SANDIA REPORT
SAND2010-4765
Unlimited Release
Printed September 2010

Assessment of Current Cybersecurity Practices in the Public Domain: Cyber Indications and Warnings Domain

Curtis M. Keliiaa and Jason R. Hamlet

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

NOTICE: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone:	(865) 576-8401
Facsimile:	(865) 576-5728
E-Mail:	reports@adonis.osti.gov
Online ordering:	http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd.
Springfield, VA 22161

Telephone:	(800) 553-6847
Facsimile:	(703) 605-6900
E-Mail:	orders@ntis.fedworld.gov
Online order:	http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2010-4765
Unlimited Release
Printed September 2010

Assessment of Current Cybersecurity Practices in the Public Domain: Cyber Indications and Warnings Domain

Curtis M. Keliiaa
Advanced Networking Integration

Jason R. Hamlet
Assurance Technology and Assessments

Sandia National Laboratories
P.O. Box 5800
Albuquerque, New Mexico 87185-MS0806

Abstract

This report assesses current public domain cyber security practices with respect to cyber indications and warnings. It describes cybersecurity industry and government activities, including cybersecurity tools, methods, practices, and international and government-wide initiatives known to be impacting current practice. Of particular note are the U.S. Government's Trusted Internet Connection (TIC) and "Einstein" programs, which are serving to consolidate the Government's internet access points and to provide some capability to monitor and mitigate cyber attacks. Next, this report catalogs activities undertaken by various industry and government entities. In addition, it assesses the benchmarks of HPC capability and other HPC attributes that may lend themselves to assist in the solution of this problem. This report draws few conclusions, as it is intended to assess current practice in preparation for future work, however, no explicit references to HPC usage for the purpose of analyzing cyber infrastructure in near-real-time were found in the current practice.

This report and a related SAND2010-4766 *National Cyber Defense High Performance Computing and Analysis: Concepts, Planning and Roadmap* report are intended to provoke discussion throughout a broad audience about developing a cohesive HPC centric solution to wide-area cybersecurity problems.

ACKNOWLEDGMENTS

Appreciation is due to the following individuals for their contributions in the development of this report.

Tan Chang Hu, 05627 Assurance Tech & Assessments
David G. Robinson, 01415 Computer Science & Informatics
Anand Ganti, 09336 Advanced Networking Integration
Vitus J. Leung, 01415 Computer Science & Informatics
Keith Vanderveen, 08961, Manager, Scalable Computing R&D
Robert A. Ballance, 09328 Scientific Computing Systems
Steve A. Gossage, 09336, Advanced Networking Integration
Kevin Pedretti, 01423, Computer Systems Analysis

CONTENTS

FIGURES

TABLES

1. EXECUTIVE SUMMARY

This report exists in the larger context as the second in a series of three related reports. The first, *A Sensor Network Study* [1] provides a detailed description of sensor networks, which will likely be a critical component for collection, filtering, and initial analysis of cybersecurity data in distributed, high performance computing (HPC) oriented wide-area cybersecurity architectures. The third, *The National Cyber Defense High Performance Computing and Analysis: Concepts, Planning & Roadmap*[2] represents an initial concepts, planning and roadmap effort toward comprehensive cybersecurity HPC analysis to alleviate the cybersecurity dilemma on a national scale. This report discusses current, unclassified cybersecurity practice and finds a lack of cybersecurity application of HPC, but identifies abundant opportunity for HPC.

This report assesses current (public domain) cyber security practices with respect to cyber indications and warnings. The information collected is in preparation for evaluation of the advantages of applying HPC technology to cybersecurity, as well as to identify other advances required to properly address this problem space.

This report provides a high level, generalized view. Starting with section 2, it describes cybersecurity industry and government activities, including cybersecurity tools, methods, practices, and international and government-wide initiatives known to be impacting current practice. Of particular note are the U.S. Government's Trusted Internet Connection (TIC) and "Einstein" programs, which are serving to consolidate the Government's internet access points and to provide some capability to monitor and mitigate cyber attacks.

Section 3 of this report catalogs activities undertaken by various industry and government entities. In addition, it assesses the benchmarks of HPC capability and other HPC attributes that may lend themselves to assist in the solution of this problem. This report draws few conclusions, as it is intended to assess current practice in preparation for future work, however, no explicit references to HPC usage for the purpose of analyzing cyber infrastructure in near-real-time were found in the current practice.

After surveying the current practice in cyber monitoring, we note that:

1. Large networks can be viewed as complex systems exhibiting emergent behavior. Without a more profound understanding of these complex systems and their behavior it is difficult to learn how to protect them. HPC can provide a platform for simulation of large-scale networks and discovery of their behaviors.

2. Deep packet inspection requires high-speed pattern matching on large datasets and, if implemented with sufficient throughput, could be used for effective IDS/IPS and filtering of http traffic.

[1] SAND2009-6671 P A Sensor Network Study, January 2010

[2] Keliiaa, C.M. and Hamlet, J.R. SAND2010-4766 *National Cyber Defense High Performance Computing and Analysis: Concepts, Planning, and Roadmap,* August 2010.

3. Large-scale computation is needed for analysis of vast amounts of cybersecurity data collected across a wide area over time. HPC systems offer a means for large-scale computational capacity and management. In tandem, high-speed sensor networks offer a means of collection across a wide area over time.

4. Correlation and analysis algorithms for wide area cybersecurity applications are lacking. Further development of these algorithms is necessary to collect, transport, store, and conduct high performance analysis of cybersecurity data.

Section 3.3 summarizes possible directions for applying high performance computing to cybersecurity. Briefly, these directions are to:

1. Identify trusted connections and automated process opportunities for collecting and analyzing data or cybersecurity applications.

2. Examine informatics and statistical TCP/IP anomalous behavior research to detect anomalous and malicious protocol characteristics; analysis of temporal and spatial characteristics of attack; harvesting data to seed Informatics visual analytics for cybersecurity subject matter.

3. Examine cybersecurity mathematical and statistical analysis research to collect, handle, and analyze large datasets for modeling, intrusion detection, attack response, and identification of multistage attacks.

4. Examine cybersecurity complexity science analysis research for studying the unpredictability in programs, machines, and networks, and large-scale modeling and simulation.

5. Examine modeling, simulation and analysis of complex networked systems, including large scale network models and models of network dynamics and cyber attack. Applications include intrusion detection, and examining the evolution of cyber threats.

6. Expand HPC analysis and correlation algorithms for identification of temporal and spatial characteristics associated with anomalous events, and detection of widespread, multistage, multiple method attacks.

7. Understand the sociology and psychology of cyber engagement for solving the problem of attributing an attack to responsible parties, and to aid in the development of anticipatory and preventative safeguards and countermeasures based on predictable behaviors.

2. CYBER INDICATION AND WARNINGS DOMAIN

2.1 Overview

This report provides a general assessment of cybersecurity, high-speed, real-time computing infrastructures and high performance computing (HPC) in the unclassified Cybersecurity Indications and Warnings Domain (CIWD). A comprehensive HPC cybersecurity solution that serves the intersection of these technology areas requires that the complete problem space be understood and addressed. Cybersecurity is introduced as a multipart problem space comprised of cybersecurity information management, technology, and sociology.

The national cyber risk management challenge to secure the national information infrastructure is critical. There is a growing need for high performance analysis of threat, high-speed sensing of events, correlation of events, and decision-making based on the adverse events seen across multiple and independent large-scale network environments.

We assess the current state of the industry to provide background information for determining the most efficient and beneficial path forward. We find that HPC large-scale computation and cybersecurity data collection are essential to address the scope and scale of a global cyber threat. Further, the feasibility of such a system is dependent on collection of data, transport of data for analysis, and large-scale computation and analysis.

To this end, high-speed sensor networks are fundamental to CIWD data collection in high-speed, real-time computing infrastructures. HPC large-scale computation is fundamental for analysis of very-large data sets. A high-speed real-time computing infrastructure for cybersecurity will consist primarily of:

1. System-of-systems network and HPC system architectures
2. Wired or wireless sensors performing data collection
3. Communication of sensor data to a (potentially distributed) analysis center
4. Database and storage
5. Algorithms for data correlation and analysis

Implicit to this infrastructure is the need for a complete system architecture that includes storage, file system, and database systems capable of simultaneously supporting multiple fast read and write operations on extremely large data sets. Additionally, the system I/O must be capable of performing large data transactions without congestion, and the architecture must support real time, interactive operations.

HPC provides the large-scale computation necessary for very large dataset and complex associative analysis needed for cybersecurity applications. Consider for example that data collected across a wide area could potentially be national or international in scope and over time frames extending into years.

There is ample evidence of opportunity for increased cybersecurity utilization of high-speed sensor networks and HPC analysis of threat, although many organizational and technical challenges exist. The foremost challenge and goal can be stated as moving from a reactionary response posture to an assured, predictive and anticipatory defense posture for a more secure national information infrastructure.

This evaluation focuses on the current state of the industry. Table 1 illustrates the primary areas of interest addressed in this report and a basic deconstruction of the problem domain.

Table 1. Cyber Indications & Warnings Domain

CYBERSECURITY	HIGH-SPEED, REAL-TIME COMPUTING INFRASTRUCTURES	HIGH PERFORMANCE COMPUTING	HIGH PERFORMANCE ANALYSIS
• Information Management • Technology • Sociology	• System-of-Systems Network and System Architectures • Sensor Data • Correlation Algorithms	• HPC System Architecture • HPC I/O Architecture • Class of Systems • Performance Benchmarks	• Codes • Analysis Algorithms • Database and Storage • Complexity • Emergent Behavior • Mathematics

2.2. Value Proposition

Advances to cybersecurity in the national information infrastructure are achievable through greater utilization of high-speed sensor networks for wide area data collection, HPC large-scale computation and database storage for high performance analysis of threat, and continued development of advanced data correlation and analysis algorithms.

High performance analysis of threat is the precursor to high performance response to attack. Coordinated response to attack across the national information infrastructure and national preparedness are needed in a world where cyber warfare is an emerging threat. HPC large-scale computation will permit studies on a global economy of scale to adequately assess cyber risk. The resulting knowledge of high performance analysis will support decision and policy makers with high value data for the development of national cyber risk management strategy, and to design, test, and build risk mitigation tools and techniques. High performance analysis will potentially lead to more effective predictive and anticipatory defense safeguards and countermeasures.

Research in these areas could provide a multitier path for achieving benefits and advantages such as:

1. Decreased time to detect intrusion and new threats
2. Shorter time to resolution of vulnerabilities and exposure
3. In-depth analysis of threats and exploitation
4. Faster development of safeguards and countermeasures
5. Faster time to test and develop new defenses

6. Faster fielded response to attack
7. Faster time to deploy new defenses
8. Coordinated response to wide area attack
9. Collaborative cybersecurity risk management
10. Increased understanding of normal and abnormal network behavior
11. Improved models for detecting anomalous events

These potential advantages could provide services needed for advancements in cybersecurity on the national and international stage.

Areas of interest we have identified include the advancement of informatics, complexity science, emergent behaviors within complex systems, and the psychology of attack and defense in cyberspace. Increased knowledge is sought in the areas of:

- Attribution of Attack
- Methods of Attack
- Internet Traffic Characterization
  - Command and Control Methods
  - Covert Channel Methods
  - Exploitation (IRC, HTTP)
- Geographic Cyberterrain Characterization
  - Origin of Attack
  - Aggregate Points of Control, Compromise, and Data Exfiltration
  - Replication and Proliferation Methods
  - Local, Regional, National, and International Systemic Vulnerabilities
- Temporal Threat Characterization
- Threat Event Correlation and Characterization
- Polymorphic, Spatial and Temporal Models and Methods of Attack
  - Widespread Attack
  - Multistage Attack
  - Sophisticated Multi-Methods of Attack
  - Low and Slow Attack

The opportunities and needs for advancement are many. More fundamental is the application of high performance resources to address the global cyber-threat. Such resources provide a feasible path on a scale that matches the problem.

Cybersecurity data collection is the starting point for analysis. The findings herein indicate that high-speed sensor networks are a fundamental element of a CIWD HPC system. Wireless and wired (i.e. typical intranet and Internet infrastructure) well resourced sensor networks are best suited for gathering and processing applications such as deep packet inspection (DPI). Furthermore, constrained resource sensor networks, such as wireless ad hoc sensor networks are better suited for lighter weight applications, such as detection and notification for situational awareness. High performance analysis is achievable with massively parallel computing platforms with storage, I/O, and architecture designed to match distributed large data set collection across a wide area.

Cyber risk management is necessary in the face of an escalating threat to the confidentiality, integrity, and availability of the national information infrastructure. In addition, excessive technical and implementation vulnerabilities, fragmented information management methods and disjointed response to attack increase the vulnerability and exposure to intrusion and compromise. Solutions to these problems are complicated by the difficulty of fragmented information management across the wide-area, the complexity of the technology involved, and sociological factors.

2.2.1. Cybersecurity Information Management

The broad challenge is to adequately protect against unauthorized disclosure and appropriately share information. A discussion of cybersecurity sensor networks and centralized analysis of data collected from a wide area and multiple organizations will inevitably have to deal with the responsibilities of data ownership and data stewardship. This is evidenced by a growing field of national and international cybersecurity law. Legal issues that may complicate the sharing of cybersecurity data across domains include the handling of export controlled information, proprietary information, third party proprietary information, international law and treaties, legal jurisdiction and investigation, personally identifying information, intellectual property, trade secrets and trademarks, and sensitive vulnerability and threat information. Interconnection security agreements also need to be considered.

Current policy, law and directive lag the need to coordinate cybersecurity information exchange across domains. A foundation for coordination of policy and practice between organizations that share information is needed to define the requirements for inter-organizational information management. Vehicles to manage the control of information between organizations include: contractual agreements, non-disclosure agreements, memorandums of understanding or agreement, and chains of trust. There are legal consequences to consider, for example a cybersecurity professional must have appropriate written authority when scanning and assessing the vulnerabilities of a system or face prosecution as an attacker. Criminal investigations that involve evidentiary cyber data must be appropriately handled with documented chains of custody or risk failure of prosecution.

2.2.2. Technology

The Internet has evolved from a trusted environment for relatively few academic and government users to an untrusted environment of global use with applications unimagined by its inventors. Over time government, business, academic, public safety, and supervisory control and data acquisition (SCADA) system interconnectivity and operation have become critically dependent on the Internet. The breadth and scope of Internet connectivity magnifies the level of threat and the massive amounts of cybersecurity data that must be managed.

Major vendors have built sophisticated features and function for the purpose of generating revenue. The economics of investment, research and design (R&D) and delivery of a product to market often mean that security is not a critical design criterion. In addition, vendors have to integrate and accommodate rapid growth in enterprise system technology, global connectivity, and mobility to remain competitive. An industry wide coordinated commitment to address secure software and trusted hardware standards remains elusive.

Enterprise networks are highly complex as each enterprise is uniquely architected with specific technical investments and implementation methods. Consider this brief description of technologies as an example of complexity: public and internal facing web services, physical machine and virtual machine datacenters, network technologies such as passive optical networks, multi-protocol label switching (MPLS), IP version 4 (IPv4) and version 6 (IPv6), mobility, and backend systems such as identity and access management (I&AM), enterprise job processing and operational support systems. All of the constituent components are applied and orchestrated toward specific business purposes of each organization and sector. There is great diversity in the design, application and use of technology and no organization is the same.

Why is this important? These technologies, systems and their implementation are fraught with vulnerability and present a significant attack surface. Automated tools are abundantly available for the reconnaissance, enumeration, and exploitation of Internet connected networks. Yet technology may be the easiest part of the problem to solve.

2.2.3. Sociology

People design, build and use information technology for the purpose of information exchange. An understanding of the social diversity of those that build, manage and use technology is essential to addressing the cybersecurity dilemma because people have everything to do with intrusion and compromise. Cybersecurity is much more than a technology problem.

Each organization represents a unique culture that brings forward a history of the organization, for example the sociology of users and managers can greatly influence risk management decisions and investments. In addition, there are subcultures within the IT communities that engineer, build, deliver, manage, operate and use technology. Mechanical, electrical, systems, and network engineering are all distinct disciplines. Computer science is divided into mathematics, R&D, programming, and more as computing advances continue, for example, toward quantum computing. System administration, database management, network management and cybersecurity are distinct disciplines as well. These communities face the challenge of delivering a common response to threat. The mission of cybersecurity raises the need for a collective response from each discipline area as innovation, automation and integration push technology to new heights.

“Cyber” is one of those terms that mean something different to almost everyone. Consider that virtual worlds such as Second Life are common. Massively multi-player online gaming and social networking are today embedded in a global social fabric. The public at large is much more trusting and careless in the protection of personal information in this new paradigm. Unfortunately, this presents an opportunity for those that prey on the innocent, such as criminal enterprises that profit from identity theft.

The sophisticated well resourced cyber-threat has emerged on a global scale. Two primary vectors of attack are at play; external cyber attack and the insider threat.

It is commonly accepted that the insider threat is the most dangerous and difficult to mitigate.

On a broad scale organizations are faced with securing information in the midst of the rapid proliferation of technology, extended interconnectivity through partner networks and the Internet, and expanded access to information for communities of interest (COI) and the public.

"*The United States may be facing the most serious economic and national security crisis of the 21st century. Our government and private sector networks are being exploited at an unprecedented scale by a growing array of state and non-state actors*" [Hathaway].

2.2.4. Summary

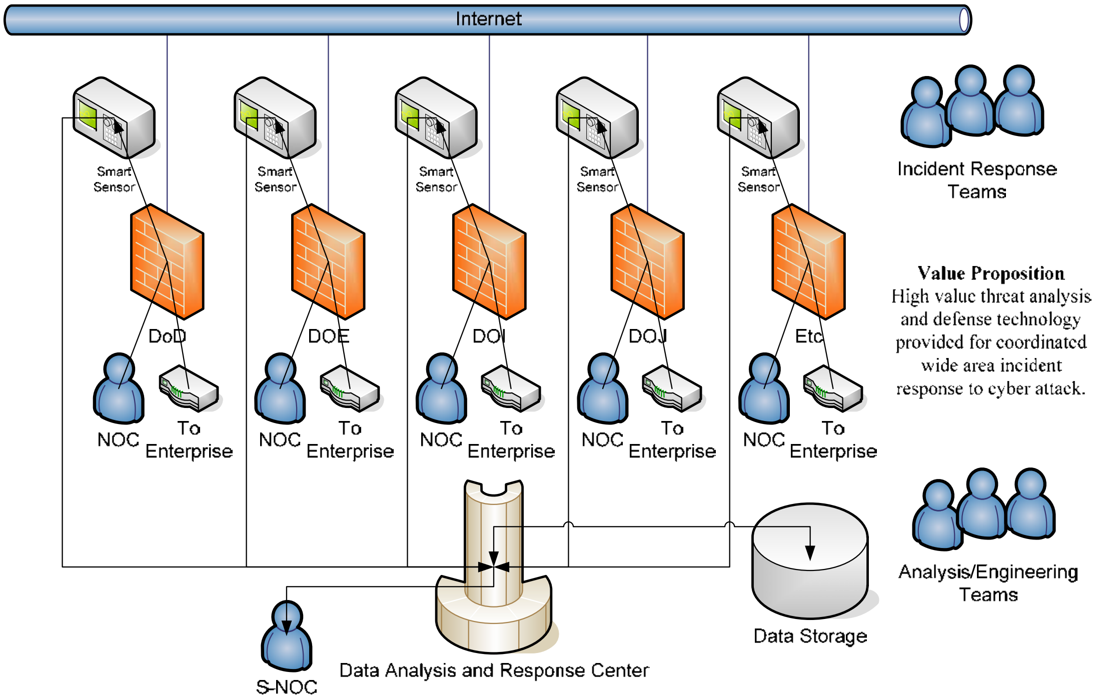
HPC systems offer a means for large-scale computation, which is needed for analysis of vast amounts of cybersecurity data collected across a wide area over time. In tandem, high-speed sensor networks offer a means of data collection across a wide-area over time.

Correlation and analysis algorithms for wide-area cybersecurity applications are lacking. Further development of these algorithms is necessary to collect, transport, store, and conduct high performance analysis of cybersecurity data.

The need for high performance analysis of threat and high-speed sensing of events, correlation of events, and decision-making based on the adverse events is evident in the level risk to digitized information. The scale and scope of threat to information is unprecedented in the face of continuous and automated cyber attack.

Wide spread operational dependence on the Internet has created shared risks to information across the wide-area. High-speed, real-time computing infrastructures, HPC and high performance analysis comprise the desired future state capability needed to address shared risk and secure the national information infrastructure.

As the national focus on cyber increases there is an evolving need for a capability to provide for near real-time sensing of events, correlation of independent events, and decision-making based on adverse events seen across multiple, independent, large-scale network environments. Security operations, pattern recognition and courses of action at a national level require technological advancements and political innovation. The long term objective of this effort is to realize a widely distributed data collection and analysis capability similar to Figure 1.



Value Proposition
High value threat analysis and defense technology provided for coordinated wide area incident response to cyber attack.

Figure 1. Conceptual Architecture

The remainder of this report presents a survey of existing cybersecurity tools, methods and practices that are relevant to the utilization of HPC cybersecurity analysis and high-speed real-time computing infrastructure in the unclassified Cybersecurity Indications and Warnings Domain (CIWD).

3. FINDINGS AND RECOMMENDATIONS

The private, public and USG unclassified CIWD includes new initiatives for distributed detection, notification, analysis and sharing of cyber security event information from multiple, independent, large scale network environments. The unclassified and collaborative CIWD is not encumbered with the requisite physical and logical safeguards and countermeasures needed for more sensitive environments. Although the unclassified realm is less stringent, information assurance and effective cybersecurity safeguards and countermeasures are none the less complex and essential to a sustainable national information infrastructure.

Additionally, the unclassified domain offers abundant opportunity to engage the academic and public sectors for increased assurance, for example in support of the National Science Foundation (NSF) National Cyber Infrastructure (NCI).

Here we focus on the feasibility and applicability of high-speed real-time computing infrastructures in the current state of the industry and government.

3.1. Cybersecurity Industry and Government Activities

3.1.1. Threats, Vulnerability and Exposure

Vulnerabilities are numerous and can be compromised in a variety of ways. For example, software distribution attacks wherein an attacker places a Trojan horse backdoor version of a commonly used tool on a web site used to distribute the tool. Domain Name Server (DNS) or other critical operational systems and software (OSS) service exploitations exist. Exploits such as a buffer overflow allow an attacker to gain root access. Rootkits allow an attacker to maintain root access. Botnets and polymorphic malware designed to avoid detection have become a common threat on the Internet. “Malicious code (or malware) has become one of the most pressing security problems on the Internet. In particular, this is true for bots, a type of malware that is written with the intent of taking control over hosts on the Internet.” [Stone-Gross].

Cyber vulnerability and exploitation information is available from many commercial and government sources. One example is the NIST Computer Security Divisions Computer Security Resource Center (CSRC) which provides a reading room of general Internet security vulnerabilities and exposures [CSRC]. Another authoritative source of information is the United States Computer Emergency Readiness Team [CERT].

The National Vulnerability Database also is an authoritative source of cybersecurity vulnerability and exploitation information [NVD]. Of several common vulnerability formats the common vulnerabilities and exposures (CVE) is the most prevalent [CVE]. Similar standards include the common configuration enumeration (CCE) standard, and the open vulnerability assessment language (OVAL) [OVAL].

Vulnerability scanning tools are also available, such as IBM Internet Security Systems (ISS) Internet Scanner [ISS]. ISS Internet scanner is integrated with CVE definitions and provides an evaluation of high, medium and low technical risks.

Other vulnerability scanning and penetration testing tools include the metasploit exploitation tools, the cheops-ng network mapper, and Nikto, a web-service scanner that identifies well known server problems. Many other tools are available [SANS].

3.1.2. Federal Safeguards and Countermeasures Guidance

Cyber safeguards and countermeasures are implemented through policy, process and technical controls. NIST guidance categorizes information security controls by family and class (management, operational and technical) [NIST800-53]. Technical controls include access control, identification and authentication, system and communications protection, and auditing and accountability. Management controls include security assessment and authorization, planning, risk assessment, and system services and acquisition. Operational controls include awareness and training, configuration management, contingency planning, incident response, maintenance, media protection, physical and environmental protection, personnel security, and system and information integrity.

3.1.3. The Cybersecurity Industry

The industry at large is responding to the need for better cybersecurity. The SysAdmin, Audit, Network, Security (SANS) [SANS] and the International Information Systems Security Certification Consortium (ISC^2) are the leading industry cybersecurity organizations [SANS, ISC2]. Both organizations serve as certification authorities for IT professionals. SANS provides training for cybersecurity tools and technology and the (ISC)^2 provides requisite understanding of cybersecurity methods and practices.

3.1.4. The SANS (SysAdmin, Audit, Network, Security) Institute

The SANS certifications offer assurance that incident first responders and investigators that perform preservation of evidentiary information, immediate damage control, and damage assessments for investigation are trained in the complexity of technology in preparation for field incident response and day-to-day maintenance and operations. System administrators and network professionals can also be certified.

3.1.5. The SANS Internet Storm Center

SANS administers the Internet Storm Center which provides threat level assessment, trends, reports and diary style blogs for security professionals. “Thousands of sensors that work with most firewalls, intrusion detection systems, home broadband devices, and nearly all operating systems are constantly collecting information about unwanted traffic arriving from the Internet. These devices feed the DShield database where human volunteers as well as machines pour through the data looking for abnormal trends and behavior. The resulting analysis is posted to the ISC's main web page where it can be automatically retrieved by simple scripts or can be viewed in near real time by any Internet user [ISC].”

3.1.6. International Information Systems Security Certification Consortium

The International Information Systems Security Certification Consortium
(ISC)^2 offers the Certified Information System Security Professional (CISSP) as its flagship certification, which is a Department of Defense (DoD) requirement (DoD Directive (DoDD) 8570) for information assurance government and government contractor employment. This training is well suited for cybersecurity decision makers, stakeholders and administrators. (ISC)^2

also offers concentrations in architecture, engineering and management and also a Certified Secure Software Lifecycle professional certification [ISC2].

The $(ISC)^2$ CISSP certification covers access control, application security, business continuity and disaster recovery planning, cryptography, information security and risk management, legal, regulations, compliance and investigations, operations security, physical and environmental security, security architecture and design, and telecommunications and network security emphasizing the broad scope of cyber security.

Recently, SANS and $(ISC)^2$ have partnered in support of comprehensive information assurance training, education and awareness. SANS and $(ISC)^2$ and cybersecurity vendors are dependent on voluntary participation and reporting of attempted and successful attacks and exploitation of cyber vulnerabilities and exposures.

3.1.7. Cybersecurity Tools Methods & Practices

Cybersecurity tools, methods and practices include a host of security devices. Perimeter security protection mechanisms include firewalls and Internet facing controlled interfaces such as proxy servers and one-way interfaces. Edge and internal safeguards and countermeasures include intrusion detection systems (IDS), intrusion prevention systems (IPS), vulnerability scanners, system log servers, and honeypots.

Honeypots are network decoys that are monitored for intrusion and record malicious or unauthorized use. Early warning system honeypots are used for early warning and recognition of malicious activity, permitting rapid response to attack. These honeypots should not have access to secure information and are used only as a security resource. Low-emulation honeypots may only capture network traffic logs, intrusion detection system logs, and honeypot log files. High-emulation honeypots may be real systems or virtual machine sessions.

In addition to other resources, such as the SANS ISC, and national vulnerability database previously noted, is the Active Threat Level Analysis System (ATLAS**).** ATLAS is a global threat analysis network maintained by Arbor Networks [Atlas]. Data is collected from a distributed network of sensors with data acquisition and analysis capabilities. The data sources include honeypot payloads, IDS and scan logs, DoS statistics, news and vulnerability reports, malware samples, and information on phishing infrastructure and botnet command and control.

Cybersecurity continues to evolve and university programs are now in place to address the educational needs of the cybersecurity profession. Examples include the information Networking Institute (INI), which is a Carnegie Mellon graduate degree program, and the Massachusetts Institute of Technology cybersecurity efforts [INI, MIT].

3.1.8. Virtualized Environment Security

A significant trend for information service delivery is virtualization. Virtualization is the software representation of discrete systems such as application servers that can be deployed quickly. Virtual architectural models encompass virtualized single systems, multiple systems, network infrastructure, and security devices. Virtualization can be described as software automation.

Virtualization of systems (virtual machines and hypervisors) has evolved into datacenter virtualization, which includes the virtualization of network and security devices. Virtualization offers significant cost benefit when compared to maintaining a datacenter with many physical servers and associated HVAC, power, and operational support infrastructure.

Automation is a sophisticated and complex tool kit and as such can be used to realize great benefit. However, automation can also rapidly propagate problems. One study enumerates security threats as total compromise, partial compromise and abnormal termination [Ormandy]. This report describes a hostile virtualized environment where untrusted code is being executed or when untrusted data is being processed by services inside the virtual machine (VM). The report discusses evaluation of various VM implementation methods and tools, and concludes with recommendations for safely deploying virtualization, such as treating each VM as a service that can be compromised and removing or disabling unnecessary emulated hardware and modules. Strict configuration control is a necessity in virtualized environments.

Standards for network management are provided by the Tele-Management Forum and the Distributed Management Task Force (DMTF). DMTF has recently built an Interoperability Program for Virtualization Management [TMF, DMTF, IPVM].

3.1.9. Secure Software and Assurance

Software assurance is a critical requirement for security collection and analysis as raw data may contain malicious code and sensitive information. NIST SP 500-268 defines a minimum capability to assist in understanding and meeting software security assurance needs [NIST]. A textual scan of source code is performed against a set of common code weaknesses to increase assurance that the code meets certain security standards.

The NIST Software Assurance Metrics and Tool Evaluation (SAMATE) project strives to quantify and improve the state of the art for different classes of software security assurance tools [SAMATE]. The SAMATE web site provides further information and taxonomy of software assurance (SA) tools.

The National Information Assurance Partnership (NIAP) is an NSA and NIST program to evaluate IT product conformance to international standards [NIAP].

Due the level of malicious activity on the internet there is an increased need for software assurance. Resilient software assurance is needed throughout industry and for custom or commodity HPC platforms.

3.1.10. International Activities and Government Wide Initiatives

Government is also responding to the growing global cyber threat. International technical guidance for cybersecurity is available through the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) [IOS/IEC]. For example, guidance for implementation and operation of an Information Security Management System (ISMS) is provided per the International Standard ISO/IEC 27001.

New international approaches to the cyber dilemma are also developing. For example on July 10, 2009 legislation was introduced in Senate by Sen. Kirsten Gillibrand (D-N.Y.) to express the sense of Congress on improving cybersecurity globally. The bill, known as 'Fostering a Global Response to Cyber Attacks Act', would require the Secretary of State to submit a report to Congress on improving cybersecurity [Bain, FGRCCA].

Other international efforts to secure cyberspace include the 2006 International Cyber Crimes Treaty and the European Electronic Crime Task Force. The European Electronic Crime Task Force is charged with preventing identity theft, computer hacking and other computer-based crime.

In the case of cybersecurity threat data collection and analysis, the USG information sharing environment (ISE) may offer insight. Examples include the Suspicious Activity Reporting (SAR) standard, the privacy guidelines issued and the information sharing architecture many have adopted. Further information is available at the ISE website [ISE].

The US Computer Emergency Readiness Team (CERT) is the operational arm of the National Cyber Security Division, NCSD, at the Department of Homeland Security, DHS. As a part of the NCSD their main objectives are:

1. Prevent cyber attacks against America's critical infrastructures
2. Reduce national vulnerability to cyber attacks
3. Minimize damage and recovery time from cyber attacks that do occur

There is collaboration between states, educational institutions (Carnegie Mellon, CERT Coordination Center), industry, and international partners. Since new vulnerabilities and attacks are discovered and created continuously, the response needs to be just as swift. CERT specific actions within the NCSD are to act as this continuous response to cyber attacks.

3.1.11. U.S. Government Requirements

Government compliance is mandated in several arenas through legislation and directive. Relevant policy and law include various acts of Congress and executive orders [NSAct47, HIPPA, GLB, SOA2002, FISMA, HSA2002, ITMRA, EO12333, EO12968, EO13388, FOIA].

Government compliance is also guided by several NIST publications and special publications [NIST1, FIPS, FIPS199, FIPS200, FIPS201, NIST800-53, NIST800-48, NIST800-113, NIST800-97, NISTITL, NISTSP500-529, NISTSP500-268].

3.1.12. The Comprehensive National Cyber Security Initiative

The Comprehensive National Cybersecurity Initiative (CNCI) is a broad initiative to proactively and more effectively secure government computer systems. The CNCI was formalized by presidential directive in 2008 and is intended to protect against foreign and domestic intruders and prepare for future threats. CNCI was initiated by HSPD23 and NSPD54.

3.1.13. Trusted Internet Connections (TIC)

The Trusted Internet Connections (TIC) initiative is a federal government reduction of external connections, including internet points of presence, to fewer than 100 connections [CJohnson]. There are expected to be approximately 79 gateways [Mosquera].

The TIC program is intended to optimize and standardize individual external network connections and to improve the federal government's security posture and incidence response capability through the reduction of external connections and via centralized gateway monitoring at a select group of TIC Access Providers (TICAPs) [TICSOC]. TICs should improve the federal government's cybersecurity by making it easier to monitor traffic and under the assumption that fewer connections imply fewer vulnerabilities which implies better networks. Einstein will be used to monitor at the gateways. A conceptual diagram of the TIC architecture is present in Figure 2 [TICSOC]. Several commercial entities are advertising their TIC solutions [Juniper, CISCOTIC].

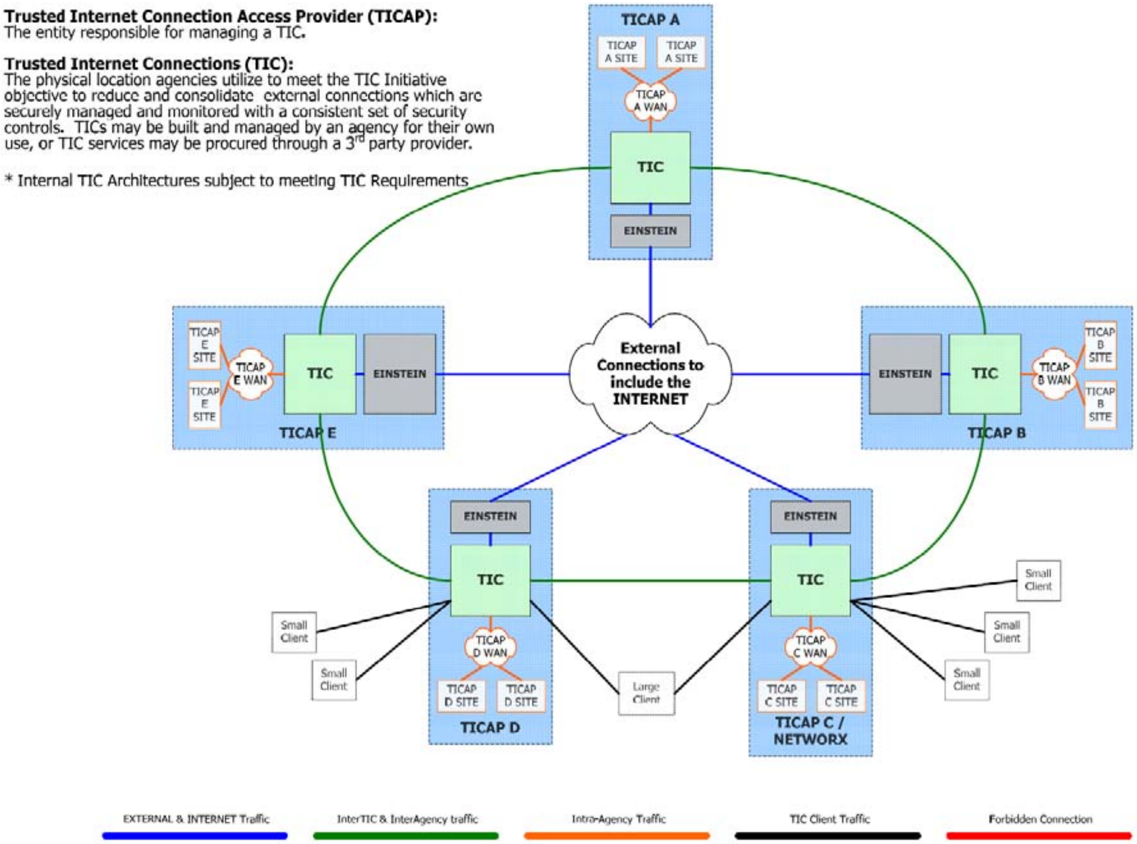

Trusted Internet Connection Access Provider (TICAP):
The entity responsible for managing a TIC.
Trusted Internet Connections (TIC):
The physical location agencies utilize to meet the TIC Initiative objective to reduce and consolidate external connections which are securely managed and monitored with a consistent set of security controls. TICs may be built and managed by an agency for their own use, or TIC services may be procured through a 3rd party provider.
* Internal TIC Architectures subject to meeting TIC Requirements
TICAP A
TICAP A SITE
TICAP A SITE
TICAP A WAN
TIC
EINSTEIN
TICAP E SITE
TICAP E SITE
TICAP E WAN
TIC
EINSTEIN
TICAP E
External Connections to include the INTERNET
EINSTEIN
TIC
TICAP B WAN
TICAP B SITE
TICAP B SITE
TICAP B
EINSTEIN
TIC
TICAP D WAN
TICAP D SITE
TICAP D SITE
TICAP D
EINSTEIN
TIC
TICAP C WAN
TICAP C SITE
TICAP C SITE
TICAP C / NETWORX
Small Client
Small Client
Large Client
Small Client
Small Client
Small Client
EXTERNAL & INTERNET Traffic
InterTIC & InterAgency traffic
Intra-Agency Traffic
TIC Client Traffic
Forbidden Connection

Figure 2. Trusted Internet Connections

3.1.14. Einstein Program

The Einstein program is a USG initiative for automated collection, correlation, analysis, and sharing of computer security information across the federal civilian government so that federal agencies will be aware, in near real-time, of the threats to their infrastructure. Einstein 1 collects data such as autonomous system numbers, ICMP type, packet length, protocol, sensor ID and connection status, source and destination IP addresses and ports, TCP flags, timestamps and duration information [EPIA].

Einstein 2 resulted from the TIC initiative. Einstein 2 sensors will be placed at all TICs and expand on the capabilities of Einstein 1 by performing signature based intrusion detection and anomaly based detection. The anomaly based detection will use statistical characteristics and behavioral, protocol, and traffic information and will monitor network flow [EPIA2].

Einstein 3 implements the same functionality as its predecessors, and has the additional ability of reading internet traffic and intercepting data content before it has a chance to reach a government system. Specific information on how Einstein 3 is used has not been forthcoming.

3.2. High Performance Computing Class of Systems and Benchmarks

The need for more sophisticated cybersecurity safeguards and countermeasures is increasing. In turn, the utility for HPC cybersecurity applications and analysis is increasing. An HPC assessment has been conducted via open-source resource research and references are included within the text.

For the purpose of this report "supercomputing" is broadly defined to include traditional massively parallel machines, distributed architectures, and application-specific hardware accelerated platforms.

Traditional HPC performance is measured by a variety of benchmarks, such as the Top500 which uses the Linpack "Highly Parallel Computing" benchmark that requires solving a dense system of linear equations, and the HPC Challenge benchmark that measures floating point execution performance, processor to processor communications, communication bandwidth and latency, and memory updates [TOP500][HPCBENCH].

The DARPA High Performance Computing Systems (HPCS) Scalable Synthetic Compact Applications Benchmarks (SSCA) is another approach to HPC performance benchmarking. [HPCS] HPCS is working towards trans-petaflop systems and decreasing the time-to-solution, which requires an assessment of processing including quantitative development time, tools, methods and metrics for comparative analysis, execution time and projected performance in an effort to determine what tradeoffs exist between execution time and development time and to leverage productivity metrics to integrate system specific capabilities with user specific needs.

The benchmarks include highly parallelizable and hard to parallelize tasks and file IO and are useful for ongoing efforts to apply linear algebra for analysis of email and keyword associations, and IP header analysis for traffic characterization and anomaly detection.

The NASA Advanced Supercomputing NAS Division [NAS] provides the NAS parallel benchmarks, which stem from the Numerical Aerodynamic Simulation (NAS) program [RNR-94-007]. The NAS benchmarks codes include sorting and linear algebra operations and highly parallelizable tasks.

3.2.1. Trends in High Performance Computing

There is an historical and ongoing increase in performance in HPC Top500 class systems. It is predicted that a 100 petaflop system will likely be realized by 2016 and an exascale system may be realized by 2019 [Gietl, GmbH Meuer]. The Gietl-Meurer paper identifies the major challenges to HPC processor requirements as: low cost, low power consumption, availability of support for parallel programming, and efficient porting of existing codes. Clustered multi-core highly parallel systems will increase in system scale but require advances in parallel software development and expanded bandwidth for memory access. Challenges also include increased power consumption as system-scale increases, software assurance, and hardware resilience and reliability.

There is a trend toward commoditization of teraflop computing capacity at the desktop [Ganapti]. In the referenced article, a four GPU Tesla personal supercomputer from Nvidia can offer 4 teraflops of parallel supercomputing performance with 960 cores and two Intel Xeon 5500 Series Nehalem processors [NVIDIA]. Another example is the Nvidia Compute Unified Device Architecture (CUDA) [CUDA]. CUDA is a software platform for massively parallel high performance computing. Personal supercomputing doesn’t provide the processing power of enterprise class supercomputing systems but provides significant computational power for smaller-scale applications of modeling and simulation. The resulting trend is that the utility of HPC is moving more mainstream.

Cloud computing is another developing high-performance platform for large-scale collaboration. In a March 2008 Wired Magazine article, NSF Director Arden Bement discussed a large-scale computing collaboration between Google, IBM, and the NSF [Madriagal]. These trends indicate continued mainstream expansion of distributed clusters or cloud computing.

In a variety of problem domains HPC offers high value as data availability mushrooms and global interconnectivity of the Internet continually expands, for example with Internet enabled cells phones, mobility, and IP version 6 (IPv6) adoption abroad.

3.2.2. Data Collection

An initial challenge to applying HPC to cybersecurity is getting the data in a consistent format for HPC analysis. In addition, cyber data collection and correlation methods need to accommodate the targeted collection of data. For example, targeted traffic could be http port 80 traffic through ingress and egress aggregation points, such as autonomous system border routers, to and from specific hosts at a specific time-of-day. Traffic data characterization and normalization will need to account for Internet communication that is inherently unreliable with traffic transported mostly via UDP and traversing a variety of infrastructure segments in route from source to destination. Infrastructure segments will also vary in availability and reliability, for example bandwidth, congestion, delay, jitter, latency, and Quality of Service (QoS) (prioritization, queuing techniques, traffic shaping and policing policies) parameters.

Interesting data comprises many views that include compute environment data, trace and log data, metadata of evidentiary and forensics value (such as timestamps, and access success or failure data), use case data, and identity (threat and defender) and entity (a service or device acting on behalf of a user) data. Data will need to be categorized and separated at the collection point and forwarded to analysis centers. Raw data may contain malicious code requiring containment for damage control and protection of sensitive information. In addition, maintaining data integrity across the wide area is a requirement for high fidelity analysis.

A common concern is that real data is needed to build sophisticated models that are representative of live-networks and communities that conduct information exchange. The issue is how to stay relevant in abstract analysis of constantly changing cyber environments. There is also a need to verify and validate models in near real time for high confidence results of analysis. Conversely, there is concern about how to generate and integrate synthetic data to enhance emulation of large scale modeling and simulation of computing and social networks. Data models will be needed to resolve how to utilize and integrate both real data collected from the wide-area and synthetic data generated for large scale modeling and simulation.

The sociology of cyber-threat introduces the need to incorporate identity and access management (I&AM) technologies such as directory services. The purpose is to provide personification data into models for simulating cybersecurity interests and emergent behaviors of complex cyber environments. Much of the unpredictable behavior of an engineered system of systems is based on human interaction and directive. The premise of personification can be further expanded to the concept of live massively multiplayer interactive networks for the study of the psychology of attack and defense behavior.

The Complexity Science Challenges in Cyber Security report [SNL2007] introduces the idea of Leadership Class Computing utilizing virtualization as a possible framework from emulating physical networks with a high fidelity. Another concept recently developed at Sandia is that of emulated analytical networks (network modeling and simulation) that could be instrumented through virtualized systems, networks and security devices to effectively couple physical networks with HPC emulated networks.

3.2.4. Applications of HPC to Cybersecurity

There is little publicly available work describing the application of HPC to the cybersecurity domain. This is beginning to change.

A very promising development in the network modeling and simulation domain for large network systems is the use of virtual machine (VM) technology. For example, computer scientists at Sandia National Laboratories in Livermore, Calif., have for the first time successfully demonstrated the ability to run more than a million Linux kernels as virtual machines. Sandia scientists used VMs and the Thunderbird supercomputing cluster for the demonstration.

The Sandia VM achievement may lead to effective simulation of botnets and eventually, of the networks of entire nations. The research may also lead to high-fidelity modeling of sections of the internet. This is important and related to the DHS’ identification of internet mapping as an area of interest. High-fidelity models of the internet on a network level, and emulation of internet functionality, will help us study poorly understood phenomena that occur on the internet [SNL09].

DARPA's National Cyber Range (NCR) is a DARPA contribution to the CNCI. The NCR is intended to provide a system that permits assessment of information assurance and survivability tools in a representative, simulated network environment [NCR]. The NCR will replicate large-scale, heterogeneous networks and enable Internet/Global Information Grid scale research that incorporates nation-state quality actors. The NCR has been described as a mechanism for formalizing the government's cyber war games [DGNCR, BLAND].

HPC potentially impacts NCR type activities by providing a means to conduct simulations of large-scale networks and to collect, store, analyze, and visualize data resulting from simulated attack and defense scenarios. Efforts such as the National Cyber Range could benefit from a multitier approach to cybersecurity data collection through high-speed sensor networks, virtualization, and HPC analysis. For example, well resourced, high performance (computation, communication, and power) sensor networks would provide deep packet inspection and signature analysis and high-speed wireless ad-hoc sensor networks (constrained computation, communication, and power) would provide situational awareness. HPC and virtualized cybersecurity instrumentation could enhance wide-area intrusion alert and notification with evaluation of system and application state, network and performance state, identity and entity data. Analysis results would conceivably be fed back to end points and sensors for response to suspected intrusion

In a 2008 presentation the DHS identified opportunities to leverage HPC for cybersecurity R&D. Examples provided in the presentation include complex, distributed simulations; static and runtime software analysis and software testing; and the analysis of malware feeds through large collections of AV engines. The DHS also identifies internet mapping as an important research area. In this context, the internet is treated as a dynamic system that we know little about [Thompson]. Mapping may help us study the stability of the system, provide situational awareness, and allow for measurement and analysis. Additionally, an internet map could impact resource utilization, measurement efficiency, and deployment of sensor and monitor systems. Mapping the internet, and the study and analysis of the resulting graphs, would provide valuable data for sustainable HPC high-speed sensor network design.

A presentation from PNNL identifies scalable string matching as fundamental to modern cybersecurity and well suited to HPC [Khaleel]. The author shows string matching at speeds in excess of 25000Gb/s on a Cray XMT [XMT].

A 2007 Sandia National Laboratories report discussing complexity science challenges in cyber security outlines several research directions that may be amenable to HPC. Amongst these is the study of unpredictability in programs, machines, and networks, complex systems modeling, and large-scale modeling and simulation. This approach treats computers and networks of computers as complex systems that demonstrate emergent behavior that have positive and negative impacts to cybersecurity [SNL0805].

Other recent work highlights the need for large-scale modeling, simulation, and analysis of complex networks. A presentation from LLNL states the need for mathematical and statistical models of complex networks that could be simulated to study dynamic network behavior at levels of fidelity ranging from individual bits and instructions to high level mathematical models

[BRASE]. Highly parallelized event-driven simulations would form the basis for these tools, and provide an avenue for development of HPC cybersecurity applications.

Two related reports describing mathematical challenges and statistical opportunities in cyber security also discuss modeling and simulation opportunities in cybersecurity that may be appropriate for HPC [SNL0805, LBNL]. In particular, *Mathematical Challenges in Cybersecurity* discusses the need for large-scale network models and models of network dynamics and cyber attacks. Applications of such models include testing algorithms for network defense, intrusion detection, modeling the spread of worms and viruses, and studying the evolution of cyber threats. *Statistical Opportunities in Cyber Security* discusses modeling, intrusion detection, attack response, and the need to collect, handle and analyze large sets of data. Particularly interesting is the reference to statistical algorithms for identifying multistage attacks [Stone-Gross2].

3.3. High Speed Sensor Networks

High-speed sensor technologies fall into two basic categories; well resourced sensor platforms such as agent-based on a laptop and constrained sensor platforms such as mobile ad hoc sensors.

There is a perception that wireless sensor networks are limited in resources and purpose. Although true in the case of wireless ad hoc sensor networks, wireless networks with wireless endpoint devices such as laptops are robust and well resourced.

Sensor networks can be characterized by their computational, communications, power, and mobility capabilities. This suggests that different platforms are useful for different purposes. For example, well resourced high-speed sensor networks can be utilized for more robust collection, processing and distribution of data, while constrained mobile ad hoc sensor networks could be utilized for rapid deployment of lightweight situational awareness sematic sensing or warning of intrusion and anomalous events.

3.3.1. Network Sensors

The network sensors and monitors described in the public domain can be categorized according to the types of information they collect and the intended purpose of this data collection. Some monitors are intended to provide a view of the overall network status. Such monitors typically measure ICMP or TCP times and include the Internet Traffic Report and Internet Health Report. Other sensors monitor network flow and potentially aggregate and analyze data from a collection of sensors. Data collected might include ASN, ICMP code, protocol, IP addresses and port numbers, packet length, TCP flags, timestamp and duration, and firewall logs [EPIA1, ISC]. Examples include Einstein 1 and the Internet Storm Center’s sensors. More sophisticated sensors also collect DoS statistics and data from signature based intrusion detection systems, honeypot payloads, scan logs. ATLAS and Einstein 2 are representative [EPIA2, ATLAS].

Areas of technology that show continued development in the International communities include IPv6, Radio Frequency Identity (RFID), and Global Positioning Systems (GPS). These technologies coupled with mobility and Web 2.0 social networking technologies such as blogs, Twitter and Facebook all have possible application in high-speed sensor network evolution.

Automated RFID sensors that are IPv6 addressed have potential in agricultural, environmental, and fauna tracking systems. Social networks have potential to alert responders in the event of eye witness observation of an incident such as with the recent Iran election social upheaval.

A brief list of sensor categories includes network flow sensors (Einstein 1 and 2, INC, Internet Traffic Report, Internet Health Report, ATLAS), Intrusion detection signature-based sensors (Einstein 1 and 2, Snort, OSSEC HIDS, Untangle, Guard Dog) and anomaly detection profile-based sensors (Einstein 2, ATLAS, Ourmon, Peakflow X, Riverbed Cascade, QRadar) [Ourmon, peakflow, riverbed,QRadar].

3.3.2. Wireless Sensor Networks

For the purpose of this report, wired and wireless sensor networks are considered a fundamental data source for cybersecurity analysis.

Wireless networks can be well resourced and are commonly deployed in intranet and Internet infrastructure. Sensors add an application layer of collecting data, which is typically based on event triggers, pattern recognition, signature data, or behavioral baseline data for anomaly detection.

Wireless ad hoc sensor networks that can consist of potentially mobile nodes that communicate with one other wirelessly, possibly without external control. These types of sensor networks have applications in area, environmental, and industrial monitoring. Examples of such networks include NOAA’s ARGOS and SEAMONSTER networks, SensorScope, PermaSense, and Glacsweb [ARGOS, SEAMONSTER, SensorScope, PermaSense, Glacsweb].

Surveys of wireless sensor networks are presented in [Ganti, Papageorgiou]. A comprehensive bibliography is maintained by [Krishnamachari].

3.3.3. Wireless Robust Security Networks

A wireless robust security network (RSN) is defined as a wireless security network that permits the creation of robust security network associations (RNSA) only. Such networks are needed for secure information exchange of cybersecurity information. The IEEE 8011i amendment to the 802.11 wireless security enhancements provides a framework for RSN. The RSN framework includes stations (STA), access points (AP) and authentication servers.

NIST SP 800-97 addresses NIST recommendations for establishing wireless RSN and provides a guide to 802.11i. The IEEE 802.11i amendment defines and discusses two mechanisms for data confidentiality and integrity for RSNA’s: the Temporal Key Integrity Protocol (TKIP) and Counter Mode Cipher Block Chaining Message Authentication Code Protocol (CCMP). The Extensible Authentication Protocol (EAP) is also recommended because it provides flexibility and integration of wireless local area networks (WLAN) into enterprise networks.

Further guidance is provided in the NIST SP 800-48 for securing legacy IEEE 802.11 wireless networks.

3.3.4. Wireless Ad Hoc Sensor Networks

The limited power, computational and communications capabilities of wireless ad-hoc sensor networks do not easily accommodate a high demand in any of the three key sensor resource areas. These limitations also affect the ability to push or pull data in a real-time fashion.

In ad hoc sensor networks, typically not all nodes are connected to one another and so the nodes dynamically route data through each another in an effort to reach their destination. Each node has limited power, memory, and processing capability. Additionally, and unlike well resourced sensor networks, wireless ad hoc sensor networks have shared, time varying, lossy channels with potentially time varying topology. Moreover, the medium access, routing, and transmission problems are coupled in sensor networks, and typically the algorithms for each must be power aware.

Wireless ad hoc sensor networks are particularly vulnerable to security threats due to their constrained computational, power, and communications resources. Additionally, insider attacks on sensor networks and potential responses to them have been described [ChrDav, WSN Security, Krishnamachari].

A more detailed discussion of sensor networks is available in [Ganti].

3.4. Algorithms

Industry review of current practice revealed no explicit references to HPC and advanced algorithm usage for the purpose of analyzing cyber infrastructure in near-real-time.

3.4.1. Deep Packet Inspection

Deep packet inspection (DPI) consists of analyzing network traffic on layers 2-7. DPI analyzes packet headers and contents, and enables analysis at all network layers across series of datagrams, which can provide insight to the source, destination, application, and intent of traffic. DPI permits stateful flow analysis and can be used to search for malware, spam, intrusions, or other predefined criteria, to collect traffic statistics, and to make determinations about what to do with a packet [ED]. DPI can be used to identify packet flows, which permits control actions to be determined from an accumulation of flow data. Heuristics can be included to identify behavior.

DPI operations include string and regular expression matching for signature identification, application layer gateway-based identification, and behavior based identification. In application layer gateway-based identification the application layer gateway identifies and parses the protocol of the control flow and then inspects the service flow. For behavior based identification historical data is used to heuristically determine intent. It is useful for analyzing activities, such as spam, that cannot be inspected by the protocol [HAUWEI].

Regular expressions can be used to define patterns of interest for signature matching. Software implementations include SNORT [snort], Bro [bro], and Linux L7-filter [L7] [AbuHmed].

For more details on the following algorithms and a comparison of their throughputs, refer to the survey paper [AbuHmed] and references therein, many of which are replicated below.

Algorithms for string matching include the brute force approach of character by character comparisons, and other more efficient approaches such as Knuth-Morris-Pratt [Knuth], Boyer-Moore [Boyer], Aho-Corasick [Aho] and variations [Alicherry Tan], the AC_BM algorithm [Coit], Wu-Manber [Wu], and Commentz Walter [Commentz-Walter]. In hardware, the most common implementations use parallel Bloom filters [Dharmapurikar], CAMs, or TCAMs in FPGAs. Additionally, hardware implementations often employ finite automata for finding regular expressions. Potential finite automata classifications used in DPI include nondeterministic finite automata (NFAs), compressed and uncompressed deterministic finite automata (DFAs), delayed input DFAs (D^2DFA) [Kumar1], and content addressed delayed input DFAs (CD^2FA) [Kumar2]. The D^2DFAs use 95% less memory than normal DFAs, while CD^2FAs have twice the throughput of uncompressed DFAs and require only 10% of the memory.

Some of the advances in automata theory may make DPI feasible for implementation in power, communications, and computationally constrained wireless sensor networks. The reduced memory requirements and increased throughputs of D^2FAs and CD^2FAs may be appropriate for implementation in low-power electronics. Additionally, the relatively large throughput of these techniques may permit them to perform real-time monitoring of the traffic in WSNs.

Performing DPI at line rates is non-trivial, making it well suited for implementation in specialized processors, dedicated hardware, network co-processors with multi-core CPUs [ED], and potentially HPC. HPC could also be used to store and process data offline, potentially performing detailed analysis of intrusion attempts and using model-building to learn the behaviors that precede attacks and aid in future intrusion detection capabilities. Such model-building could use data from immediately before an attack to allow real-time response, but could conceivably include a long history, potentially helping to identify “low and slow” attacks.

3.4.2. Statistical Algorithms

[Leland] studies the statistical self-similarity of Ethernet traffic, and [Kannan] uses Poisson processes to analyze large collections of connection logs and extract the structure of application sessions within the connections. That work is based on [Paxson], [Nuzman]. Moreover, Kannan’s work aims to understand the causal relationships between network activities and references a collection of other related work that is applicable to DPI for IDS/IPS. These include relating traffic received by a host to code later executed by the host [Costa, Crandall, Newsome], tracking an attacker moving amongst nodes [Staniford], determining which hosts infected other hosts [Kumar, Xie] and studying the behavior of attackers obscuring their identity by relaying traffic through compromised machines while attacking other machines [Blum, Staniford-Chen, Yoda, Zhang].

3.5. Overall System of Systems Architecture

An overall system architecture will require reference architectures, data models, and event correlation for effective information management across the wide area. Reference architectures and data models will provide a means of consistent system design and implementation for multiple autonomous systems. Inter-organizational event correlation will provide a consistent means of recognizing and categorizing an intrusion.

A notional systems-of-systems architecture for wide area cybersecurity data collection and analysis has emerged from the research. System elements include:

1. High Speed Sensor Networks
2. Correlation Algorithms
3. Distributed Data Transport
4. HPC Large-Scale Computation
5. Analysis Algorithms
6. HPC I/O Communications
7. High Performance Database and Storage

3.5.1. Reference Architectures and Data Models

The development of widely distributed sensor networks and HPC analysis for cybersecurity applications is a new area of work requiring further development. The TMF and DMTF offer insight into possible reference architectures and data models, for example the Common Information Model (CIM) "is an open standard that defines how managed elements in an IT environment are represented as a common set of objects and relationships between them. This is intended to allow consistent management of these managed elements, independent of their manufacturer or provider. One way to describe CIM is to say that it allows multiple parties to exchange management information about these managed elements. However, this falls short in expressing that CIM not only represents these managed elements and the management information, but also provides means to actively control and manage these elements. By using a common model of information, management software can be written once and work with many implementations of the common model without complex and costly conversion operations or loss of information." [CIM2]

3.5.2. Event Correlation

Event correlation is a significant challenge requiring a common understanding of the meaning and definition of events and the actions required when an event occurs.

Technology areas that provide insight into large scale event correlation include provisioning and de-provisioning of entitlements and permissions, an enterprise service bus (ESB), and Policy-based network management (PBNM) [ESB].

Provisioning/de-provisioning requires authoritative sources of information and decision logic based on an action that propagates change through the I&AM domain of information resources and assets. Corporate organizations such as human resources serve as a point of workflow initiation, for example for personnel changes to information repositories. The intention is for provisioning of entitlements to information resources and assets to be automated for new employees and for entitlements to be de-provisioned automatically or through work-flow approval when an employee is terminated or when a visitor no longer needs temporary access. "Provisioning" often appears in the context of virtualization, orchestration, utility computing, cloud computing, and open configuration concepts and projects. For instance, the OASIS Provisioning Services Technical Committee (PSTC) defines an XML-based framework for exchanging user, resource, and service provisioning information, e.g. SPML (Service Provisioning Markup Language) for "managing the provisioning and allocation of identity information and system resources within and between organizations" [Provisioning].

An ESB is an internal enterprise transport for event correlation signaling to facilitate automated processes. An ESB is commonly associated with Service Oriented Architecture (SOA) and is a software architecture construct that provides foundational services for more complex architectures via an event-driven and standards-based messaging engine (the bus) [ESB]. An ESB does not implement a SOA, but provides the features with which one may be implemented.

3.5.3. Trends in Automation

The virtualization of datacenter systems and infrastructure has emerged as a significant cost savings measure for hosting information services. Benefits include rapid deployment and application load balancing while hosting virtualized machines on physical systems and hardware. Virtualization increases the complexity of information service delivery and requires stringent configuration control to deploy securely. Virtualization offers the advantage of software based systems, network and security devices that may prove very useful in the development of large scale interactive networks. Another advantage of virtualization is the potential instrumentation for HPC modeling and simulation. An instrumented virtual environment could serve as a front end to HPC / ASC analysis of cybersecurity threat case scenarios.

Attackers rely on automation as well. Botnets have become prevalent on the Internet as a means of automated control and compromise of Internet connected systems. The Torpig botnet uses a sophisticated network infrastructure and multistage attack to compromise and exfiltrate financial and personally indentifying information [Stone-Gross].

3.5.4. Expanded Broadband Infrastructure

Well resourced connectivity is fundamental to a high speed real-time computing infrastructure. There is potentially broad adoption of distributed Cyber indication and warning systems in the open and unclassified HPC/ASC regime. Examples of potential growth in the private/public sector involving the national information infrastructure include the American Reinvestment and Recovery Act (ARRA) expansion of broadband access [BBUSA]. Two programs are rapidly developing to expand service to un-served and underserved areas across the United States. The Department of Agriculture Rural Utility Services (RUS) Broadband Infrastructure Program (BIP) and the Department of Commerce National Telecommunication and Infrastructure Agency (NTIA) Broadband Telecommunication Opportunity Program (BTOP) have reviewed project proposals from state, tribal, municipal and private sector parties for the ARRA notification of funds availability (NOFA). The programs include provisions for anchor institutions, such as universities, hospitals and libraries to increase public service and access to information with public computer centers.

The ARRA programs potentially increase the utility of public HPC resources. Examples include the TeraGrid, a multiple university HPC grid available for collaborative research, and the New Mexico Computing Applications Center‘s (NMACC), Encento, an SGI HPC resource [TeraGrid, NMCAC, Encento].

National initiatives provide opportunity in the CIWD. An example of ongoing progress in the CIWD is the Argonne National Laboratory program for Cyber Security “Neighborhood Watch” [Cooper].

3.5.5. Trends and Opportunities from Current Practice

Distributed monitoring systems can not only provide security at the sensor location, but can also be used to aggregate data for analysis and correlation to help uncover large scale behavior. ICMP and TCP times can be used to identify network status and provide a view of the current environment. Globally, firewall, IDS, and scan logs, honeypot and spamtrap payloads, and darknet data can all be collected, aggregated, and analyzed. Locally such information can be used to identify and block suspect IP addresses and to supplement implementation of IPS.

Social engineering is an increasingly important aspect of cybersecurity, since malevolent actors often target users, rather than systems. Increased understanding of social activities, actions, and behaviors in a networked environment could provide game-changing advancement in cybersecurity.

A need for large scale simulation and visualization has been identified and could be used for internet-scale network mapping, identification of changes in network topology, visualization of the spread of malware or other attacks, and could be combined with identification of optimal locations to deploy defenses for the containment of the attack.

Large networks can be viewed as complex systems exhibiting emergent behavior. Without a more profound understanding of these complex systems and their behavior it is difficult to learn how to protect them. HPC can provide a platform for simulation of large-scale networks and discovery of their behaviors.

Deep packet inspection requires high-speed pattern matching on large datasets and, if implemented with sufficient throughput, could be used for effective IDS/IPS and filtering of http traffic.

As a society we share the common cyber risks of intrusion to information systems and compromise of information. Common goals also bind us in the face of common cyber threat.

3.6. Conclusions

There is ample evidence of opportunity for increased cybersecurity utilization of HPC/ASC to facilitate shared goals for cybersecurity. A fulsome understanding of offensive and defensive actor behavior and emergent behavior of complex systems will lead to the development of resilient safeguards and countermeasures to address shared cyber risks in the national information infrastructure.

Large-scale computation is needed for analysis of vast amounts of data collected across a wide-area over time. HPC systems offer a means for large-scale computational capacity and management. In tandem, high-speed sensor networks offer a means of data collection across a wide area over time.

Correlation and analysis algorithms for wide area cybersecurity applications are lacking in the public domain. Further development of these algorithms is necessary to inform the collection, transportation, storage of data and to conduct high performance analysis. For distributed data collection wireless or wired well resourced sensor networks are well suited for gathering, processing, and performing first level analysis, such as deep packet inspection.

We share common cyber risks and goals in the face of a global cyber threat. High confidence knowledge is needed to:

- Address the evolving threat
- Mitigate vulnerability and many types of attacks
- Reduce the threat exposure of major vendors
- Resolve configuration problems and human error
- Reduce the threat exposure and impact of mobile devices and the convergence of web applications
- Solve hard problems such as attribution of attack to responsible parties
- Effectively respond to new vulnerabilities such as zero day exploits
- Rapidly respond to attack

Our assessment of the current state of the industry provides background information and a set of recommended areas of research as a means to the most efficient and beneficial path forward. We find that HPC large-scale computation and cybersecurity data collection are essential to address the scope and scale of a global cyber threat. Further, the feasibility of such a system is dependent on collection of data, transport of data for analysis, and large-scale computation and analysis.

3.6.1. Influencing Factors of Distributed Cybersecurity Systems

The national focus on cyber is driving the need for high value HPC analysis and coordinated high-speed sensing and correlation of events. Influencing factors include:

- Industry dependence on voluntary reporting of attacks
- Information assurance for distributed cybersecurity systems:
    - Sharing cybersecurity information across domains

- o Division of roles and responsibilities
- o Information ownership and custodial responsibilities
- o Software assurance to mitigate the risk of malicious code
- o Strong access controls to mitigate the risk of compromise
- o Defined measures of assurance for distributed system elements
- o Inter-organizational coordination of communications, data handling, and trusted hardware and software
- o Divergence of inter-connected organizations in policy, and technology architecture
- o Easily exploitable IT infrastructures

Significant organizational, social and technical barriers exist that inhibit the advancement of distributed cybersecurity systems and coordinated response. Coordination challenges include the contractual and technical implementation of chains of trust, consistent policy, and achieving an acceptable balance between information protection and information sharing.

3.7. Recommendations

HPC large-scale computation will enable a new era of understanding in cybersecurity. Increased temporal and spatial understanding of the cyberterrain will yield technological and political advancements at a national level.

A temporal and spatial view of the cyberterrain in terms of attack and defense activity is needed to better assess and secure the national information infrastructure. Increased understanding of the progression of sophisticated and multistage attacks will yield stronger system and network defenses. Increased understanding of spatial or geographic distribution characteristics of attack will yield better coordinated defense across the wide area and multiple autonomous systems, as well as better understanding of malicious attack propagation and command and control.

Recommended actions are expressed as potential areas of research to advance HPC utility in cybersecurity applications and in an approach to building a foundation for a coordinated effort. Recommended Areas of Research are:

1. Identify trusted connection and automated process opportunities for collecting, correlating, analyzing, and sharing computer security information for HPC cybersecurity applications.

2. Examine informatics and statistical TCP/IP anomalous behavior research to: trend dataflow and protocol characteristics for identifying anomalous and malicious patterns; analyze temporal and spatial characteristics of attack for understanding of new attack and defense techniques; harvest data to seed Informatics visual analytics based on heuristics, cognitive psychology, and text analytics comprising graphical representation and visual vocabularies for cybersecurity subject matter.

3. Examine cybersecurity mathematical and statistical analysis research to collect, handle, and analyze large datasets for modeling, intrusion detection, attack response, and identification of multistage attacks.

4. Examine cybersecurity complexity science analysis research for studying the unpredictability in programs, machines, and networks, complex systems modeling, and large-scale modeling and simulation. Increased knowledge and insight concerning the behavior of complex networked systems will improve our understanding of their behavior and permit us to more readily secure them.

5. Examine modeling, simulation and analysis of complex networked systems, including large scale network models and models of network dynamics and cyber attack. Applications include intrusion detection, studying the spread of malware, and examining the evolution of cyber threats.

6. Expand HPC analysis and correlation algorithms for identification of temporal and spatial characteristics associated with anomalous events, modeling of normal network behavior, and detection of widespread, multistage, multiple method attacks.

7. Understand the sociology and psychology of cyber engagement. One of the most prevalent problems is attributing an attack to responsible parties. For example an IP address does not conclusively identify an organization or individual. Additionally, it is increasingly common for attacks to target people, rather than technology. Understanding the sociology and psychology of attack would benefit the development of anticipatory and preventative safeguards and countermeasures based on predictable behaviors.

These areas of research offer game changing opportunity in the effort to secure the national information infrastructure. The ordering should not imply that work proceed sequentially and is not ranked by importance; rather, work should be pursued in parallel to address the Nation’s cyber dilemma.

Developing a solution will require a coordinated effort. Interested parties should establish partnerships to work in parallel to develop analysis capabilities for event correlation and coordinated information protection and sharing. Such capabilities will serve to communicate new vulnerabilities, exposures, and emerging threats.

Coordination of efforts across various domains and authorities will require a framework for communication with defined roles and responsibilities, and an understanding of critical processes and dependencies is necessary to identify and prioritize appropriate risk mitigations. Business continuity planning (BCP) is the foundation for emergency preparedness and management, IT disaster recovery, and continuity of operations planning. A cyber recovery component of BCP at a national scale could benefit this effort.

Work should be done to complete a set of reference architectures that would provide consistency and understanding to all parties participating in HPC analysis and high-speed sensor networks. Reference architectures could include cybersecurity information management, access control, controlled interfaces, interconnectivity and transport systems, instrumentation, and sensor, network, and HPC topology.

Road mapping and planning will also provide a common approach and path for participating organizations to support a collaborative HPC analysis high-speed sensor network. Formulation of data models and formats, a common definition of terms, roles and responsibilities, communication plans, and escalation and order of succession will be required.

Finally, we need to foster a national commitment to build a sustainable HPC analysis high-speed sensor network and work to develop software assured HPC codes, informatics and algorithms for analysis of complexity and emergent behavior, anomaly analysis with mathematics and statistics, associative behavioral analysis, and network modeling and simulation.

4. REFERENCES

[AbuHmed] T. AbuHmed, A. Mohaisen and D. Nyang, A survey on Deep Packet Inspection for Intrusion Detection Systems, eprint arXiv:0803.0037

[Aho] A. V. Aho and M. J. Corasick, Efficient string matching: An aid to bibliographic search. Communications of the ACM, 18(6):333-340, 1975.

[Alicherry] M. Alicherry, M. Muthuprasanna and V. Kumar, High speed pattern matching for network IDS/IPS. In ICNP, pp. 187-196, 2006.

[ARGOS] ARGOS.
http://www.argos-system.org/?nocache=0.6055142466906847

[ASCIRed] R. Thomas, ASCI Red: The World's First TeraOps SuperComputer 2003, http://www.sandia.gov/ASCI/Red/

[Atlas] Active Threat Level Analysis System, http://atlas.arbor.net/

[Bain, FGRCCA] Federal Computer Week, Ben Bain July, 14, 2009 and the Fostering a Global Response to Cyber Attacks Act (Introduced in Senate). http://thomas.loc.gov/cgi-bin/query/z?c111:S.1438:

[BBUSA] Broadband USA : The Portal to Apply for Broadband Funding Under the American Recovery and Reinvestment Act of 2009, 2009. http://www.broadbandusa.gov/

[BEFF] The algorithm of b_{eff} (version 3.6), https://fs.hlrs.de/projects/par/mpi//b_eff/

[BLAND] E. Bland, Pentagon Funds Cyber Range For Web Warriors, Discovery News, Feb. 2009. http://dsc.discovery.com/news/2009/02/24/cyber-range-military.html

[Boyer] R. S. Boyer and J. S. Moore, A Fast String Searching Algorithm, Communications of the ACM, 20(10):76-172, 1977

[BRASE] J. Brase and D. Brown, Modeling, Simulation and Analysis of
Complex Networked Systems: A Program Plan, May 2009. https://wiki.cac.washington.edu/download/attachments/7478403/ComplexNetworkedSystemsProgram-final.pdf?version=1

[Bro] Bro Intrusion Detection System, http://www.bro-ids.org/
http://www.nsf.gov/

[CCE] Common Configuration Enumeration, http://cce.mitre.org/

[CCIM] Computation, Computers, Information, and Mathematics Center Homepage, http://www.cs.sandia.gov/.

[CERT] US-CERT, http://www.us-cert.gov/.

[CHEOPS-NG] Cheops-ng, http://www.cheops-ng.sourceforge.net

[CISCOTIC] Trusted Internet Connection Architecture for Single Service Providers, April 2009. http://www.cisco.com/web/strategy/docs/gov/TICAParchit_wp.pdf

[CJohsnon] C. Johnson III. "Implementation of Trusted Internet Connections (TIC), Memorandum for the Heads of Executive Departments and Agencies (M-08-05)" (PDF). Office of Management and Budget. November 20, 2007.
http://www.whitehouse.gov/omb/memoranda/fy2008/m08-05.pdf. Retrieved 2009-08-17.

[ChrDav] C. Karlof and D. Wagner, “Secure routing in wireless sensor networks: Attacks and Countermeasures.”

[CIM] DMTF, Common Information Model (CIM) Standards, 2009.
http://www.dmtf.org/standards/cim/

[CIM2] Common Information Model (computing), 2009.
http://en.wikipedia.org/wiki/Common_Information_Model_(computing)

[Coit] C. J. Coit, S. Staniford and J. McAlerney, Towards faster string matching for intrusion detection or exceeding the speed of snort. In DARPA Information Survivability Conference & Exposition II, pp. 367-373, 2001.

[Commentz-Walter] B. Commentz-Walter, A string matching algorithm fast on the average, In Proceedings of ICALP, pp. 118-132, 1979

[Cooper] B. Cooper, Argonne develops program for cyber security “Neighborhood Watch”: Cyber security team wins 2009 DOE innovation, technology achievement award, July, 2009. http://www.anl.gov/Media_Center/News/2009/news090716.html

[CorEph]: M.S. Corson and A. Ephremides, A distributed routing algorithm for mobile wireless networks, ACM/Baltzer J. Wireless Networks, vol 1, no. 1, feb 1995, pp. 61-82.

[CPLANT] Computational Plant, http://www.cs.sandia.gov/cplant/

[CSRC] NIST Computer Security Divisions Computer Security Resource Center, http://csrc.nist.gov/

[CUDA] CUDA Zone The Resource for CUDA Devlopers, 2009. http://www.nvidia.com/object/cuda_home.html#

[CVE] Common Vulnerabilities and Exposures, http://cve.mitre.org/

[DGNCR] DARPA National Cyber Range

Broad Agency Announcement (BAA), May, 2008.
http://www.darkgovernment.com/news/darpa-national-cyber-range/

[Dharmapurikar] S. Dharmapurikar, P. Drishnamurthy, T. S. Sproull and J.W. Lockwood, Deep packet inspection using parallel bloom filters, IEEE Micro, 24(1):52-61, 2004.

[DMTF] Interoperability Program for Virtualization Management, Common Information Model (CIM), http://www.dmtf.org/newsroom/pr/

[DRKJ]: D. Petrovic, R. C. Shah, K. Ramachandran, J. Rabaey. Data Funneling: Routing with Aggregation and Compression for Wireless Sensor Networks.

[Encento] NMCAC Encanto: New Mexico's Supercomputer,
http://www.newmexicosupercomputer.com

[ED] D. Proch, and R. Truesdell, Plumb The Depths Of Deep Packet Inspection. ED Online ID #21562, Aug. 13, 2009.
http://electronicdesign.com/Articles/Index.cfm?AD=1&ArticleID=21562

[EO12333] Executive Order 12333. http://www.archives.gov/federal-register/codification/executive-order/12333.html

[EO12968] Executive Order 12968.
http://www.opm.gov/extra/investigate/eo12968.asp

[EO13388] Executive Order 13388. www.ise.gov/docs/guidance/eo13388.pdf

[E2PIA] Reviewing Official Hugo Teufel III. Privacy Impact Assessment for
EINSTEIN 2, May 19, 2008.
http://www.dhs.gov/xlibrary/assets/privacy/privacy_pia_einstein2.pdf

[EPIA] DHS NSCD, US-CERT, Privacy Impact Assessment EINSTEIN Program: Collecting Analyzing, and Sharing Computer Security Information Across the Federal Civilian Government, Sept. 2004.
http://www.dhs.gov/xlibrary/assets/privacy/privacy_pia_eisntein.pdf

[ESB] Enterprise Service Bus, http://en.wikipedia.org/wiki/Enterprise_Service_Bus.

[FFT] FFTE: A Fast Fourier Transform Package, http://www.ffte.jp/

[FIPS] Federal Information Processing Standards Publications.
http://www.itl.nist.gov/fipspubs/

[FIPS199] FIPS PUB 199 FEDERAL INFORAMTION PROCESSES
STANDARDS PUBLICATION – Standards for Security Categorization of
Federal Information and Information Systems, Feb. 2004.

http://csrc.nist.gov/publications/fips/fips199/FIPS-PUB-199-final.pdf

[FIPS200] FIPS PUB 200 FEDERAL INFORAMTION PROCESSES STANDARDS PUBLICATION – Minimum Security Requirements for Federal Information and Information Systems, March 2006.
http://csrc.nist.gov/publications/fips/fips200/FIPS-200-final-march.pdf

[FIPS201] FIPS PUB 201 FEDERAL INFORAMTION PROCESSES STANDARDS PUBLICATION - Personal Identity Verification of Federal Employees and Contractors, Feb. 2005.
http://csrc.nist.gov/publications/fips/fips201-1/FIPS-201-1-chng1.pdf

[FISMA] The Federal Information Security Management Act of 2002.
http://csrc.nist.gov/groups/SMA/fisma/index.html

[FlexiSec] FlexiSec. http://www.mirlabs.org/jias/jinwala.pdf

[FOIA] Freedom of Information Act http://www.usdoj.gov/oip/index.html

[Ganapati] P. Ganapati, Personal Supercomputers Promise Teraflops on Your Desk, Aug. 2009. http://www.wired.com/gadgetlab/2009/08/personal-supercomputers/

[Ganti] Ganti, A. Study on Sensor Networks. Internal Sandia Report, 2009.

[GLB] The Gramm-Leach-Bliley Act of 1999. http://banking.senate.gov/conf/

[Gietl, GmbH Meuer] The Top Trends in High Performance Computing, 2009-05-20, Horst Gietl, Executive Consultant, Prometeus GmbH, Germany and Hans Meuer, ISC'09 General Chair, University of Manheim & Proteteus GmbH, Germany

[GlacsWeb] GlacsWeb. http://envisense.org/glacsweb/index.html.

[Hathaway] Cyber Security, An Economic and National Security Crisis, Fall 2008, by Melissa Hathaway, Former Senior Advisor to the Director of National Intelligence and Coordination Executive

[HAUWEI] Hauwei Data Communication. The DPI Technology and Functions.
http://www.huawei.com/products/datacomm/catalog.do?id=1235.

[HIPAA] The Health Insurance Portability and Accountability Act of 1996
http://www.hhs.gov/ocr/privacy/

[HPCBENCH] HPC Challenge Benchmark, http://icl.cs.utk.edu/hpcc/.

[HPCP] High Performance Computing Projects, Sandia Sees Data Management Challenges Spiral, http://www.hpcprojects.com/news/news_story.php?news_id=922, August 2009.

[HPCS] High Performance Computing Systems, http://www.highproductivity.org/, http://www.highproductivity.org/SSCABmks.htm

[HPL] A. Petitet, R. C. Whaley, J. Dongarra, A. Cleary, Ver. 2.0, HPL - A Portable Implementation of the High-Performance Linpack Benchmark for Distributed-Memory Computers, http://www.netlib.org/benchmark/hpl/

[HSA2002] The Homeland Security Act of 2002, Public law 107-296. http://www.ntia.doc.gov/ntiahome/infrastructure/index.html

[IGE]: C. Intanagonwiwat, R. Govindan and D. Estrin, "Directed Diffusion: A scalable and robust communication paradigm for sensor networks", IEEE/ACM Mobicom 2000 , pp. 56-67

[INF] Informatics, http://en.wikipedia.org/wiki/Informatics

[INI] INI – Information Networking Institute at Carnegie Mellon University http://www.ini.cmu.edu/

[IPVM] Interoperability Program for Virtualization Management, Common Information Model (CIM), http://www.dmtf.org/newsroom/pr/

[ISC] SANS Internet Storm Center, http://isc.sans.org/about.html

[ISC2] International Information Systems Security Certification Consortium, www.isc2.org

[ISE] Information Sharing Environment, http://www.ise.gov/

[ISO/IEC] The International Standard ISO/IEC 27001, Information technology – Security techniques – Information security management systems – Requirements, First edition 2005-10-15

[ISS] IBM Internet Security Systems Internet Scanner, http://www-935.ibm.com/services/us/index.wss/offering/iss/a1027208

[ITMRA] The Clinger Cohen Act of 1996. www.army.mil/ArmyBTKC/docs/CCA-Book-Final.pdf

[JohMal]: D.B. Johnson and D.A. Maltz, "Dynamic source routing in ad hoc wireless networks", in Mobile Computing, Kluwer Academic Publishers, 1996

[JUNIPER] Solution Brief - Juniper Networks Trusted Internet Connection (TIC) Solution: Enhancing the Federal Government Cyber Security Perimeter.

http://audio.federalnewsradio.com/temp/JuniperNetworksTrustedInternetConnectionSolutionSB.pdf

[Kannan] Kannan, J., Jung, J., Paxson, V., and Koksal, C. E. 2006. Semi-automated discovery of application session structure. In *Proceedings of the 6th ACM SIGCOMM Conference on internet Measurement* (Rio de Janeriro, Brazil, October 25 – 27, 2006). IMC '06. ACM, New York, NY, 119-132.

[Khaleel] M. Khaleel Beyond the Desktop: The role of computational architectures in accelerating discovery,
http://www.pnl.gov/science/images/newsmakers/TechAlliancePresentation.pdf

[Knuth] D. Knuth, The Art of Computer Programming: Semi-numerical Algorithms, Vol. 2, 1997.

[Krishnamachari] B. Krishnamachari, A Wireless Sensor Networks Bibliography, 2007. http://ceng.usc.edu/~anrg/SensorNetBib.html.

[Kumar1] S. Kumar, S. Dharmapurikar, F. Yu, P. Crowley and J.S. Turner, Algorithms to accelerate multiple regular expressions matching for deep packet inspection, In SIGCOMM, pp. 229-250, 2006.

[Kumar2] S. Kumar, J.S. Turner, and J. Williams, Advanced algorithms for fast and scalable deep packet inspection, In ANCS, pp.81-92, 2006.

[L7] L7-filter Application Layer Packet Classifier for Linux,
http://l7-filter.sourceforge.net/

[LBNL] LBNL-1667E, Mathematical and Statistical Opportunities in Cybersecurity, March 2009

[Leland] W. Leland, M. Taqqu, W. Willinger, and D. Wilson". "on the self-similar nature of Ethernet traffic". *IEEE/ACM Transactions on Networking*, 2(1), 1994.

[LexisNexis] LexisNexis Risk Solutions Technology,
http://www.lexisnexis.com/risk/about/technology.aspx

[Lighfoot] Lighfoot.
http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=04374458

[Madriagal] A. Madrigal, Wired Science News for Your Neurons NSF Head: All Hail the Cluster, March 2008. http://www.wired.com/wiredscience/2008/03/nsf-head-cluste/

[MHALJ]: M. Kubisch, H karl, A Wolisz, L.C. Zhong, and J Rabaey, "Distributed Algorithms for Transmission Power Control in Wireless Sensor Networks."

[MIT] Massachusetts Institute of Technology. http://web.mit.edu/

[Mosquera] M. Mosquera. (Agencies make headway in reducing gateways, Jul 10, 2008. http://gcn.com/Articles/2008/07/10/Agencies-make-headway-in-reducing-Internet-gateways.aspx)

[NAS] NASA Advanced Supercomputing Division, http://www.nas.nasa.gov/Resources/Software/npb.html

[NCR] National Cyber Range Questions and Answers, http://www.darpa.mil/sto/ia/pdfs/NCR_Qs_and_As.pdf

[Netezza] http://www.netezza.com/

[NIAP] The National Information Assurance Partnership, http://www.niap-ccevs.org/

[NIAP2] New NIAP CCEVS Strategy for FY10, March 16, 2009. http://www.niap-ccevs.org/cc-scheme/

[NIST] National Institute of Standards and Technology. http://www.nist.gov/index.html

[NIST] SP 500-268, Source Code Security Analysis Tool Functional Specification Version 1.0, May 20007, Series 800 Special Publications, Series 500 Special Publications

[NISTITL] NIST Information Technology Laboratory (ITL) SP 500, http://www.itl.nist.gov/lab/specpubs/sp500.htm

[NISTSP500-268] P. Black, M.Kass, and M. Koo.NIST SP 500-268, Source Code Security Analysis Tool Functional Specification Version 1.0, May 2007. http://samate.nist.gov/docs/source_code_security_analysis_spec_SP500-268.pdf

[NISTSP500-529] K. Mills, NIST SP500-529, Networking for Pervasive Computing, July 2005. http://www.antd.nist.gov/pubs/NIST500259.pdf

[NIST800-48] NIST SP 800-48, Guide to Securing Legacy IEEE 802.11 Wireless Networks, Revision 1, July 2008. http://csrc.nist.gov/publications/nistpubs/800-48-rev1/SP800-48r1.pdf

[NISTR800-53] NIST Recommended Security Controls for Federal Information Systems and Organizations, INFORMATION SECURITY, Initial Public Draft Feb. 2009. http://csrc.nist.gov/publications/PubsSPs.html

[NIST800-97] NIST SP 800-97, Establishing Wireless Robust Security Networks, A Guide to IEEE.11i, February 2007. http://csrc.nist.gov/publications/nistpubs/800-97/SP800-97.pdf

[NIST800-113] NIST SP 800-113, Establishing Guide to SSL VPNs, Recommendations of the National Institute of Standards and Technology, July 2008. http://csrc.nist.gov/publications/nistpubs/800-113/SP800-113.pdf

[NMCAC] New Mexico Department of Information Technology, New Mexico Computing Applications Center, http://www.doit.state.nm.us/supercomputer_project.html

[NSAct47] The National Security Act of 1947. http://www.intelligence.gov/0-natsecact_1947.shtml

[NVD] The National Vulnerability Database, http://web.nvd.nist.gov/

[OMB: M-08-156] (OMB: M-08-16, Guidance for Trusted Internet Connection Statement of Capability Form (SOC) (PDF), The White House, April 4, 2008. Retrieved on August 17, 2009.)

[NVIDIA] NVIDIA High Performance Computing (HPC), 2009.

[Ormandy] An Empirical Study into the Security Exposure to Hosts of Hostile Virtualized Environments, Tavis Ormandy, Google, Inc.

[Ourmon] Ourmon. http://ourmon.sourceforge.net/

[OVAL] Open Vulnerability Assessment Language, http://oval.mitre.org

[Papageorgiou] P. Papageorgiou. Literature Survey on Wireless Sensor Networks http://www.cs.umd.edu/~pavlos/papers/unpublished/papageorgiou03sensors.pdf

[Paxson] V. Paxson and S. Floyd. Wide area traffic: the failure of Poisson modeling. IEEE/ACM Transactions on Networking, 3(3):226–244, 1995.

[peakflow] Peakflow X. http://www.arbornetworks.com/en/peakflow-x.html

[PermaSense] PermaSense. http://www.permasense.ch/

[PiyKum1]: P. Gupta and P.R. Kumar, “The capacity of wireless networks”, IEEE Transactions on Information Theory, IT-46(2):388-404, March 2000.

[PiyKum2]: P. Gupta and P.R. Kumar. A System and Traffic Dependent Adaptive Routing Algorithm for Ad Hoc Networks

[Provisioning] Provisioning, http://en.wikipedia.org/wiki/Provisioning

[PSS] N. Pundit, Parallel Systems Software, 2008.
http://www.cs.sandia.gov/capabilities/ParallelSystemsSoftware/index.html

[PTRANS] PARKBENCH MATRIX KERNEL BENCHMARKS,
http://www.netlib.org/parkbench/html/matrix-kernels.html

[R40427] R40427Comprehensive National Cybersecurity Initiative: Legal Authorities and Policy Considerations, March, 2009 http://opencrs.com/document/R40427/

[QRadar] Qradar. http://www.q1labs.com/

[RA] RandomAccess
GUPS (Giga Updates Per Second),
http://icl.cs.utk.edu/projectsfiles/hpcc/RandomAccess/

[RahJan]; R.C. Shah and J.M. Rabaey, "Energy Aware Routing for Low Energy Ad Hoc Sensor Networks"

[riverbed] Riverbed Cascade. http://www.riverbed.com/products/cascade/

[RNR-94-007] THE NAS PARALLEL BENCHMARKS, D. Bailey, E. Barszcz, J. Barton, D Browning, R. Carter, L. Dagum, R. Fatoohi, S. Fineberg, P. Frederickson, T. Lasinsksi, R. Schreiber, H. Simon, V. Venkatakrishnan and S. Weeratunga, RNR Technical Report RNR-94-007, March 1994

[Robinson] D. G. Robinson, Statistical Language Analysis for Automatic Exfiltration Event Detection, Sandia Report SAND2010-2179, April 2010.

[Royer]: E.M.B. Royer, "Multi-Level Hierarchies for Scalable Ad Hoc Networking", Wireless Networks 9, pp. 461-478, 2003.

[RozKum]: R. Rozovsky and P.R. Kumar, "SEEDEX: A MAC protocol for ad hoc networks."

[RS] J. Tomkins and S. Kellly, Red Storm,
http://www.cs.sandia.gov/platforms/RedStorm.html

[RS2] Red Storm, Sandia National Laboratories News Release, 2004. http://www.cs.sandia.gov/platforms/RedStorm_072704NewsRelease.html

[RSQUALL] D. Doerfler, Red Squall,
http://www.cs.sandia.gov/platforms/RedSquall.html

[RSR] Red Storm Rising, Technology Magazine, Summer 2006 Issue. http://www.cs.sandia.gov/platforms/RedStorm_Rising.pdf

[SAMATE] NIST Software Assurance Metrics and Tool Evaluation http://samate.nist.gov/index.php/Main_Page.html

[SANS] The SANS Institute, http://www.sans.org

[SCARCH] N. Pundit, Supercomputing Architecture, 2008. http://www.cs.sandia.gov/capabilities/SupercomputingArchitecture/index.html

[SEAMONSTER] SEAMONSTERSouth East Alaska Monitoring Network for Science, Telecommunications, Education, and Research
http://seamonster.jun.alaska.edu/lemon

[SensorScope] SensorScope http://sensorscope.epfl.ch/index.php/Main_Page

[SNL0805] SAND 2009-0805 Mathematical Challenges in Cybersecurity

[SIO] L. Ward, Scalable IO: Current Projects http://www.cs.sandia.gov/Scalable_IO/

[SNL09]. Sandia computer scientists successfully boot one million Linux kernels as virtual machines, July 2009,
http://www.sandia.gov/news/resources/releases/2009/linux.html

[SNL2007] SAND2009-2007 Complexity Science Challenges in Cyber Security,
March 2009

[SNORT] SNORT open source network intrusion prevention and detection system. http://www.snort.org/

[SOA2002] Spotlight on Sarbanes-Oxley Rulemaking and Reports, 2002. http://www.sec.gov/spotlight/sarbanes-oxley.htm

[Stone-Gross] Stone-Gross, B., Cova, M., Cavallaro, Gilbert, L., Syzdlowski, B., Kremmerer, M., Kruegel, C., and Vigna, G. Your Botnet is My Botnet: Analysis of a Botnet Takeover, Security Group, Department of Computer Science, University of California, Santa Barbara

[Stone-Gross2] Modeling network intrusion detection alerts for correlation. ACM Trans. Inf. Syst. Secur., 10(1):4, 2007.

[Strassner] Policy Based Network Management, Solutions for the Next Generation, Strassner, 2004, ISBN 1-55860-859-1

[STREAM] J. McCalpin, STREAM: Sustainable Memory Bandwidth in High Performance Computers, http://www.cs.virginia.edu

[Tan] L. Tan, B. Brotherton and T. Sherwood, Bit-split string-matching engines for intrusion detection and prevention. TACO, ACM, 3(1):3-34, 2006.

[TBIRD] Thunderbird Linux Cluster, 2008.
http://www.cs.sandia.gov/platforms/Thunderbird.html

[TDOE] Transforming DOE CyberSecurity Wiki, 2009.
https://wiki.cac.washington.edu/display/doe/Home

[TeraGrid] TeraGrid, 2009. http://www.teragrid.org/

[Thompson] Cyber Security, June, 2009.
http://www.teragrid.org/tg09/files/TG09_KevinThompson.pdf

[TICSOC] Trusted internet connections (TIC) initiative statement of capability evaluation report, June 4, 2008. Available:
http://www.whitehouse.gov/omb/assets/egov_docs/2008_TIC_SOC_EvaluationReport.pdf

[TinySec] TinySec www.cs.berkeley.edu/~daw/papers/tinysec-sensys04.pdf

[TMF] Interoperability Program for Virtualization Management, Common Information Model (CIM), http://www.dmtf.org/newsroom/pr/

[TOP500] Top500 Supercomputing Sites, www.top500.org/

[TOP500ARCH] Architecture share for 06/2009
http://www.top500.org/stats/list/33/archtype

[TOP500ARCH2] The Main Architectural Classes, 2006.
http://www.top500.org/orsc/2006/architecture
[VAST] IEEE VAST Challenge
www.cs.umd.edu/hcil/VASTchallenge09/http://www.cs.umd.edu/hcil/vaschallenge09 2009

[VSRRP] V. Kawadia, S. Narayanswamy, R. Rozovsky, R.R. Sreenivas, and P.R.Kumar: Protocols for Media Access Control and Power Control in Wireless Networks

[WSN Security] Wireless Sensor Networks Security.
http://www.wsn-security.info/

[Wu] S. Wu and U. Manber, A fast algorithm for multi-pattern searching, Department of Computer Science, University of Arizona, 1994.

[XMT] Cray XMT Massively Multithreaded Platform,
http://www.tera.com/products/XMT.aspx

DISTRIBUTION

1	MS0321	Robert W. Leland	1400
1	MS1221	James S. Peery	5600
1	MS0801	Tom Klitsner	9300
1	MS1231	Ann N. Campbell	1950
1	MS0620	Edward J. Nava	5620
1	MS0620	Mark G. Terhune	5630
1	MS0620	Rebecca Darnell Horton	5640
1	MS0823	John D. Zepper	9320
1	MS0838	G. Kelly Rogers	9330
1	MS1318	Vitus J. Leung	1415
1	MS1318	David G. Robinson	1415
1	MS1319	Suzanne M. Kelly	1423
1	MS1319	Kevin Pedretti	1423
1	MS0672	David P. Duggan	5621
1	MS0671	Fredrick M. McCrory	5627
1	MS0671	Tan (Richard) Chang Hu	5627
2	MS0672	Jason R. Hamlet	5627
1	MS9158	Keith Vanderveen	8961
1	MS0823	William R. Claycomb	9326
1	MS0807	John P. Noe	9328
1	MS0807	Robert A Ballance	9328
1	MS0823	Geoffrey (Geoff) F. McGirt	9328
1	MS0806	Timothy (Tim) M. Berg	9336
1	MS0806	Anand Ganti	9336
1	MS0806	Steve A. Gossage	9336
2	MS0806	Curtis M. Keliiaa	9336
1	MS0806	John H. Naegle	9336
1	MS0806	Marie-Elena (Laney) Kidd	9336
1	MS0806	Jimmie V. Wolf	9336
1	MS0932	Timothy L. MacAlpine	9514
1	MS0899	Technical Library	9536 (electronic copy)

Sandia National Laboratories