NetSage Network Data Privacy Policy
February 28, 2017

I. Introduction

NetSage is an open privacy-aware network measurement, analysis, and visualization service designed to address the needs of today's international networks. Modern science is increasingly data-driven and collaborative in nature, producing petabytes of data that can be shared by tens to thousands of scientists all over the world. The NSF-supported International Research Network Connection (IRNC) links have been essential to performing these science experiments. NetSage is a partnership between Indiana University, University of California Davis, Lawrence Berkeley National Laboratory, and University of Hawaii Mānoa.

NetSage was funded by the NSF's IRNC program to better understand the use of the IRNC funded backbone networks and exchange points. In much the same way as other large-scale NSF facilities track their end users, NetSage was created to understand the use of the networks. The XSEDE (https://www.xsede.org/) high performance computing platform tracks end users by institution, science domain, and project, as does the Open Science Grid computing consortium (https://www.opensciencegrid.org/). NCAR/UCAR (https://library.ucar.edu/) track the use of their data resources in similar ways.

The primary interest of the NetSage project is to understand and visualize large data flows associated with research, education, and science projects. The data that is collected from the IRNC participants is de-identified and used primarily to understand the network behaviors of large flows and to better understand the general use and functionality of IRNC-funded networks and exchange points.

NetSage works with *Partner Hosting Sites*, which are IRNC-funded projects including the IRNC NOC, exchange points, and backbone providers. NetSage's data privacy policy strives to balance the privacy interests of users whose data transits the networks that NetSage monitors, the operational needs of the Partner Hosting Sites, and the need to demonstrate the broader benefit of the NSF-funded resources. We are committed to protecting privacy and informing interested parties about our policies and practices.

II. Scope of this Policy

This policy identifies:

- The information NetSage collects about data transferred by its infrastructure;
- The ways in which this information may be used and disclosed to third parties; and

- The security measures adopted to prevent unauthorized access to this information.

III. What Information Is Collected?

NetSage captures and collects *active* networking data (for example, latency and throughput from a tool such as perfSONAR) and *passive* network metadata (for example, SNMP and flow data). This data may consist of packet headers in addition to performance data, but will never contain payload data from flows. Data sets are de-identified at the source before being stored. This data is highly aggregated and does not contain information about traffic flows specific to individual users. IRNC participants may choose to release additional data unrelated to NetSage, however that data is not part of this policy.

IV. Disclosure of Data

NetSage is the steward of all the network data it collects. NetSage, at the direction of the NetSage PI, may share network data under the following circumstances:

1. NetSage plans to make summaries of de-identified Partner Hosting Site network traffic data public on the NetSage portal (http://portal.netsage.global).
2. Upon request, Hosting Site Owners can have access to the full, de-identified data sets for their site.
3. In rare cases where there is an ongoing performance issue for a specified flow, access to raw data may be needed to debug performance issues. If both endpoints of a flow agree, the IRNC NOC may, for a limited time, use raw data at a collection point to help identify ongoing performance problems between two sites. Internal procedures exist to ensure this is done securely. NetSage and the IRNC NOC take the following actions when collecting raw data:
    a. Permission is obtained from the authorized representative at Partner Hosting Site Owner.
    b. Affected organizations are notified.
    c. Internal logs are maintained documenting what data is collected, the time period covered, who collected the data, and why the data was collected.
    d. When the issue has been resolved, the Hosting Site Owner will also be informed and the raw data files will be subsequently destroyed using industry best practice data sanitization techniques.

V. How Data Is Collected, Retained, and Protected

All network data is managed under the control of NetSage project members authorized by the NetSage Principal Investigator.

NetSage takes appropriate steps to protect collected network data from unauthorized access or disclosure. Additionally, NetSage employs industry standard

security measures to protect against the disclosure, loss, misuse, and alteration of the information under our control.

VI. Notice for Updates and Changes to Policy

This document is derived from ESnet's privacy policy (available at https://www.es.net/about/governance/data-privacy-policy/), which itself is derived from the Internet2's policy on privacy of network flow data (available at http://www.internet2.edu/policies/network-flow-data-privacy-policy/). NetSage reserves the right to update this privacy policy at any time to reflect changes in the manner in which it deals with traffic, whether to comply with applicable regulations and self-regulatory standards, or otherwise. Then Privacy Policy posted here will always be current. We encourage you to review this statement regularly.

VII. Who to Contact if You Have Questions

If you have any questions about this privacy policy, please contact Dr. Jennifer M. Schopf, the PI of the NetSage project, at jmschopf@iu.edu.

VIII. Glossary

Networking data: Active — Network data collected using tools such as perfSONAR

Networking data: Passive — Network traffic data, such as netflow or sFlow data, or data collected using passive monitoring tools such as Tstat

Partner Hosting Site — IRNC-funded projects including the NOC, exchange points, and backbone providers. This includes AmLight (PI Ibarra), TransPAC4 (PI Schopf), PIREN (PI Lassner), StarLight (PI Mambretti), AmPATH (PI Ibarra), the PacificWave Exchange (PI Fox), and the IRNC NOC (PI Jent)

Partner Hosting Site Owners — the PIs of the IRNC-funded Partner Hosting Sites