CSI2 Workshop
March 5-6, 2007
Cambridge, Massachusetts
Rev. 31-July-07

Comments to Steve Olshansky <steveo@internet2.edu>

Attending:
Chris Misra, University of Massachusetts
Daniel Adinolfi, Cornell
Brian Smith-Sweeney, NYU
Mark Poepping, Carnegie-Mellon
Doug Pearson, Indiana University/REN-ISAC
Phil Denault, Worcester Polytechnic
Kevin Amorin, Harvard
Ken Klingenstein, Internet2
Joel Rosenbaum, Columbia Univ.
Beth Young, MOREnet
Dave LaPorte, Harvard
Steve Olshansky, Internet2
Dean Woodbeck, Internet2

1. Executive Summary

The Internet2 Salsa Computer Security Incidents (CSI2) Working Group convened a meeting and discussed three main topics: RENOIR, the Shared Darknet, and the evaluation and development of open source security tools. Demonstrations and discussion focused on progress made to-date and the next steps in each of these areas.

- **RENOIR** (Research and Educational Networking Operational Information Retrieval system)
  The need for reporting and extracting information yielded these next steps: 1) develop a Web-based user interface; 2) develop a SOAP interface for queries; 3) investigate whether RSS feeds would be useful for alerts/relevant information; and 4) define the relationship to EDDY in terms of data transport. Data retention issues will focus on whether different levels of tickets (encrypted, restricted, open, etc.) require different lengths of time for retention.

- **Shared Darknet**
  The Shared Darknet discussion focused on the need to determine the reports/ information that will be available to REN-ISAC members as this service expands. While the current institutions reporting to the Shared Darknet provide daily batch files, there is a need to investigate different input, such as NetFlow and GRE tunnels, and to determine if a raw real-time data transport would be useful.

- **Open Source Security Tools**
  The Salsa-CSI2 working group will discuss whether to create a reference to list and evaluate current open source security tools and their uses, and evaluate the possibility of creating useful tools that do not currently exist (if the resources are available).

2. Welcome from Internet2

Ken Klingenstein, director of Internet2 Middleware and Security, gave a welcome and overview of his hopes for the working group meeting. He asked that participants discuss how campuses can move beyond individual policy issues and attitudes to develop collaborative ways to enhance security. Some of the main existing resources are the existence of a trusted community (REN-ISAC) and a federated trust fabric – such as the InCommon Federation for the research and education communities in the US.

While there are many federations internationally with similar missions to InCommon, they have grown much faster, mainly because they are oftentimes mandated and supported by governments.

3. REN-ISAC Review

REN-ISAC is the Research and Education Networking Information Sharing and Analysis Center, located at the University of Indiana. Doug Pearson, REN-ISAC director, provided an update and an overview of the topics to be discussed at the upcoming REN-ISAC member meeting. REN-ISAC has a malware sandbox under development and plans to implement some commercial security applications and make their functionality available (as possible) at no cost to the member institutions. The executive advisory group plans to provide an orientation at the REN-ISAC meeting and will take up the question of how membership might expand.

4. RENOIR

RENOIR is the Research and Educational Networking Operational Information Retrieval system developed by Phil Deneault of Worcester Polytechnic Institute, under the umbrella of the Internet2 Salsa-CSI2 working group. RENOIR provides a common mechanism and format for submission and transmission of security incident reports to a central repository, for analysis and access by other authorized organizations. It provides a structured way to gather the same information about each incident. This standardization allows for aggregation and analysis of data from many sources. Reports can be created and shared with authorized users with varying levels of access.

Some of RENOIR's features include:

- Standardized Report format - IODEF: Incident Object Description and Exchange Format
  - XML format developed by the IETF
  - Used for human-reported data in a machine readable, extensible format
  - Can be used as a container for other data

- Varying levels of encryption and access control
  - Encrypted
  - Restricted
  - Open
  - Semi-Anonymous
  - Informational

By using the standard IODEF format, RENOIR will collect specific data, including incident identifiers, related activity, a description of the event, contact person, various time stamps, assessments, and other relevant items.

There are still some questions to answer concerning RENOIR, which was one of the purposes of this meeting session. These questions include:

1. Determining the best methods for getting incident reports into the system and extracting information (such as reports and analyses)

Given the purpose of RENOIR – providing a centralized system for collecting incident reports and performing analyses of aggregate data – the goal is to make reporting incidents to REN-ISAC as easy as possible, while maintaining the security and privacy of data. Reports need to be timely, tracked, organized, and easily retrievable. They will be useful for determining the scope of a problem by understanding how disparate reports may in fact correlate or otherwise relate to each other, assessment of damage, sharing knowledge of the problem and sharing effective solutions.

In terms of reporting and user friendliness, it would be better to store the necessary encryption keys on the server side. It would also facilitate developing a web interface for inputting data to RENOIR. The consensus was that a simple web interface to collect incident data, without Java or other programming for more sophisticated functionality, would be the logical approach. Other features would include a SOAP interface (Simple Object Access Protocol - for exchanging XML-based messages over networks) and the ability to support email alerts to appropriate users when new events are added or information is added to existing events.

On the client side, users would be able to view and create tickets, view reports and save tickets. Using EDDY (End-to-end Diagnostics DiscoverY effort at Carnegie Mellon University) as the transport would provide a way to normalize tickets and move or categorize them with relative ease.

2. Determining the default level of access – whether new “tickets” should automatically be encrypted or otherwise restricted to authorized users.

   The general consensus was that the default for RENOIR tickets should be “restricted.” Creating a lot of encrypted tickets may hinder compiling and correlating aggregated data. When multiple people have access to an encrypted ticket, for example, they all have a piece of the secret (key) required to unlock the encryption and access the data. This means a user’s access can only be removed with their permission. One suggestion was to explore whether RENOIR could provide more granularity in individual fields, allowing individual encrypted fields in any type of ticket, as appropriate.

3. Data retention policies

   The length of data retention may be related to the encryption level of the ticket. An entry with general information may be stored forever, while an encrypted ticket with time-sensitive information may be stored for just a few days. The challenge is determining what types of data might relevant for later review and analysis. Some of the options might be:

   - Deletion after X number days
   - Deletion based on type of data (encrypted vs. open for example)
   - Archiving status of data
   - Purging only parts of a message

One of the next steps for RENOIR will be to make the appropriate changes in the near term, then recruit early target users to evaluate the system. One potential application is integration with the Shared Darknet project underway at the REN-ISAC, with incident reports created in

RENOIR, moved to the IODEF incident database, and then distributed to appropriate places and parties.

At the same time, the fields to be included in RENOIR, and those that will be mandatory, need to be defined (e.g. time stamp, IP address, description of event).

5. Shared Darknet

A darknet is a type of network security sensor comprised of a “collector” which “listens to” or monitors one or more blocks of routed, allocated, but unused IP address space. It is used to collect information about those who may be engaging in potentially malicious behavior. The shared darknet aggregates the information from all participating institutions’ local darknets. The shared darknet is operated through REN-ISAC, with data currently provided by Indiana University, New York University and the University of Auckland.

Current data submission workflow to the shared darknet

- Local institution gathers darknet flows
- Local institution anonymizes and filters the data and converts it to structured data (SD) (e.g. CSV) format
- Local institution makes SD/CSV file available to REN-ISAC via Secure Copy (SCP)
- Once per day, about 1:00 am local time, REN-ISAC retrieves SD/CSV file from local institution
- REN-ISAC stores SD/CSV data into database

Currently supported data types – (all batch processing methodologies)

- IMS
- Netflow
- Argus
- PCAP

The current REN-ISAC model requires members to provide their information in a daily batch. There is some consideration underway about accepting a dataflow from an institution in real time, perhaps using GRE (Generic Routing Encapsulation) tunnels or NetFlow. EDDY may also be a useful transport mechanism for such submissions.

REN-ISAC will allow members to anonymize data, but encourages them not to do so. One of REN-ISAC’s purposes is to provide analyses and correlation of aggregate data and anonymization will limit the information that is collected and thus the extent and utility of these analyses. REN-ISAC is also working toward reducing the “noise” among the shared darknet data, such as caused by P2P (e.g. LimeWire and Aries) NAT and firewall traversal protocols.

REN-ISAC develops two sets of reports from the collected data. (1) Notifications of specific machines that scanned into the darknet are sent to the abuse/security contacts at the machine-owning organization. The notifications include IP address, timestamp, and observed activity, i.e. protocol and ports scanned. (2) Reports on aggregate observed behavior in the Shared Darknet, i.e. top active protocols and ports, reported over days, weeks, months.

In general, participants felt uncomfortable receiving raw data from other institutions. However, a detailed report for their own institution, combined with an aggregate summary of what is happening at other institutions, would be helpful. Automated batch reports and/or RSS feeds would also be useful.

6. REN-ISAC Cybersecurity Registry

Doug Pearson of REN-ISAC/Indiana University provided an overview of REN-ISAC's planning for a cybersecurity registry. The objective is to build a registry that will provide a way to find the contact information for the proper person at an organization to contact about a given security issue. The challenge is to map organizations and their relationships with the contact person, domains, netblocks and, in some cases, other organizations (in a multi-campus system, for example). The mapping chart might look like this:

Org (N) – Contact (N)
Org (N) – domain (N)
Org (N) – netblocks (N)
Org (N) – Org (N)

An organization has N number of contacts and N number of domains and an N number of netblocks. An organization can also have a relationship to other organizations, such as Indiana University-Bloomington (iub.edu) and Indiana-University Kokomo (iuk.edu). Contacts at organizations will be REN-ISAC functional contacts and REN-ISAC members. There is also an escalation procedure so if the first contact person does not respond, there are subsequent contacts.

7. Open Source Security Tools

One question for the Internet2 Salsa-CSI2 working group to consider is what can be done in the area of supporting the development and adoption of new, or extension of existing, open source security tools. Examples might be web application assessment toolkits, event and incident management toolkits, or agent-based endpoint security tools. The point is to provide tools which address acknowledged needs but may not currently exist.

In addition, there was discuss of collecting and posting information about tools that currently exist, their strengths and weaknesses and, perhaps, some sort of rating system, knowledge base of experiences, and cost/benefit analyses. This would serve as a useful reference for the community. For example, if an institution is upgrading their security program, where are effective places to put limited resources?

Brian Smith-Sweeney, NYU, will develop a document to include an analysis of existing security tools, the effort and complexity required to install and configure them, the time it takes to maintain them, and the value/benefit of the tools.

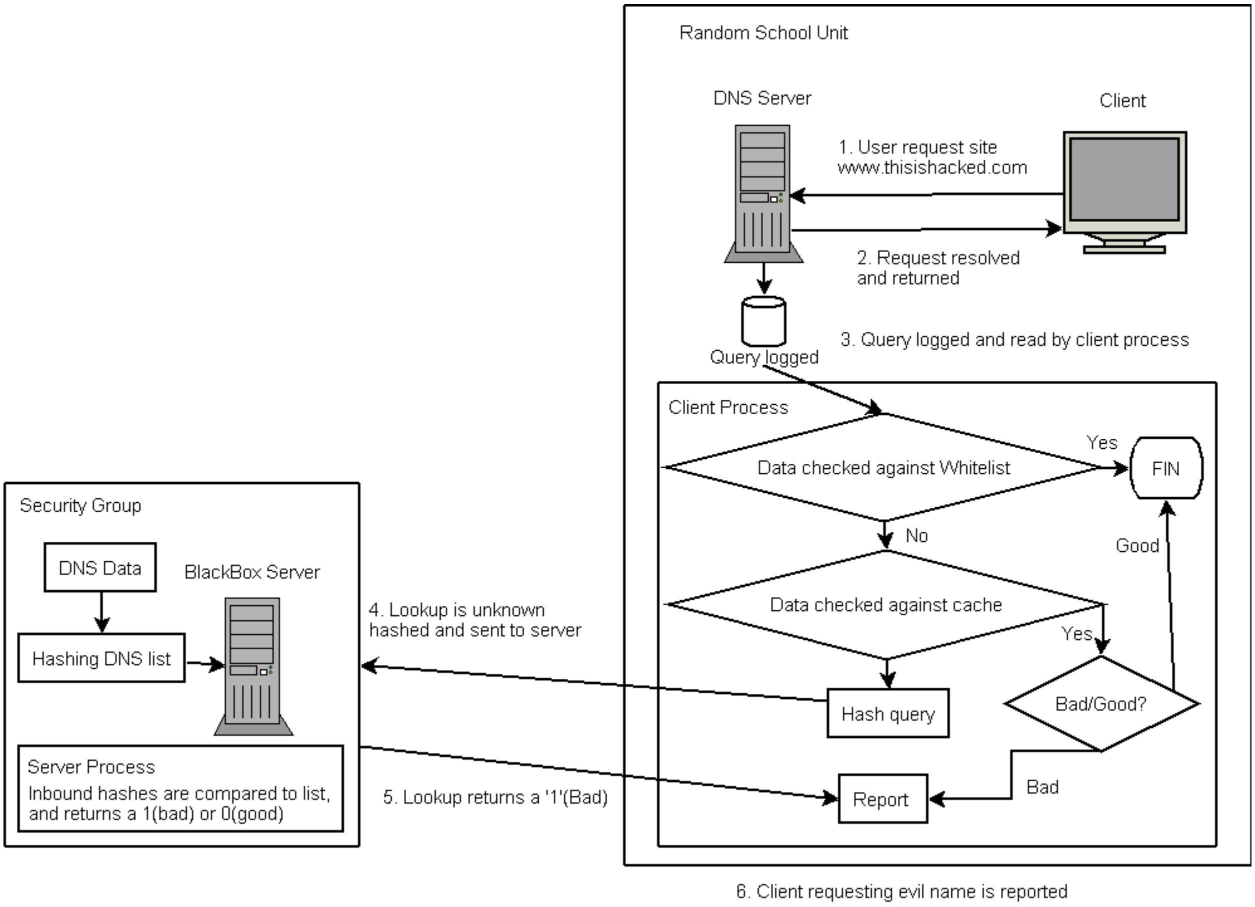
Besides evaluating current tools, Salsa-CSI2 is interested in determining if other software is needed and, if so, participating in developing such software. The key will be determining the problem set, defining a managed process for requirements analysis and development, and creating an environment that allows developers to create code specificallytargeting these requirements. Some examples include:

1. Tool development in support of REN-ISAC
2. General purpose tools like RENOIR
3. Generalizing the hashing model

4. Identifying projects already in process that would benefit from some funding and/or development support from the Higher Education community
5. Defining how to support such tools as they move into production

8. Hashing

Hashing is one possibility for protecting the security and privacy of individual institutions' data, while providing a way for the institution to retrieve that data. It would be a way to allow REN-ISAC to collect and retain the data to provide security alerts and perform its aggregation and analysis functions, while keeping the data secure. Institutions would still be able to query their own data. Hashing solves the problem of privacy for each member. A sample workflow is shown below.


Random School Unit
DNS Server
Client
1. User request site
www.thisishacked.com
2. Request resolved
and returned
Query logged
3. Query logged and read by client process
Client Process
Data checked against Whitelist
Yes
FIN
No
Good
Data checked against cache
Yes
Bad/Good?
Hash query
Bad
Report
Security Group
DNS Data
BlackBox Server
Hashing DNS list
Server Process
Inbound hashes are compared to list, and returns a 1(bad) or 0(good)
4. Lookup is unknown
hashed and sent to server
5. Lookup returns a '1'(Bad)
6. Client requesting evil name is reported

9. End-to-end Diagnostic DiscoverY (EDDY)

A brief presentation provided background information on EDDY. EDDY was developed to provide a structure to manage and analyze diagnostic data. The key to EDDY is the Common Event Record (CER). By having CERs contain standard information, events can be normalized and moved to a diagnostic "backplane." A CER includes such information as timestamp, observation point, normalizer, location, event type, GUID and severity. EDDY is extensible and can accommodate different data formats.

APPENDIX A: References/Glossary

1. End-to-end Diagnostics DiscoverY (EDDY)
   http://www.cmu.edu/ eddy/
2. IMS
   http://ims.eecs.umich.edu/
3. Netflow
   http://www.cisco.com/en/US/products/ps6601/products_ios_protocol_group_home.html
4. Argus
   http://www.qosient.com/argus/
5. PCAP
   http://sourceforge.net/projects/libpcap/
6. Generic Routing Encapsulation (GRE)
   http://www.faqs.org/rfcs/rfc2784.html
7. LimeWire
   http://www.limewire.com/