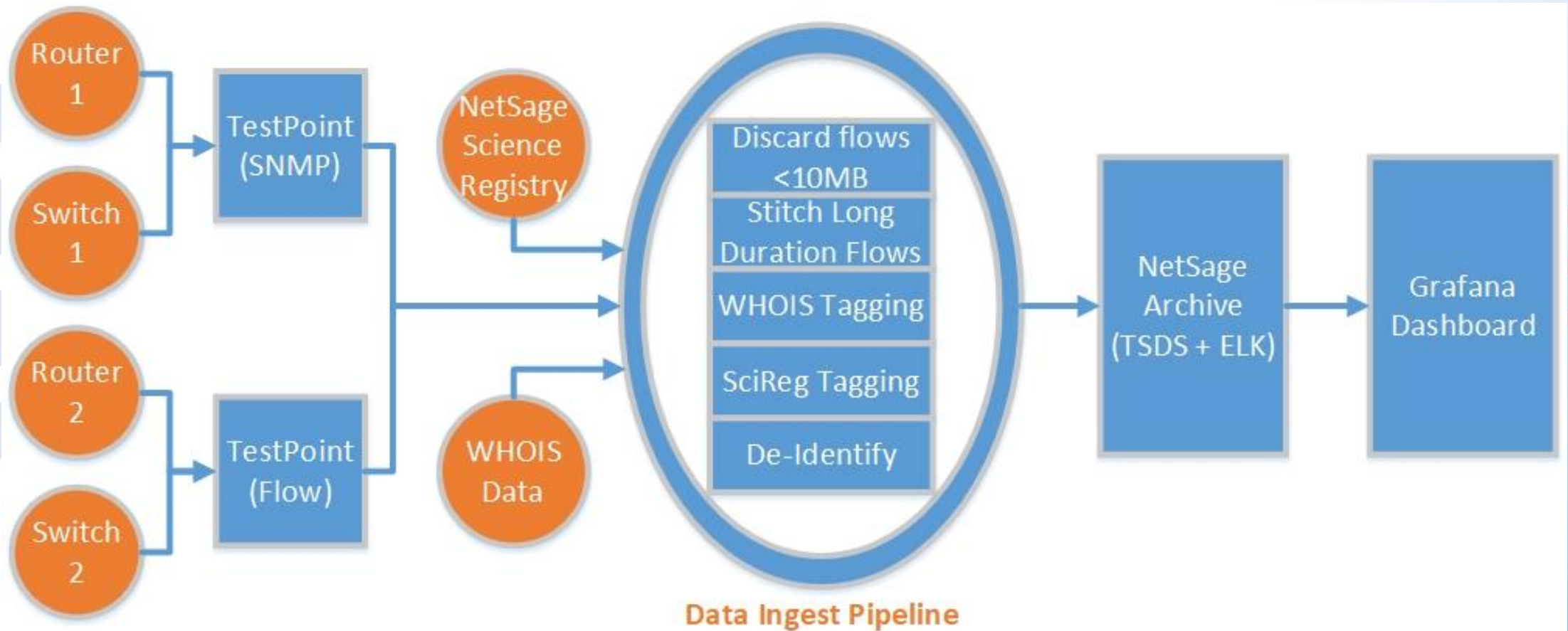
The screenshot shows the Global Science Registry website interface. At the top, there is a navigation bar with the logo, the text "Global Science Registry", a search icon, and links for "Login" and "Contact Us". Below the navigation bar, there are tabs for "Resources", "Organizations", and "Projects". The "Resources" tab is active, displaying a table with columns for "Org", "Resource Name", "CIDR", "Organization", "Discipline", "Role", "Country", and "Lat, Long". The table contains several rows of data, including entries for A-STAR.sg, AARNet.au, and A. Sinica.

Org	Resource Name	CIDR	Organization	Discipline	Role	Country	Lat, Long
A-STAR.sg	A-STAR - Biomedical Sciences Institutes	103.37.196.11/32	Agency for Science, Technology and Research (A-STAR)	BIO.Medical	Unknown	Singapore	1.3667, 103.8000
A-STAR.sg	A-STAR - Biomedical Sciences Institutes - SingAREN Database Mirroring Service	103.37.198.99/32, 103.37.198.100/32	Agency for Science, Technology and Research (A-STAR)	BIO.Medical	Storage	Singapore	1.3667, 103.8000
AARNet.au	AARNet - PerfSonar - act-actn	138.44.162.66/32	Australian Academic and Research Network (AARNet)	CS.Network Testing and Monitoring	Network Testing	Australia	-31.9875, 115.7861
AARNet.au	AARNet - PerfSonar - sa-prka	138.44.192.50/32	Australian Academic and Research Network (AARNet)	CS.Network Testing and Monitoring	Network Testing	Australia	-34.9287, 138.5986
AARNet.au	AARNet - PerfSonar - vic-wmlb	138.44.65.166/32	Australian Academic and Research Network (AARNet)	CS.Network Testing and Monitoring	Network Testing	Australia	-38.1014, 144.3542
A. Sinica	Academia Sinica	140.109.28.111/32, 140.109.32.9/32, 140.109.54.2/32, 140.109.55.8/32, 140.109.55.224/32, 140.109.55.234/32, 140.109.72.115/32, 140.109.73.77/32, 140.109.74.101/32, 140.109.254.74/32, 140.109.143.227/32	Academia Sinica	Unknown	Unknown	Taiwan	25.0320, 121.5490
A. Sinica	Academia Sinica - Institute for Information Science - Bioinformatics Lab	140.109.16.30/32, 140.109.23.100/32	Academia Sinica	BIO.Genomics and Bioinformatics	Unknown	Taiwan	25.0320, 121.5490

NetSage Privacy

- NetSage is committed to privacy, and preemptively addressing any security or data sharing concerns
 - No PII collected
 - Remove the last octet from IP address
 - Only keep data on flows over 10M for circuits
 - 1M for archives
- Data Privacy Policy
 - <http://www.netsage.global/home/netsage-privacy-policy>
- Data Flow Data Retention (De-Identification) Policy
 - <https://tinyurl.com/netsage-deid>

NetSage Ingest Pipeline



Take aways

- Circuits help support science and research
- “If you can’t measure, you can’t improve”
- Additional capacity offers opportunities
- Questions or comments?
 - Doug Southworth, dojosout@iu.edu