

NetSage

Award #1540933

Year 1 Annual Report

1 May 2015 through 31-January-2016

PIs: Jennifer Schopf (IU), Sean Peisert (UCD), Jason Leigh (UHM)

Summary

The goal of the NetSage project is to collect data from the NSF International Research Network Connections (IRNC)-funded backbones and exchange points to better understand the use of the circuits. In addition, this collected data is also made available for use by the NOC to better understand day-to-day operations. In the first year of the project, the main accomplishments were setting up the shared archive with the NOC, the beginning of data collection from the backbones, and a prototype of the visualization system.

1. NetSage Overview

NetSage is focused on building and deploying advanced measurement services that will benefit science and engineering communities, focusing on:

- Better understanding of current traffic patterns across IRNC links, and the ability to better understand growth trends for capacity-planning purposes;
- Better understanding of the main sources and sinks of large, elephant flows in order to know where to focus attention on outreach and training; and
- Better understanding of where packet loss is occurring, whether or not the loss is caused by congestion or other issues, and the impact of loss on end-to-end performance.

When fully operational, NetSage services will provide a combination of passive measurements (including SNMP data, flow data, and traffic-header analysis), and active measurements (mainly perfSONAR) with longitudinal network performance data visualization. Year 1 of the project focused on initial deployments of measurement end points on the IRNC-funded backbones, collection of data into a common archive, and preliminary visualization tools.

This report details the staffing, collaboration, tool development, deployment, and planning for the project.

2. Staffing

The staff currently funded by the project includes:

- Jennifer Schopf, IU, PI - overall project director
- Andrew Lee, IU, senior network engineer - measurement project management, IRNC project coordination
- Ed Balas, IU, system architect - collection and reporting
- Mitch McCracken, IU, software developer - perfSonar, Data processing pipeline
- Dan Doyle, IU, TSDS development and support
- Alice Jackson, IU, administration
- Sean Peisert, UC Davis and LBNL, co-PI - security, privacy, performance experimental design
- Brian Tierney, UC Davis and LBNL, staff scientist - monitoring architecture, performance experimental design, privacy
- Jonathan Ganz, UC Davis, graduate research assistant - configuration of and experimentation with performance monitoring tools
- Jason Leigh, UH Mānoa, co-PI - visualization oversight
- Alan Whinery, UH Mānoa, senior personnel - perfSONAR and coordination with backbones
- Alberto Gonzalez, UH Mānoa, graduate research assistant - visualization developer
- Eric Wu, UH Mānoa, graduate research assistant - visualization developer

In Quarter 2 we identified the need to shift project staffing to include resources at LBNL to better meet deliverables. At that time, we began the process of allocating a third sub-award to LBNL.

3. Collaborations, Travel, and Training

NetSage staff participated in various meetings to support ongoing deployment, collaboration, and training. Over the first year of the project, these included:

- Schopf, Lee, Tierney - Terena Networking Conference, June 15-18, Porto, Portugal, <https://tnc15.terena.org/>. Schopf, Lee, and Tierney attended the Terena Networking Conference. During the week, meetings were held with several other IRNC awardees, including Julio Ibarra and Joe Mambretti.
- Full staff -NetSage Kickoff Meeting, July 17-18, Berkeley, CA. In attendance were Schopf, Lee, Doyle, McCrackin, Peisert, Tierney, Ganz, and Leigh. Attending remotely from UH were Gonzalez and Whinery. Steve Lau, proposed head of the Privacy committee attended part time, as did Johanna Amann from the Bro project and Christopher Tracy, UCD. The project staff walked through the overall plan for the project and detailed what was feasible in Year 1. Staff assignments were agreed upon, and initial management tools were put in place. The next All Hands Meeting was proposed to take place in February 2016 in Hawaii.

- Schopf - AMIS PI meeting, July 21, 2015 Boston, MA. Schopf met with Yan Luo, UMass Lowell, to discuss AMIS-NetSage interactions.
- Ganz - BroCon '15, Aug. 4-6, 2015, MIT, Cambridge, MA, <https://www.bro.org/community/brocon2015.html>. UC Davis PhD student Jonathan Ganz attended tutorials at BroCon to learn about the viability and use of Bro for various aspects of network measurement as related to the NetSage project, including the use of the Bro TCP analyzer for measuring retransmits (dropped packets) and the SumStats module for measuring other key statistics such as top talkers.
- Lee - AMIS team meeting, August 6-7, Boston, MA. Lee attended a meeting in Boston with the AMIS team to introduce the projects to each other, discuss overlap and collaboration opportunities and the initial AMIS deployment at AMPath in Florida. Tierney attended remotely and discussed perfSONAR.
- Peisert - Oct. 2, San Francisco, CA. Met with John Wroclawski from USC ISI to discuss security and privacy issues related to network measurement, including privacy-preserving techniques used and developed at ISI for data used by the DETERlab project.
- Peisert - Stanford, Oct. 12, Palo Alto, CA. Met with Bob Gleichauf from Lab41 to discuss data anonymization, and in particular, the balance of sanitization pre- and post-query.
- Lee - Internet2 TechX, Oct 4-8, Cleveland, OH. Attended the conference to keep abreast of the activities of the wider research and education community, and to meet with the PIs of other IRNC projects to discuss monitoring strategies. Attended the perfSONAR developer planning meeting. Schopf, Tierney, Balas, and Doyle also attended under other funding sources.
- Gonzalez - IEEE Visualization 2015, Oct 26-30, Chicago, IL, <http://ieevis.org> Attended as continuing education and training in data visualization for grad students. Gonzalez also met with Jon Dugan, developer of my.es.net, to understand the framework's relevance to ongoing NetSage visualization work.
- Schopf, Lee - SuperComputing 2015, Nov 15-19, Austin, TX, <http://sc15.supercomputing.org/> Attended IRNC PI meeting, presented preliminary results to NSF PO, gave talk in IU Booth, had extensive conversations in person with other IRNC project PIs and international collaborators. Tierney, Balas, and Leigh also attended under other funding sources.
- Schopf, PTC 2016 and associated side meetings, January 15-21, 2016, Honolulu, Hawaii, <https://www.ptc.org/ptc16.html>. Joined collaborative planning meeting for Guam exchange point at University of Hawaii, as well as attended meetings with vendors and potential collaborators. Additional side meeting with PI Leigh. Trip partially funded by TransPAC and other sources.

4. Project Coordination

Following up from our All Hands Kickoff meeting in Quarter 1, we emphasized working with the IRNC-funded Backbone operators and the NOC in the first year of the project. We established that each backbone (TransPAC, AmLight, PIREN) is currently collecting open archives of SNMP and perfSONAR data.

Joint with the NOC, we hosted all IRNC PI meeting at SuperComputing on November 17, 2015. This meeting enabled us to establish basic guidelines and steps forward with all projects in a more formal way than we have, and especially to lay the groundwork to understanding what measurement data can be collected in Year 1 from the backbones and in Year 2 to collect from the IRNC-funded exchange point operators.

A second NetSage All Hands meeting is being planned for the week of February 1, 2016 in Hawaii. The timing will allow APAN travelers to stop on their way back to the continental US from that meeting. This session will focus on Year 2 planning.

We continue to coordinate with the AMI research projects. Lee attended the kickoff meeting for the AMIS project, established some coordination guidelines, and was added to several mailing lists for the project. For the InSight project, we had some email exchanges with Greg Cole, who in January retired, and the project has been handed over to a new PI, Gregor Jens. We have invited both projects to attend the Third All-Hands meeting in May 2016.

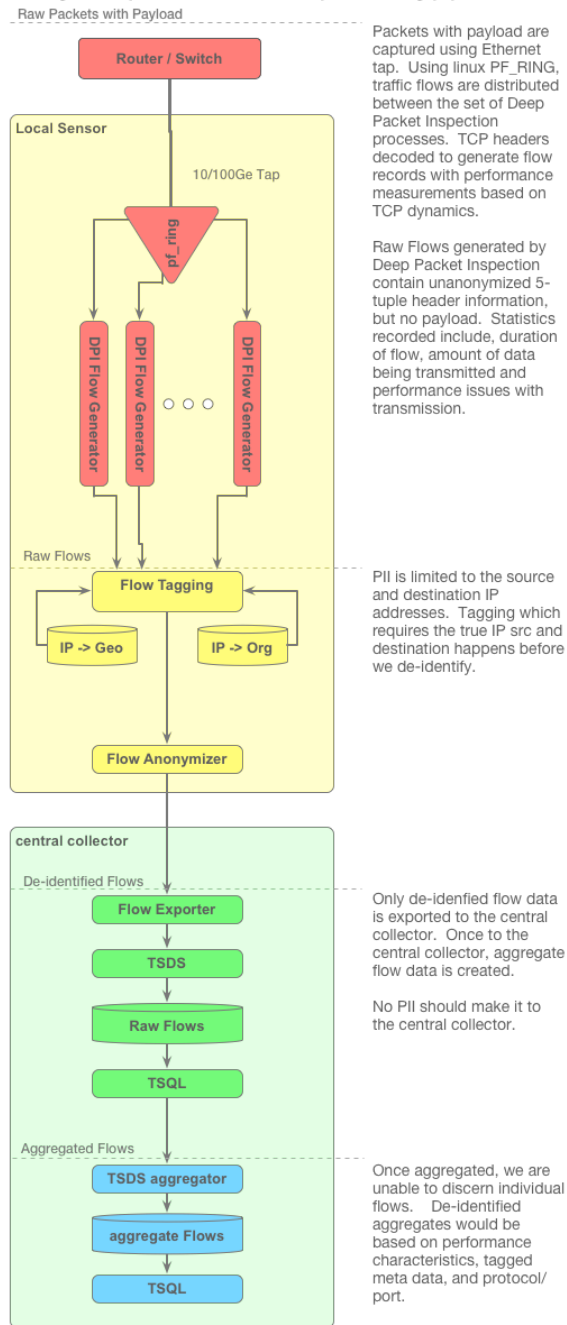
5. Software Development

Over the course of the project, NetSage staff will be involved in the development and deployment of various pieces of software to support active and passive measurements, monitoring, archiving, analysis, and visualization. This sections details those projects.

5.1 System Architecture

Based on feedback from the Berkeley kickoff meeting, the packet processing pipelines were expanded in detail and now cover two expected use cases, depicted in Figures 1 and 2.

NetSage raw packet to flow data processing pipeline



Created by Edward Balas 01/04/2016 11:59 AM

Figure 1: NetSage data architecture for raw packet to flow data processing.

NetSage flow data processing pipeline

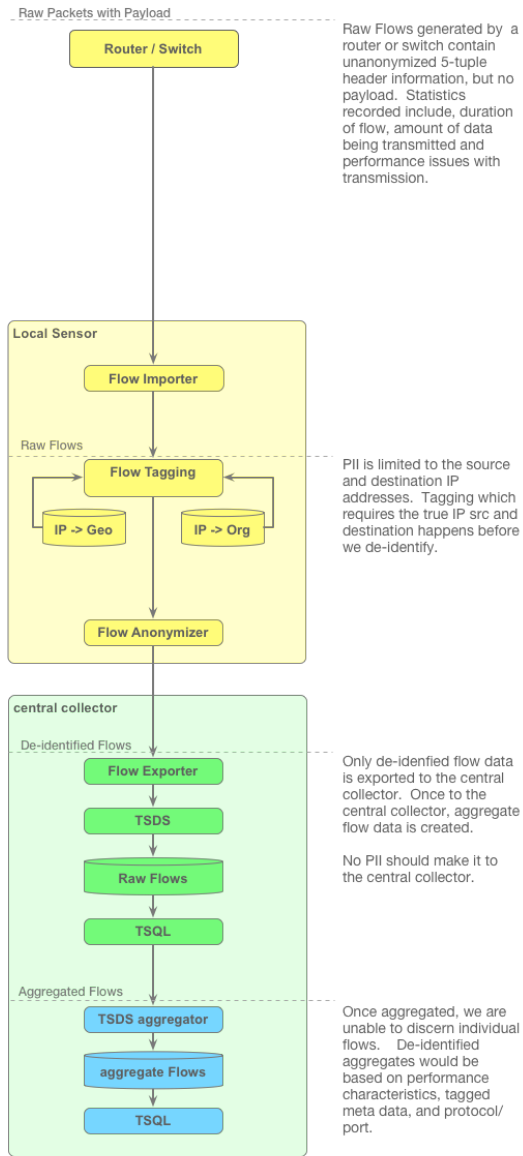


Figure 2: NetSage processing pipeline for flow data.

To meet the needs for storing and analyzing detailed / non-aggregated flow data, preliminary design work was conducted into the requirements and design for a sparse storage mode in TSDS. Review of existing storage of aggregated flow data was also conducted and determined to be adequate for initial use.

Addition design details for the flow processing pipeline were developed, including a proposed canonical internal flow format that can include TCP performance metrics and a convention for storing network topology data in TSDS as metadata.

5.2 Time Series Data System (TSDS)

The Time Series Data System (TSDS)

(<http://globalnoc.iu.edu/software/measurement/tsds.html>) is a software suite that provides well structured and high performance storage and retrieval of time series data, including interface throughput rates, flow data, CPU utilization, and number of peers on a router. Along with the raw data, the TSDS suite is capable of tracking and reporting based on metadata, for example viewing interface throughput from the viewpoint of a VLAN or BGP peer sessions from a particular ASN. In addition to existing SNMP support within TSDS, integrations for perfSONAR loss and latency data were added, and plans made for storing both aggregate and raw network flow data. Prior to the initial open sourcing, there were 2 releases of the TSDS backend and frontend components in the early part of the reporting period.

In October, the first open source release of TSDS took place. The code is hosted on github (<https://github.com/GlobalNOC>) and uses the Apache2 license. Developers from Hawaii and IU identified how to use TSDS to provide data to address the use cases.

In November, TSDS Services version 1.4.2 was released, adding several query language enhancements and performance improvements. TSDS Frontend 1.0.5 was also released, which provides an improved searching and filtering capability to enhance exploratory analysis.

In December, a new data aggregation subsystem was designed and implementation began in order to increase generalized scalability and system resiliency. A release is planned for in early 2016.

5.3 Simple Network Management Protocol (SNMP)

The Simple Network Management Protocol (SNMP) is an application-layer protocol defined in RFC1157 for collecting and organizing information about managed devices on IP networks. SNMP is commonly used by routers and switches to monitor networks for conditions that warrant administrative attention. This data is commonly collected and openly archived by most R&E networks.

In Quarter 3, we completed inclusion of the three backbone SNMP archives into TSDS. This was joint work with the IRNC NOC. The only remaining circuits to include are PIREN circuits, planned for January 2016, and the newly established Seattle-Tokyo TransPac-PacWave 100G circuit, which will be addressed in the first quarter of Year 2.

5.4 perfSONAR

perfSONAR (<http://www.perfsonar.net/>) is a network measurement toolkit designed to provide federated coverage of paths and help to establish end-to-end usage expectations. The NetSage project will be using perfSONAR for its active measurements of bandwidth and throughput. This data will be part of the NetSage

archive using TSDS. We have developed a perfSONAR exporter tool that pulls data from an open perfSONAR MA and inserts it into TSDS.

In Year 2, we will fully deploy an IRNC set of perfSONAR tests across the identified sites. We will coordinate with the IRNC NOC to resolve any performance issues. In Year 1 we examined several approaches to getting unified perfSONAR data across the backbones and exchange points, with the best solution resulting in a unified upgrade to perfSONAR 3.5 and the issuing of IRNC-specific tests.

5.5 Bro

Bro (<http://www.bro.org>) is a network capture, measurement, and analysis framework that is useful for a variety of measurement and analysis end uses, including security and performance monitoring. NetSage planned to use Bro to gather more indepth network measurements, including information about packet loss.

In Year 1, project members installed Bro on a network data analysis cluster at LBNL and established that the Bro TCP analyzer had significant difficulties in keeping up with multi-Gb/sec flows and was particularly problematic for asymmetric traffic flows, as reported in Quarter 2. Ongoing testing indicated that this software package was unsuitable for the needs of this project, so alternatives were explored.

5.6 Tstat

Tstat (<http://tstat.polito.it/>) is part of the EUMeasurement Plane (mplane) FP7 project developed by Munafó and Mellia at Politecnico di Torino. Tstat can be used to analyze either real-time or captured packet traces, and rebuilds each TCP connection by looking at the TCP header in the forward and reverse direction. Tstat reports a number of useful TCP statistics, including congestion window size and number of packets retransmitted, which can be used to analyze the health and performance of the link.

We are able analyze 10G TCP flows using less than 50% of a single core. Given that tstat also includes the ability to de-identify IP using Crypto-Pan prefix-preserving IP anonymizer, this makes tstat a promising tool for us going forward.

Our current plan is to use the tstat for both flow analysis and for TCP retransmission analysis. We set up a test site at the Los Angeles TransPAC node, and will be incorporating tstat data into TSDS in year 2, as well as deploying tstat more widely.

5.7 SFLOW data

SFlow is an industry standard for packet export at Layer 2. It provides a means for exporting truncated packets, together with interface counters, and can be used to better understand the flows over a circuit.

While it is recognized that this data is highly useful, there can be privacy concerns around its collection and use. In Year 1, we worked closely with the TransPAC team,

which was already collecting anonymized sflow data on their LA link, to understand the issues surrounding this data source and to put in place our anonymization pipeline. In Year 2, we will investigate collecting this data for the other backbones.

5.8 Data Collection from DTNs

In Project Year 2, the NetSage team plans to contact the operators of some DTNs used for large data transfers over IRNC links to see if they are willing to install tstat on their DTNs. This would include installing a tool that parses the tstat results and sends it to the NetSage central collector. Our current plan is to start with NCAR and NERSC, and then investigate additional data sources being used by the large flows.

5.9 Visualization Tools for NetSage

In Year 1, the visualization team defined several NetSage use case queries to determine what would be feasible in the short term and in the long term. We used this list to define needed data sources and privacy concerns, as well as the visualization approach. A list can be found at :

<https://docs.google.com/spreadsheets/d/1cnjrYGoUqWXzm-MNz1ly7Bu7yRpCHJU71jSl2OeltS0/edit#gid=1801119320&vpid=A1>

The visualization team designed a first prototype sketch of NetSage visualization approach based on initial use case queries. A simplified version is shown in Figure 3, and the full version can be found here:

<https://drive.google.com/open?id=0B7Dp4nQ3CpApbzAxVy1UOHgtamc>

A second prototype, based on feedback from the other members of the NetSage project, was developed and presented at SC'15 in November for additional feedback. Initial visualization codes were developed using offline sample TSDS data in anticipation of the availability of the TSDS server with live data.

In Year 2 we will continue the development of the prototype, plugging into live TSDS data. We will work with the IRNC PI's for feedback and identifying needed additional functionality.

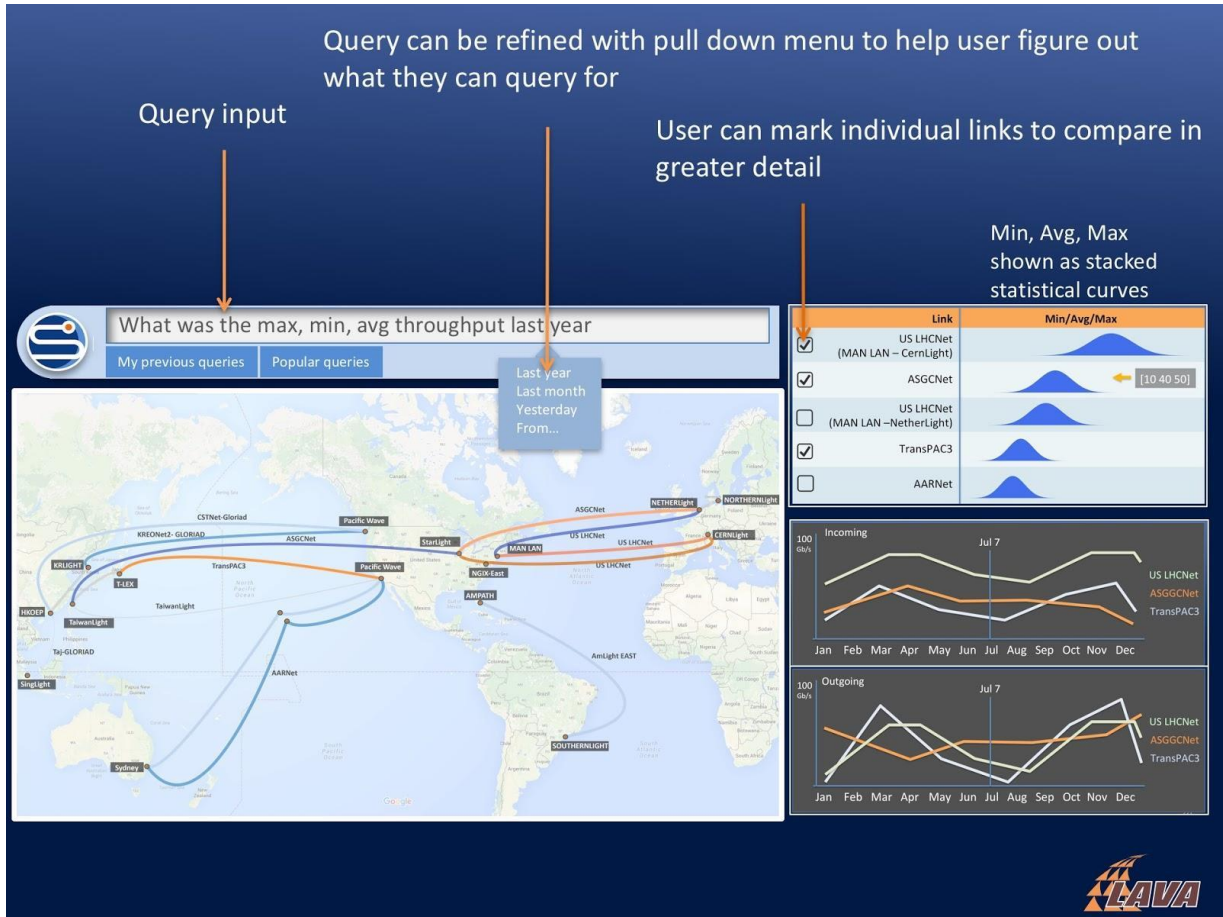


Figure 3: Sample visualization prototype.

6. Deployment, Data Collection, and Analysis

In Year 1, the project focused on the three IRNC backbone projects, and the collection of SNMP and perfSONAR from them. Quarter 2 saw the base build for the NetSage-NOC archiver system being put in place, and by the end of the project year we had SNMP data from all three backbones and were working toward inclusion of perfSonar data sets.

Most of the analysis will need to wait until data is available, however, we were able to move forward with development of scripts to analyze Top Talkers in a time frame, using data collected from the TransPAC 10G LA-Tokyo link. This work is jointly funded with TransPAC.

The work on capacity planning techniques has pushed back to Year 2 in part due to the UC Davis/LBNL staffing issues mentioned in Section 2.

7. Data Privacy and Security

Basic security measures are being maintained, and there were no security incidents to report for this project year.

The Data Privacy and Security committee did not move forward as hoped in the first year. We had originally identified Steve Lau, from the University of California CIO's office, but he was unable to follow through on this. We are working to establish an alternative group lead and next steps forward. We have contacted members of the CACR and CAIDA for assistance, but it has proven difficult to find another candidate to lead this group. We are considering an alternative model where we work with a more loosely organized set of experts to verify documents that we work on internally.

8. Summary of Financial Reporting

The Table 2 shows the month-by-month expenditures across the project through the end of Project Year 1. The project is underspent for by approximately \$335,000. Over the first 9 months, we should have spent approximately \$750K. This underspend is in part due to a slower than planned startup, especially with respect to the subawards, which will be addressed in the following years. We had an approximately \$220K underspend in salaries, in part due to the slow subaward startup and in part due to us postponing the hiring of an analyst at IU, since that work shifted to Year 2. The travel budget saw a savings of \$15K, in part due to the smaller staff. The other significant savings was in the equipment budget. We had originally thought that we would need to purchase significant equipment to deploy at the backbone sites in Year 1, but this has not been the case. This resulted in a \$75K approximate savings in Year 1. We are also underspent on "other" expenses (publications, consultant services, etc). The underspend at IU enabled us to put in place a third subaward to LBNL to use staff expertise there to meet deliverables.

In Year 2, we plan to be fully staffed across all four project teams. Some of the savings in Year 1 will be used in Year 2 to deploy testpoints on DTNs, as described in Section 5.8. We will be hiring on 3-5 summer students at IU to help with analysis tasks, and seeking to hire a full time analyst as part of the team in Year 2 as well.

9. Financial Reporting Details 1-May-2015 through 31-Jan-2016

	Univ	May 15	Jun 15	Jul 15	Aug 15	Sep 15	Oct 15	Nov 15	Dec 15	Jan 16	TOTAL
STAFF COSTS (INCLUDING BENEFITS, F&A)											
Schopf, Jennifer-PI, 8%	IU	1,030	1,030	1,672	1,672	1,672	1,672	1,672	1,672	1,672	14,967
Jackson, Alice- Admin Support, 10%	IU	1,726	1,726	1,754	1,754	1,754	1,754	1,754	1,754	1,754	15,736
Lee, Andrew - Senior Architect, 25%	IU	3,759	3,759	3,147	3,147	3,147	3,147	3,147	3,147	4,428	30,794
Balas, Ed - Sen or SW Developer, 10%	IU	1,292	1,292	1,292	1,292	1,292	1,292	1,312	1,312	1,312	11,690
McCracken, Mitch - Software Developer, 100%	IU	0	0	35,610	0	0	35,281			35,610	106,561
Sean Peisert-UC Davis Leac, 5-15%	UCD			1,891	1,891	1,891	1,891	1,891	1,891	1,891	13,238
Brian Tierney-Staff Scientist, 6-15%	UCD			3,041	3,041	3,041	3,041	3,041	3,041	3,041	21,289
Jonathan Ginz - GSR, 45% AY, 100% Surur	UCD			4,532	4,532	4,532	4,532	4,532	4,532	4,532	31,724
Alan Whinary (2 months)	UH			4,559	4,559	4,559	4,559	4,559	4,559	4,559	31,911
Alberto Gonzalez (11 month grad student) [Step 13 salary]	UH				3,831	3,831	3,831	3,831	3,831	3,831	22,985
Eric Wu (11 month grad student) [Step 11 salary]	UH				3,542	3,542	3,542	3,542	3,542	3,542	21,250
TOTAL STAFFING		8,410	8,410	67,624	29,256	29,256	64,537	29,276	29,276	66,200	322,146
Travel and Other Costs (including overhead)											
Schopf - NetSage kickoff meeting, July 15-17, Berkeley	IU			1,137		1,533					2,669
Lee - NetSage kickoff meeting, July 15-17, Berkeley	IU			1,113				887			2,030
McCracken - NetSage kickoff meeting, July 15-17, Berkeley	IU			1,053			418				1,472
Doyle - NetSage kickoff meeting, July 15-17, Berkeley	IU			1,828	1,368						3,196
Lee - AMIS meeting, Aug 5-7, Boston	IU			458				665			1,123
Leigh - NetSage kickoff meeting, July 15-17, Berkeley	UH			2,335							2,335
Canz - BroCon, Aug. 4-8, Cambridge, MA	UCD				3,714						3,714
Peisert - Oct. 2, San Francisco, CA	UCD							107			107
Peisert - Stanford, Oct. 12, Palo Alto, CA	UCD							115			115
Travel - Lee - Internet2 net Cleveland	IU		2,582			990					3,572
Travel - Lee Terena Portugal June 2015	IU		1,148			3,691					4,839
Travel - Schopf - AMIS Boston July 2015	IU					164					164
Travel - Schopf SC Austin, TX Nov 2015	IU								603		603
Travel - Lee SC Nov Austin, TX Nov 2015	IU								396		396
Equip - Lee - Laptop	IU				5,351						5,351
GoLaddy Dornain Registration	IU							649			649
Travel - Schopf Hawaii Netsage Mtg Feb 2016	IU									462	462
Travel - Balas Hawaii Netsage Mtg Feb 2016	IU									2,277	2,274
Travel - Gonzalez - IEEE Visualization Conference - Oct 2016	UH						5,032				5,032
TOTAL TRAVEL		0	3,730	7,954	10,462	6,377	5,678	2,202	1,089	2,736	40,227
PARTICIPANT SUPPORT											
SC Meeting										2,549	2,549
TOTAL PARTICIPANT SUPPORT		0	0	0	0	0	0	0	0	2,549	2,549
EQUIP OVER \$5K											
Testbed Setup						19,940	5,783				
TOTAL EQUIP		0	0	0	0	19,940	5,783	0	0	0	25,723
IU Overhead on Subawards (32% of 75K per year)											
TOTAL SUBAWARD OVERHEAD		24,000									24,000
TOTAL EXPENDITURES		32,410	12,140	65,478	39,719	55,574	75,985	31,478	30,365	71,486	414,645

10. Progress Against WBS

We submit an updated WBS for Year 2 below. Note that this has combined the original WBS with restructuring and the updated Year 1 WBS into a version that should be able to be used (and expanded) more easily in the coming years.

	WBS full sub.	Year 2 updated	Status
Data Collection	1.1-1.5	1	
Prep work for data from backbones		1.1	
Meet with XP Pis to understand current practices		1.1.1	Completed - Meet at Tech Ex (Oct 2015) and SC (Nov 2015)
Meet with Viz folks to understand data needs for viz questions		1.1.2	Ongoing
Establish what data is to be collected		1.1.3	Completed Y1Q1 - SNMP, PerfSONAR for year 1; flow/tstat to follow
Prepwork for data from exchange points	1.2.2/1.2.3	1.2	
Meet with XP Pis to understand current practices		1.2.1	Started Y1Q3 - SC Nov 2015, Ongoing
Meet with Viz folks to understand data needs for viz questions		1.2.2	Ongoing
Establish what data is to be collected		1.2.3	Ongoing
Signoff on CENIC for Data collection		1.2.4	Y2 goal
Signoff on Miami for Data collection		1.2.5	Y2 goal
Signoff on StarLight for Data collection		1.2.6	Y3 goal
Collector setup	1.1.1	1.3	
Purchase equipment (joint with NOC)		1.3.1	Completed Y1Q2
Initial build		1.3.2	Completed Y1Q2
Open source TSDS	1.1.2	1.3.3	Completed Y1Q2
PerfSonar Related Tasks	1.1.3	1.4	
Evaluate/refine PS to TSDS integration - needs collector turned on for current Maddash	1.1.3.2	1.4.1	Ongoing
Define and deploy new perfSONAR test mesh For BackBones	1.1.3.3	1.4.2	Ongoing
Input TransPAC LA data	1.1.3.5	1.4.2.1	Expected Y2Q1
Input Ampath data	1.1.3.6	1.4.2.2	Expected Y2Q1
Input PIREN data	1.1.3.7	1.4.2.3	Expected Y2Q1
Work to make the boxes not all orange		1.4.2.4	Expected Y2Q1
Define and deploy perfSONAR nodes for Exchange points	1.1.3.3	1.4.3	

Deploy new server to Starlight	1.2.2	1.4.3.1	Y2 goal
Deploy new server to CENIC possibly		1.4.3.2	Y2 goal
Set up perfSONAR Measurement Archive at IU for data collection	1.1.3.4	1.4.4	Completed Y1 Q3
SNMP related tasks	1.1.4	1.5	
Evaluate and tune SNMP to TSDS integration	1.1.4.1	1.5.1	Completed Y1 Q3
Input TransPAC LA SNMP Data	1.1.4.2	1.5.2	Completed Y1 Q3
Input Miami SNMP Data	1.1.4.3	1.5.3	Completed Y1 Q3
Input ACE WIX SNMP Data		1.5.4	Completed Y1 Q3
Input Hawaii SNMP Data	1.1.4.4	1.5.5	Expected Y2Q1
Input Transpac 100G link		1.5.6	Expected Y2Q1
Input StarLight SNMP data	1.2.2/1.2.3	1.5.7	Y2 goal
Input exchange point SNMP data-Miami	1.2.3	1.5.8	Y2 goal
Input exchange point SNMP data-CENIC	1.2.3	1.5.9	Y2 goal
General FlowData Initial Handling	1.1.5	1.6	
TSDS Flow: add sparse storage model to TSDS for flow data with per flow metadata tagging		1.6.1	Y2 goal
TSDS Flow: non aggregated,first Proof of concept		1.6.2	Y2 goal
TSDS Flow: storage model stress testing and refinement		1.6.3	Y2 goal
TSDS Flow: design refinement and scaling		1.6.4	Y2 goal
Pipeline: design modular de-identification rabbitmq based processing pipeline		1.6.5	Y2 goal
Pipeline: reference flow de-identification processor		1.6.6	Y2 goal
TSDS Flow: flow stitching for both histograms and flows over a date line		1.6.7	Y2 goal
TSDS Flow: flow stitching for parallel flow stitching, ex: gridftp		1.6.8	Y2 goal
Tstat/Flow Deployment	1.2.1	1.7	
tstat logs into pipeline	2.1.1.4	1.7.1	Y2Q1 goal
Write up a document for Anonimization pipeline including data delete on source, etc		1.7.2	Y2Q1 goal
Bro and TSTAT analysis	2.1.1	1.7.3	
Evaluate tstate for scalability (TCP retransmit)	2.1.1.4	1.7.4	Completed Y1Q2
Develop initial configuration for Bro	2.1.1.5	1.7.5	Completed Y1Q2 - not feasible for project use

Develop initial configuration for tstat	2.1.1.6	1.7.6	Deployed in TP LA Y1Q3, needs additional testing
Deploy and test tstat at TransPac LA	2.1.1.8	1.7.7	Y1Q3 - initial deployment, needs evaluation
Tstat from TransPAC/Seattle (dep on LA tstat testing)		1.7.8	Y2Q1 goal
Input ACE/WIX SFlow Data	1.1.5.2	1.7.9	Y2 goal
Input ACE/WIX Tstat Flow Data	1.1.5.2	1.7.10	Y2 goal
Input Ampath Tstat Flow Data	1.2.1	1.7.11	Y2 goal
Talk to Julio about Tstat vs sFlow data being available		1.7.11.1	Y2 goal
Evaluate using AMIS data for input		1.7.11.2	Y2 goal
Purchase/Deploy equipment if needed		1.7.11.3	Y2 goal
Incorporate data into TSDS		1.7.11.3	Y2 goal
Input PIREN Flow Data	1.3.2	1.7.12	Year 3 goal
Talk to Lassner/David Wilde about Tstat vs sFlow data being available		1.7.12.1	Year 3 goal
Purchase/Deploy equipment if needed	1.3.3	1.7.12.2	Year 3 goal
Incorporate data into TSDS		1.7.12.3	Year 3 goal
Input CENIC Flow Data (Year 3)	1.3.2	1.7.13	Year 3 goal
Input StarLight Flow Data (Year 3)	1.3.2	1.7.14	Year 3 goal
Instrumentation of DTNs	1.3.4/1.5.5	1.8	Y2 goal
Find the DTNs to instrument using top talkers list		1.8.1	Y2 goal
Add in nesc data as a test data set for internal use (?)		1.8.2	Y2 goal
NCAR DTN as endpoint - Don Middleton		1.8.3	Y2 goal
Add tstat to DTN #3		1.8.4	Y3 goal
Add tstat to DTN #4		1.8.4	Y3 goal
Input SDN Data (Year 4)	1.4.1/1.5.1/1.5.2	1.9	Y4 goal
Analysis and Experimentation	2.1-2.5	2	
Topology publication Service		2.1	Completed Y1
Capacity Planning tools	2.2.4	2.2	Delayed to Year 2
Design capacity planning tools	2.1.2	2.2.1	Delayed to Year 2
Implement and refine capacity planning tools- report generation hopefully starting for Q2	2.2.4	2.2.1	Delayed to Year 2
Recreate AS to Science Project data base		2.4	Year 2 goal
Develop basic database framework		2.4.1	Year 2 goal

Make sure framework has easily updated front end		2.4.2	Year 2 goal
Input data from Gloriad database		2.4.3	Year 2 goal
Need to get routing table dumps		2.5	Year 2 goal
Top 10 talkers			Prototype completed in Y1
Largest transfer per month analysis	2.1.3.2	2.6	Y2Q1
Tstat Analysis scripts (non flow, retransmits etc)	2.1.1/2.2.3	2.7	Y2Q2-3
Work with Viz guys to adapt TSDS framework for analysis - start with top 10 talkers		2.8	Y2Q2-3
Work on caching analysis results in TSDS (storage of derived metrics)		2.9	Y3 goal
Traffic characterization based on HNTES (of Elephant flows) - sort of part of Ed's work too	2.2.5	2.10.	Y3 goal
Flow analysis capability based on esnet tool	2.2.4	2.11	Y3 goal
Design detailed flow analysis capability based on esnet tool		2.11.1	Y3 goal
Implement detailed flow analysis capability based on esnet tool (Year 3)		2.11.2	Y3 goal
Troubleshooting tools for NOC (Year 3-5)	2.2.6/2.3.2/2.4.5/2.5.3	2.12	Y3-5 goal
Traffic analysis in SDN and multi-tennat networks (Year 3-4)	2.3.3/2.4.1	2.13	Year 3-4 goal
Use of BGP metrics with analysis of flow systems (Year 3)	2.3.5	2.14	Year 3 goal
Evaluation of PS tests and sampling (Year 3-4)	2.3.6/2.4.4	2.15	Year 3-4 goal
Incorporate BGP information (Year 4-5)	2.4.2/2.5.1	2.16	Year 4-5 goal
Develop SDN monitoring prototype (Year 4)	2.4.3	2.17	Year 4 goal
Support for Test Lab at IU	2.1.4.3	2.18	Ongoing
Visualization Tasks		3	
create a google sheet with all network details needed to generate vis		3.1	Completed Y1Q2
Integrate TSDS database queries into prototype.	3.1.6	3.2	Ongoing
Default Summary View Visualizaiton		3.3	Ongoing
Test prototype against real available data (waiting on CORS)	3.1.7	3.4	Ongoing
Develop viz prototypes based (flow data)	3.2.1	3.5	
Develop hierarchical visualization column interface & widgets		3.5.1	Completed Y1Q2
Create basics of visualization		3.5.2	Completed Y1Q2

Research how to save user preferences, visualization configuration...(cookies, login users?)		3.5.3	Completed Y1Q2
Develop Interactivity- eg drilling down into the data		3.5.4	Ongoing
Aesthetic improvements		3.5.5	Ongoing
Bind visualization (interface) to TSDS database (waiting on CORS)		3.5.6	Ongoing
Define for visualization team https://docs.google.com/a/lbl.gov/document/d/1F4nodKtHQHE3ZKAGmrxz-QxYmQ7FVsOTyCptb6FeodU/edit?usp=sharing	3.2.2	3.6	ongoing
Define what an elephant flow is		3.6.1	Initial defintion completed Y1 Q3
Concatenate TransPac flow data, and generate histogram of that data		3.6.1.1	Completed Y1Q2
Define what "loss" is		3.6.2	Ongoing
Define what "top talkers" means		3.6.3	ongoing
Show prototype to IRNC community & gather feedback (perhaps via Youtube video)	3.2.4	3.7	Planned Y2Q3
Incorporate feedback and revise visualization	3.2.5	3.8	Planned Y2Q4
Development of viz prototype for Tstat data (Year 3)	3.3.1	3.9	Goal Y3
Refine/refactor year 3 prototype	3.3.2	3.10.	Goal Y3
evaluate 3rd set of prioritized queries (Year 3)	3.3.3	3.11	Goal Y3
Release and evaluate prototype (Year 3)	3.3.4	3.12	Goal Y3
Development of viz prototype (longitudinal data analysis) (Year 4)	3.4.1	3.13	Goal Y4
Refine/refactor year 4 prototype	3.4.2	3.14	Goal Y4
evaluate 3rd set of prioritized queries (Year 4)	3.4.3	3.15	Goal Y4
Release and evaluate prototype (Year 4)	3.4.4	3.16	Goal Y4
Refine previous prototype, bug fixing (Year 5)	3.5.1	3.17	Goal Y5
Release and evaluate prototpep (Year 5)	3.5.2	3.18	Goal Y5
Project Coordination	4.2	4	
Project management and coordination	4.2.2	4.1	Ongoing
Weekly project meetings		4.1.1	Ongoing
Coordinate with NOC		4.2	Ongoing
Coordinate with AMIS research project		4.3	Ongoing
Help with TSDS integration		4.3.1	Y2 goal
Integrate other data with TSDS (Year 3)		4.3.2	Y3 goal

Coordinate with InSight research project		4.4	Ongoing
Evaluate integration of data sources		4.4.1	Y2 goal
Integrate data if feasible		4.4.2	Y3 goal
NetSage Data Privacy Policy	4.2.1	4.5	ongoing
Draft web page message		4.5.1	Y2Q1 goal
Draft partner policy		4.5.2	Y2Q1 goal
Get feed back from IRNC PIs		4.5.3	Y2Q1 goal
Get feed back from Kim Milford		4.5.4	Y2Q1 goal
Get feedback from Erin from CAIDA		4.5.5	Y2Q1 goal
Year 1 reporting		4.6	completed
Y1 annual report (with Q3)		4.6.1	Completed
Year 2 reporting		4.7	Ongoing
Updated budgets for Year 2		4.7.1	Completed
4 year 2 reports		4.7.2	Ongoing
Year 3 reporting		4.8	Y3
Updated budgets for Year 3		4.8.1	Y3
4 Year 3 reports		4.8.2	y4
Year 4 reporting		4.9	y4
Updated budgets for Year 4		4.9.1	y4
4 Year 4 reports		4.9.2	y4
Year 5 reporting		4.10.	y5
Updated budgets for Year 5		4.10.1	y5
4 Year 5 quarterly reports		4.10.2	y5
Final report for project		4.10.3	y5
Domestic Travel Year 2	4.2.4	4.11	Y2
All Hands Meeting - February 2016 U Hawaii Manoa		4.11.1	Y2
AMIS PI Meeting (February 2016 El Paso)		4.11.2	Y2
May 12,13 2016 All Hands Meeting (Bloomington)		4.11.3	Y2
Invite AMIS (Yan Luo) and InSight (Greg Cole)		4.11.3.1	Y2
Global Summit (May 2016 Chicago)		4.11.4	Y2
IRNC PI Meeting		4.11.4.1	Y2
SuperComputing Nov 2016		4.11.5	Y2
Requirements workshop with IRNC PI	4.2.3	4.11.5.1	Y2
consider doing a BOF submission on Network monitoring (Due July 30)		4.11.5.2	Y2
Demonstration of initial capabilities to Kevin at SC2016		4.11.5.3	Y2
International Travel Year 2	4.2.5	4.12	Y2
Terena (Prague June)		4.12.1	Y2

Conference paper documenting NetSage (terena paper)	4.2.6	4.13	Y2
--	-------	------	----