Security at the Cyberborder Workshop Report

Exploring the relationship of International Research Network Connections and Cybersecurity

February 22-23, 2012

University Place Conference Center and Hotel

850 W. Michigan St., Indianapolis, IN 46202

Published March 28th, 2012

Von Welch, Doug Pearson, Brian Tierney, James Williams

Report editors and workshop organizers

Please cite as:

Von Welch, Doug Pearson, Brian Tierney, and James Williams (eds). Security at the Cyberborder Workshop Report. March 28th, 2012. http://hdl.handle.net/2022/14070

Contact author: Von Welch (vwelch@indiana.edu)

Acknowledgements

The editors would like to thank:

- Bridget Sutton, Dara Eckart and James Boyd at Indiana University's Center for Applied Cybersecurity Research (CACR), under funding from the Lilly Endowment, Inc., for logistical support.
- The workshop presenters: Michael Bailey, Robert Brammer, Robert Danford, Erin Kenneally, Doug Pearson, Josh Polterock, Michael Sinatra, Wayne Routly, Robin Sommer, Brian Tierney, and Kevin Thompson.
- All the workshop participants for their contributions during discussions.
- KC Claffy of CAIDA for advice during workshop preparation and a review of an early draft of this report.

We thank the National Science Foundation for supporting the workshop and the writing of this report. This material is based upon work supported by the National Science Foundation under Grant No. 1158796. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Executive Summary

The NSF's international research network connections (IRNC) program provides five high-performance (multiple 10Gbps circuits), production network connections ("ProNets") linking U.S. research with peer networks in other parts of the world. This document reports on a two-day workshop held on February 22nd and 23rd in Indianapolis, IN that explored the relationship between the IRNC ProNets and cybersecurity. The workshop had two goals:

(1) Exploring the responsibilities, opportunities, and challenges for IRNC ProNets to contribute to the cybersecurity of the research communities that they serve.

(2) Exploring opportunities and challenges for the IRNC connections to foster cybersecurity and networking research.

The scope of the workshop included technical, legal, policy, and social expectations regarding these two goals. While all aspects of cybersecurity are important, the CyberBorder workshop focused on cybersecurity in terms of the impact of the traffic flowing over a ProNet on the cybersecurity of the customers of that network (i.e., the research communities) and the broader Internet community.

A set of 10 findings emerged from discussions, which are listed immediately following this summary. The main observations from the findings, are summarized as follows:

- Responsibility for cybersecurity lies ultimately with organizations that are the IRNC ProNet's customers, however there is potential for ProNets, due to their view of a large amount of aggregated traffic, to contribute to their cybersecurity through monitoring (of, for example, DNS traffic) and attribution of spoofed malicious traffic, and, at the request of organizations, activity mitigation of certain threats.
- Data for cybersecurity research is important; an analysis by one of the presenters (Michael Bailey, U. Michigan) showed that having data played a key role in over half the papers presented at five recent cybersecurity conferences. However, sharing data must address concerns from a legal (including privacy) standpoint, requiring a combination of technical and policy controls to do so appropriately. Those undertaking sharing are advised to leverage experiences and services offered by, for example, the DHS PREDICT and CAIDA programs.

A key challenge in expanding the mission of IRNC ProNets to include cybersecurity or data sharing is that while a U.S.-centric view of IRNC ProNets suggests them as NSF projects, they are complex partnerships between NSF, grantee institutions, carriers, and international partners. All of these partners have legal, sociological, or policy constraints to consider.

Workshop Findings

1. IRNC ProNets are part of a complex ecosystem, and cannot provide a "silver bullet" in terms of cybersecurity to any of their research communities, but instead can contribute to the cybersecurity process. IRNC ProNets should look to contribute where they are especially qualified due to their perspective in that ecosystem as an aggregator of significant traffic and bridge between communities.
2. Sharing of monitored network data by IRNC ProNets is beneficial to both the broader community as well as the ProNets themselves. Even sharing just a subset of all the available data (e.g., flow information regarding a single application) can be beneficial.
3. IRNC ProNets should only undertake active measures, such as blocking traffic, at the request of affected sites in order to avoid operational issues and confusion.
4. It would be beneficial if IRNC ProNets undertook preparation (planning and simulated event exercises) to deal with major incidents such as denial of service attacks that transverse their infrastructure. Denial of service attacks are a prime example of an incident for which IRNC ProNets could prepare.
5. IRNC ProNets have challenges in providing cybersecurity contributions to their research communities. Those challenges include:
    - IRNC ProNets provide high-performance, highly available end-to-end connectivity to support science. The operational impacts of any cybersecurity measures need to be understood in the context of that mission.
    - IRNC ProNets are partnerships between U.S. and international participants, and changes to their mission should be coordinated across their international communities.
    - To properly manage expectations and avoid operational confusion, IRNC ProNets need to be transparent and clear in what their cybersecurity contributions are.
    - Due diligence needs to be taken regarding the collection, use, and disclosure of data to ensure that privacy and confidentiality implications are considered within an overall risk management approach to data stewardship.
6. Network data can be personally identifiable, creating risks to both data providers and recipients when sharing data for research, including privacy and confidentiality implications for individuals and organizations. Hence, implementing a data-sharing program to support research is a critical legal and technical undertaking. Existing data sharing frameworks, models, and experiences (e.g., PREDICT, CAIDA) should be considered to advance data sharing effectiveness and efficiency.

7. Anonymization of data may impede its utility or preclude its benefit for research. It is important to keep in mind that anonymization is just one type of data disclosure control that should be part of a data sharing framework. It may be appropriate and necessary in certain circumstances, but it is often accompanied by a trade-off with respect to utility for both research and operations.
8. Sharing data broadly is a significant effort that should be approached from a risk management perspective. It may be strategically prudent to undertake such efforts incrementally. Sharing data initially in a limited manner can serve to validate the usefulness of the data before investing significant resources, engender trust in the sharing framework, and reinforce its value once the sharing is scaled.
9. There are benefits to sharing data (and costs to not doing so, such as hindering reproducibility of research), but these are hard to objectively measure beyond a few simple metrics such as requests, citations, etc.
10. IRNC ProNets have a position in the ecosystem as an aggregator of large amounts of traffic that give them a rare perspective. In undertaking any data-sharing activity, they will provide the most benefit when leveraging that position.

TABLE OF CONTENTS

WORKSHOP PARTICIPANTS 1

1 INTRODUCTION 2

2 WORKSHOP GOALS 2

3 RELATED WORKSHOPS AND REPORTS 3

4 DAY ONE: THE ROLE OF IRNCS IN CYBERSECURITY 4

4.1 BACKGROUND: NSF IRNC PROGRAM 4

4.2 DAY ONE PRESENTATIONS 5

4.3 DAY ONE DISCUSSIONS AND FINDINGS 6

5 DAY TWO: THE ROLE IRNCS IN CYBERSECURITY AND NETWORKING RESEARCH ... 9

5.1 DAY TWO PRESENTATIONS 9

5.2 DAY TWO DISCUSSION AND FINDINGS 11

6 CONCLUSION 14

7 ORGANIZER COMMENTS 14

8 REFERENCES 15

APPENDIX A WORKSHOP AGENDA 16

APPENDIX B SUGGESTED READING 18

APPENDIX C DATA AND ACCESS WISH LIST 19

APPENDIX D WORKSHOP PRESENTATIONS 20

Workshop Participants

Invited participants included the IRNC PIs and staff, cybersecurity staff from research projects and institutions of higher-education, and cybersecurity and network researchers. The 33 participants (including the 4 organizers) at the workshop were:

- Mike Bailey, U. Michigan
- Robert F. Brammer, Internet2
- Carter Bullard, Argus Project
- Scott Campbell, NERSC
- Greg Cole, U. Tennessee-Knoxville/GLORIAD
- Chip Cox, Florida International U./America's Lightpath
- Robert Danford, Verizon
- Tom Davis, Indiana U.
- John Gerber, U. Tennessee-Knoxville/GLORIAD
- Minaxi Gupta, Indiana U.
- John Hicks, Indiana U./ACE
- Julio Ibarra, Florida International U./America's Lightpath
- Gabriel G. Iovino, REN-ISAC
- Dave Jent, Indiana U.
- Erin Kenneally, CAIDA
- David Manz, PNNL
- David Monnier, Team Cymru
- Doug Pearson, REN-ISAC (organizer)
- Josh Polterock, CAIDA
- Wayne Routly, GEANT/DANTE
- Daniel Schwalbe, U. of Washington
- Michael Sinatra, ESnet
- Abe Singer, LIGO
- Steve Smith, U. Hawaii/TransLight
- Brian Smith-Sweeny, NYU
- Robin Sommer, ICSI and Bro Project
- Martin Swany, Indiana U.
- Kevin Thompson, NSF
- Brian Tierney, ESnet (organizer)
- Steven Wallace, Indiana U.
- Von Welch, Indiana U. (organizer)
- Alan Whinery, U. Hawaii/TransLight
- Jim Williams, Indiana U./ACE (organizer)

1 Introduction

The United States research and education community collaborates with colleagues in the global community. As part of these collaborations, members of this community access remote instruments, data, and computational resources located throughout the world. Similarly, major NSF investments in large-scale science and engineering facilities, located both inside and outside the U.S., are utilized by multi-national research and education collaborations.

NSF's international research network connections (IRNC) program supports these international scientific collaborations. The IRNC program connects U.S. scientists to every part of the globe. Supported scientific endeavors range from astronomy (for optical and radio telescopes) to physics (LHC data exchange) to climate studies (supporting oceanic sensor networks).

IRNC networks connections carry a broad variety of traffic to and from research and educational institutions across the U.S. borders. Hence the IRNC networks are unusual in their operation of network connections that cross the U.S. border. This places their networks in a position to see a large amount of aggregated traffic crossing national boundaries, between customers under different cultural and legal frameworks.

On February22-23, 2012, Indiana University hosted a two-day workshop in Indianapolis, IN that explored the relationship between the IRNC ProNets and cybersecurity. This document is the report from that workshop.

2 Workshop Goals

The workshop had two goals, one of which was explored on each day of the two-day workshop:

(1) The first day explored the responsibilities, opportunities, and challenges for IRNC ProNets to contribute to the cybersecurity of the research communities that they serve.

(2) The second day explored opportunities and challenges for the IRNC connections to foster cybersecurity and networking research.

The scope of the workshop included technical, legal, policy, and social expectations regarding these two goals.

Cybersecurity is a broad topic, and in the context of IRNC ProNets, the organizers determined at least three aspects:

1. Security of the network: Ensuring routers and other equipment are protected from unauthorized tampering and eavesdropping, both remotely and physically.

2. Routing security: Ensuring routing, both internal to the network and with its interactions with peer networks, happens in a manner that is resistant to unauthorized manipulation.
3. Cybersecurity impacts of the traffic flowing over the network on users of the network. Traffic flowing over an IRNC, as with any network, can be malicious, from simplistic brute force scanning, SPAM, and denial of service attacks to sophisticated spear phishing, targeted exploits, and exfiltration.

While all three aspects are important, the CyberBorder workshop focused on the third aspect of cybersecurity in the context of IRNC ProNets—that is the impact of the traffic flowing over a ProNet on the cybersecurity of the customers of that network (i.e., the research communities) and the broader Internet community. This choice to focus on the third aspect was in part to scope conversations so that progress could be made, but also because the organizers perceived that this area has not received the same level attention to-date (e.g., the second is being tackled by the DHS SPRI program [6]).

And while IRNC ProNets are international projects serving international research communities, the workshop took a primarily U.S. focus in both the invited participants and in the legal and policy issues it considered. This was done to allow for a manageable scope of discussions in terms of policies and legal aspects. As discussed in the findings and in possible next steps, actions taken by IRNC ProNets or subsequent deeper investigations should consider international IRNC collaborators as equal partners.

3 Related Workshops and Reports

We briefly describe some prior workshops and reports the editors are aware of and are most related to topics and goals of the Cyberborder workshop. These differ primarily from the Security at the Cyberborder workshop in that they are all scoped more broadly, with only the Inernet2 workshops being specific to networking. In general, they provide motivating examples for the Cyberborder workshop. Where discussions and findings in the subsequent sections of this report relate to any of these, we include a note in the text.

In August 2010, Georgia Tech hosted an NSF workshop on Cyber Security Data for Experimentation [1]. This workshop brought together people from companies, academia, government agencies, and other organizations to discuss "(1) models of engagement that will allow the research community to conduct experiments with real-world data sets, (2) how to share research results, and (3) how funding agencies can facilitate the process."

In 2009, a number of cybersecurity researchers authored "Data for Cybersecurity Research: Process and 'Wish List'" [10]. This paper describes a set of needs for data for cybersecurity research (more generally than networking), modalities and delivery mechanisms, touches on privacy and legal concerns, and then provides a detailed wish list for data.

In 2009 and 2010, Internet2 held two workshops on Collaborative Data-Driven Security for High Performance Networks [2,3]. These workshops shared experiences between various projects involved in data-driven cybersecurity projects.

The National Academies Press published a report "Expanding Access to Research Data: Reconciling Risks and Opportunities" [11] which includes a discussion of accessing data for research in a general context.

4 Day One: The Role of IRNCs in Cybersecurity

The first day explored the opportunities and challenges for IRNC ProNets to contribute to the cybersecurity of the research communities that they serve. This included exploration of active traffic blocking, monitoring and notification, and passive monitoring.

The day started with a welcome by the organizers and Indiana University's CIO and Vice President for Information Technology Bradley Wheeler. Kevin Thompson from NSF opened the technical program with a presentation on the NSF IRNC program summarized in the following section[1].

4.1 Background: NSF IRNC Program

The opening presentation of the workshop was an overview of NSF's IRNC program [4] given by Kevin Thompson. The goals of the IRNC program are:

- Provide network connections linking U.S. research with peer networks in other parts of the world
- Stimulate the deployment and operational understanding of emerging network technology and standards in an international context
- Support science and engineering research and education applications

The program supports three different types of projects: Production Networks, a.k.a. "ProNets", Experimental Networks, and Special Projects. This workshop was primarily focused on the IRNC ProNets[2], and how they relate to cybersecurity, hence we scope discussions in this section to IRNC ProNets.

There are five ProNet awards, each providing high-performance connectivity (typically multiple 10Gbps circuits) between the U.S. and research networks in other parts of the world. Details of these five awards can be found at [5].

A key fact that emerged during the workshop is that while a U.S.-centric view of IRNCs is as NSF projects, they are in actuality complex partnerships between NSF, grantee institutions, carriers, and international partners. IRNC ProNets typically have significant contributions from the international partners, many times larger than the contributions from NSF.

---

[1] Copies of all workshop presentations are provided in Appendix D.
[2] The term "an IRNC" was sometimes used during the workshop as shorthand for "an IRNC ProNet."

4.2 Day One Presentations

Overview of threats and options for defense:

Robin Sommer, of the ICSI and the Bro project, gave a presentation on network-based cybersecurity threats with emphasis on commercialization of attacks, highly-targeted attacks (e.g., advanced persistent threats), and insider attacks. He followed up that discussion of threats with a defense strategy based on (1) establishing visibility, (2) analyzing semantics, (3) sharing intelligence, and (4) undertaking active response[3].

Robert Danford of Verizon followed with a talk on his network defense experiences. He described a key goal being to make attackers "noisy" so that their tunneling and beaconing activities are more easily detectable by network monitoring. The question came up as to what level of NetFlow sampling was appropriate for monitoring to balance resources required for that monitoring versus having enough data to obtain context; his response was that he has found sampling needs to be more frequent than 1:100 to be useful.

Experiences of similar networks:

Representatives from Internet2 and ESnet, networks similar to IRNCs though without the international focus, were invited to present relevant experiences in regards to providing cybersecurity to their customers.

ESnet experiences: Brian Tierney and Michael Sinatra from ESnet followed with a presentation they co-authored with William Johnston. Their presentation described ESnet's stance that cybersecurity is the responsibility of the DOE laboratories and sites that comprise their customer base. However, they did go on to describe how ESnet has intervened in rare circumstances where they were particularly well situated to do so, for example, by withdrawing routes, changing traffic priority and signature-based filtering during a distributed denial of service (DDOS) attack. On an ongoing basis, ESnet collects NetFlow data from all ESnet core routers and share that data with sites (sites are allowed to see data for which they are the source or destination).

They also presented a threat analysis of IRNC ProNets, which concluded that the primary risk enabled by those ProNets was due to the high bandwidth they provide. Targeted attacks are typically low bandwidth and hence IRNC ProNets don't change the risk profile for such attacks, but DDOS attacks often do benefit from higher-bandwidth networks and hence IRNC ProNets may change the risk profile for such threats. Provisioned paths (VLANS, MPLS paths, light paths, etc.) provided by IRNC ProNets were identified as a secondary risk in that they could provide a mechanism to bypass normal site protections.

GÉANT experiences with handling incidents: Wayne Routly from DANTE gave a presentation on the GÉANT Network Security Handling and Response Process

[3] Similar to the "data, analysis, collaboration and action" elements identified by [2].

(NSHaRP), an automated incident handling and notification system used by GÉANT. NSHaRP is used by GÉANT to automate 85% of the handling for the 2,500 incidents/month they handle. He also briefly touched on some of the different sharing constraints that their European customers have (e.g., Scandinavian NRENs are not allowed to share NetFlow data without a request from law enforcement) and constraints on GÉANT itself (they cannot share European IP addresses suspected of being the source of malicious behavior outside of the European Union).

CAIDA Privacy-Sensitive Sharing Model: Erin Kenneally from CAIDA gave a presentation on challenges of data sharing to support network and security research. Those challenges include (1) an evolving and sometimes uncertain legal regime (e.g., ambiguity surrounding the sensitivity of IP addresses), (2) compartmentalized technology solutions for addressing information risks, (3) inconsistency and ambiguity in understanding data sensitivity, and (4) ill-defined and under-valued benefits networking research. She then proceeded to discuss that despite these challenges, data sharing is possible when both policy/legal and technical stakeholders are committed to that objective, and described the Privacy-Sensitive Sharing (PS2) model [12] for sharing that combines privacy-enhancing technology with privacy-principled policies.

Internet2 Cybersecurity Initiative: The final presentation from day one was from Robert Brammer of Internet2, on Internet2's vision for a new Internet2 Cybersecurity Initiative. The vision included potential opportunities that included partnerships to leverage resources, attracting industry partners for Internet2's new Net+ program, furthering Internet2 and REN-ISAC collaboration, and using the Internet2 network as a platform for cybersecurity research and development.

4.3 Day One Discussions and Findings

Discussions during and following the day one presentations explored several aspects of how IRNC ProNets could and should contribute to the cybersecurity of the research communities they serve. A fair amount of the discussion focused on distinguishing between what IRNC ProNets could do and what they should do.

One thing that was strongly agreed to was that sites still have ultimate responsibility for their cybersecurity and often have very different policies and tolerances in that regard. Combined with other technical difficulties (e.g., encrypted web and SSH content), this indicates that IRNC ProNets can only contribute to a site's cybersecurity, not provide a complete solution. Further discussion concluded that IRNC ProNets should focus on aspects of cybersecurity that they are especially qualified for, namely aspects that take advantage of the aggregated traffic that transverses their links and the fact they bridge what would normally be operationally separate communities.

Finding 1: IRNC ProNets are part of a complex ecosystem, and cannot provide a "silver bullet" in terms of cybersecurity to any of their research communities, but instead can contribute to the cybersecurity process. IRNC ProNets should look to contribute where they are especially qualified due to their perspective in that ecosystem as an aggregator of significant traffic and bridge between communities.

The participants discussed a number of specific examples of contributions to cybersecurity that IRNC ProNets could undertake (it should be noted that these examples, like most of the possibilities discussed in the workshop, are not currently part of the mission or funding of IRNC ProNets). These examples included:

- Being prepared to leverage their position in having visibility in a large amount of aggregate traffic to monitor that traffic, even if they don't undertake regular monitoring. This could assist in their customers in debugging, or cybersecurity incident investigation, or working with organizations that monitor Internet threats to identify those threats. Examples include helping victims of a DDOS attack that utilized spoofed source addresses trace the source of that traffic to the next hop., and tracking command and control systems for botnets.
- Passive DNS Monitoring (pDNS) [7] to help monitor Botnet and other malware command and control traffic.
- Helping establish best practices across the communities they connect.
- IRNC ProNets, in fulfilling their mission, interact with a number of broadly distributed entities, e.g., NRENS, autonomous systems (AS's), networks such as ESnet and Internet2, science research projects, and organizations hosting those projects; there may be a role here for IRNC ProNets to facilitate communication of cybersecurity-relevant information among those entities.
- Sharing information e.g., pDNS monitoring, with Internet-wide entities, e.g., ISC.org, REN-ISAC, Route Views.
- Explore the possibility of providing services in aggregate across their customer sites, when that could be done more cheaply or effectively (e.g., given limited skilled workforces) in the aggregate.

With regards to the last point on sharing information, there were discussions that such sharing was mutually beneficial as it often generates reciprocal information sharing.

Sharing information with the IRNC ProNet's research communities directly was also discussed as beneficial as well, though the point was made that such sharing should be filtered based on the interest of those communities to avoid "spamming" them with information that isn't relevant to them.

Finding 2: Sharing of monitored network data by IRNC ProNets is beneficial to both the broader community as well as the ProNets themselves. Even sharing just a subset of all the available information (e.g., flow information regarding a single application) can be beneficial.

There was strong agreement among the participants that there are some circumstances in which it makes sense for an IRNC ProNet to take an active role in cybersecurity, e.g., blocking DDOS traffic. The threats discussed in this context were coherent with the observation from the ESnet presentation that the main impact on the threat landscape from IRNCs is their bandwidth and hence possible enabling of DDOS and similar high-traffic attacks.

However the participants agreed that it was important for IRNC ProNets not to create operational problems and confusion for its customers, and such active measures should only be done at the request of the effected sites.

Finding 3: IRNC ProNets should only undertake active measures, such as blocking traffic, at the request of affected sites in order to avoid operational issues and confusion.

Exploring active measures more deeply, the example of DDOS mitigation and the post mortem from Tierney and Sinatra's ESnet presentation was discussed. Given that their analysis showed that IRNC ProNet's major potential change to the risk profile of the community is their increased bandwidth, a denial of service attack seems like a prudent incident for which to prepare. The participants agreed that such crises are prime situations during which IRNC ProNets could help their communities. However, it is beneficial to be prepared for such events through planning and simulation exercises.

Finding 4: It would be beneficial if IRNC ProNets undertook preparation (planning and simulated event exercises) to deal with major incidents such as denial of service attacks that transverse their infrastructure. Denial of service attacks are a prime example of an incident for which IRNC ProNets could prepare.

The question was also raised, given the collaborative, international nature of IRNC ProNets, what the process would be for a ProNet to modify its mission to include active mitigation activities. This question is one of a number of challenges discussed with regards to IRNC ProNets providing the cybersecurity contributions discussed in this section; these challenges are described in Finding 5.

Finding 5: IRNC ProNets have challenges in providing cybersecurity contributions to their research communities. These challenges include:

- IRNC ProNets provide high-performance, highly available end-to-end connectivity to support science. The operational impacts of any cybersecurity measures need to be understood in the context of that mission.
- IRNC ProNets are partnerships between U.S. and international participants, and changes to their mission should be coordinated across their international communities.
- To properly manage expectations and avoid operational confusion, IRNC ProNets need to be transparent and clear in what their cybersecurity contributions are.
- Due diligence needs to be taken regarding the collection, use and disclosure of data to ensure that privacy and confidentiality implications are considered within an overall risk management approach to data stewardship.

5 Day Two: The Role IRNCs in Cybersecurity and Networking Research

The second day of the workshop shifted the topic to explore the opportunities and challenges of IRNC making contributions to cybersecurity and networking research[4]. The scope of contributions included the providing of data as well as opportunities for network access to test new algorithms or technologies.

5.1 Day Two Presentations

Presentation of draft day one findings: The organizers started the day off by presenting a set of draft findings from the first day of the workshop and accepting feedback from the participants on those findings.

Researcher Data Needs: Michael Bailey of the University of Michigan started the presentations addressing the goals of the second day with a presentation on researchers' need for data. He pointed out an analysis he had conducted of papers submitted to major cybersecurity research conferences that concluded that while

[4] For a discussion of the differences between data for operational cybersecurity and data for research, see Section II-1 of [2].

the most common categories of papers were demonstrations, attacks, and proofs, the "long tail" of remaining categories all depended on data.

Dr. Bailey described how obtaining data used to be solely an ad hoc process through personal relationships, but now services such as PREDICT provided formal, well-publicized mechanisms. He discussed how the type of desirable data has evolved to keep up with the evolving threats, from worms and viruses, to botnets, and how a challenge with providing data is keeping up with the "hot cybersecurity topic."

He presented two main ways of mitigating the privacy risks in sharing data for research, namely (1) restricting access to the data and (2) restricting the data (that it filtering or anonymizing it). He discussed that restricting the data while preserving its utility was a possibly insurmountable problem and then turned to the challenge of restricting access and presented four approaches for doing so: (1) licensing, or requiring researchers to legally agree to obligations on how they handle data they are given, (2) remote execution systems where data is maintained by the owner and researchers are allowed to submit queries against the data, (3) data enclaves, physical rooms in which the researchers is allowed to access the data and limited in what they can take from the enclave, and (4) virtual data enclaves, which are a mix of the other three approaches, allowing for remote access to licensed researchers with a restricted ability to act on the data.

CAIDA Data Sharing Case Study: Josh Polterock gave a presentation of CAIDA's experiences in data sharing. He said that CAIDA has received 3,405 requests for their data over the past nine years, with 2,297 of those requests being approved, and 1,868 of the approved requests being acted on by data being accessed. For the past three years, Polterock presented a breakdown of requests by dataset type, with two most popular datasets being OC192 and OC48 traces, and topology data.

Polterock discussed an ongoing effort to create a CAIDA master acceptable use policy (AUP) that covered their data, categorized at four different levels of sensitivity. Polterock concluded with a slide giving a "wish list" of data that CAIDA staff believes would be of interest to researchers.

REN-ISAC Darknet Data Case Study: Doug Pearson gave a presentation on REN-ISAC's experiences sharing data from the Shared Darknet program (monitored traffic destined for unallocated IP addresses, and hence with no legitimate purpose). He conveyed there had been a disappointing level of use of the data by researchers and a number challenges in operating the sharing program, including the effort required in establishing and operating the sharing framework, finding funding for the activity, and the fact isn't an obvious channel to promote awareness of the data in the research community.

Internet2 Data Sharing Case Study: Robert Brammer gave a presentation on Internet2's experiences in sharing data through the Internet2 Observatory. While there was good use of these data a few years ago, primarily for networking as opposed to cybersecurity, that usage has decreased in recent years for reasons that have not been discerned.

5.2 Day Two Discussion and Findings

A fair amount of the day two discussions were not particular to the IRNC program, or even networking, in that they discussed generic problems with sharing data for research. We cover those discussions first and conclude this section with discussions particular to the IRNC program.

It was agreed, based on the experiences shared regarding the PREDICT, CAIDA, REN-ISAC and other efforts, that setting up a data sharing program is definitely a legal and technical challenge. Projects attempting such as task should, at a minimum, understand those experiences, and give strong consideration to leveraging or learning from existing sharing models and frameworks (though we note that it arose in discussion that that getting data into PREDICT is more effort than getting access to data in PREDICT). Models for handling sensitive in other contexts (e.g., medical data) and non-open research (classified data). can benefit similar efforts in cybersecurity research.

Network data can be personally identifiable, which motivated discussion about the risks of sharing data, and consensus that such risks can and do negatively affect individual network users and organizations in terms of legal, financial, operational and reputational impacts. A number of risk management strategies for these risks were discussed, primarily based on the Kenneally and Bailey presentations, and included:

- Anonymization of the data to protect privacy. However, there was discussion regarding the challenge of anonymizing data sufficiently to protect privacy may render that data of limited-to-no use in research. Further, anonymization done by Internet2 and REN-ISAC (zeroing of lower order bits of IP addresses) is strongly suspected to have contributed to the lack of use of their datasets. This reinforces the contention that anonymization should be viewed as one component to an overarching framework.
- Mediating access to data. This may involve limiting data access to trusted recipient researchers via technical means (e.g., data enclaves) and/or policy controls (e.g., licensing).
- Only sharing data within an institution eases some of the legal burdens, but such limits means the data only has limited value in allowing for the reproducibility of results. Realize the trade-offs that follow from limiting the type or scope of data to be shared, e.g., restricting data disclosure to a single institution can limit legal risk but simultaneously constrain scientific reproducibility goals.

Also discussed was that organizations that gather threat information, particularly as part of their mission, may be sensitive to sharing that information under the concern it may alert malicious actors as to their monitoring methods.

Given the discussions, there was general agreement that a combination of technical and policy procedures, as described in Kenneally's PS2 framework and exemplified by something like the Virtual Data Enclave described in Bailey's talk, seemed most

appropriate models for data risk management. This discussion of multiple modalities for sharing data is also expressed in [11].

Finding 6: Network data can be personally identifiable, creating risks to both data providers and recipients when sharing data for research, including privacy and confidentiality implications for individuals and organizations. Hence, implementing a data-sharing program to support research is a critical legal and technical undertaking. Existing data sharing frameworks, models and experiences (e.g., PREDICT, CAIDA) should be considered to advance data sharing effectiveness and efficiency.

Finding 7: Anonymization of data may impede its utility or preclude its benefit for research. It is important to keep in mind that anonymization is just one type of data disclosure control that should be part of a data sharing framework. It may be appropriate and necessary in certain circumstances, but it is often accompanied by a trade-off with respect to utility for both research and operations.

Fostering usage of the data is more than "build it and they will come." The most effective mechanism for advertising a new data source is its inclusion in research papers. But this creates a chicken-and-egg situation where the data needs to be used in order to attract users. A related challenge that was discussed was ensuring data is actually useful. These issues combine to suggest that data sharing should be done initially with individual researchers and then elevated to a public framework once it has proven useful.

Finding 8: Sharing data broadly is a significant effort that should be approached from a risk management perspective. It may be strategically prudent to undertake such efforts incrementally. Sharing data initially in a limited manner can serve to validate the usefulness of the data before investing significant resources, engender trust in the sharing framework, and reinforce its value once the sharing is scaled.

There was also significant discussion regarding the benefits of sharing data for research. While it was generally agreed that some metrics exist for measuring benefit, e.g., number of requests for data, number of citations of the data in publications, it is hard to quantify the overall objective benefits in a manner that allows one to weigh those benefits against the risks.

Related to this, measuring the opportunity costs of not sharing data is also difficult. For example, without access to data, once cannot reproduce research based on it

(echoed by Finding 2 of [1]). This lack of metrics against more tangible risks can make an organization's decision to share a challenge.

One aspect of sharing that was discussed is that researchers can often use personal relationships with data sources to obtain data they need, but these relationships typically take time to establish and may not be mechanisms open to new researchers, hence the importance of general data-sharing mechanisms (e.g., CAIDA, PREDICT) (echoed in [10]).

> Finding 9: There are benefits to sharing data (and costs to not doing so, such as hindering reproducibility of research), but these are hard to objectively measure beyond a few simple metrics such as requests, citations, etc.

The final general topic discussed was that while there are significant legal and non-technical challenges to data sharing, even after all those are addressed, technical challenges still remain, such as managing data storage, the availability and expense of network taps, and having suitable computational algorithms and performance for analyzing the data (this is also expressed in [2]).

Turning to IRNC-specific discussions, the discussions came to a similar conclusion with regards to what IRNC ProNets could contribute to cybersecurity and network research that they did on day one, that is the ProNets should seek out opportunities that take advantage of their position in the ecosystem. Some specific examples of data that IRNC ProNets could provide that would take advantage of their attributes were:

- Passive DNS monitoring data
- IPv6 flows; it was discussed that IPv6 data is harder to obtain than IPv4
- Data from SSL handshakes regarding signing certificate authorities (e.g., similar to what is provided by the SSL Observatory [8])

> Finding 10: IRNC ProNets have a position in the ecosystem as an aggregator of large amounts of traffic that gives that gives them a rare perspective. In undertaking any data-sharing activity, they will provide the most benefit when leveraging that position.

Returning to the subject of risks, given that IRNC ProNets are complex partnerships with international partners, discussions at the workshop made it clear that understanding their legal constraints and social expectations is a complex process. Coming to a thorough understanding of this process would require further investigation and may be unique to each IRNC ProNet.

Near the end of the workshop, participants were asked to set aside all concern regarding risks and enumerate a "wish list" of data and other services they would like from IRNC ProNets; the list is captured in Appendix C.

6 Conclusion

IRNC ProNets have a rare position in the network ecosystem to view a great deal of aggregate traffic that could be beneficial to both cybersecurity and researchers. In addition, their position could allow them to take active measures to defend their customers such as blocking denial of service attacks.

However, IRNC ProNets are complicated projects that represent partnerships between NSF, grantee institutions, carriers and international partners, with significant investments coming from those international partners (demonstrating a strong leverage of NSF funding). This means there are a variety of sensitivities to be addressed to undertake passive monitoring, either for the purpose of cybersecurity or to foster research.

In order to avoid operational issues, active cybersecurity measures need to be done with due care, transparency, and at the behest of the organizations and projects that constitute their customers. Additionally, IRNC ProNets provide high-speed, high-reliability network connectivity to serve the science community and must consider any changes in the context of how it impacts service to that community.

Future work on investigating these issues should increase the breadth of the participants to include greater representation from technical and legal representatives from IRNC ProNet international partners and the research and scientific communities. Increasing the depth of exploration could be done by selecting a specific IRNC ProNet and examining its specific challenges.

7 Organizer Comments

The following are post mortem comments from Von Welch as an organizer and are solely his opinions.

The strongest outcomes of the workshop are the findings related to what IRNC ProNets should do for cybersecurity and researchers out of a broad array of activities they could undertake. Specifically recognizing their rare position in the network ecosystem and the need to leverage that position in any activities was the most beneficial result, and undertaking active measures carefully and only at the request of their customers.

Turning to the organizational process of the workshop, we followed the model of the Internet2 workshops [2] and adopted a public forum. This caused only two changes that we are aware of: a decline from one participant to speak due to an organizational policy not to speak in public, and a request from another presenter to withhold their slides from the report due to similar sensitivities.

Invitations to the workshop were fairly successful in soliciting attendance, except that participation from academic researchers was disappointing. Reasons given by the declines were all due to schedule conflict or too much existing travel, but it must be admitted it is unclear if we were just unlucky or not compelling in our choice of topics.

Given that we were bringing together members of multiple communities (IRNC PIs, cybersecurity researchers, and practitioners), we felt background presentations were important to provide all participants with common understanding, but also felt sufficient time for open discussion was critical. We attempted to split workshop time roughly 50/50 between presentations and discussion. This balance seemed to work well. Presentations tended to run long, though this was often due to discussion they elicited, so we believe we stayed close to the split in spirit.

A speaker's dinner held the night before the workshop was not well attended since many speakers arrived too late to attend. However a dinner for the participants between the two days of the workshop was well attended and seemed to be well received.

8 References

1. Workshop for Cyber Security Data for Experimentation: http://www.gtisc.gatech.edu/nsf_workshop10.html
2. Workshop on Collaborative Data-Driven Security for High Performance Networks: http://security.internet2.edu/ddcsw/
3. Second Workshop on Collaborative Data-Driven Security for High Performance Networks: http://security.internet2.edu/ddcsw2/
4. NSF International Research Network Connections (IRNC) program. http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503382
5. International Research Network Connections. www.irnclinks.net
6. Secure Protocols for the Routing Infrastructure (SPRI). http://www.cyber.st.dhs.gov/spri/
7. Passive DNS Replication. Florian Weimer. April 2005. http://www.enyo.de/fw/software/dnslogger/first2005-paper.pdf
8. The EFF SSL Observatory. https://www.eff.org/observatory
9. D. Kaminsky. Staring into the Abysss. 2009. http://s3.amazonaws.com/dmk/Staring%20Into%20The%20Abyss.pdf
10. Camp, J., et al. Data for Cybersecurity Research: Process and "Wish List." June 2009. http://www.cc.gatech.edu/~feamster/papers/data-wishlist.pdf
11. Expanding Access to Research Data: Reconciling Risks and Opportunities. The National Academies Press. http://www.nap.edu/catalog.php?record_id=11434
12. Kenneally, Erin and Kimberly Claffy, "Dialing privacy and utility: A proposed data-sharing framework to advance Internet research", IEEE Security and Privacy, vol. 8, no. 4, pp. 31-39, July/Aug 2010. http://www.caida.org/publications/papers/2010/dialing_privacy_utility/

Appendix A Workshop Agenda

Day 1, February 22: What does "securing an IRNC" mean?

- 8:00am Continental Breakfast provided
- 9:00am Welcome, Rules of engagement, introductions, goals and non-goals. (Organizers)
- 9:30am Review of IRNC program and existing security activities (Kevin Thompson, NSF)
- 10:00am Network security today (Robin Sommer, Bro project)
- 10:30am Break
- 11:00am Robert Danforth
- 11:20am Experiences of other networks
    - ESNet (Brian Tierney, Michael Sinatra)
    - DANTE (Wayne Routly)
- Noon Lunch, provided
- 1:00pm Legal Issues (Erin Kenneally, CAIDA)
- 1:30pm Internet2 Cybersecurity Perspective (Robert Brammer, Internet2)
- 2:00pm Open Discussion (facilitated by organizers). Talking points:
    - What are our obligations here? What are the risks of doing nothing?
    - How would customers be impacted/feel about an IRNC doing DPI? Filtering? Logging?
    - What are the side effects and risks of the discussed applied security solutions?
    - How do IRNCs different from other network backbones?
- 3:00pm Break
- 3:30pm Open discussion continued.
- 4:30pm Summarize days discussions and conclusions
- 5:00pm Adjourn
- 6:00pm Workshop dinner, provided

Day 2: How can the IRNC networks foster relevant NSF (and other) research into security for IRNC and similar networks?

- 8:00am Continental Breakfast provided
- 9:00 am Further thoughts from Day 1 (Organizers)
- 9:30am Introduction to day 2 topic (Organizers; 30 minutes).
- 10:00am Cybersecurity Researcher Data Needs (Mike Bailey, U. of Michigan)
- 10:30am Break
- 11:00am Case studies for making network data available to researchers (20 minutes each)
    - CAIDA (Josh Polterock, CAIDA)
    - REN-ISAC Research Access to Shared Darknet Data (Doug Pearson, REN-ISAC)
    - Internet2 (Robert Brammer, Internet2)
- Noon Lunch, provided
- 1pm Open Discussion (facilitated by organizers). Talking points:
    - What are the unusual aspects of the IRNC networks that would bring benefit to network security research?
    - What are challenges for IRNCs to foster research and how have some other organizations overcome those?
    - What are some options for next steps?
- 3pm Break
- 3:30pm Summarize days discussions and conclusions (Organizers)
- 4pm Adjourn

Appendix B Suggested Reading

The following was suggested as preparatory reading for the workshop to the participants:

- The NSF IRNC projects: http://irnclinks.net/
- The Workshop on Cyber Security Data for Experimentation (CSDE) materials (http://www.gtisc.gatech.edu/nsf_workshop10.html), specifically:
    - Camp, J., et al. Data for Cybersecurity Research: Process and "Wish List" June, 2009. http://www.cc.gatech.edu/~feamster/papers/data-wishlist.pdf
    - Erin Kenneally 's CSDE presentation: http://www.gtisc.gatech.edu/files_nsf10/csde.ppt
- Report from the Internet2 Data Driven Collaborative Security Workshop (DDCSW) for High Performance Networks: http://security.internet2.edu/ddcsw/
- DETER: http://deter-project.org/
- PREDICT: http://www.caida.org/projects/predict/

Appendix C Data and Access Wish List

During the second day of discussions, participants were asked to set aside concerns regarding risks of releasing data and brainstorm on what sorts of data from and access to networks they would find useful if risk was not an issue. The list follows. Note that there was no attempt to gain consensus on or deeply explore any of the suggestions, they are simply recorded as made.

- If an IRNC ProNet were to support OpenFlow, having the ability to mirror of traffic, allowing one to see one's own traffic at a vantage point would be useful for diagnostics.
- Sharing of pDNS monitoring data with an Internet-wide database. Particularly, pDNS data from Asia is hard to come by.
- Data allowing for the comparison of IPv6 and IPv4 traffic. For example, how prevalent is each, performance differences. It was pointed out that researchers do not, in general, have access to much IPv6 data.
- Information about SSL usage such as collected by the EFF SSL observatory (https://www.eff.org/observatory).
- Detection and monitoring of malicious botnet command and control traffic.
- Data allowing for a comparison of dynamic circuits and static shared IP connections in terms of performance, overall usage, etc.
- Data allowing for measurement of usage by IRNC ProNet customers (e.g., science projects).
- Non-anonymous NetFlows for operational debugging..
- Increased metadata or a small amount of the application payload, e.g. DNS hostname in queries and replies, HTTP headers. (A request also put forward by [2], [10] and Finding 4 of [1].)
- Mitigating risks of data sharing includes restricting who data is shared with to "trusted researchers." Having a way to identify trusted researchers would be valuable.
- Data fingerprinting operating systems and applications [9].
- Providing a proving ground for hardware.
- Mapnet (http://www.caida.org/tools/visualization/mapnet/) for IRNC networks, which would need IRNC ProNet links and bandwidths[5].

[5] Added at suggestion of KC Claffy post-workshop, during review of the workshop report.

Appendix D Workshop Presentations

The workshop presentations follow. The presentation of Robert Danford is omitted by his request. The presentation of Wayne Routly has redaction of screen captures.

Security at the CyberBorder

Feb 22-23, 2012

Exploring the relationship between International Research Network Connections (IRNCs) and cybersecurity

INDIANA UNIVERSITY

INDIANA UNIVERSITY

Welcome

Bradley C. Wheeler

Vice President for Information Technology & CIO, Professor of Information Systems, IU Kelley School of Business

INDIANA UNIVERSITY

Thank you to NSF

This material is based upon work supported by the National Science Foundation under Grant No. 1158796. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

INDIANA UNIVERSITY

Thank you for logistics

Bridget Sutton

Dara Eckart

INDIANA UNIVERSITY

Introductions

Organizers:

- Doug Pearson (REN-ISAC)
- Brian Tierney (LBNL/ESnet)
- Von Welch (Indiana U./CACR)
- Jim Williams (Indiana U.)

Introduction of participants.

INDIANA UNIVERSITY

Ground Rules: Public Report

While the workshop is invite only, the product of the workshop will be a public report. Discussions at the workshop will be considered public and quotable. We will allow participants to preface remarks with a request for the remark to be kept out of the public report.

INDIANA UNIVERSITY

Workshop Notes

https://docs.google.com/a/vonwelch.com/document/d/1zgMXcD-bu3dXjAk9SgoG8Fc8phs5bH82hKYiG9RDo2I/edit

Or:

http://tinyurl.com/cyberborder

INDIANA UNIVERSITY

Workshop Report

- Organizers will capture findings in a report.
- Will present Findings before we leave
- Circulate draft in 1-2 week

INDIANA UNIVERSITY

Goals

- Day 1: What does "securing an IRNC" mean?
- Day 2: What can IRNCs do to foster cybersecurity (and networking) research?

INDIANA UNIVERSITY

Day 1: Non-Goals

- Security of networking equipment, links, other hardware
  Important! But out of scope.
- BGP, routing security
  Important! But of of scope
- Deep dives, Geopolitics, time burners…

INDIANA UNIVERSITY

Day 1 Goal - Refined

- What could/should IRNCs, given their unique perspectives, do for their global user communities?
  In terms of cybersecurity and privacy.
- How is an IRNC any different than any other backbone operator?
- What's practical versus research?

INDIANA UNIVERSITY

Agenda: Day 1, morning

- IRNC Overview
  Kevin Thompson (NSF)
- Network security
  Robin Sommer (ICSI, Bro project)
- Network security
  Robert Danford (Verizon)
- ESnet Experiences
  Brian Tierney, Michael Sinatra (ESnet)
- DANTE Experiences
  Wayne Routly (DANTE)
- Lunch

INDIANA UNIVERSITY

Agenda: Day 1, afternoon

- Legal issues
  Erin Kenneally (CAIDA)
- I2 Cybersecurity Perspective
  Robert Brammer (I2)
- Open Discussion
- 5pm Adjourn
- 6pm Workshop Dinner

INDIANA UNIVERSITY

Open Discussion

INDIANA UNIVERSITY

Day 1 Goal - Refined

- What could/should IRNCs, given their unique perspectives, do for their global user communities?
  In terms of cybersecurity and privacy.
- How is an IRNC any different than any other backbone operator?
- What's practical versus research?

INDIANA UNIVERSITY

Questions

- What can IRNCs do that sites cannot? What can they do better?
- Active blocking seems challenging given different site policies, is there a LCD?
- What are the unique challenges/ opportunities of IRNC vs standard bankbone operators?

INDIANA UNIVERSITY

Adjourn

- 6pm - Dinner @ Chancellors in Hotel
- Stop by registration for parking pass.
- Back here tomorrow
  8am Breakfast
  9am Start – Findings from today

Security at the Cyberborder

Day 2

February 22-23, 2012

INDIANA UNIVERSITY

INDIANA UNIVERSITY

Ground Rules: Public Report

While the workshop is invite only, the product of the workshop will be a public report. Discussions at the workshop will be considered public and quotable. We will allow participants to preface remarks with a request for the remark to be kept out of the public report.

INDIANA UNIVERSITY

Workshop Notes

https://docs.google.com/a/vonwelch.com/document/d/1zgMXcD-bu3dXjAk9SgoG8Fc8phs5bH82hKYiG9RDo2I/edit

Or:

http://tinyurl.com/cyberborder

INDIANA UNIVERSITY

Introductions

- Any new participants?

INDIANA UNIVERSITY

Any further Day 1 thoughts?

INDIANA UNIVERSITY

Day 1 Draft Findings

- DRAFT
- Did we get the important points?
- We're not wordsmithing now.

INDIANA UNIVERSITY

Day 1 Draft Findings

- IRNCs are part of a complex ecosystem, as part of that they can contribute to the cybersecurity process.
  - They are not a "silver bullet"
- It makes the most sense for IRNCs to contribute where they are especially qualified due to their unique perspective and location in the network

INDIANA UNIVERSITY

Day 1 Draft Findings

- IRNCs are partnerships between U.S. and International participants, and actions should be coordinated across the their International community.
- IRNC "users" are most directly NRENS, and networks such as ESnet and I2, but are also secondarily organizations, science projects, and researchers.

INDIANA UNIVERSITY

Day 1 Draft Findings

- IRNCs seem to be well poised to undertake Identification and Attribution of incidents:
  - Specifically:
    - passive DNS monitoring
    - tracing traffic to next hop, e.g. for DDOS

INDIANA UNIVERSITY

Day 1 Draft Findings

- IRNCs would seem to make good communication vehicles between ASs
- IRNCs may be well poised to foster best practices across the International community.

INDIANA UNIVERSITY

Day 1 Draft Findings

- In general, sharing of cybersecurity information is to be encouraged.
- IRNCs should undertake due diligence in privacy for collected monitoring data

INDIANA UNIVERSITY

Day 1 Draft Findings

- Sharing information in the larger community has the benefit of not only raising the boat but receiving information in return
  - E.g. sharing passive DNS monitoring data with ISC.
- Intelligence to customers should be filtered for relevance to based on customer interest.
  - Customer and IRNC may not agree on "bad"

INDIANA UNIVERSITY

Day 1 Draft Findings

- Any active mitigation should be done only at the request of customers
  - E.g. Blocking DDOS, spoofing, etc.
  - Sites may be doing things (e.g. Honeypots) that could easily trigger false positives.
- It would be beneficial if IRNCs undertook preparatory exercises for dealing with major incidents

INDIANA UNIVERSITY

Day 1 Draft Findings

- IRNCs may be poised to provide some services more efficient across their customer base.
  - Possible economy of scale, e.g., in accessing to netflow information
  - Leveraging of limited workforce

INDIANA UNIVERSITY

Day 1 Draft Findings

- IRNC's fortuitous view of aggregated netflow data could be very valuable to cybersecurity incident investigators
- IRNCs have strong international relationships which may make them valuable resources in International investigations

INDIANA UNIVERSITY

Day 1 Draft Findings

- IRNCs have a critical role, due to their funded mission, to play in providing a highly available end-to-end network
- IRNCs should work to ensure proper expectation management in what they provide in terms of cybersecurity

INDIANA UNIVERSITY

Day 2 Goal

What can IRNCs do to foster cybersecurity (and networking) research?

INDIANA UNIVERSITY

Day 2 Goal - refinement

- What sorts of data could IRNC provide?
- What sorts of access to the network?
- What would be useful?
- What are the challenges with doing so?

INDIANA UNIVERSITY

Agenda: Day 2, Morning

- Cybersecurity Researcher Data Needs
  Mike Bailey (U. of Michigan)
- CAIDA case study
  Josh Polterock (CAIDA)
- REN-ISAC Shared Darknet Data
  Doug Pearson (REN-ISAC)
- Internet2 case study
  Robert Brammer (Internet2)

INDIANA UNIVERSITY

Agenda: Day 2, Afternoon

- Lunch
- Open Discussion
- 3pm Break
- 3:30pm Wrap up
- 4pm Adjourn

INDIANA UNIVERSITY

Open Discussion

- What sorts of research data are IRNC in a unique position to provide?
- What lessons learned/services could IRNC leverage from the community in providing research data?
- How does one measure the ROI in providing research data?

INDIANA UNIVERSITY

New Rules

- One statement from Carter
- And then we are “Monarchs of the World” and can share any data we want… what do we do?

INDIANA UNIVERSITY

Workshop Report next steps

- Look for draft report in your inbox in the next 1-2 weeks.
- Will allow a week for comments.

INDIANA UNIVERSITY

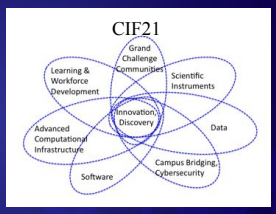
Thank you

NSF's International Research Network Connections (IRNC) Program

Kevin Thompson
NSF Office of CyberInfrastructure
February 2012

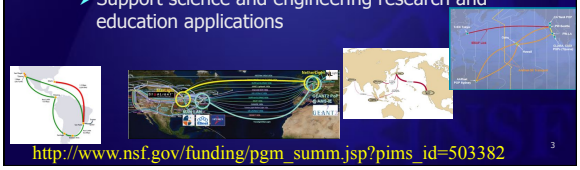
Office of Cyberinfrastructure

- HPC/ACI
- Software
- Data
- Campus Bridging/ Networking
- Learning and Workforce Development
- VOs and building Research Communities



International Research Network Connections (IRNC)

- Goals:
  - Provide network connections linking U.S. research with peer networks in other parts of the world
  - Stimulate the deployment and operational understanding of emerging network technology and standards in an international context
  - Support science and engineering research and education applications



http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503382

IRNC

- Increase capacity and reach into new and existing regions
- Strive to support to the scientist/engineer/educator/ student by targeting an end-to-end approach in appropriate activities
- Maintain and build strong international relationships with colleagues and partners
- Leverage resources (capacity, tools, ideas) where feasible
- Collaborate closely with U.S. domestic R&E network providers and leaders
- Technical emphasis areas include measurement/ monitoring
- **www.irnclinks.net**

IRNC 2010 Awards

- Three funded areas
  - Production Networking
  - Experimental Networking
  - Special Projects
- ~$32.5M over 5 years
- Awards in FY10

http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503382

IRNC 2010-2014: Production Networking ("ProNet")

- 5 awards over 5 years
- Continue primary goal to connect US researchers with broad communities
  - Multi-Gbps connections
  - Balancing current/projected needs with limited funding availability
- Continue to grow, emphasize, leverage strong partnerships with regional and continental r&e network efforts
  - Opportunities for new connections
  - Shared IP and Hybrid networking as production services., Dynamic circuit services
- Improving quality of network operations and engineering while promoting interoperability
  - Advanced measurement
  - Platform for early stage deployment of new technologies

IRNC: Experimental Networking (Exp)

- Experimental networking area
  - Activities anchored to identified science application drivers
- Integration and End-to-end networking
  - Scientific Apps and distributed environments
  - Connectivity to unique instruments
  - Seed money for network support for eScience applications
- 1 award (PI Tom DeFanti) related to StarLight
- See www.irnclinks.net

IRNC: Special Projects (SP)

- Education/outreach and training (NSRC)
- Dynamic circuit services (DyGir)
- Measurement (IRIS, CAIDA, Routeviews)
- More details at www.irnclinks.net

ACE: American Connects to Europe

PI: Jim Williams, Indiana Univ.

Partners: DANTE, NYSERNet, Internet2, DLT, and others

- Multiple 10G connections - US and Europe

A short ACE overview

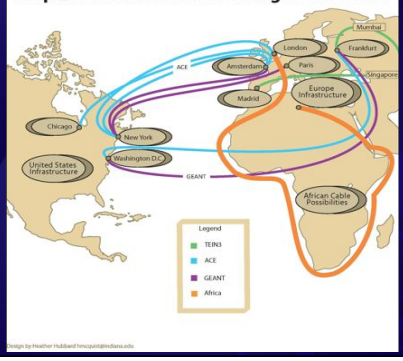
Cooperative partnership among Indiana University, DANTE, NYSERNet and others

60Gbps deployed between EU-US

2x10Gbps	**AMS-NYC**	**bonded, shared IP**
2x10Gbps	**FKK-WIX (DC)**	**bonded, shared IP**
1x10Gbps	**AMS-CHI**	**ptp**
1x10Gbps	**PAR-NYC**	**ptp**

International Connectivity

http://internationalnetworking.indiana.edu



TransPAC3: Connecting Asia and the US

PI: Jim Williams, Indiana Univ.

Partners: APAN, TEIN3, NICT-Japan, NII-Japan, CERNET – China, DLT, others

- Multiple 10G connections from the US to Asia
- Connections continued to South East Asia and South Asia
- Through EU-supported TEIN3 network, provides connectivity to 19 countries

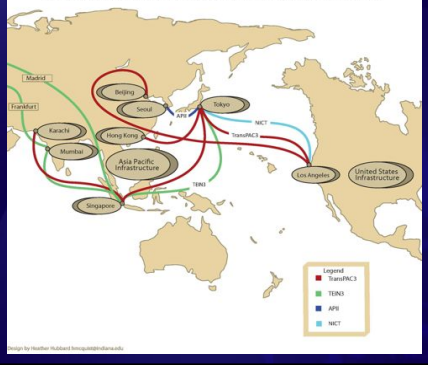
A short TP3 overview

Cooperative partnership among Indiana University, APAN, CERNET, NICT-Japan and others

30Gbps deployed between US and Asia

10Gbps	**LA-Tokyo**	**NSF**
10Gbps	**Tokyo-LA**	**NICT**
10Gbps	**LA-Beijing**	**CERNET+NSF**
155Mbps	**SG-Karachi**	**PK+NSF**
Various	**Various**	**TEIN3-EC (soon to change)**

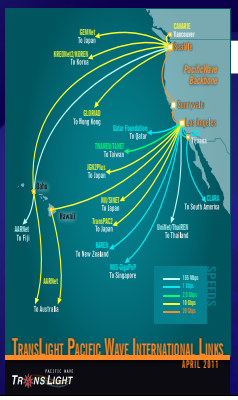
International Connectivity
http://internationalnetworking.indiana.edu



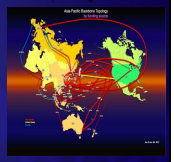
TransLight/PacificWave

PI : David Lassner, University of Hawaii
Partners: AARNet & others

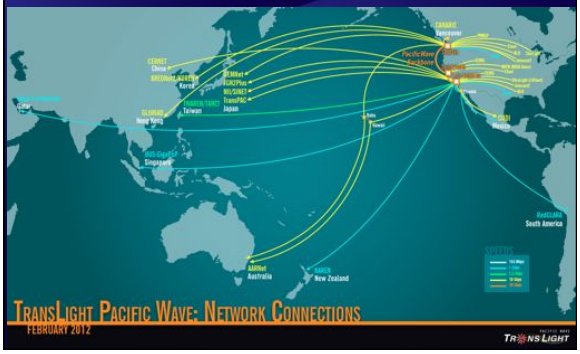
- Two 10Gbps circuits from Australia to Hawaii (including Mauna Kea) to the West Coast
- Support numerous additional international R&E circuits that terminate at West Coast
- Ensure avail. of new services, incl. monitoring infrastructure and dynamic circuit services
- Install and maintain new international links, including emerging links from Pacific Islands, and bandwidth upgrades to existing links as funding and feasibility permit



Asia-Pacific Backbone Topology by Funding Source



TL/PW Feb 2012

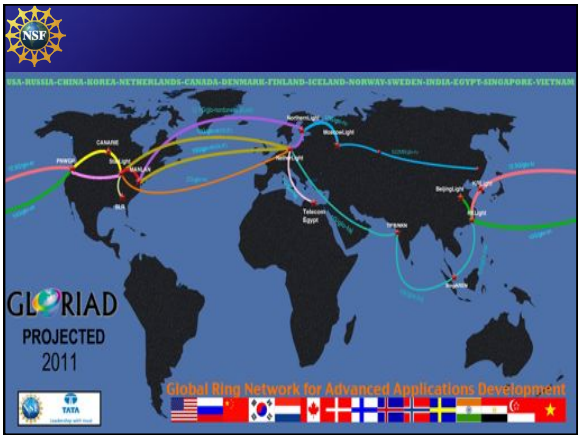


Americas Lightpaths (AmLight): Linking the Americas and Beyond

- AMLIGHT East:
  - Miami-Sao Paulo
    - 2 x 10G connections
- AMLIGHT West:
  - Tijuana-San Diego-Los Angeles
    - 10G connection
- AMLIGHT Andes:
  - Sao Paulo-Santiago:
    - 1G connections
- AMLIGHT Central:
  - San Antonio-Mexico City:
    - 1G connection
- Linking R&E networks in the Americas via key exchange points

Global Ring Network for Advanced Applications Development (GLORIAD)

- PI: Greg Cole, University Tennessee
- Int'l Partners: CNIC; KISTI; SURFnet; NORDUNet, e-ARENA, CANARIE, ENSTINET, Tata, SingAREN, VINAREN
- US to: Russia (to 10G), China, Korea (to 12.5G), Canada, Netherlands, Nordic countries (to 20G)
- New capacity/services in SE Asia, Africa, Middle East, Caucasus and Central Asia
- Deploy improved tools, standards, procedures for distributed, decentralized global operations via GLORIAD's "dvNOC"



SDCI Awards from FY11

- #1127210 – "Distributed Web Security for Science Gateways" – Jim Basney (UIUC)
- #1127500 – "Metadata Management Software Tools to support R&D of Cyberinfrastructure" – kc claffy (UCSD) and Minaxi Gupta (Indiana)
- #1127361 – "New Software Platforms for Supporting Network-wide Detection of Code Injection Attacks" – Fabian Monrose (UNC)
- #1127506 – "Transport-Layer Abusive Traffic Detection and Mitigation" – Rob Beverly (NPS)
- #1127388 – "Traffic Modeling and Generation with Custom Fidelity for Cyber Security Experimentation" – Jelena Mirkovic (USC)
- #1127425 – "SESv3 (Security Event System – Version 3)" – Doug Pearson (Indiana)
- #1127195 – "Passive and Active DNS Monitoring Tools for Detecting and Tracking the Evolution of Malicious Domain Names" – Roberto Perdisci (Georgia)
- #1127396 – "Integrated Simulation and Emulation Platform for Security Experimentation" – Yuan Xue (Vanderbilt)

Backup SDCI talk/slides

NSF's Cyber Security Awards FY11 in the Software Development for Cyberinfrastructure (SDCI) program

SDCI End-to-end High Performance Networking Focus

- **Measurement and monitoring tools and services: supporting network research and operations. Includes both active and passive measurement**
- **Network protocols: includes alternatives/mods to TCP, native IPv6 application support, transport protocols supporting 10's Gbps, and interoperability with new protocols**
- **Network software for Sensor and wireless networks**

SDCI Cyber Security Focus

- Malware detection and prevention: areas include polymorphism, countering obfuscation, command&control
- Situational understanding: includes varying time scales, attack attribution, cross-boundary data sharing, and handling massive data.
- Assurance techniques in sharing data: anonymization and protections for data collected, in transit, and at rest
- Software assurance: tools and services for software testing and vulnerability analysis
- HPC resources and distributed computing: secure access to, leverage of, and improving security of existing compute resources.
- Research infrastructure: supporting cyber security r&d

SDCI Proposal Requirements

- Partnering with a scientific research project (users) and/or deployment environment (CI infrastructure)
- Working with an existing software base or initial prototype capability of some kind
- Use of NSF-funded software build-and-test services or alternative
- Application to NSF community, in terms of user identification, missing/needed capability, and impact
- Project plan including release, deployment info and eval plan
- open source license to be used
- Collaborations with industry encouraged, not required
- "Strong preference will be shown for efforts that provide near-term benefit to a broad user base in the NSF community"

Awards from FY11

- 8 projects
- Award activities 3 years in duration
- Total funding ~$5M
- Award abstracts available at www.fastlane.gov

#1127210 – Distributed Web Security for Science Gateways

- PI: James Basney, UIUC
- Provides authz and delegation s/w for science gateways
  - Oauth server implementation
  - Client libraries and authN modules
  - Full integration with popular gateways addressing common use cases

#1127500/1127406 – Metadata Management Software Tools to Support Research and Development of Cyberinfrastructure

- PI: Kim Claffy (Lead), SDSC
  - Minaxi Gupta, Indiana
- Integrates IMDC and DatCat
- Expand capabilities to support cyber security r&d
  - Improving AUP-free access to data
  - Automated indexing and easier metadata annotation
  - Starting emphasis on DNS tld zone files

#1127361 – New Software Platforms for Supporting Network-wide Detection of Code Injection Attacks

- PI: Fabian Monrose, UNC
- Design/implement/deploy ShellOS prototype
- Essentially a h/w sandbox for malware injection and shell code execution.
- Continuously analyze network streams and program buffers for harmful executable code
- h/w virtualization based allowing direct execution of instruction sequences on the cpu
- Planned deployment on UNC campus
- See recent Usenix paper/presentation on ShellOS

#1127506 – Transport-Layer Abusive Traffic Detection and Mitigation

- PI: Rob Beverly, NPS
- Derived from previous r&d on TCP-layer anomalies and signal analysis
- operational prototyping to detect bots and botnet infrastructure using content agnostic traffic signal characterization methods
- Transport traffic classifier to be built and deployed made available initially as a spam assassin plugin

#1127388 – Traffic Modeling and Generation with Custom Fidelity for Cyber Security Experimentation

- PI: Jelena Mirkovic, USC/ISI
- Towards a more realistic traffic generation capability to support cyber defense testing in network testbeds
- Building a traffic generator that
  - Mines values for user-specified dimensions of traffic logs
  - Generates random traffic fitting the mined model
  - Replays traffic from a log matching user-specified dimensions
- Explores user customized definitions of realism
- Integration into DETER

#1127425 – SESv3 (Security Event System – Version 3)

- PI: Doug Pearson, Indiana Univ.
- SES collects aggregated security event information among partners in REN-ISAC, correlates, and provides high confidence threat intelligence back
- SESv3 adds BGP, passive DNS, and free form data types e.g. email, IRC, tweets
- Redesign of underlying data repository and system architecture for scaling
- SES access integrated into common incident analyst and responder tools

#1127195 – Passive and Active DNS Monitoring Tools for Detecting and Tracking the Evolution of Malicious Domain Names

- PI: Roberto Perdisci, Univ. of Georgia
- Develop tools to detect and track malicious domains as they evolve in time.
  - FlexBuster – detect fast flux networks "in the wild" through monitoring points from ISC/SIE
  - PassTrack and Digger – detect and track malware-related domain names
- Industrial partner Damballa
- As with all other awards, tools released as open source

#1127396 – Integrated Simulation and Emulation Platform for Security Experimentation

- PI: Yuan Xue, Vanderbilt
- iSEE platform for **CPS** modeling, evaluation, experimentation
  - Modeling environment for system specification and experiment config
  - Run-time environment for experiment execution
  - Component problems – sim/emu time synch, domain-specific modeling languages, runtime monitoring instruments
- Integration with DETER

Network Security Today

Robin Sommer
International Computer Science Institute, &
Lawrence Berkeley National Laboratory

robin@icsi.berkeley.edu
http://www.icir.org/robin

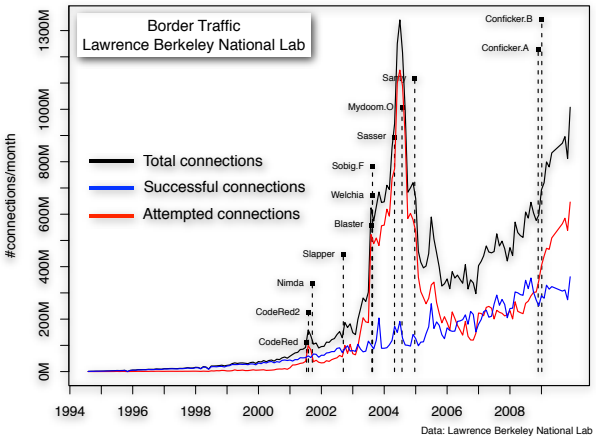
Security at the CyberBorder
February 2012, Indiana University

Outline

Part 1: Today's Network Threats.

Part 2: Defender Strategies.

The Old Days ...



Part 1: Today's Threats

Trend 1: Commercialization of Attacks

Trend 2: Highly Targeted Attacks

Trend 3: Insider Attacks

Trend 1: Commercialization of Attacks

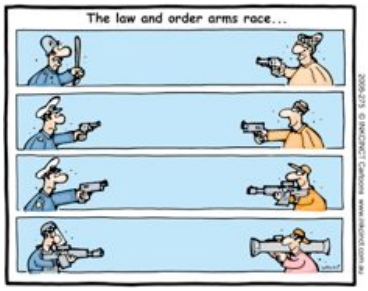
Attacks aimed at making a profit.
- Selling (illegal) goods and services.
- Exfiltrate information.

Thriving underground economy.
- Empowered by virtually endless supply of "bots".
- Everything is on sale ("crime-as-a-service").

"Pay Per Install" Services

Crime Economics

Accelerated arms race.
Innovative, fast moving attackers.



Bear race.
If attack pays, it's good enough.

Trend 2: Highly Targeted Attacks

High-skill / high-resource attacks.
- Targeting **you.**
- Extremely hard to defend against.
- Attribution virtually impossible.

Typical Instances
- "Advanced Persistent Threats".
- Activist hacking.

Targeted Attacks: APTs

Open Letter to RSA Customers

Like any large company, EMC experiences and successfully repels multiple cyber attacks on its IT infrastructure every day. Recently, our security systems identified an extremely sophisticated cyber attack in progress being mounted against RSA. We took a variety of aggressive measures against the threat to protect our business and our customers, including further hardening of our IT infrastructure. We also immediately began an extensive investigation of the attack and are working closely with the appropriate authorities.

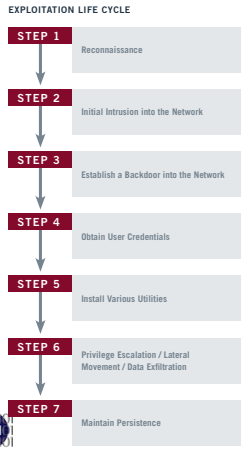
Arthur W. Coviello, J

Source: RSA

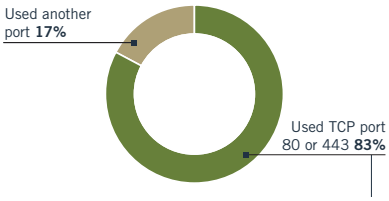
Advanced Persistent Threat (APT). MANDIANT defines the APT as a group of sophisticated, determined and coordinated attackers that have been systematically compromising U.S. government and commercial computer networks for years. The vast majority of

Source: MANDIANT

Targeted Attacks: APTs (2)



APT MALWARE COMMUNICATION
100% of APT backdoors made only outbound connections



In no instance was any APT malware written or configured to listen for inbound connections.

Source: MANDIANT

Targeted Attacks: Activist Hacking

Source: Wikipedia

Defender Strategies

Challenges

Varying threat models.
No ring rules them all.

Semantic complexity.
The action is really at the application-layer.

Volume and variability.
Network traffic is an enormous haystack.

Legal and ethical frameworks.
Not everything you can do, you may.

Defender Strategies

Creating visibility.
Instrument the network comprehensively.

Analyze semantics.
Not bytes.

Share intelligence.
"The good guys share, too!"

Active response.
Blacklist, or whitelist, what you know.

Creating Visibility with Bro

```
> bro -i en0
[ ... wait ...]
> cat conn.log
```

```
#fields ts        id.orig_h        id.orig_p  id.resp_h        id.resp_p proto  service  duration
1144876741.1198  192.150.186.169 53115      82.94.237.218    80        tcp    http     16.14929
1144876612.6063  192.150.186.169 53090      198.189.255.82   80        tcp    http     4.437460
```

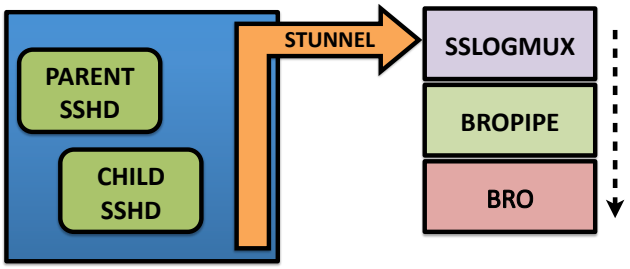
```
[...] host              uri                  status_code  user_agent [...]
      docs.python.org   /lib/lib.css         200          Mozilla/5.0
      docs.python.org   /icons/previous.png  304          Mozilla/5.0
      docs.python.org   /lib/lib.html        200          Mozilla/5.0
      docs.python.org   /icons/up.png        304          Mozilla/5.0
      docs.python.org   /icons/next.png      304          Mozilla/5.0
      docs.python.org   /icons/contents.png  304          Mozilla/5.0
      docs.python.org   /icons/modules.png   304          Mozilla/5.0
      docs.python.org   /icons/index.png     304          Mozilla/5.0
       www.google.com   /                    200          Mozilla/5.0
```

```
1144876742.3338  192.150.186.169  53116      docs.python.org  /icons/index.png  304  Mozilla/5.0
1144876745.6144  192.150.186.169  53117      www.google.com   /                 200  Mozilla/5.0
```


Creating Visibility: Encryption

NERSC

"Auditing SSHD"



Source: Scott Campbell / NERSC

NERSC Computer Use Policies Form

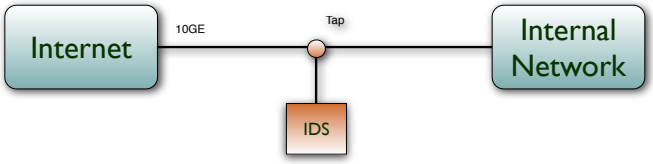
Monitoring and Privacy

Users have no explicit or implicit expectation of privacy. NERSC retains the right to monitor the content of all activities on NERSC systems and networks and access any computer files without prior knowledge or consent of users, senders or recipients. NERSC may retain copies of any network traffic, computer files or messages indefinitely without prior knowledge or consent.

The Security Fence

Cartoon Courtesy Clay Bennett / The Christian Science Monitor

Analyzing Semantics

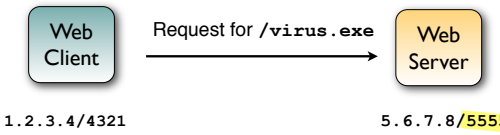


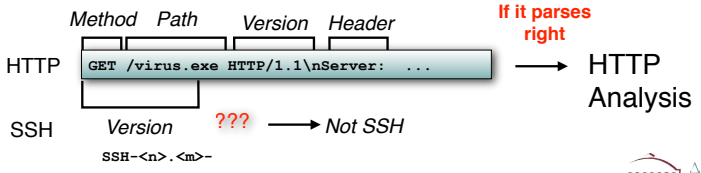
Example: Finding downloads of known malware.

1. Find and parse all Web traffic.
2. Find and extract binaries.
3. Compute hash and compare with database.
4. Report, and potentially kill, if found.

Port-independent Application Analysis

Bro's Dynamic Protocol Detection





Identifying HTTP Servers

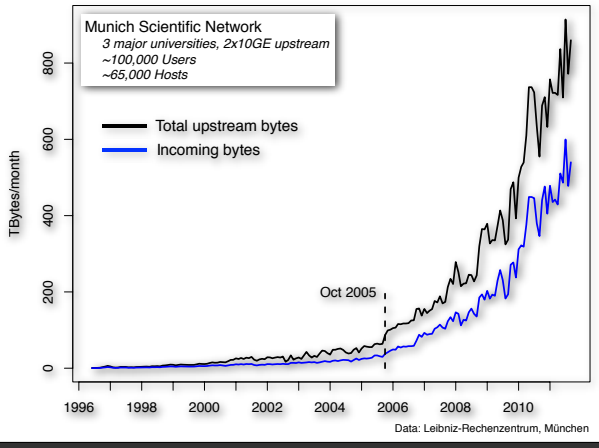
Server Addresses

```
a198-189-255-200.deploy.akamaitechnolgies.com
a198-189-255-216.deploy.akamaitechnolgies.com
a198-189-255-217.deploy.akamaitechnolgies.com
a198-189-255-230.deploy.akamaitechnolgies.com
a198-189-255-225.deploy.akamaitechnolgies.com
a198-189-255-206.deploy.akamaitechnolgies.com
a198-189-255-201.deploy.akamaitechnolgies.com
a198-189-255-223.deploy.akamaitechnolgies.com
                72.21.91.19
a198-189-255-208.deploy.akamaitechnolgies.com
a198-189-255-207.deploy.akamaitechnolgies.com
          nuq04s07-in-f27.1e100.net
 a184-28-157-55.deploy.akamaitechnologies.com
a198-189-255-224.deploy.akamaitechnolgies.com
a198-189-255-209.deploy.akamaitechnolgies.com
a198-189-255-222.deploy.akamaitechnolgies.com
a198-189-255-214.deploy.akamaitechnolgies.com
          nuq04s06-in-f27.1e100.net
       upload-lb.pmtpa.wikimedia.org
          nuq04s08-in-f27.1e100.net
```

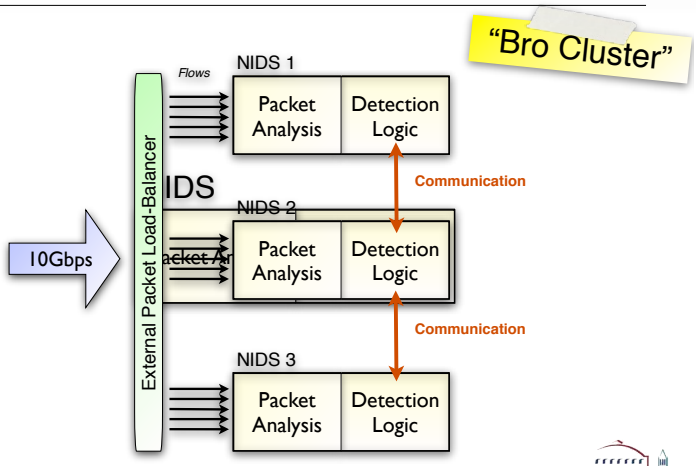
HTTP Host Headers

```
ad.doubleclick.net
ad.yieldmanager.com
b.scorecardresearch.com
clients1.google.com
googleads.g.doubleclick.net
graphics8.nytimes.com
l.yimg.com
liveupdate.symanteclive update.com
mt0.google.com
pixel.quantserve.com
platform.twitter.com
profile.ak.fbcdn.net
s0.2mdn.net
safebrowsing-cache.google.com
static.ak.fbcdn.net
swcdn.apple.com
upload.wikimedia.org
www.facebook.com
www.google-analytics.com
www.google.com
```


Performance Challenges

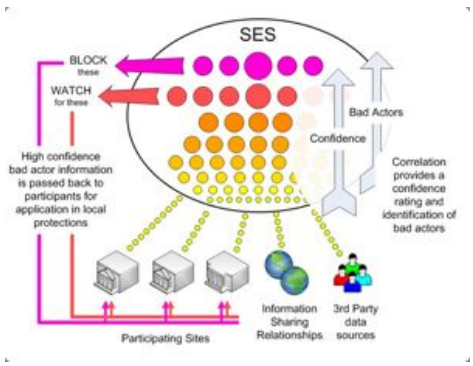


Load-balancing Architecture



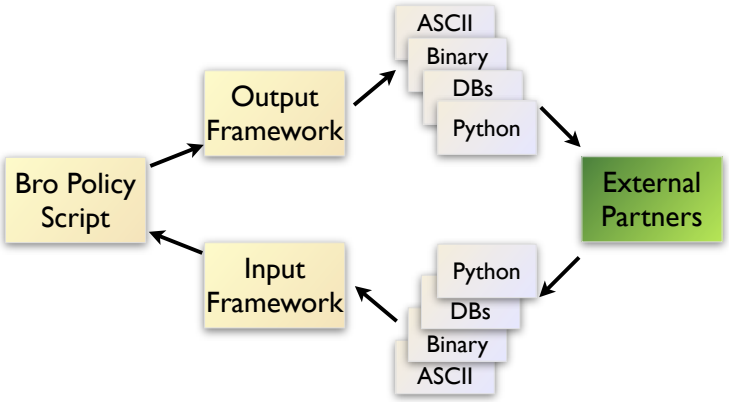
Sharing Intelligence

REN-ISAC's Security Event System



Source: REN-ISACs

Real-time Intelligence with Bro



Active Response

Interface to the network layer.
- Kill sessions.
- Block hosts (local, remote).
- Block applications (static, dynamically).

The Extreme: White-list activity.
- Destinations, applications, services.
- Technically challenging, need full proxy.

Conclusion

Today's Threats
- Commercialization of Attacks
- Highly Targeted Attacks
- Insider Attacks

Defender Strategies
- Creating visibility.
- Analyze semantics.
- Share intelligence.
- Active response.

Understand semantics, put activity into broader context, react, and share your knowledge.

Thanks for your attention.

Robin Sommer
International Computer Science Institute, &
Lawrence Berkeley National Laboratory

robin@icsi.berkeley.edu
http://www.icir.org/robin

Security at the CyberBorder
February 2012, Indiana University

ESnet and Cybersecurity

Brian Tierney, Michael Sinatra, William Johnston

Security at the Cyberborder Workshop
Indianapolis, IN
February 22, 2011

U.S. DEPARTMENT OF ENERGY Office of Science
BERKELEY LAB

Outline

1. ESnet Overview
2. ESnet Cybersecurity Policy
3. Example DDOS Incident and Response
4. Threat Analysis for International Interconnects

ESnet is a **Mission Organization**

DOE Office of Science mission: to deliver scientific discoveries and major scientific tools for transforming our understanding of nature, and to advance the energy security, economic security, and national security of the United States.

ESnet mission: to accelerate scientific discovery for the DOE Office of Science by delivering unparalleled network infrastructure, services, tools, and innovation.

ESnet **Supports DOE Office of Science**

SC provides broad support for Labs, Facilities, people

- almost $5B/year in funding
- 45% of Federal support for physical sciences research
- key funding for basic research in biology, computing, energy, climate
- over 100 Nobel Prizes in past 60 years

Supporting > 27,000 PhDs, grad students, engineers at >300 institutions.

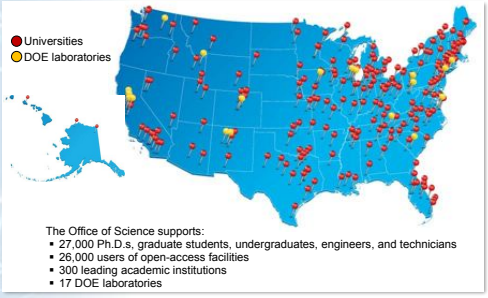
Provides world's largest collection (32) of scientific user facilities:

- supercomputer centers, accelerators, light sources, neutron sources, electron microscopes, nano-scale centers, a sequencing center, fusion facilities.

ESnet is one of them – connecting DOE sites, facilities, scientists and collaborators.

- optimized for science data transport
- every service exists to support scientific discovery

ESnet **Supports DOE Office of Science**



The Office of Science supports:
- 27,000 Ph.D.s, graduate students, undergraduates, engineers, and technicians
- 26,000 users of open-access facilities
- 300 leading academic institutions
- 17 DOE laboratories

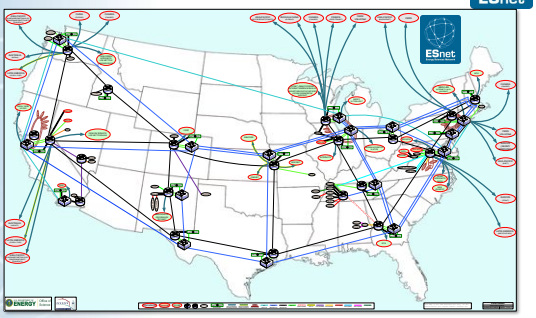
ESnet is a Division of Berkeley Lab

Recognition of strategic importance of networking. ESnet also has easier access to scientific leadership.

Oppenheimer, Seaborg, Lawrence 1946

- multi-program, strong in many fields
- 12 Nobel Prizes, 55 Laureates trained or had significant collaborations at LBNL
- 13 National Medal of Science winners
- 4200 Employees, ~$800M budget
- in the process of designing a major second campus

ESnet Topology (n x 10G core)



ESnet and **Cyber-Security**

In a 'big data' environment, cyber-security is a challenge.

- common techniques: firewall avoidance, intrusion detection, real-time injection of ACLs or null routes, instrumented sshd, OTP

Labs and sites are responsible for their own security, per DOE policy.

- However ESnet plays an active role, accommodating sites however it can, during events.
- ESnet is extremely cautious about its own infrastructure.
- Our challenge: secure the commons while supporting open science.

Michael Sinatra, ESnet

ESnet Policy on DDOS attacks

ESnet Policy (DDoS)

- ESnet doesn't provide *day to day* site security.
- But ESnet can and will intervene in a DDoS situation, where a site may need help.
- Several mechanisms in place to help mitigate DDoS attacks.
  - ability to selectively withdraw routes
  - blackhole routes inside and outside of ESnet
  - separate from commodity peers—selectively or wholesale
- ESnet also has ability to monitor netflow take packet traces and perform other types of monitoring and mitigation.
- ESnet NOC and engineering staff are available 24x7 to deal with issues.

ESnet Response to DDOS attack

ESnet Maintaining Science Mission Critical Infrastructure in the Face of a Cyberattack

ESnet has tools in place to allow ESnet engineers to respond to extraordinary situations

The tools involve things like

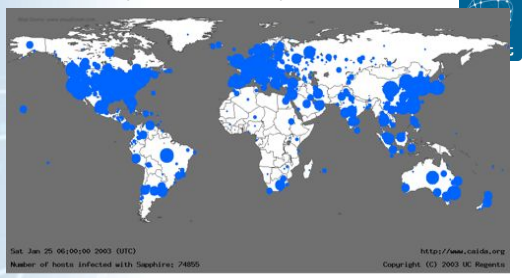
- Separating ESnet core routing functionality from external commercial Internet connections by means of a "peering" router that can have a policy different from the core routers
- Reducing the priority of traffic from the Internet so that it will not compete with site-to-site traffic
- Applying filter based signatures to incoming Internet traffic depending on the nature of the attack

One example of an "extraordinary situation" is the large-scale denial-of-service attack that occurred in early 2003

ESnet engineers used a phased response that ranged from blocking the "bad" traffic at the ingress points to blocking infected sites as they (autonomously) start to attack other ESnet sites

- A complete isolation of ESnet from the commercial Internet is possible

Historical Example of ESnet Response to DDOS Attack

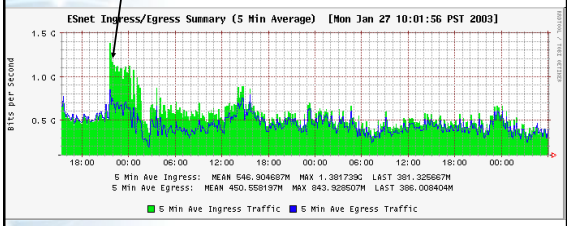


30 minutes after the Sapphire/Slammer worm was released, 75,000 hosts running Microsoft's SQL Server (port 1434) were infected.

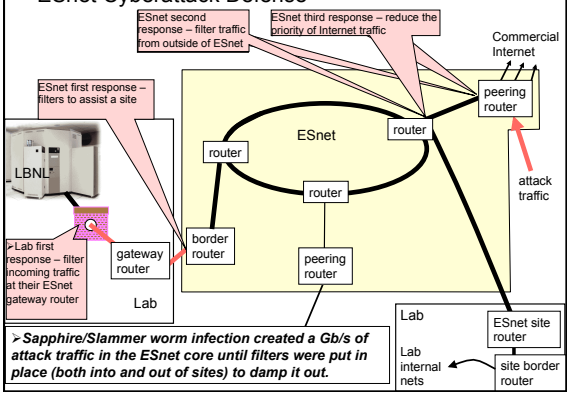
(From "The Spread of the Sapphire/Slammer Worm," David Moore (CAIDA & UCSD CSE), Vern Paxson (ICIR & LBNL), Stefan Savage (UCSD CSE), Colleen Shannon (CAIDA), Stuart Staniford (Silicon Defense), and Nicholas Weaver (Silicon Defense & UC Berkeley EECS). Jan., 2003. (http://www.caida.org/publications/papers/2003/sapphire/))

Historical Example of ESnet Response to DDOS Attack

Sapphire/Slammer worm infection hits creating almost a full Gb/s (1000 megabit/sec.) traffic spike on the ESnet backbone



ESnet Cyberattack Defense



➢ ***Sapphire/Slammer worm infection created a Gb/s of attack traffic in the ESnet core until filters were put in place (both into and out of sites) to damp it out.***

Threat Analysis for International Interconnects

Possible Security Issues

A general risk is created for research and education (R&E) institutions

- High-speed access from areas of the world that have not previously had such access

Threats may arise from those who may view IRNCs as a new "opportunity"

- Organized crime groups in, e.g., the former Soviet Union
- Ideologically motivated groups
- Mischief makers worldwide

Possible Security Issues

Risks and their mitigation

- Risks (*as perceived by policy makers*)
  - Targeted attacks by individual adversaries who are skilled and motivated
- Mitigation
  - *Opportunities for targeted attack are largely unchanged by IRNCs*
    - they are typically not bandwidth dependent
    - most of the areas served by IRNCs already have access to R&E institutions over low and medium speed networks
  - R&E institutions with sensitive code and data already have policy and procedures in place to manage access
    - *This is a site responsibility, not the responsibility of the WAN networks*
    - Even in a mission agency network like ESnet there is considerable variation of site policy that governs preparedness and response which argues against WAN involvement

IRNC – Possible Security Issues

- Risks
  - High-speed distributed denial of service (DDOS) attacks cutting off network access to R&E institutions
- Mitigation
  - This is the responsibility of WAN networks – sites cannot protect their network connectivity from distributed denial of service attacks
  - Detect when DDOS attacks are happening
    - Intrusion detection systems should be run on the WAN backbone and at the peering routers (ingress/egress points)
    - Both US DOE and US DHS are funding work in high-speed intrusion detection (the Bro system)
    - More simple (but nonetheless sophisticated) threshold monitoring can also help (e.g. Arbor)
    - Provide selective filtering of routed IP traffic
    - A traffic control architecture should be in place to deal with DDOS and worm related attacks to protect sites
      - see ESnet Phased Response Security Architecture, below

IRNC – Possible Security Issues

- Risks
  - High-speed intrusion into R&E institutions by malicious agents
  - Provisioned circuits (VLANs, MPLS paths, lightpaths, etc.) can reach inside of sites, bypassing normal site protection
- Mitigation
  - This is indirectly the responsibility of WAN networks because the WANs provide the mechanism of attack (virtual circuit services)
  - Provide virtual circuit setup mechanisms (management plane) that use strong endpoint authentication when establishing circuits
  - Facilitate the use of strong authentication
    - For example, WAN organizations can support high quality X.509 identity certificate issuing authorities (CAs) that have science oriented policy - e.g. the DOE Grids CA (doegrids.org), European NREN CAs, or similar CAs

What are some other risks?

- Attacks on the critical Internet infrastructure (e.g. root DNS servers)
  - ESnet (and probably CPS) peer with root DNS servers; so do non-US R&E networks.
  - How do we avoid being a vehicle for attack (or what can we do to help the rest of the Internet cope?)?
- Route/traffic hijacking
  - We have played a leadership role in the past with DNSSEC—at least in terms of helping sites deploy it. Time to do the same with RPKI and BGPSEC?
- How does reliance on "cloud" or probably more accurately, outsourced network services affect our ability to react?

Additional thoughts on opportunities (for tomorrow)

- Possibility of Bro and/or netflow at IRNC peering points
  - Both raise a lot of questions. How much packet inspection to do? How much netflow to get (versions, sampled)?
- Passive DNS sensors within IRNC peering points (and importance of botnets)
- Routing data?
- ➔ Need to distinguish data which is beneficial for operational security vs. basic research. Need to define scopes!

Questions?

Thanks You.

Brian Tierney, bltierney@es.net
Michael Sinatra, ms@es.net
http://www.es.net/
http://fasterdata.es.net/

U.S. DEPARTMENT OF ENERGY Office of Science
BERKELEY LAB

GEANT – Protecting The Network

Wayne Routly, DANTE
Security @ Cyber Border Meeting
Indianapolis.us
22 February, 2012

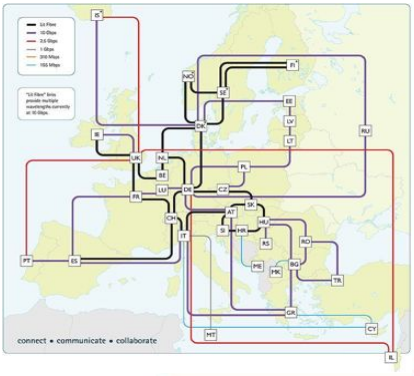
Contents

- GEANT : Who What How
- What Are The Issues
  - Protecting GEANT Users
  - A Security Conundrum
- Security is a Multi Faceted Approach
  - GEANT Way
  - NSHaRP System
- Looking Ahead
  - What Do We Want
  - What Do We Not Want
- Q & A

Why are we doing this?
What problem are we solving?
Is this actually useful?
Are we adding value?
Will this change behavior?
Is there an easier way?
What's the opportunity cost?
Is it really worth it?

'GEANT : Who What How'

- State of the Art Pan-European Network
  - …..Transit Network….ISP
  - 18 Physical Pops
- 40 GB links -> 100GB
  - TB of Data shifted
  - 10 million+ IPs
  - 100 Workstations
  - Unusual Traffic
- Truly Global
  - Interconnects – 5+
  - NRENs - 38
  - Commercial & Commodity Traffic - 8/2



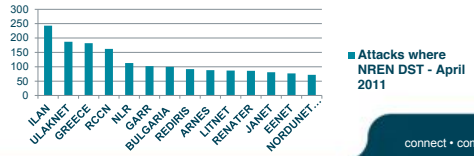
'Protecting GEANT Users'

- In an age of ever faster networks, greater connectivity, networks and users are under even greater risk of attack.
- Network Based Attacks
  - Paypal, VISA
  - Amazon
  - Wikileaks
  - Political
- Maintain service levels by proactively monitoring and mitigating against potential attacks?

'A Security Conundrum'

- How do we notify potential victims / sources & assist in solving those incidents for dozens of situations?

- How do we double the number of tickets we can handle without doubling the manpower? Automate it?



'Security is a Multi Faceted Approach'

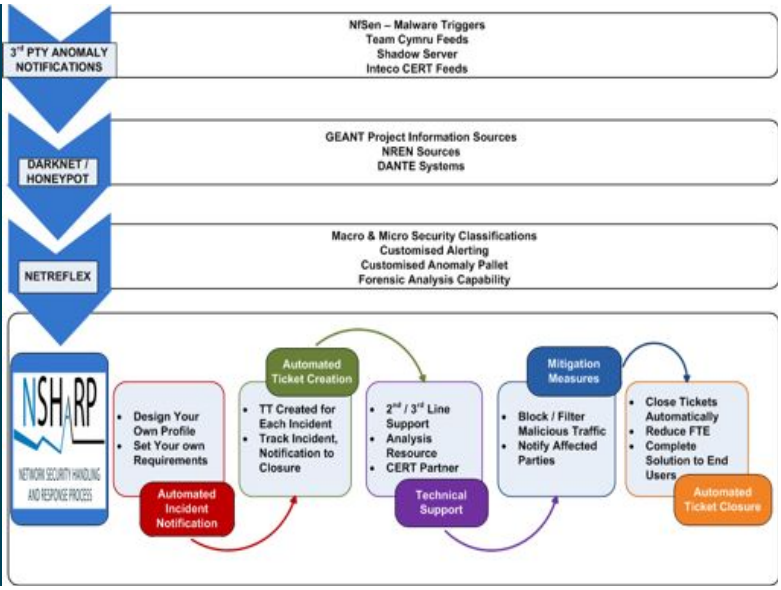
- Backbone
  - Broken into Zones
  - Generic Firewall Templates
  - VM / Radius
- The Solutions
  - Netreflex
  - NFSen
  - NSHaRP

‘NSHaRP System: What Is It?’

- Complete security solution
- Provides mechanism to quickly and effectively inform affected users
- Adds Value - Serves as an extension to NRENs CERT
- An Automated Incident Notification & Handling System
- Extends NRENs detection and mitigation capability to GEANT borders
- Innovative and Unique - Caters for different types of requirements
- Supported with GEANT NOC TTS

NETWORK SECURITY HANDLING AND RESPONSE PROCESS

connect • communicate • collaborate



‘Looking Ahead’

- **What Do We Want**
  - Vision
    - *Strategy*
    - *Roadmap*
  - Collaboration
  - Make Our Networks Safer
    - *Benefit Our Users*
  - More Structured Information
    - *Scalable*
- **What Do We Not Want**
  - Another Mailing List
  - Politics ☺

connect • communicate • collaborate

Questions & Answers

connect • communicate • collaborate

Thank-You

Thank You

Wayne Routly
wayne.routly@dante.net

DANTE

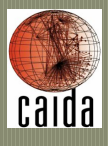
NETWORK SECURITY HANDLING AND RESPONSE PROCESS

connect • communicate • collaborate

Legal Aikido: A Data-Sharing Framework to Advance Network & Security Research

Erin Kenneally, M.F.S., J.D.
Erin @ caida.org

Cooperative Association for Internet Data Analysis
University of California San Diego



Talk Map

- Defining the Issue & Solution Space
- Challenges & Motivations
  - Uncertain Legal Regime
  - Incomplete Technology Solution Models
  - Data Risks
  - Under-valued Benefits of NetSec Research
- Applying Akido:
  - Self-Reg Opportunity
  - Operational model: PS2 Framework

The Issue Space Defining the Solution

- Current posture:
  - defensive, default-deny sharing network traffic data
- (Misinformed) assumptions
  - Privacy risks and legal restrictions >>> benefits of sharing
  - Unprecedented data availability = plethora of network infrastructure information
  - ISE directives post-911 → incent network data exchange
- Muted legislative, judicial, policy drivers
  - Threat model from NOT sharing data = vague
  - No body count / $billion losses (at least no explicit, causal)
- No widespread, standard procedures for exchange
  - Ad-hoc, nod & wink
- Dynamic and normative-deficient understanding of privacy risk and research utility
  - No cost-accounting for privacy risk
  - No ROI for investment in empirical network measurement

- **confusion = window of opportunity**

Challenges & Motivations (1) Uncertain Legal Regime

- No legal framework explicitly prescribes, incentivizes, or forbids sharing network data for security research
- Ambiguity between tech & legal discourse re: fundamental concepts driving risk
  - PII, REP, content, URLs, IPAs, packet headers, payload … oh my!
  - Law inconsistent- functional equivalent of PII
  - E.g., is IPA 'content' and URL 'addressing' data (ECPA, 4th A. purposes)?
    - *Johnson v. Microsoft* (2008) - IPA does not identify persons
    - *State v. Reid* (2007) - REP in subscriber information attached to IPA
    - *US v. Forrester* (2007) - URLs may have REP - reveal communication content
    - *HIPAA Privacy Rule* – IPA is protected PII
    - States' data breach laws – IPA not in definition of personal information
- Social normative expectations: my IPA, URLs + search terms are digital fingerprints?
  - E.g. Tor, automated in-browser cookie and URL deletion

Challenges & Motivations (2) Incomplete Technology Solution Models

- Point solutions fail to address context-dependent risks
  - Cases-in-point: de-anonymization attacks success
    - Prefix-preserving anonymization subject to re-identification
    - Poster cases (Netflix, Yahoo!, Traffic injection attacks)
- Trade-off: Purely technical approaches v. research utility goals
  - Data minimization intentionally obfuscate essential data (network management, countering security threats, evaluating algorithms, apps, architectures)
  - E.g., Conficker

Challenges & Motivations (3) Data Sensitivity Risks

- Sources: legal compliance, ethical obligations, norms/court of public opinion
- Main categories
  - Disclosure risk
    - Public
    - Accidental/malicious
    - Compelled – Private (RIAA subpoenas) , Gov't (NSA, Telco releases)
  - Misuse risk
    * increasing quant & qual - Evolving tech > capabilities and < costs:
    - False inference (match linking 1st/2nd order identifiers)
    - Data confidentiality (network topology, health)
    - Privacy
      - linking network data to individuals
      - de-anon / re-identification commoditization
        - Tension- protect (law, policy) AND motivation to uncover PII (profit, avoid legal liability triggers, attribution)
        - Cat & mouse gamers = LE investigations, biz intel, legal dispute resolution, security incident response, gov't infrastructure protection

Challenges & Motivations
(4) Under-valued Benefits of Network Research

- Justify the Ask:
  - Understanding structure, function of CI
  - (topology, workload, traffic routing, performance, threats & vulnerabilities)
- Network Data sharing utility criteria
  - Objective is positively related to social welfare
  - Need for empirical research
  - Research purpose not being pursued
  - Research could not be conducted without
  - No sufficiently similar data already being collected that could be shared
  - Uses of shared data are transparent, objective, scientific, control for risk
  - Results using shared data can be acted upon meaningfully
  - Results are capable of being integrated into ops or biz processes (security improvements, situational awareness)

Solution Space: Using Aikido on Net Sec Researcher's BFF

- Options:
  - Amend law (ECPA research exception)
  - Aikido the law via self-reg regime (i.e. inform norms & legal precedent)
- **/1st/ ID the attack** – on voluntary disclosure of non-content to researchers
  - ECPA's invariables- From Whom, To Whom, What, Where
  - ECPA's variables- Consent, Provider Exception & Relevance, Gov't, Content
- **/2nd/ Blend & harmlessly redirect attack** - use the law itself to clarify the gray and embolden the exceptions
  - **Consent Exception**
    - interpreted broadly, esp if internal to Provider, so define for network performance engineering and research
    - unclear USE scope, so define specific uses viz. ToS & Privacy Policies, banners
  - **Provider Exception** (outsourcing research under cyber security justification)
    - ECPA allows EE of an ECS to "intercept, disclose, or use" communications when such activity "is a necessary incident to the rendition of his service or to the protection of the rights or property of the provider of that service."
    - ECPA does not, per se, prohibit outsourcing research
    - Internal use is largely unregulated /eg/ telcos and SE retention of traffic
    - necessary incident..protect rts & property- largely untested
    - So → define "research" explicitly, make nexus between collection and sharing for protection of rts/ property substantially clear; ensure that provider procedures are consistent with one of the statute's exceptions; SCA no apply if sharing occurs w/i service provider (ie, researcher employed by provider)
  - **Provider Relevance**
    - Non-public provider may disclose non-content records to a governmental entity (state university researcher); or
  - **Gov't Entity**:
    - ECPA no define, unclear if State U. Researcher is "gov't"
    - Press whether public sector entity must have compulsory powers to = "gov't"
  - **Content** -
    - Does ECPA apply data if data anonymized beyond being recognizd as "content"? (substance, purport or meaning)

Implementing Akido

- **Privacy-Sensitive Sharing (PS2) model**
  = Privacy-enhancing technology + privacy-principled policies
- Risk – Benefit methodology
  - Bridges risk – utility perception gap
- Enables transparency as touchstone of data sharing
  - counter to subjective, opaque evaluations
  - engender trust, beyond "trust me"
- Considers practical challenges of stakeholders
  - network researchers, sys operators, security professionals, legal advisors, policymakers
- Bottom-up & Proactive
  - Anchor point to demo community norms, inform law & policy

PS2 Framework
Policy Components

- Core underpinnings:
  - Make risks 'contagious' (sharing= data AND responsibilities & obligations)
  - Components rooted in principles and practices of national & global laws, policies
    1. Authorization
    2. Transparency
    3. Compliance with applicable laws
    4. Purpose adherence
    5. Access limitations
    6. Use specifications and limitations
    7. Redress mechanisms
    8. Oversight
    9. Security
    10. Audit tools
    11. Data quality assurances
    12. Training
    13. Transfer to 3rd parties
    14. Ethical impact assessment
    15. Disclosure minimization

PS2 Framework
Technology Component

- **Disclosure Minimization/Controls**
  a) Deleting all / part of sensitive data
  b) Generalizing
  c) Perturbing
  d) Pseudonomizing all/ parts
  e) Aggregation or sampling techniques
  f) Mediation techniques (sending code-to-data)
  g) Aging the data
  h) Limiting quantity
  i) Synthesizing
  j) Layering anonymization
- **Implementing Vehicles :**
  - MOU/MOA/MOC, model K's, binding organizational policy, NDA, AUP

Evaluating PS2
Addressing Data Risk & Utility Goals

- **Criteria**:
  - 1. How well PS2 addresses data risks (table 1)
    - Policy controls, alone = coverage gaps
    - Tech controls, alone = seemingly control for privacy risks (? policy control components superfluous ?)
  - 2. To what extent PS2 impedes utility goal (table 2)
    - Technical controls, alone = impedes utility
- **Conclusion**:
  - Only tech solution breaks down along utility dimension
  - Only policy solution leaves too high privacy risk exposure
  - Therefore, hybrid strategy - dial down tech controls for utility objectives AND dial up policy controls to cover risk
  - Framework is both
    - Evaluation of hybrid model
    - Policy Dev tool for data sharing

Evaluating PS2
Addressing Privacy Risk & Utility Goals

PS2 / Privacy Risk	Public Disclosure	Compelled Disclosure	Malicious Disclosure	Government Disclosure	Misuse	Inference Risk	Re-ID Risk
Authorization		X	X		X	X	X
Transparency	X	X	X	X	X		
Law Compliance			X			X	X
Access Limitation		X			X	X	X
Use Specification		X	X		X	X	
Minimization							X
Audit Trails	X	X	X	X	X	X	X
Redress	X	X	X	X	X	X	X
Oversight		X	X			X	X
Data Quality	X	X	X	X			X
Security		X				X	X
Training/Education		X	X			X	X
Impact Assessment	X	X	X	X	X		

Table 1: Privacy risks evaluated against the PS2 privacy protection components. (Minimization refers to the techniques evaluated in Table 1.)

Min. Tech. / Utility	Purpose Worthwhile?	Is there a need?	Is it already being done?	Are there alternatives?	Is there a scientific basis?	Can results be acted upon?	Can DS & DP implement?	Reasonable education costs?	Forward & backward controls?	No new privacy risks created?	No few other problem created?
Not Sharing	X	X	X	X	X	X	X				
Delete All	X	X	X	X	X	X	X		X		
Delete Part	X	X		X	X		X		X	X	
Anonymize	X	X	X	X	X		X	X	X	X	
Aggregate	X	X	X	X	X				X	X	
Mediate (SCID)	X						X	X			X
Age Data	X	X	X	X	X		X			X	
Limit Quantity	X	X	X	X	X	X	X		X	X	
Layer Anonymization	X	X	X		X	X	X	X	X		

Table 2: PS2 minimization (of collection and disclosure) techniques evaluated against utility.

(C) 2012 CAIDA | Kenneally

Security at the Cyberborder, 22 and 23 February 2012

Robert F. Brammer, Ph.D., Senior Advisor to the Internet2 CEO

Internet2 Cybersecurity Perspective

INTERNET2'S HUMAN AND COMPUTER NETWORKS

Internet2's human network

- More than 350 member institutions
- More than 65,000 community institutions connected
- 10+ major initiatives, working groups and special interest groups – networks, middleware, cloud computing, security, and applications
- 100+ International partners
- Expert core staff

Internet2's computer network

- World's highest performance network for advanced R&D – next generation now becoming operational with full deployment in 2013
- Significant differences from commercial networks
- Integrated services and tools

Key Points for This Presentation

- We will be expanding Internet2's role and activities in cybersecurity
  - We are addressing priorities within our staffing and financing ability
- Increasing cyber threats and the value of our new very high performance network make cybersecurity more significant to Internet2, members, and other participants (e.g., selected federal agencies, US UCAN) than it has been in the past
- Major threats include misuse of the very high performance capacity, denial of service, and compromise of sensitive information of our members
- We are in the planning stages for an expanded cybersecurity initiative and will not be able to afford all of the possibilities discussed in this presentation
- Your feedback is very significant for our planning

Presentation Organization

- Statement of strategic priorities and member expectations for security
- Summary of current services
  - REN-ISAC activities
  - InCommon and related identity and access management developments
  - Member education and information dissemination
  - EDUCAUSE and the Higher Education Information Security Council, SALSA, DNSSIG, …
  - Net+ security-related service providers
- Internet2 is planning a new Cybersecurity Initiative
  - In early stages and considering alternatives
  - Member consultation and feedback
  - Your inputs are encouraged

From the Internet2 Strategic Plan (2008)

- "Develop and promote cost-effective methodologies, standards, and best practices for security and end-to-end application performance. Implementation must be possible under real-world conditions across campus, regional, national and international networks."
- "Building upon the current success of the Internet2 middleware and security initiatives, develop and implement a plan (including a comprehensive funding model) for Internet2-sponsored and supported research and development of middleware for identity services and security that continues to advance these initiatives while addressing the need to support broad deployment within our community. The plan should articulate the life cycle and outcomes and create a sustainable funding model for ongoing support. The funding model should examine a combination of approaches, such as explicit agreement to fund activities within current financial capabilities, soliciting directed funding, partnerships with other organizations, spinning off functions to separate organizations, or other mechanisms."
- "Work with EDUCAUSE, government agencies, and other domestic and international partners to identify and address the security challenges unique to advanced networks and higher education; expand that collaboration to include carriers, standards bodies such as IETF, other corporations and government for the broadest discussion and understanding of solutions to this fundamental challenge."

Internet2's Current Security Services for Members

- Most of Internet2 focus on cybersecurity has been on enabling the REN-ISAC, identity and access management, and member education, …
- Contributions to the REN-ISAC
  - Participation in the Executive and Technical Advisory Groups and financial support
  - Collaboration on network issues
  - Potential for further collaboration
- InCommon Federation and related developments
  - InCommon – federation, assurance, certificates, …
  - Shibboleth – single sign-on software, Grouper – toolkit for groups, CoManage – collaborative management platform
- Member education and information dissemination
  - Representative in numerous security-related groups – SALSA, Educause and HEISC, MAAWG, DNSSIG, …
  - Security sessions at Member Meetings, JointTechs, topical meetings (e.g., Collaborative Data-Driven Security for High Performance Networks, 1 and 2)
  - Metrics, reports, web sites, …

Priorities for an Internet2 Cybersecurity Initiative

- We must be recognized by all as being good stewards of the significant investment that the federal government has made in our new very high performance network
  - "Reasonable person" theory applied to security
  - Must recognize current environment and growing threats
- We must do what we can to help our members address their threats
  - Real-time security, compliance, privacy, sensitive research information, …
- We must position us to address cybersecurity issues that may limit our growth
  - Federal agency and commercial member needs
- We must enable Internet2 to be a recognized thought leader for all issues in advanced networking and cloud services, including security
  - Implies regular communications with our members and with external groups, including the Congress, the Administration, and industry
- We have limited staff resources and can't bust our budget
  - Developing a staffing and financing strategy, including collaborations

Some Potential Opportunities to Expand Internet2 Cybersecurity Activities

- Partnerships and external funding required due to resource constraints
- Address very diverse security (cyber and physical) requirements for selected federal operations
- Attract new industry members into our Net+ program as cloud service providers
  - New products and services (e.g., access management as a service)
  - Commercialize developments that originate in the Internet2 community
- Potential areas for further collaboration with the REN-ISAC
  - Joint Internet2-REN ISAC industry membership program
  - Increase value for industry members – actionable information, recruiting
  - Increase functionality of SES-3 information and the abilities of our members to use it
  - Improved policies, processes, and tools for monitoring usage of our network
- The new Internet2 network as a platform for cybersecurity R&D
  - Assessment of high performance cybersecurity architectures and products (e.g., DPI at terabits/sec)
  - Collection and analysis of network data (e.g., anonymized IPv4 and IPv6 data, other types?)
  - Cybersecurity issues for software-defined networks

DISCUSSION

Workshop on Security at the Cyberborder

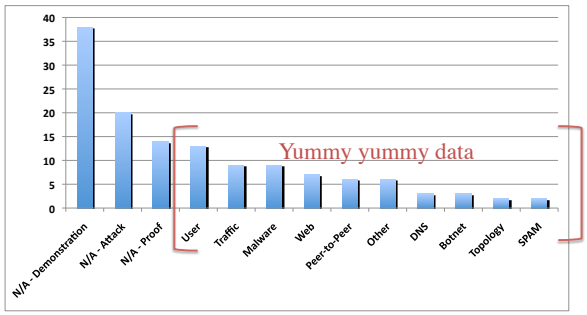
Michael Bailey
University of Michigan

Let's not beat around the bush…

I am here because I want your **DATA**. All of it. Every byte. On every interface. I want it in real time. I want to save it to disk for years. I want to link it with all kinds of external data.

Why?

- I write papers … errr create knowledge
- Maths are just fine, but make my brain hurt. I like experimental systems work!
- Look at all the papers in this year's "Bailey Top 5" (your mileage may vary)
  - Usenix Security, LEET, CCS, S&P, NDSS
  - 132 papers



How I used to get data

More recent ways of getting data

- Virtual Center for Network and Security Data is a collaborative effort among 4 institutions with the University of Michigan as lead. As part of PREDICT, Its goals are to:
  - Develop a virtual data repository of rich, correlated, security datasets
  - Providing distributed services to simplify accessing and using
  - Address the ethical and legal issues associated with the sharing data

More recent ways of getting data

- Classes of data that are interesting, people want collected, and seem reasonable to collect
  - Netflow
  - Critical infrastructure BGP and DNS data
  - IDS / firewall logs
  - Performance data
  - Network management data
  - Network telescope, honeynet data
  - SPAM data

Dataset	Size (TB)
Netflow (2009-01 to present)	46
35/8* Darknet (2005-10 to present)	100
Other Darknets	9
BGP (2008-09 to present)	2
DNS (just begun)	0.6
TOTAL:	**~157**

What I used to do with it

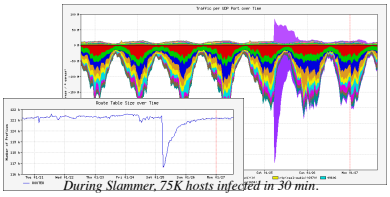
- Primary motivation of hackers was bragging rights
- Worms and viruses intended to simply wreak havoc on the infrastructure
- These were availability attacks: impacting network access and services, and often, reputations

BBC NEWS UK EDITION

Bookies suffer online onslaught

By Mark Ward
BBC News Online technology correspondent

The extent to which British betting websites are being attacked by criminals using the net to bring down a site unless a ransom is paid has been revealed by a BBC News Online investigation.



During Slammer, 75K hosts infected in 30 min.

Technology

The New York Times

February 8, 2000

Yahoo Attributes a Lengthy Service Failure to an Attack

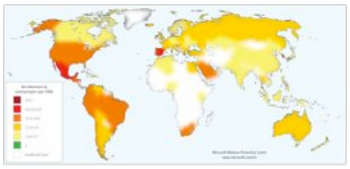
By MATT RICHTEL

SAN FRANCISCO, Feb. 7 – Yahoo Inc. blamed a "planned attack" by computer hackers for a service failure that lasted nearly three hours today, in a rare interruption of one of the most popular and best performing sites on the World Wide Web.

What I do with it now

- Unlike traditional worms and virus, botnets acknowledge
  - A compromised system is more useful alive than dead
  - A network of compromised hosts provides a powerful delivery platform
- Botnets represent today's attack platform

Rank	Family	Primary Control Mechanism	Computers Cleaned (1Q10)	Computers Cleaned (2Q10)	Change
1	Win32/Rimecud	Other	1,807,773	1,748,260	-3.3% ▼
2	Win32/Alureon	HTTP	1,463,885	1,035,079	-29.3% ▼
3	Win32/ Hamweq	IRC	1,117,380	779,731	-30.2% ▼
4	Win32/Pushbot	IRC	474,761	589,248	24.1% ▲
5	Win32/IRCbot	IRC	597,654	388,749	-35.0% ▼

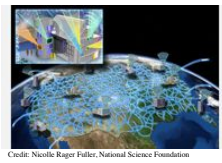


Microsoft desktop anti-malware products removed bots from 6.5 million computers around the world in 2Q10 [Microsoft SIR v9]

Where should we go?

- Guess the next hot topic (hard, magic!)
- More "risky" data, closer to raw, more PII, more legal challenges, more ethical challenges
- Better annotated data

The Next Hot Topic

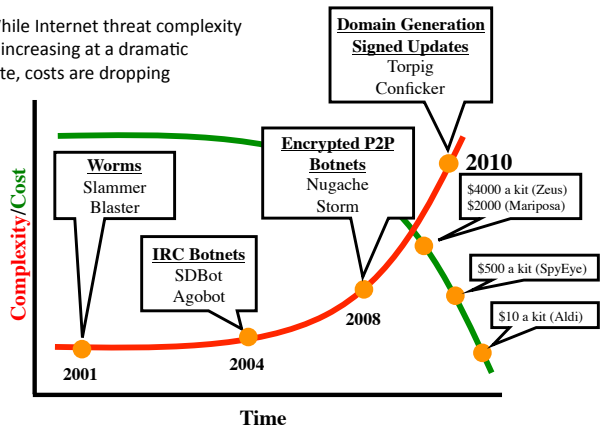


Credit: Nicolle Rager Fuller, National Science Foundation

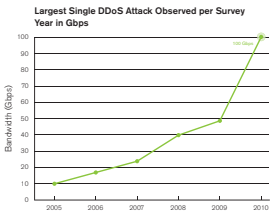
- Means: Distributed attacks increasing in size and sophistication, targeting specific applications.
- Motives: Proliferation of attacks spurred by financial gains and now political motives.
- Future security challenges will follow Internet adoption patterns:
  - Proliferation of wireless devices and social media platforms open new avenues for hackers.
  - Protecting cloud infrastructure key to long-term adoption.
  - The trend toward increasingly cyber-enabled systems expands the scope of attacks to physical infrastructure – manufacturing, energy production, healthcare and transportation.

Cheaper, Better, Faster

- While Internet threat complexity is increasing at a dramatic rate, costs are dropping

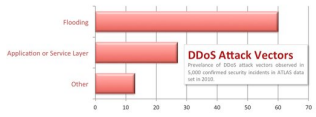


Better, Faster



- Exploits moving "up" as well – infections now delivered via web sites through drive-by installs
  - Projected 1 in 10 web sites hosts malicious content
  - Web-based delivery means outpacing email, viruses, etc.

- Increasing size and sophistication
- 10-100Gbps DDoS attacks seen by ISPs
- Attacks moving "up" – Attacking services rather than infrastructure



New York Times Malware: Bad Ad On NYTimes.com

Cheaper

Exploit packs

Spam it com Spam Services

Botnet Toolkits

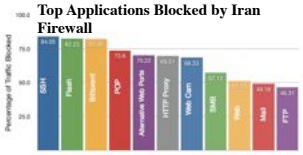
DeCaptcha and Packing Services

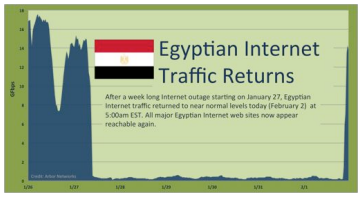
DEATH BY CAPTCHA
FASTEST DISCOUNT CAPTCHA SOLVERS

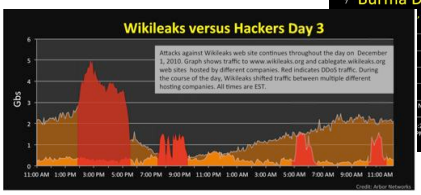
Pay per Install

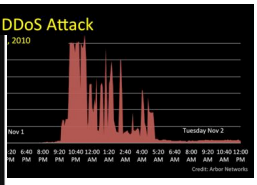
Cloud Type	Legitimate	Crimeware
IaaS	Amazon EC2, Mosso	Renting out infected bots
PaaS	Google App Engine, Azure	Botnet-backed spam services
SaaS	SalesForce, SAP ByDesign	Packing services, Decaptcha!

New Motives: Cyber-war, Censorship, Activism, and the Rise of Politically Motivated Attacks







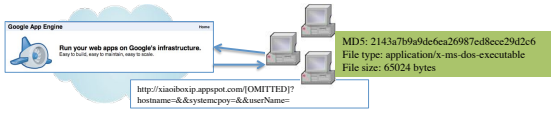


Cellular and Mobile Security

- Well over 5 billion users have access to cellular networks today.
  - Represents the only access available to the majority of people on the planet.
  - Mobile devices will be the predominant global access method by 2014.
- A wide variety of mobile platforms each with their own security model
- As mobile devices inherit the capabilities and connectivity of traditional computing, they also inherit the security risk (botnets).
- Significant sensitive and private data stored & entered on devices

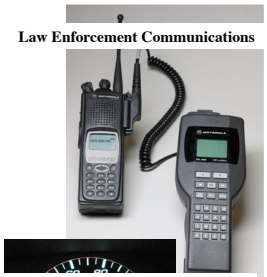
Cloud Security

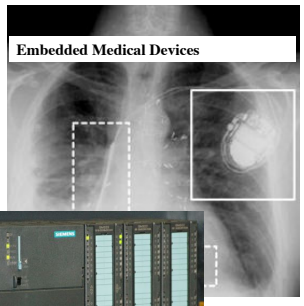
- Clouds enable attackers to have the "The reliability, performance and security of Google's infrastructure"



- Clouds are critical infrastructure that need protection.
- Interesting open problems in identifying **unique** problems in this space and their solutions
- Homomorphic encryption

Cyber-physical Security





Risky Sharing

- The literature on data sharing describes two broad approaches to protecting the confidentiality of individuals whose records appear in data collections:
  - restricting access
  - restricting data [NRC 2010]

ICPSR INTER-UNIVERSITY CONSORTIUM FOR POLITICAL AND SOCIAL RESEARCH
A PARTNER IN SOCIAL SCIENCE RESEARCH

NORC data enclave
the key to unlocking data securely

Statistics Sweden

UKDA UK Data Archive

DANMARKS STATISTIK

Statistisk sentralbyrå
Statistics Norway

Office for National Statistics

Statistics Netherlands

U.S. DEPARTMENT OF COMMERCE BUREAU OF THE CENSUS

Insurmountable (?) challenges of restricted data

- Providing general purpose, high utility networking and security data while protecting confidentiality via statistic disclose control is difficult
  - Security in particular wants to examine outliers
  - Fields intra and inter record are conditionally dependant on each other
  - Widely available sources of linking or secondary data, and in fact, we very well may want to link
  - Outcomes (results) are neither structured nor predictable
  - Massive heterogeneity in the data at all levels of abstraction means inputs are not well defined

Network	Type of organization	Hosts	Web-servers	TCP	HTTP	TCP+ Ports	TCP+ HTTP	TCP+ Ports+ HTTP
A/16	university network	5512	1237	352	241	1210	699	1386
B/16	university network	1289	169	156	73	392	249	463
C#1/17	web-server farm	11342	10080	256	862	1625	1764	3394
C#2/19	web-server farm	2438	2208	93	293	394	559	811
D#1/19	web-server farm	1859	1513	118	221	330	451	590
D#2/19	web-server farm	1652	1171	137	208	266	417	487

Restricted Access

- There are four primary strategies for restricted access:
  - Licensing
  - Remote execution systems
  - Data enclaves
  - Virtual data enclaves.

Licensing

- Licensing allows researchers to obtain highly detailed data and thus facilitates secondary analyses of the data.
  - However, this approach relies on researchers not violating the terms of the license.
  - Enforcement is generally the responsibility of the funding agency.
- A number of authors have noted that the licensing process can be difficult, time-consuming, and even costly and thus may limit the number of researchers who use the specimens and data in a storage facility
  - *"Statistical and other agencies* that provide data for research should work with data users to develop flexible, consistent standards for licensing agreements and implementation procedures for access to confidential data" (National Research Council, 2005, p. 79).

Remote Execution Systems

- In this approach, confidential data are maintained in a computer system owned by the data disseminator, and a secondary researcher who wishes to perform a study with the data submits a query to the system
- The National Center for Health Statistics and the U.S. Bureau of the Census maintain such systems.
- Limited scope of queries
- These systems are not foolproof, however, as intruders can use judicious queries to glean sensitive information about data subjects.

Data Enclaves

- With this approach, an investigator works in a room dedicated to accessing the data.
  - Only approved researchers are allowed in the enclave.
  - Computers in the room are not connected to the Internet or to other external resources.
  - Researchers cannot take individual-level data from the enclave, and all results they obtain are checked by the data disseminator for potential confidentiality breaches before they can be taken out of the enclave.
- It also carries a variety of costs, however.
  - Data enclaves are inconvenient for investigators who must travel to them.
  - Moreover, most enclaves require that data users pay a substantial fee to cover the costs of maintaining the facility and its staff; they also require potentially time-consuming background checks and proposal approvals.
  - Hence, the amount of analysis likely to be carried out on the data in an enclave is limited (National Institutes of Health, 2006).

Virtual Data Enclaves

- Virtual data enclaves combine features of licensing, remote execution systems, and data enclaves.
  - The data are housed in a system owned by the data disseminator.
  - Licensed secondary researchers access the data remotely.
  - Users of a virtual data enclave cannot store the data on their computers, and certain functions on their computers, such as printing and the ability to copy data to removable media (including disks and micro-vault storage media), are disabled.
- Virtual data enclaves thus allow researchers to access the data without traveling to a secure data enclave.
- They also avoid some of the disclosure risks posed by licensing researchers to store data on their own machines, such as the accidental loss of CD-ROMs or sharing of data with unapproved investigators or students.
- As with licensing, however, confidentiality protection depends on researchers not violating terms of the data use agreement.

How Do You Pick?

- Data release can be seen as a 'spectrum' of access points balancing
  - value of data
  - ease of use
  - disclosure risk
- The aim is, for a given level of confidentiality, to maximize data use and convenience

Richie 2009

Develop a Set of Interlocking Controls

- Safe Projects
  - Access needs to be for valid statistical purposes
- Safe People
  - Researchers can be trusted to use the data appropriately and follow procedures
- Safe Data
  - The data itself is inherently non-disclosive
- Safe Settings
  - The technical controls surrounding access prevent the unauthorized removal of data
- Safe Outputs
  - The statistical results produced do not contain any disclosive results

Richie 2008

Better Data

- All about ground truth
  - **Actual knowledge** of a characteristic or event or policy by virtue of the fact that an individual has direct access to the actual information of interest
  - **Expert knowledge** of a characteristic or event by someone with vast operational experience in a particular area
  - **Artificial data** consisting of models of the activity of interest. An often blended dataset with combinations of real and generated background and event traffic
  - **Artificial knowledge** built by applying known algorithms or methods to data in order to automate the annotation of a dataset at scale.
- Why ground truth? Varies greatly with the dataset (of course), but some broad classes of use come to mind:
  - Building models and simulation from this known truth
  - Evaluating the effectiveness of a collection or observation technique
  - Evaluating the effectiveness of an algorithm

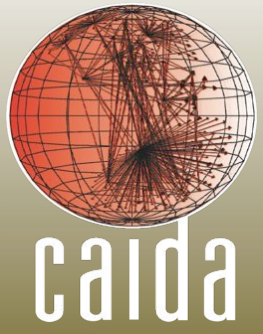
Questions?

Michael Bailey
University of Michigan

~~Fostering Cybersecurity Research Through Data Sharing~~

~~Data: You Show Me Yours, I'll Show You Mine~~

CAIDA: A Data Sharing Case Study



Josh Polterock, CAIDA
NSF Security at the CyberBorder Workshop
February 23, 2012

Overview

- CAIDA:
- our mission
- conducting research
- building infrastructure
- collecting and curating data
- developing tools
- modeling economics
- informing policy
- coordinating and hosting workshops

Our Mission

CAIDA The Cooperative Association for Internet Data Analysis (CAIDA) is an independent analysis and research group based at the University of California's San Diego Supercomputer Center. CAIDA investigates both practical and theoretical aspects of the Internet, with particular focus on:

- collection, curation, analysis, visualization, dissemination of sets of the best available Internet data,
- providing macroscopic insight into the behavior of Internet infrastructure worldwide,
- improving the integrity of the field of Internet science,
- improving the integrity of operational Internet measurement and management,
- informing science, technology, and communications public policies.

Conducting Research

- Macroscopic Topology Project
  - IPv4 and IPv6 topology discovery
  - hostname collection
  - Alias Resolution
  - Router to AS assignment
  - dual graph
  - AS relationships
  - IPv4 and IPv6 topology discovery
  - AS Rank
- Internet Interconnection Economics
- Darkspace Analysis
- Named Data Networking
- Modeling Complex Networks
  - Internet evolution
  - Routing complex networks
- DatCat Returns!

Infrastructure

- Archipelago
  - CAIDA's active measurement infrastructure
  - 57 monitors – growing 1 or 2 per month
  - 28 w/ IPv6 connectivity
  - currently used for:
    - Team-probing experiment to collect IPv4 and IPv6 topology
    - alias resolution measurements
    - Spoofer experiment
- Passive Monitors
  - Tier1 10GE backbone link packet header capture

Infrastructure

- UCSD Network Telescope
  - Near real-time access to raw telescope data
  - 2 days of telescope dataset http://www.caida.org/data/passive/telescope-2days-2008_dataset.xml
  - 3 days of Conficker dataset http://www.caida.org/data/passive/telescope-3days-conficker_dataset.xml
- CAIDA working with SDSC to use cloud storage services (http://cloud.sdsc.edu/)
  - plans to purchase 25 TB of cloud storage for last 60 days of system backups
  - granted 20TB from SDSC Executive Team for scientific data
    - telescope (103TB)
    - packet headers (18.8TB)
    - skitter/ark topology (4TB)

Data collection - passive

- OC192 backbone: March 2008 - December 2011
  - 18.8 TB compressed, 35.7 TB uncompressed
  - unanonymized: 10.9 TB compressed, 21.3 TB uncompressed
  - anonymized: 7.9 TB compressed, 14.4 TB uncompressed
  - Doing cleanup towards retaining only quarterly traces
  - Completed 2011 Passive Dataset
- Problems:
  - Hardware failures at collection sites: Chicago monitors have been offline since September. We are working with remote hands to troubleshoot.
- Plans:
  - New 2012 annual dataset will start with upcoming trace on January 19
  - strip payload/L1/L2, transfer, anonymize, archive
  - collect 1 hour trace per month = 200-250 GB (compressed)
  - keep a quarterly sample - select the best quality

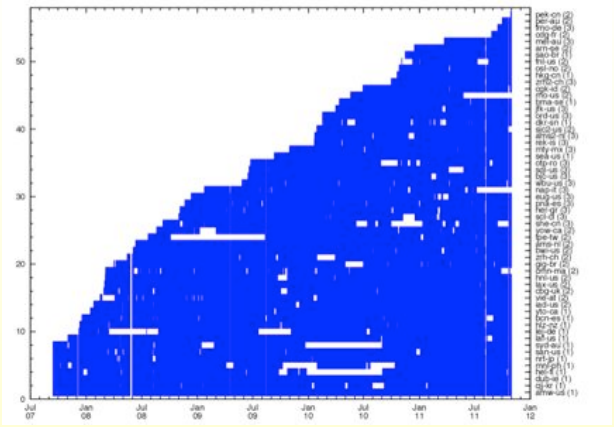
Data collection - passive

- UCSD telescope:
  - data from most recent 30-days (really five weeks) "live" on disk
    - typically 2.9 TiB compressed, 5.5 TiB uncompressed
  - the previous months - backed up on tape (samqfs)
    - current: 2008/04/12 - 2012/01/12
    - 102 TB (compressed), 192 TB (uncompressed)
    - received new NSF award "CRI-Telescope: A Real-time Lens into Dark Address Space of the Internet"
- OC48 traces:
  - 964.5 GB (compressed), 1.7 TB (uncompressed)
  - unanonymized: 815.7 GB (compressed), 1.5 TB (uncompressed)
  - anonymized: 148.8 GB (compressed), 285.2 GB (uncompressed) (in PREDICT)

Data collection - active

- old skitter data (in PREDICT):
  - 1.47 TB (compressed), 4.02 TB (uncompressed)
    - discontinued in February 2008
- current Ark data:
  - IPv4 topology: 1.8 TB (compressed), 5.8 TB (uncompressed)
  - IPv6 topology: 2.8 GB (compressed), 9.7 GB (uncompressed)
  - 57 monitors in 30 countries, 28 IPv6 capable
  - continues to expand
- data curation:
  - create derivative data sets
  - aggregate in ITDK
    - router-level topologies: nodes and links
    - host names
    - router-to-AS assignment
    - geographical information
      - http://www.caida.org/data/active/internet-topology-data-kit/
- NSF award to curate/analyze/annotate IPv6 data

Archipelago Monitors and Data

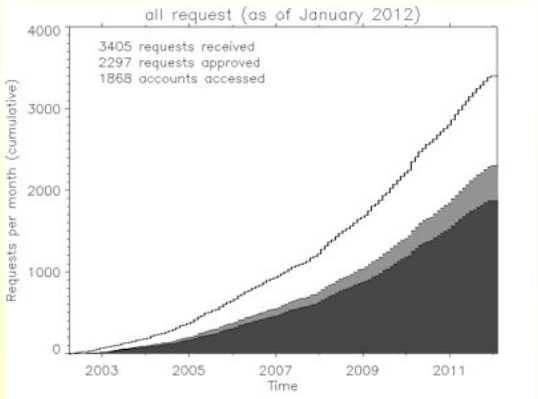


Requests for the data, 2011/2010/2009

Dataset	Requests	Approved	Accessed	Served
Backscatter	51/73/95	34/47/60	28/36/46	Feb 2003
Passive	275/185/233	210/150/179	170/126/157	Feb 2004
Topology	155/163/129	129/113/83	85/80/63	Jul 2004
Witty	16/16/27	12/13/17	10/11/14	Mar 2008
Telescope	29/34/37	22/23/21	18/19/17	Jul 2009
DNS-RTT	10/7/7	8/5/2	6/4/2	Aug 2006
DDoS	92/108/NA	62/74/NA	51/66/NA	Mar 2010
Total	**628/586/528**	**477/425/362**	**368/342/299**	

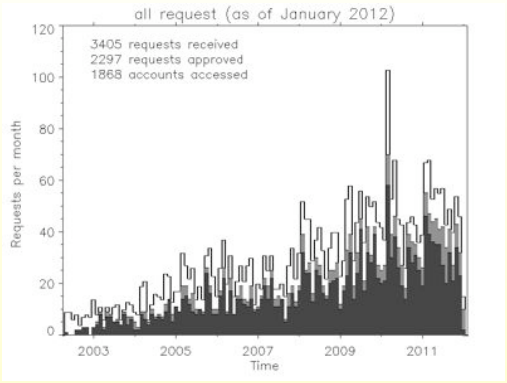
Data request stats

- all requests (cumulative)



Data request stats (cont)

- All requests (monthly)
  - spike (40 requests) in first month of DDoS dataset



Data Set Popularity

- 1st best - OC192 and OC48 traces
  - requested 693 times, accessed 454 times (since 2009)
  - who used it: 259 .edu, 141 .cn, 43 .uk, 28 .com (since 2004) …
    - and 56 more domains
    - of 839 total accounts: 270 from U.S.
- 2nd best - topology data
  - requested 447 times, accessed 228 times (since 2009)
  - who used it: 256 .edu, 119 .cn, 41 .uk, 31 .kr, 29 .com, 26 .jp (since 2004) …
    - and 52 more domains
    - of 785 total accounts: 191 from U.S.

Data availability

- PREDICT (OC48 traces, topology from skitter, telescope)
- Derived data sets are publicly available (i.e., AS-links)
  - sample use: http://semilattice.net/projects/map-of-the-internet/
- Academics who sign AUP (OC192, topology from Ark, telescope)
- Commercial researchers
  - a small sample of data to entice interest
  - join CAIDA, various membership levels are offered

Data statistics - online

- Aggregated, (near) real time
- OC192 backbone
  - report generator
  - http://www.caida.org/data/realtime/passive/?monitor=equinix-chicago-dirA
- topology
  - Ark statistics: http://www.caida.org/projects/ark/statistics/index.xml
  - path dispersion (AS and IP), path length distribution, RTT distribution, RTT vs. distance, median RTT per country, ...

Meta-data for packet traces

- OC192 data: 2008-2011
  - an hour-long trace every month
  - usually, 3rd Thursday, 13:00 - 14:00 UTC
- OC48 data: 2002-2003
- Statistics:
  - Date, start time, stop time
  - Numbers of IPv4, IPv6, unknown packets
  - Transmission rate in pkts/s, bits/s
  - Link utilization (%)
  - Average packet size
  - Graph of packet size distribution (IPv4 and IPv6)

http://www.caida.org/data/passive/trace_stats/

Recent publications

- A. Dainotti, R Amman, E. Aben, kc claffy, *Extracting Benefit from Harm: Using Malware Pollution to Analyze the Impact of Policital and Geophysical Events on the Internet,* ACM SIGCOMM CCR accepted for publication.
- Erin Kenneally *A Refined Ethical Impact Assessment Tool and a Case Study of its Applications,* submitted to WECSR 2012.
- kc claffy, *Tracking IPv6 Evolution: Data we have and Data We Need,* ACM SIGCOMM CCR V. 41, p. 43-48, 2011.
- kc claffy, *The 3rd Workshop on Active Internet Measurements (AIMS-3) Report,* ACM SIGCOMM CCR V. 41, p. 37-42, 2011.

Recent publications

- A. Dianotti, C. Squarcella, E. Aben, kc claffy, M. Chiesa, M. Russo, A. Pescape *Analysis of country-wide Internet outages caused by censorship,* accepted to IMC 2011.
  - national level outages in Egypt and Libya
  - data used:
    - public BGP
    - CAIDA telescope
    - Ark (could have done more)
  - analyzed methods used for traffic blocking, duration, testing
- B. Huffaker, M. Fomenkov, kc claffy *Geocompare - a comparison of public and commercial geolocation databses,* CAIDA tech report, 2011.
  - cross-analyzed multiple databases
  - used available ground truth data (PlanetLab, French networks, Tier 1 provider)
  - Ark RTT data

PREDICT Phase II Data Sets

- UCSD telescope: near Real-Time Telescope Dataset (RTTD)
- topology: Ark data (ongoing)
  - IPv4 Routed /24 Topology dataset
  - IPv4 Routed /24 DNS Names dataset
  - IPv6 Routed Topology dataset
- topology: updated ITDK 2010
- OC192 backbone: 2007-2011

Updates of CAIDA policies

- Telescope data (RTTD)
  - different from previous packaged data
  - simplified and streamlined the AUP language
  - Immediate use by postdoc A. Dainotti and his student
  - analysis of macroscopic events (e.g earthquakes) on the Internet, collaborating with RIPE-NCC on publication.
- ARK hosting sites
  - Now using updated MoC for all new hosting sites
- Passive data collection MOC
  - Recently completed
    http://www.caida.org/data/collection/aup/internet_traffic_collection_moc.xml

CAIDA Master AUP

- 4 categories of data - different levels of sensitivity
  - real-time telescope data
  - passive traces
  - active traces
  - derived topology
- Document proliferation
  - 7 data request forms
  - 22 data set web pages
  - 22 README files
- Master AUP 1.0 for all CAIDA data sets
  - Factor out common conditions
  - Remove inconsistencies
  - Sent out to PI list for feedback

General Principles of AUPs?

- Access conditions
  - Accreditation, validation, transparency
- Use restriction
  - Purpose, probing, other
- Disclosure obligations
  - Publication, 3rd party transfer, attribution
- Enforcement
  - Compliance, attestation
- Corrections / amendments
  - Measurement error notifications
- Disposition
  - Account closure, renewal
- Policy Vehicle: AUP, MOA, MOC…

Other activities

- what are we protecting?
  - PII (including IP addresses)
  - organization proprietary data
  - Privacy: Individual vs. Organization
- relevant for Best Practice documentation efforts

Other activities

- the (disastrous) flood of digital data
- DMPTool
  - helpful but does not add much for our purposes
- no ready-to-use guidelines
  - NSF-required Data Management Plan
  - who bears the cost?
  - how much is the cost?
    - thousands of $ per TB per year - commercial clouds
    - $390 per TB per year - SDSC preferred rate
    - $3,000 per TB to store **forever** - Princeton offer
  - NSF position: communities should develop acceptable guidelines
    - what to store?
    - for how long?
- Lots of work on how to better store telescope data; not much to show yet

Other activities

- Hired Ark system administrator, Parisa Nahavandi
- Hired Telescope research programmer, Alistair King
- Hosting Visiting Scholar, Tanja Zseby, working with telescope group
- Hired two postdocs, Matthew Luckie and Alberto Dainotti

Storage Update

Ark IPv4
Total stored data: 1.8 TiB
Total stored no of files: 79600
Total free space: 3.4 TiB (shared with Ark IPv6)
Yesterday growth: 1.7 GiB

Ark IPv6
Total stored data: 2.8 GiB
Total stored no of files: 7444
Total free space: 3.4 TiB (shared with Ark IPv4)
Yesterday growth: 8.2 MiB

Passive high-speed equinix traces
Total stored data: 7.9 TiB
Total stored no of files: 10094
Total free space: 5.9 TiB
Yesterday growth: 6.0 GiB

Live telescope data (ogma)
Total stored data: 33.9 TiB
Total stored no of files: 8813
Total free space: 3.7 TiB
Yesterday growth: 100.1 GiB

Long-term Telescope storage on tape:
Total stored data: 82.9 TiB
Total stored no of files: 20428
Total free space: N/A
Yesterday growth: N/A

Overall Cummulative Stats
Total stored data: 126.5 TiB
Total stored no of files: 126379
Total free space: 13.0 TiB
Yesterday growth: 108 GiB

Data We Want

- interdomain traffic data to look at interdomain traffic patterns
- financial data
- peering relationship ground truth
  > BGP data
  > router-level topology ground truth
- whole maps of ISPs to use as ground truth
- path performance ground truth
- maps at various levels
- anything v6
- ASes and IP addresses by org
- actual organization tree
- much better IP alias resolution blocked on router-level topologies
- hop-by-hop performance several hops away
- anything that shows network neutrality issues

Possible Collaboration

- How the community might interact with CAIDA
- Contribute interdomain traffic statistics to help build a model of interdomain interconnection and dynamics
- AS Rank feedback interface to provide corrections to our AS relationships model

For more information please contact:
info@caida.org

http://www.caida.org/

Sharing Quotes

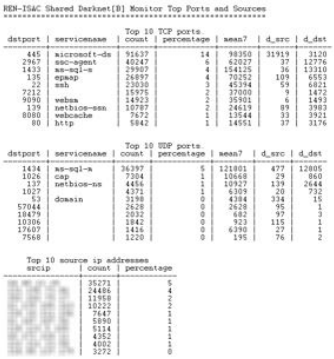
- "Love only grows by sharing. You can only have more for yourself by giving it away to others." – Brian Tracy
- "And a new philosophy emerged called quantum physics, which suggest that the individual's function is to inform and be informed. You really exist only when you're in a field sharing and exchanging information. You create the realities you inhabit." – Timothy Leary, *Chaos & Cyber Culture*
- "Share your knowledge. It is a way to achieve immortality." – Dalai Lama XIV

REN-ISAC

Research Access to REN-ISAC Shared Darknet Data

CyberBorder Workshop

Feb 2012


REN-ISAC Shared Darknet[B] Monitor Top Ports and Sources

Top 10 TCP ports

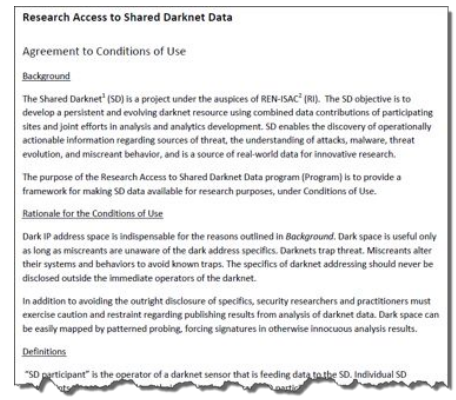
dstport	servicename	count	percentage	mean?	d_src	d_dst
445	microsoft-ds	91637	14	98350	31919	3120
2967	ssc-agent	40247	6	62027	37	12776
1433	ms-sql-s	29807	4	154125	36	13310
135	epmap	26897	4	70252	109	6553
22	ssh	23030	3	45394	99	6821
7212		15975	2	37000	9	1472
9090	websm	14923	2	35901	6	1493
139	netbios-ssn	10787	2	24619	89	3983
8080	webcache	7672	1	13544	33	3921
80	http	5842	1	14951	37	3176

Top 10 UDP ports

dstport	servicename	count	percentage	mean?	d_src	d_dst
1434	ms-sql-m	36397	5	121801	477	12805
1026	cap	7304	1	10668	29	860
137	netbios-ns	4456	1	10927	139	2644
1027		4371	1	6309	20	732
53	domain	3198	0	4384	334	15
57044		2628	0	2628	95	1
18479		2032	0	682	97	3
10306		1842	0	923	115	1
17607		1416	0	6390	27	1
7568		1220	0	195	76	2

Top 10 source ip addresses

srcip	count	percentage
[redacted]	35271	5
[redacted]	24486	4
[redacted]	11958	2
[redacted]	10222	2
[redacted]	7647	1
[redacted]	5890	1
[redacted]	5134	1
[redacted]	4352	1
[redacted]	4002	1
[redacted]	3272	0


Research Access to Shared Darknet Data

Agreement to Conditions of Use

Background

The Shared Darknet[1] (SD) is a project under the auspices of REN-ISAC[2] (RI). The SD objective is to develop a persistent and evolving darknet resource using combined data contributions of participating sites and joint efforts in analysis and analytics development. SD enables the discovery of operationally actionable information regarding sources of threat, the understanding of attacks, malware, threat evolution, and miscreant behavior, and is a source of real-world data for innovative research.

The purpose of the Research Access to Shared Darknet Data program (Program) is to provide a framework for making SD data available for research purposes, under Conditions of Use.

Rationale for the Conditions of Use

Dark IP address space is indispensable for the reasons outlined in *Background*. Dark space is useful only as long as miscreants are unaware of the dark address specifics. Darknets trap threat. Miscreants alter their systems and behaviors to avoid known traps. The specifics of darknet addressing should never be disclosed outside the immediate operators of the darknet.

In addition to avoiding the outright disclosure of specifics, security researchers and practitioners must exercise caution and restraint regarding publishing results from analysis of darknet data. Dark space can be easily mapped by patterned probing, forcing signatures in otherwise innocuous analysis results.

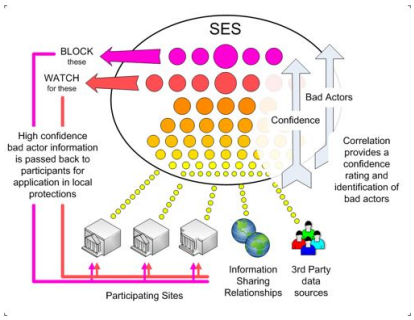
Definitions

"SD participant" is the operator of a darknet sensor that is feeding data to the SD. Individual SD

what's been our experience?

- little uptake
- unclear marketing channel
- time consuming
- not trivial to establish the sharing framework
- onetime and ongoing cost that doesn't directly contribute to operational bottom line (mission) of the organization

Improve operational use and reduce costs by incorporating into REN-ISAC SES framework



continued challenges

- little uptake
- unclear marketing channel
- time consuming
- not trivial to establish the sharing framework
- a “researcher SES” or other distribution means

is it all bad then?

- No!
- solvable problems
- a reusable framework would go a long way to making 2..n much easier

Security at the Cyberborder, 22 and 23 February 2012

Robert F. Brammer, Ph.D., Senior Advisor to the Internet2 CEO

Making Internet2 Data Available for Researchers

INTERNET2'S HUMAN AND COMPUTER NETWORKS

Internet2's human network

- More than 350 member institutions
- More than 65,000 community institutions connected
- 10+ major initiatives, working groups and special interest groups – networks, middleware, cloud computing, security, and applications
- 100+ International partners
- Expert core staff

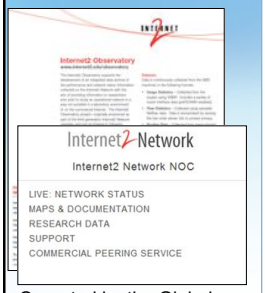
Internet2's computer network

- World's highest performance network for advanced R&D – next generation now becoming operational with full deployment in 2013
- Significant differences from commercial networks
- Integrated services and tools

Key Points for This Presentation

- Internet2 has developed processes and systems for collecting and distributing network data for research and operations analysis
- There was a relatively high level of interest in the data 4-5 years ago, and the Internet2 Observatory was redesigned to accommodate this interest
- Although there is much good research being done today with Internet2 data, in the past few years, requests have decreased and some collections have been reduced or eliminated
- We have had some recent inquiries about routing infrastructure data, notably DNS and BGP
  - Developing a plan to include new network – need to attract funding

The Internet2 Observatory



Operated by the Global NOC at IU

- Usage statistics
- Flow statistics
- Routing data (BGP and IGP (ISIS))
- Latency measurements
- Throughput measurements
- Router data and configurations
- Syslog data
- Some tools (e.g., PerfSonar) available to support research

Some Internet2 Support for Cybersecurity Research (2008)

A Look At The Unidentified Half of Netflow (With an Additional Tutorial On How to Use The Internet2 Netflow Data Archives)

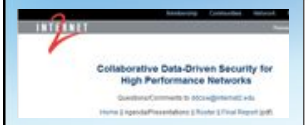
ESCC/Internet2 Joint Techs Workshop
University of Hawaii, January 20-24, 2008

Joe St Sauver, Ph.D.
(joe@uoregon.edu or joe@internet2.edu)
Internet2 Security Programs Manager
Internet2 and the University of Oregon
http://www.uoregon.edu/~joe/missing-half/

Notes: All opinions expressed in this talk are strictly those of the author. These slides are provided in detailed format for ease of indexing, for the convenience of those who can't attend today's session in person, and to insure accessibility for both the hearing impaired and for those for whom English is a secondary language.

- Other references to relevant research in are also available

Data-Driven Security



Report From the Internet2 Data Driven Collaborative Security Workshop for High Performance Networks

Joe St Sauver, Ph.D.
joe@internet2.edu or joe@uoregon.edu
Security Programs Manager, Internet2

Internet2 Fall Member Meeting
San Antonio Texas

http://www.uoregon.edu/~joe/ddcsw-fmm

- Internet2 organized two workshops on "Collaborative Data-Driven Security for High Performance Networks" (May 2009 and Aug 2010)
- Workshop material available at security.internet2.edu
- Helped to stimulate some further interest in data-driven security

Topics for Discussion

- There has been some very good research done with Internet2 data and collaborations with a variety or research teams
  - Results reported in previous references and at various conferences
  - Research primarily focused on networking and not on security
- However, given the increase in cybersecurity research, there should be more interest in our data (particularly for high performance areas)
- Is there something more that Internet2 should be doing to stimulate further research interest in our data?
- Additional data types? Modified data access policies, consistent with applicable legal and regulatory measures? Create more awareness in the research community?

DISCUSSION