National Science and Technology Council

Federal Plan for Cyber Security and Information Assurance Research and Development

Report by the Interagency Working Group on Cyber Security and Information Assurance

April 2006

About the National Science and Technology Council

The National Science and Technology Council (NSTC) was established by Executive Order on November 23, 1993. This Cabinet-level Council is the principal means for the President to coordinate science and technology across the diverse parts of the Federal research and development enterprise. Chaired by the President, the NSTC membership consists of the Vice President, the Science Advisor to the President, Cabinet secretaries, agency heads with significant science and technology responsibilities, and other White House officials.

An important objective of the NSTC is the establishment of clear national goals for Federal science and technology investments in areas ranging from information technologies and health research, to improving transportation systems and strengthening fundamental research. The Council prepares research and development strategies that are coordinated across Federal agencies to form investment packages aimed at accomplishing multiple national goals.

About this Report

This report was developed by the Interagency Working Group (IWG) on Cyber Security and Information Assurance (CSIA), an organization under the NSTC. The CSIA IWG reports jointly to the Subcommittee on Infrastructure of the NSTC's Committee on National and Homeland Security, and the Subcommittee on Networking and Information Technology Research and Development (NITRD) of the NSTC's Committee on Technology.

The report is published by the National Coordination Office for Networking and Information Technology Research and Development (NCO/NITRD). The NCO/NITRD supports overall planning, budget, and assessment activities for the multiagency NITRD Program under the auspices of the NSTC's NITRD Subcommittee.

To Request Additional Copies

To request additional copies of the *Federal Plan for Cyber Security and Information Assurance Research and Development* or other NITRD Program publications, please contact: NCO/NITRD, Suite II-405, 4201 Wilson Boulevard, Arlington, Virginia 22230; (703) 292-4873; fax: (703) 292-9097; e-mail: nco@nitrd.gov. Electronic versions of this report and other NITRD documents are also available on the NITRD Web site: *http://www.nitrd.gov*.

Copyright Information

NATIONAL SCIENCE AND TECHNOLOGY COUNCIL

FEDERAL PLAN

FOR

CYBER SECURITY AND INFORMATION ASSURANCE RESEARCH AND DEVELOPMENT

A Report by the
Interagency Working Group on Cyber Security and Information Assurance

Subcommittee on Infrastructure
and
Subcommittee on Networking and Information Technology Research and Development

April 2006

page intentionally left blank

EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF SCIENCE AND TECHNOLOGY POLICY
WASHINGTON, D.C. 20502

Dear colleague:

The Nation's information technology (IT) infrastructure – the seamless fabric of interconnected computing and storage systems, mobile devices, software, wired and wireless networks, and related technologies – has become indispensable to public- and private-sector activities throughout our society and around the globe. Pervasive, cost-effective communication enables a vast, constant flow of information that has transformed work environments and processes in government, business and industry, and advanced research, health care, and many other fields.

This IT infrastructure also supports other critical U.S. infrastructures, such as those that supply our food, water, energy, financial transactions, and transportation, as well as public health, emergency response, and other vital services. The interconnectivity that makes seamless delivery of essential information and services possible, however, also exposes many previously isolated critical infrastructures to the risk of cyber attacks mounted through the IT infrastructure by hostile adversaries. The exposure of critical infrastructure to cyber-based attacks is expected to increase, as convergence of network and device technologies accelerates, and as systems increasingly connect to the Internet to provide added functionality or greater efficiency.

Safeguarding the Nation's IT infrastructure and critical infrastructure sectors for the future is a matter of national and homeland security. Developed by the Cyber Security and Information Assurance Interagency Working Group under the auspices of the National Science and Technology Council, this *Federal Plan for Cyber Security and Information Assurance Research and Development* presents a coordinated interagency framework for addressing critical gaps in current cyber security and information assurance capabilities and technologies. The Plan focuses on interagency research and development (R&D) priorities and is intended to complement agency-specific prioritization and R&D planning efforts in cyber security and information assurance. The Plan also describes the key Federal role in supporting R&D to strengthen the overall security of the IT infrastructure through development of fundamentally more secure next-generation technologies.

I commend this Plan as an important step in addressing the Administration's national and cyber security priorities, and I look forward to working with Federal agencies and the private sector to develop the research roadmap called for in the Plan.

Sincerely,

John H. Marburger, III
Director


EXECUTIVE OFFICE OF THE PRESIDENT OF THE UNITED STATES
NATIONAL SCIENCE AND TECHNOLOGY COUNCIL

Table of Contents

EXECUTIVE OFFICE OF THE PRESIDENT OF THE UNITED STATES
NATIONAL SCIENCE AND TECHNOLOGY COUNCIL

Executive Summary

Powerful personal computers, high-bandwidth and wireless networking technologies, and the widespread use of the Internet have transformed stand-alone computing systems and predominantly closed networks into the virtually seamless fabric of today's information technology (IT) infrastructure. This infrastructure provides for the processing, transmission, and storage of vast amounts of vital information used in virtually every facet of society, and it enables Federal agencies to routinely interact with each other as well as with industry, private citizens, state and local governments, and the governments of other nations. As the IT infrastructure has broadened to global scale, the volume of electronic information exchanged through what is popularly known as "cyberspace" has grown dramatically and new applications and services proliferate.

The IT infrastructure supports critical U.S. infrastructures such as power grids, emergency communications systems, financial systems, and air-traffic-control networks. While the vast majority of these critical infrastructures (including their IT components) are owned and operated by the private sector, ensuring their operational stability and security is vital to U.S. national, homeland, and economic security interests.

Cyber threats are asymmetric, surreptitious, and constantly evolving – a single individual or a small group anywhere in the world can inexpensively and secretly attempt to penetrate systems containing vital information or mount damaging attacks on critical infrastructures. Attack tools and resources are readily available on the Internet and new vulnerabilities are continually discovered and exploited. Moreover, the pervasive interconnectivity of the IT infrastructure makes cyber attack an increasingly attractive prospect for adversaries that include terrorists as well as malicious hackers and criminals.

The Federal Role

In this environment of heightened risk, the Federal government has an essential role to play in cyber security and information assurance (CSIA) research and development (R&D). As in other science, technology, and engineering fields of critical importance to the Nation, Federal leadership should energize a broad collaboration with private-sector partners and stakeholders in academia and the national and industry laboratories where the bulk of Federal research is carried out. Such a partnership can chart a national R&D agenda for strengthening the security of the Nation's IT infrastructure.

This *Federal Plan for Cyber Security and Information Assurance Research and Development* takes the first step toward developing that agenda. The Plan also responds to recent calls for improved Federal cyber security and information assurance R&D, as outlined in the following documents: the OSTP/OMB Memorandum on Administration FY 2007 R&D Budget Priorities; *Cyber Security: A Crisis of Prioritization,* the 2005 report of the President's Information Technology Advisory Committee (PITAC); the 2003 *National Strategy to Secure Cyberspace;* and the 2002 Cyber Security Research and Development Act (P.L. 107-305).

Developed by the Cyber Security and Information Assurance Interagency Working Group (CSIA IWG), an organization under the National Science and Technology Council (NSTC), the Plan provides baseline information and a technical framework for coordinated multi-agency R&D in cyber security and information assurance. Other areas – including policy making (e.g., legislation, regulation, funding, intellectual property, Internet governance), economic issues, IT workforce education and training, and operational IT security approaches and best practices – also have substantial roles to play in

improving cyber security and information assurance. However, these subjects are outside the scope of the Plan, which addresses only the role of Federal R&D.

Likewise, the Plan is not a budget document and thus does not include current or proposed agency spending levels for cyber security and information assurance R&D. Agencies determine their individual budget priorities according to their mission needs and requirements.

Strategic Federal R&D Objectives

The following strategic Federal objectives for cyber security and information assurance R&D are derived from a review of current legislative and regulatory policy requirements, analyses of cyber security threats and infrastructure vulnerabilities, and agency mission requirements:

1. Support research, development, testing, and evaluation of cyber security and information assurance technologies aimed at preventing, protecting against, detecting, responding to, and recovering from cyber attacks that may have large-scale consequences.

2. Address cyber security and information assurance R&D needs that are unique to critical infrastructures.

3. Develop and accelerate the deployment of new communication protocols that better assure the security of information transmitted over networks.

4. Support the establishment of experimental environments such as testbeds that allow government, academic, and industry researchers to conduct a broad range of cyber security and information assurance development and assessment activities.

5. Provide a foundation for the long-term goal of economically informed, risk-based cyber security and information assurance decision making.

6. Provide novel and next-generation secure IT concepts and architectures through long-term research.

7. Facilitate technology transition and diffusion of Federally funded R&D results into commercial products and services and private-sector use.

Plan's Baseline Information

To provide a starting point and framework for coordinated interagency R&D to improve the stability and security of the IT infrastructure, the CSIA IWG agencies developed the following baseline information about ongoing Federal R&D activities in cyber security and information assurance:

R&D categories and technical topics – a list of cyber security and information assurance R&D technical topics grouped in broad categories. While not intended to be definitive, the list provides a structure for a survey and analysis of agency technical and funding priorities.

R&D technical and funding priorities – a set of interagency priorities derived by aggregating agency information. These interagency priorities differ from individual agency priorities, which vary based on agency missions and R&D requirements.

Investment analysis – a comparison of interagency technical and investment priorities to identify topics that are interagency technical priorities and in which there might be investment opportunities. (Table 1, pages 18-19.)

R&D technical topic perspectives – commentaries on the status of R&D in each topic. Prepared and reviewed by agency representatives with expertise in specific topics, these commentaries describe the topic and its importance, the state of the art, and gaps in current capabilities. (Part II, beginning on page 31.)

Together, the technical and funding priorities, investment analysis, and technical commentaries set the stage for the development of a cyber security and information assurance R&D roadmap and other coordinated activities proposed in the Plan's recommendations.

Findings and Recommendations

Strategic interagency R&D is needed to strengthen the cyber security and information assurance of the Nation's IT infrastructure. Planning and conducting such R&D will require concerted Federal activities on several fronts as well as collaboration with the private sector. The specifics of the strategy proposed in this Plan are articulated in a set of findings and recommendations. Presented in greater detail in the report, these findings and recommendations are summarized as follows:

1. Target Federal R&D investments to strategic cyber security and information assurance needs

Federal cyber security and information assurance R&D managers should reassess the Nation's strategic and longer-term cyber security and information assurance needs to ensure that Federal R&D addresses those needs and complements areas in which the private sector is productively engaged.

2. Focus on threats with the greatest potential impact

Federal agencies should focus cyber security and information assurance R&D investments on high-impact threats as well as on investigation of innovative approaches to increasing the overall security and information assurance of IT systems.

3. Make cyber security and information assurance R&D both an individual agency and an interagency budget priority

Agencies should consider cyber security and information assurance R&D policy guidance as they address their mission-related R&D requirements. To achieve the greatest possible benefit from investments throughout the Federal government, cyber security and information assurance R&D should have high priority for individual agencies as well as for coordinated interagency efforts.

4. Support sustained interagency coordination and collaboration on cyber security and information assurance R&D

Sustained coordination and collaboration among agencies will be required to accomplish the goals identified in this Plan. Agencies should participate in interagency R&D coordination and collaboration on an ongoing basis.

5. Build security in from the beginning

The Federal cyber security and information assurance R&D portfolio should support fundamental R&D exploring inherently more secure next-generation technologies that will replace today's patching of the current insecure infrastructure.

6. Assess security implications of emerging information technologies

The Federal government should assess the security implications and the potential impact of R&D results in new information technologies as they emerge in such fields as optical computing, quantum computing, and pervasively embedded computing.

7. Develop a roadmap for Federal cyber security and information assurance R&D

Agencies should use this Plan's technical priorities and investment analyses to work with the private sector to develop a roadmap of cyber security and information assurance R&D priorities. This effort should emphasize coordinated agency activities that address technical and investment gaps and should accelerate development of strategic capabilities.

8. Develop and apply new metrics to assess cyber security and information assurance

As part of roadmapping, Federal agencies should develop and implement a multi-agency plan to support the R&D for a new generation of methods and technologies for cost-effectively measuring IT component, network, and system security. These methods should evolve with time.

9. Institute more effective coordination with the private sector

The Federal government should review private-sector cyber security and information assurance practices and countermeasures to help identify capability gaps in existing technologies, and should engage the private sector in efforts to better understand each other's views on cyber security and information assurance R&D needs, priorities, and investments. Federal agencies supporting cyber security and information assurance R&D should improve communication and coordination with operators of both Federal and private-sector critical infrastructures with shared interests. Information exchange and outreach activities that accelerate technology transition should be integral parts of Federal cyber security and information assurance R&D activities.

10. Strengthen R&D partnerships, including those with international partners

The Federal government should foster a broad partnership of government, the IT industry, researchers, and private-sector users to develop, test, and deploy a more secure next-generation Internet. The Federal government should initiate this partnership by holding a national workshop to solicit views and guidance on cyber security and information assurance R&D needs from stakeholders outside of the Federal research community. In addition, impediments to collaborative international R&D should be identified and addressed in order to facilitate joint activities that support the common interests of the United States and international partners.

Part I:

Federal Plan for Cyber Security and Information Assurance R&D


EXECUTIVE OFFICE OF THE PRESIDENT OF THE UNITED STATES
NATIONAL SCIENCE AND TECHNOLOGY COUNCIL

Overview

In less than two decades, advances in information and communications technologies have revolutionized government, scientific, educational, and commercial infrastructures. Powerful personal computers, high-bandwidth and wireless networking technologies, and the widespread use of the Internet have transformed stand-alone systems and predominantly closed networks into a virtually seamless fabric of interconnectivity. The types of devices that can connect to this vast information technology (IT) infrastructure have multiplied to include not only fixed wired devices but mobile wireless ones. A growing percentage of access is through always-on connections, and users and organizations are increasingly interconnected across physical and logical networks, organizational boundaries, and national borders. As the fabric of connectivity has broadened, the volume of electronic information exchanged through what is popularly known as "cyberspace" has grown dramatically and expanded beyond traditional traffic to include multimedia data, process control signals, and other forms of data. New applications and services that use IT infrastructure capabilities are constantly emerging.

The IT infrastructure has become an integral part of the critical infrastructures of the United States. The IT infrastructure's interconnected computers, servers, storage devices, routers, switches, and wireline, wireless, and hybrid links increasingly support the functioning of such critical U.S. capabilities as power grids, emergency communications systems, financial systems, and air-traffic-control networks. While the vast majority of the critical infrastructures (including the IT components of those infrastructures) are owned and operated by the private sector, ensuring their operational stability and security is vital to U.S. national, homeland, and economic security interests.

In addition to its underlying role in critical U.S. infrastructures, the IT infrastructure enables large-scale processes throughout the economy, facilitating complex interactions among systems of systems across global networks. Their split-second interactions propel innovation in industrial design and manufacturing, e-commerce, communications, and many other economic sectors. The IT infrastructure provides for the processing, transmission, and storage of vast amounts of vital information used in every domain of society, and it enables Federal agencies to rapidly interact with each other as well as with industry, private citizens, state and local governments, and the governments of other nations.

Technology Trends

The risks associated with current and anticipated vulnerabilities of, threats to, and attacks against the IT infrastructure provide the rationale for this report. Fast-shifting trends in both technologies and threats make it likely that the security issues of the IT infrastructure will only intensify over the next decade. Key areas for concern include:

- The increasing complexity of IT systems and networks, which will present mounting security challenges for both the developers and consumers
- The evolving nature of the telecommunications infrastructure, as the traditional phone system and IT networks converge into a more unified architecture
- The expanding wireless connectivity to individual computers and networks, which increases their exposure to attack. In hybrid or all-wireless network environments, the traditional defensive approach of "securing the

perimeter" is not effective because it is increasingly difficult to determine the physical and logical boundaries of networks.

- The increasing interconnectivity and accessibility of (and consequently, risk to) computer-based systems that are critical to the U.S. economy, including supply chain management systems, financial sector networks, and distributed control systems for factories and utilities
- The breadth and increasingly global nature of the IT supply chain, which will increase opportunities for subversion by adversaries, both foreign and domestic

The Federal Role

The IT infrastructure's significance to the Nation has gained visibility in the aftermath of the September 11, 2001 terrorist attacks, large-scale cyber attacks, and rapid growth in identity theft. These events have made it increasingly clear that the security of the IT infrastructure is no longer simply a problem for the private sector and private citizens but also one of strategic interest to the Federal government.

Although computer and software companies are now making investments in security-related research and product development, their work is directed primarily at short-term efforts driven by market demands to address immediate security problems. The Federal government has a different but equally important role to play in cyber security and information assurance R&D. In its February 2005 report on cyber security R&D, *Cyber Security: A Crisis of Prioritization,* the President's Information Technology Advisory Committee (PITAC) stated that one of the Federal government's responsibilities is to invest in long-term, fundamental research "that will fill the pipeline with new concepts, technologies, infrastructure prototypes, and trained personnel" needed to spur on next-generation security solutions that industry can turn into widely available products.

The Federal R&D portfolio historically has represented the broad public interest in long-term science, technology, and engineering advances. Such advances in support of Federal agency missions sustain U.S. scientific preeminence and generate the discoveries and innovations necessary to fuel economic development and a rising standard of living. The Federal portfolio also includes many investment areas deemed critical for national defense and national and homeland security in the near- and mid-term time frames. Cyber security and information assurance R&D efforts are contributing to both major purposes and all time horizons of Federal R&D investment. Moreover, without such investment, aspects of the Nation's industrial and service sectors will be unable to move toward benefiting from the IT infrastructure, curbing their growth and compromising economic competitiveness.

Federal leadership catalyzes activities in scientific realms of strategic importance to the Nation. In cyber security and information assurance R&D, such leadership can energize a broad collaboration with private-sector partners and stakeholders to generate fundamental technological advances in the security of the Nation's IT infrastructure. First, in support of national and economic security, the Federal government can identify the most dangerous classes of cyber security and information assurance threats to the Nation, the most critical IT infrastructure vulnerabilities, and the most difficult cyber security and information assurance R&D problems. Second, the Government can use these findings to develop and implement a coordinated Federal R&D effort focused on the key research needs that can only be addressed with Federal leadership. While these needs will evolve over time, this *Federal Plan for Cyber Security and Information Assurance Research and Development* provides a starting point for such an effort.

The Federal Plan in Summary

In this Plan, the terms *cyber security* and *information assurance* refer to measures for protecting computer systems, networks, and information from disruption or unauthorized access, use, disclosure, modification, or destruction. The purpose of cyber security and information assurance is to provide for:

Integrity – protection against unauthorized modification or destruction of systems, networks, and information, and system and information authentication

Confidentiality – protection against unauthorized access to and disclosure of information

Availability – assurance of timely and reliable access to and use of systems, networks, and information

The Plan comprises the following sections:

- Types of vulnerabilities, threats, and risk
- Analysis of recent calls for Federal R&D
- Strategic Federal objectives
- Technical topics in cyber security and information assurance R&D
- Current technical and investment priorities of Federal agencies in cyber security and information assurance R&D
- Results of technical and funding gaps analysis
- Findings and recommendations
- R&D technical topic perspectives, including assessments of the state of the art and key technical challenges
- CSIA IWG agencies' roles and responsibilities

The Plan recommends that cyber security and information assurance be accorded high priority at all levels of the Government and be integral to the design, implementation, and use of all components of the IT infrastructure. The work begun in this document of identifying and prioritizing Federal cyber security and information assurance R&D efforts must be an ongoing process. Continuation of ongoing interagency coordination is needed to focus Federal R&D activities on the most significant threats to critical infrastructures and Federal agency missions and to maximize the gains from these investments. In particular, the Plan points to the need for coordinated Federal R&D to solve the hard technical problems that are barriers to fundamental advances in next-generation cyber security and information assurance technologies; such R&D is typically multidisciplinary, long-term, and high-risk.

Other areas – including policy making (e.g., legislation, regulation, funding, intellectual property, Internet governance), economic issues, IT workforce education and training, and operational IT security approaches and best practices – are also germane and have substantial roles to play in improving cyber security and information assurance. However, these subjects are outside the scope of the Plan, which addresses only the role of Federal R&D.

Likewise, the Plan is not a budget document and thus does not include current or proposed agency spending levels for cyber security and information assurance R&D. Agencies determine their individual budget priorities according to their mission needs and requirements.

Plan Background

In December 2003, the National Science and Technology Council (NSTC) chartered a new Interagency Working Group (IWG) on Critical Information Infrastructure Protection (CIIP), reporting to the Subcommittee on Infrastructure of the NSTC's Committee on Homeland and National Security and its Committee on Technology. Co-Chaired by the Department of Homeland Security (DHS) and the White House Office of Science and Technology Policy (OSTP), the CIIP IWG included participants from more than a dozen Federal departments and agencies.

In August 2005, the group was rechartered to report jointly to the NSTC Subcommittee on Networking and Information Technology Research and Development (NITRD) as well as to the Subcommittee on Infrastructure, in order to improve the integration of CSIA R&D efforts with other NITRD program component areas and coordination activities (see Appendix B). In conjunction with the rechartering, the group was renamed the Cyber Security and Information Assurance (CSIA) IWG to better characterize the scope of the IWG's activities and to reflect the fact that cyber security and information assurance are essential to critical information infrastructure protection but also have a broader impact.

The IWG assumed the responsibility for gathering information about agencies' cyber security and information assurance R&D programmatic activities and challenges, and for developing an interagency Federal plan for cyber security and information assurance R&D. This document, which represents a collaborative effort of the CSIA IWG agencies, sets forth a baseline framework for coordinated, multi-agency activities that continue to develop and implement the Federal Plan.

The framework is derived from a CSIA IWG analysis that identified and prioritized cyber security and information assurance R&D needs across Federal agencies. The framework also includes extensive documentation of the current state of the art and major technical challenges across a spectrum of R&D areas of importance in the development of cyber security and information assurance technologies.

The *Federal Plan for Cyber Security and Information Assurance Research and Development* also serves as a foundational document for the *National Critical Infrastructure Protection Research and Development Plan* (NCIP R&D Plan), which is required by Homeland Security Presidential Directive (HSPD) 7. Developed by the NSTC's Subcommittee on Infrastructure, this latter plan focuses on R&D needs in support of protecting the Nation's critical infrastructures. The CSIA Plan focuses on R&D to help meet IT needs outlined in the NCIP Plan, supporting CSIA elements of key NCIP strategic goals, including a national common operating picture, a secure national communication network, and a resilient, self-healing, self-diagnosing infrastructure.

The CSIA IWG has begun to implement the coordination of agency R&D activities related to this Plan. The coordinated activities and CSIA budget are now reported in the annual Supplement to the President's Budget for the NITRD Program, beginning with the FY 2007 Supplement released in February 2006.

Vulnerabilities, Threats, and Risk

A *vulnerability* is a flaw or weakness in the design or implementation of hardware, software, networks, or computer-based systems, including security procedures and controls associated with the systems. Vulnerabilities can be intentionally or unintentionally exploited to adversely affect an organization's operations (including missions, functions, and public confidence), assets, or personnel.

A *threat* is any circumstance or event with the potential to intentionally or unintentionally exploit one or more vulnerabilities in a system, resulting in a loss of confidentiality, integrity, or availability. Threats are implemented by threat agents. Examples of threat agents are malicious hackers, organized crime, insiders (including system administrators and developers), terrorists, and nation states.

Risk is a combination of the likelihood that a particular vulnerability in an organization's systems will be either intentionally or unintentionally exploited by a particular threat agent and the magnitude of the potential harm to the organization's operations, assets, or personnel that could result from the loss of confidentiality, integrity, or availability.

In the current climate of elevated risk created by the vulnerabilities of and threats to the Nation's IT infrastructure, cyber security is not just a paperwork drill. Adversaries are capable of launching harmful attacks on U.S. systems, networks, and information assets. Such attacks could damage both the IT infrastructure and other critical infrastructures.

Cyber security has largely failed to gain wide adoption in many consumer products for a variety of reasons, including a lack of appreciation for consequences of insecurity, the difficulty of developing secure products, performance and cost penalties, user inconvenience, logistical problems for organizations in implementing and consistently maintaining security practices, and the difficulty of assessing the value of security improvements. But consumer and enterprise concerns have been heightened by increasingly sophisticated hacker attacks and identity thefts, warnings of "cyber terrorism," and the pervasiveness of IT uses.

Consequently, many in the computer industry have come to recognize that the industry's continued ability to gain consumer confidence in new, more capable applications will depend on improved software development and systems engineering practices and the adoption of strengthened security models. Thus, industry leaders, trade and professional associations, and advocacy groups support a robust Federal role in the long-term fundamental R&D needed to provide the foundations for next-generation security technologies.

Types of Threats and Threat Agents

Because organizations and agencies now rely so heavily on networked IT systems, day-to-day operations are significantly hindered when systems are out of service or performance is degraded. Today, many vulnerabilities are easy to exploit, and individuals and organizations worldwide can access systems and networks connected to the Internet across geographic and national boundaries. Current technology also makes it easy to hide or disguise the origin and identity of the individuals or organizations that exploit these vulnerabilities.

In addition, cyber security vulnerabilities are volatile; even as existing vulnerabilities are patched, new ones are discovered. Even when vulnerabilities are discovered and patched by security professionals prior to an attack, hackers are increasingly reverse-engineering patches in order to discover the vulnerabilities and develop attacks that exploit them. Hostile actors are deriving attacks from new patches with increasing speed, often launching attacks before these patches are widely tested and deployed to secure vulnerable systems. The result of these trends is a vicious cycle in which there is a constant need for new countermeasures.

While the Internet receives the most attention in press coverage of cyber incidents, from a national security perspective the playing field for potential cyber attack operations is much broader. Sensitive information tends to be isolated from the Internet, but the various gateways that exist to facilitate the transfer of information from the outside into a closed network provide many openings for possible attack.

Moreover, though substantial progress has been made in raising levels of awareness about cyber security across industry and government, securing critical infrastructures remains a significant national challenge. Many critical industries, previously isolated from Internet security problems because they used older mainframe computing systems and leased telephone lines in dedicated networks, are reaching the time when this legacy infrastructure is being retired. They are adopting modern networks using personal computers, workstations, and servers with mainstream operating systems, interconnected through local-area networks, and connected to the Internet. In addition, the telecommunications industry itself is going through a systemic transformation caused by deregulation, economic change, and technological evolution, which may also leave these networks more vulnerable to attack.

Attackers' Asymmetric Advantages

A number of factors in the current security environment provide would-be attackers with significant advantages over those trying to protect the large-scale networks and interconnected IT systems on which society increasingly depends. An attacker needs to find only one vulnerability; the defender must try to eliminate all vulnerabilities. Powerful attack tools, including automated tools for malicious actions, are now freely available for downloading over the Internet to anyone who wants them, and little skill is required to use them. The resources – including training and equipment – needed to launch potentially harmful attacks are not only readily available but relatively inexpensive compared to the costs of securing systems, networks, and information and responding to attacks.

As a result, some classes of attacks can be initiated with little sophistication. Although these attacks are not generally significant threats to systems that are kept patched and well secured, they are effective against the many unpatched and poorly secured systems connected to the Internet, and contribute to a background level of ongoing malicious network activity. The automated tools that can be used by people with relatively little skill or knowledge continue to multiply, and are gradually increasing in capability in step with improvements in cyber security and information assurance technologies. Attackers also have the ability to exploit vulnerable third-party machines to launch their attacks.

Classes of attacks that require much greater expertise pose significantly greater threats. But while the sophistication required to mount such attacks limits them to a smaller set of adversaries, the capabilities of these high-threat adversaries also continue to advance.

These trends offer a wide range of individuals and entities – from malicious hackers to nation states –

the opportunity to support or directly engage in cyber attacks. The following profiles suggest some of the possible threat agents and their motivations.

Malicious Hackers

The earliest computer hackers often were individuals with sophisticated computer skills who simply enjoyed exploring programming and stretching their computer's capabilities. Hackers of this type still exist. Others, however, use their skills to write damaging code that propagates over the Internet or to break into private networks for malicious or criminal purposes. While many malicious hacker attacks rank as nuisances rather than being harmful, other hackers have moved into more damaging hostile or criminal activities, producing increasingly sophisticated malicious technologies and tools that proliferate across the Internet. Some of these hackers are taking advantage of their skills to earn money through information theft, identity theft, fraud, denial-of-service (DoS) attacks, and extortion. The impact of hackers may expand even further if nation states and others, such as terrorist or organized criminal groups, hire the talent or exploit the hacker-developed technologies.

Organized Crime

Organized crime is increasingly using the Internet to exploit any online opportunity to make money through the avenues mentioned above (information and identity theft, fraud, extortion) as well as illegal gambling, pornography, and other methods. Moreover, organized criminal elements are generally more structured and can draw on more extensive funding resources than loosely knit hacker communities, enabling them to hire expert hacker talent or bribe insiders to gain access to more sensitive systems.

Terrorists

Hacking could be used by terrorist groups to harvest information for planning physical or cyber attacks. Audit logs from Web sites, infrastructure owners, and national laboratories have recorded extensive, systematic information gathering originating from countries that serve as home bases for terrorist groups. Terrorist groups also are using the Internet for covert communications, and sympathetic hacker groups have launched various "e-jihads," consisting primarily of Web page defacements and DoS attacks.

Terrorist groups are known to have included not only engineers, computer scientists, and business people with backgrounds in computers, networks, and computer-based systems but also people with access to hardware and software producers. Terrorist groups have even sold computer products, which could in principle include malicious software. One known terrorist group is notable because it assembles and sells computer systems. Although law enforcement has not uncovered information pointing to subversion of software products, the potential for such activity exists. The evidence indicates that terrorist groups now have or can acquire the necessary expertise for identifying targets and conducting cyber attacks with serious or catastrophic consequences.

Nation States

Within their home territories, many nations have some offensive cyber capabilities derived from such defensive technologies as forensics, network protections, and software implants (code added to system or application software for a purpose such as monitoring usage or collecting data about users) as well as from their regulatory control over, and ability to gain physical access to, local telecommunications and Internet systems. Relatively few nation states, however, have the technical and operational capabilities (including resources, logistical support, expertise, and willingness to take risks) to orchestrate the full range of adversarial cyber operations through a combination of such means as recruiting insiders, setting up front companies, establishing signals collection systems, implanting damaging hardware or software in communications networks, and subverting telecommunications switches, cryptographic defenses, and supply chains.

Threat and Vulnerability Trends

In addition to the exploitation of Internet vulnerabilities, adversaries seeking to gather sensitive information, commit crimes, or attack critical U.S. infrastructures can employ other means, such as:

Insiders

The key to malicious or hostile activities in cyberspace is access to networked systems and information. Facilitating this access through the use of insiders can greatly reduce the technological sophistication necessary to mount an attack, because authenticated and authorized insiders may be able to circumvent barriers to external access, or may have legitimate access rights and privileges that would be denied to unauthorized users. So while obtaining network access via hacking provides one potential path for malicious activity, insider (physical or logical) access to the network reduces, and can in some cases eliminate, the difficulties associated with hacking through network defenses. With the right insider, an offensive operation may involve simply copying information to a portable medium that can be carried from the premises. A single well-placed, knowledgeable insider can also exploit IT systems to disrupt local infrastructure.

Outsourcing

The IT outsourcing trend – affecting activities ranging from computer help desks and data processing to R&D – can increase the exposure of an organization's systems and information to subversion. Outsourcing of services, to either foreign or domestic suppliers, increases risk by reducing control over access to systems and information. For example, apparently legitimate paths into an organization's networks and access to network resources can be established that can be exploited for illegitimate purposes. In this environment, effective information assurance technologies are imperative.

Supply Chain Attacks

Potential attacks through subversion of hardware or software supply chains can be viewed as another type of insider threat. Access through a hardware supply chain may require development and manufacture of a subverted version of a microelectronic component and a complicated operation to insert the device into the targeted computer, possibly through use of insiders in the supply chain. A software supply chain attack might involve, for example, a subversion embedded in lower-level system software not likely to be evaluated during testing. Another approach is to subvert the master copy of software used for broad distribution, which hackers recently attempted to do with a mainstream operating system. Even if software is tested, subversions may be difficult to detect since they would typically be revealed only under circumstances difficult for a defender to discover.

Industrial Espionage

Technically savvy companies have the potential to capitalize on inadequate IT system security to engage in cyber espionage against the U.S. government and domestic corporations, primarily to collect science and technology information that could provide economic benefits. Some of these companies have considerable technical expertise and signals intelligence capabilities and have a strong presence in U.S. IT product markets – including microchips, telecommunications systems, and encryption products. One consequence of the current espionage climate is that travelers with laptops and other electronic devices risk having information stolen in such locations as airports and hotels.

State-Sponsored Espionage

Gaining access to well-protected information or systems in closed networks remains a resource-intensive effort involving traditional espionage tradecraft. Such operations do not require the simultaneous access to large numbers of systems needed for a strategic attack and thus are within reach of a much larger array of foreign adversaries.

Foreign governments for decades have successfully recruited agents in the U.S. government with access to computer systems and cryptographic information. Foreign agents have also established technology companies in this country and served as subcontractors on U.S. defense contracts to obtain access to technology. Some governments now have the operational and technical expertise for more aggressive and sophisticated cyber espionage. U.S. counterintelligence efforts have uncovered an increasing number of such activities by foreign intelligence services, including past and ongoing espionage operations directed against critical U.S. military and other government systems.

Other Trends

Many malicious code attacks are "blended threats" that exploit multiple vulnerabilities or propagate via multiple means. Among these new classes of threats are adaptive or mutating threats, which are able to change their characteristics and appearance in order to avoid detection. Attacks can exploit operating systems, other software applications, software running on hardware components (e.g., routers and firewalls), or more infrequently, the hardware components themselves. Cryptographic attacks to undermine encryption-based security processes might attempt to exploit one or more of these avenues of attack.

The trends discussed in this document are supported by the report *Cybersecurity for the Homeland,* issued in December 2004 by the Subcommittee on Cybersecurity, Science, and Research & Development of the U. S. House of Representatives Select Committee on Homeland Security. The report concluded that:

- Hacking crews and individuals are increasingly working together around the globe in virtual, anonymous networks of specialists in different types and parts of attacks, such as propagation speed, denial of service, password logging, and data theft.
- An increasing number of adversaries are developing new options for exerting leverage over the U.S. through cyberspace, creating damage as well as conducting espionage. Cyberspace provides clear avenues and the prospect of anonymity.
- Foreign governments, hackers, and industrial spies are constantly attempting to obtain information and access through clandestine entry into computer networks and systems. This is not just "surfing" the open Internet for information voluntarily placed in the public domain, but intruding into closed and protected systems to steal secrets and proprietary information.
- Because many cyber attacks are not discovered or, if discovered, are not reported, hostile actors in cyberspace act with the knowledge that they are highly unlikely to be caught, let alone prosecuted and imprisoned. Attackers discovered in other countries cannot easily be brought to justice under U.S. laws, and their conduct may not even be illegal in the jurisdiction in which they are operating.

The report made the point that these trends are exacerbated because the network and system redundancy, diversity, and excess capacity that traditionally contributed to IT infrastructure resilience are decreasing with time, in part due to economic pressures. Federal agency personnel concerned with cyber security and information assurance view this factor as a key contributor to increased cyber vulnerability.

Immediate Concerns

IT infrastructure components are also potential targets of physical attacks, such as destruction by explosives and disruption caused by high-energy radio frequency or electromagnetic pulses. Given the spectrum of potential adversaries and their goals, a list of immediate concerns for the U.S. IT infrastructure begins with physical attacks against key data centers and communications nodes, particularly by terrorists.

However, immediate concerns also include the use of cyberspace for covert communications,

particularly by terrorists but also by foreign intelligence services; espionage against sensitive but poorly defended data in government and industry systems; subversion by insiders, including vendors and contractors; criminal activity, primarily involving fraud and theft of financial or identity information, by hackers and organized crime groups; attacks on the Internet infrastructure, particularly on the routers and domain name servers critical to its operation; and coordinated physical and cyber attacks, where the emergency response is hindered by unreliable or unavailable network communications.

The two sections that follow highlight some of these concerns in two specific domains, process control systems in critical infrastructures and the IT infrastructure of the banking and finance sector.

Industrial Process Control Systems

Computerized industrial process control systems (PCSs) are integrated hardware and software systems specifically engineered to monitor, evaluate, and regulate complex, large-scale processes. They often are embedded and hybrid systems, since computers are integral parts of such systems. Examples include the Supervisory Control and Data Acquisition (SCADA) systems that manage the electric power grid and the PCSs that control the timing and volume of processes in the chemical industry. PCS technologies also control the distributed sensor and actuator elements of pipeline systems for gas, oil, and water distribution. They manage supply chains and associated transportation systems, and they increasingly control building security, fire protection, environmental systems, lighting, and communications. Automated manufacturing processes often depend on PCS networks to improve quality control and enable response to crises as well as to reduce costs.

Because attacks interrupting or damaging key PCSs could have rippling impacts across the economy, these systems may increasingly be viewed by adversaries as attractive targets that can be exploited to weaken or incapacitate U.S. industry and infrastructure. Critical infrastructure sectors debate whether or not an exclusively electronic attack on control technologies could indeed have significant impact, given the industries' backup power systems and investment in "fail safe" or otherwise resilient designs for physical systems. But trends in the application of IT in these sectors point to increasing rather than decreasing levels of vulnerability and exposure in their infrastructures.

In the past, many PCS technologies used proprietary designs. Today, in the interest of reducing cost and improving maintainability, these systems mainly rely on standardized equipment and technologies, including general-purpose computers, mainstream operating systems, and standard Internet protocols, which are more vulnerable to attack. Many organizations view increasing use of the Internet as well as wireless and Web-based control systems as not only cost-effective but inevitable developments. Furthermore, cost-reduction measures are resulting in growing linking of networks that support control systems with internal and external corporate networks that support ordinary business operations, further increasing the exposure of control systems to external attacks.

For example, wireless control systems reduce cabling and installation costs. These systems typically use short-range wireless technologies, but signals still may be susceptible to attack from outside a building's perimeter if transmission patterns are not designed carefully. Engineers from a cyber security firm recently used standard wireless software to access networks at an electric power substation. Within 15 minutes, they were able to map out the entire operational control network of the substation, without having left their car.

The trends outlined above suggest that assaults on computerized control systems will be increasingly within reach of a wide array of attackers. The main uncertainty is the extent to which systems are already at risk due to a combination of direct or indirect Internet connectivity and security vulnerabilities such as inadequately secured wireless access, unpatched

software, or insufficient authentication or access control policies and mechanisms.

Banking and Finance Sector

In speeches after the September 11, 2001 attacks, Osama bin Laden identified the U.S. economy as a key target for terrorism. Foreign military strategists also have identified the U.S. economy as a logical target in strategic warfare. With networked computer systems now playing a central role in the financial sector, such systems are provocative lures for adversaries of all kinds. Indeed, terrorists have cased financial institutions in New York City, Newark, and Washington, D.C.

However, because many financial records reside in electronic form inside computer databases, cyber security and information assurance is a core value and a vital element of the financial industry's business model. Few industries have invested as much in technology, policies, and procedures to protect their networks, systems, and data. Indeed, because of its high assurance requirements, the banking and finance sector has put additional security measures in place and hardened systems beyond traditional levels of computer security.

Today's routine cyber threats to financial systems involve identity theft and consumer-level fraud, most often as a result of phishing attacks, keylogging, spyware, Trojan horses, or the theft of sensitive information from third parties. This type of theft is so common that it has been absorbed into the industry's risk model, with the costs shared by all consumers. Criminals have been known to conduct tests to ascertain whether fraud-detection software is active and, if not, to take advantage of the downtime to transfer money using stolen account information. Quickly noticing when one bank set a particular monetary threshold for fraud investigation, criminals made a large number of transactions below the threshold.

Computer systems used within banks or for bank-to-bank transactions offer a more lucrative target, but the computer security and accounting measures used with these systems are significantly tighter. The most serious cyber incidents tend to involve insiders. For example, a group with reported mob connections used insiders to try to launder hundreds of millions of Euros belonging to the European Union that were diverted from the Bank of Sicily. More recently, investigators foiled an attempt at stealing more than $400 million from a London branch of a Japanese bank through illicit electronic transfers reportedly enabled through the use of stolen passwords and access information obtained by insiders who made use of keystroke logging devices.


EXECUTIVE OFFICE OF THE PRESIDENT OF THE UNITED STATES
NATIONAL SCIENCE AND TECHNOLOGY COUNCIL

Analysis and Plan Framework

Recent Calls for Cyber Security and Information Assurance R&D

In addition to the historic Federal role in supporting long-term R&D, significant drivers for Federal cyber security and information assurance R&D arise from current national circumstances and Federal priorities. These drivers are identified in a number of Federal documents.

OSTP/OMB Memorandum on FY 2007 Administration R&D Budget Priorities

In a July 2005 memorandum entitled "Administration R&D Budget Priorities for FY 2007," the Directors of the Office of Management and Budget (OMB) and the Office of Science and Technology Policy (OSTP) identified cyber security R&D as an FY 2007 budget priority that should receive special focus at the interagency level. Cyber security and information assurance R&D falls squarely at the intersection of homeland security R&D and networking and information technology R&D, which are both highlighted as broad interagency R&D priorities.

The budget guidance memo cites cyber security R&D as one of three priority areas in the $3-billion Federal Networking and Information Technology Research and Development (NITRD) Program, along with high-end computing and advanced networking, "due to their potential for broad impact." (See Appendix B.) The memo states: "Reflecting the importance of cyber security, agencies should continue to work through the NSTC to generate a detailed gap analysis of R&D funding in this area." While not called out explicitly under Homeland Security R&D, cyber security and information assurance are also technological requirements of many priority homeland security capabilities cited in the memorandum.

PITAC Cyber Security Report

In *Cyber Security: A Crisis of Prioritization,* a February 2005 PITAC report to the President, the independent Presidential advisory panel warns that the Nation's IT infrastructure is highly vulnerable to attacks that could damage not only the economy but national defense and national security systems as well. Noting that "market forces direct private-sector investment away from research and toward the application of existing technologies to develop marketable products," the report calls on the Federal government to fundamentally improve its approach to cyber security R&D by increasing investments in unclassified cyber security R&D; intensifying its efforts to expand the size of today's small cyber security research community; improving technology transfer to the private sector; and increasing the focus and efficiency of Federal R&D through better coordination and oversight.

The report lists 10 areas as R&D priorities, based on a PITAC analysis of more than 30 documents and reports on cyber security R&D. The report concludes that the Nation will not be able to secure its IT infrastructure without significant advances in the following areas:

- Authentication technologies
- Secure fundamental protocols
- Secure software engineering and software assurance
- Holistic system security
- Monitoring and detection

- Mitigation and recovery methodologies
- Cyber forensics
- Modeling and testbeds for new technologies
- Metrics, benchmarks, and best practices
- Non-technology issues that can compromise cyber security

The National Strategy to Secure Cyberspace

The February 2003 *National Strategy to Secure Cyberspace* calls for Federal R&D leadership in certain circumstances, such as to address an increasing number of vulnerabilities and to provide continuity of government. In the latter situation, the document states, the role of the Federal government is to ensure the safety of its cyber infrastructure and those assets required for essential missions and services. Cyber security R&D areas that support this goal, according to the report, include: forensics and attack attribution; protection of systems, networks, and information critical to national security; indications and warnings; and protection against organized attacks capable of inflicting debilitating damage to the economy.

Cyber Security Research and Development Act

Specific research activities aimed at securing cyberspace are identified in the Cyber Security Research and Development Act of 2002 (P.L. 107-305). The law calls for significantly increased Federal investment in computer and network security R&D to improve vulnerability assessment and technology and systems solutions; expand and improve the pool of information security professionals, including researchers, in the U.S. workforce; and better coordinate information sharing and collaboration among industry, government, and academic research projects.

The Act also calls for basic research on innovative approaches to the structure of computer and network hardware and software that are aimed at enhancing computer security. Cited research areas include: authentication and cryptography; computer forensics and intrusion detection; reliability of computer and network applications, middleware, operating systems, and communications infrastructure; and privacy and confidentiality.

INFOSEC Research Council (IRC) Hard Problem List

In 1999, the IRC, a group of Federal research managers representing agencies involved in information security (INFOSEC) R&D related to their missions, issued a draft list of the most difficult INFOSEC research challenges, or "hard problems," they then faced. In November 2005, the IRC released an updated *Hard Problem List.*

The new hard problem list and the recommendations of the PITAC cyber security report are compared to the technical topics identified in this document in "Cyber Security and Information Assurance R&D Priorities: Comparison with PITAC and IRC" on page 20.

Strategic Federal Objectives

This *Federal Plan for Cyber Security and Information Assurance Research and Development* responds to the imperatives in the calls for Federal action. The Plan provides a cross-agency assessment of current Federal R&D activities and priorities and a set of strategic objectives for Federal cyber security and information assurance R&D to serve as baseline information for improved agency activities and multi-agency coordination. The following strategic objectives are derived from a review of policy and legislative drivers and analyses of cyber security threats and infrastructure vulnerabilities as well as Federal agency mission requirements:

1. Support research, development, testing, and evaluation of cyber security and information assurance technologies aimed at preventing, protecting against, detecting, responding to, and recovering from cyber attacks that may have large-scale consequences.

2. Address cyber security and information assurance R&D needs that are unique to critical infrastructures.

3. Develop and accelerate the deployment of new communication protocols that better assure the security of information transmitted over networks.

4. Support the establishment of experimental environments such as testbeds that allow government, academic, and industry researchers to conduct a broad range of cyber security and information assurance development and assessment activities.

5. Provide a foundation for the long-term goal of economically informed, risk-based cyber security and information assurance decision making.

6. Provide novel and next-generation secure IT concepts and architectures through long-term research.

7. Facilitate technology transition and diffusion of Federally funded R&D results into commercial products and services and private-sector use.

Development of Baseline Information

The Plan's baseline information was developed through several interrelated activities undertaken by the CSIA IWG agencies. The following sections describe these activities.

Cyber Security and Information Assurance R&D Categories and Technical Topics

The first step in establishing the baseline information was developing a list of cyber security and information assurance technical topics and an associated categorization (see Table 1 on pages 18-19). There was general agreement among agencies that there is no unique or "correct" classification of technical topics in this domain. Legitimate arguments could be made for changes in the topic names, the classification under broad categories, or even in the categories themselves. Thus, the list and classification of technical topics should be viewed as a convenient launching point for an analysis of agencies' cyber security and information assurance R&D priorities, rather than a firm statement about how technical topics should be organized.

Prioritization of Technical Topics

In the next step of the baseline development process, the technical topics were ranked in priority order. Agencies were asked to identify their R&D priorities irrespective of their past, current, or planned funding investments – i.e., based solely on gaps between the existing state of the art and anticipated requirements or desired capabilities, and the level of importance of those gaps. In assessing their priorities, the agencies applied criteria developed informally through CSIA IWG discussion that included such indicators as the relevance of the work to agency missions as well as broader government needs, requirements, and risk; current, emerging, and anticipated threats and levels of risk; and higher-level requirements driven or informed by policy or legislation. The degree to which private sector R&D was engaged in these topics was also taken into account in the criteria, reducing such topics' level of priority. The aim of this criterion was to avoid driving government investments into topics in which the state of the art is advancing effectively without Federal funding.

The priority rankings were then aggregated across all of the agencies, resulting in a set of *interagency* technical priorities. These interagency technical priorities represent some degree of consensus because they were deemed of importance across a significant number of agencies. However, the interagency technical priorities should not be interpreted as determining the highest priorities for *all* agencies. Some interagency technical priorities may be of limited interest to some agencies, while in other cases mission priorities for a given agency may differ from those identified as interagency technical priorities. Any agency may have some mission-related technical topics that are of particularly high priority even if they are not priorities for multiple agencies.

Investment Analysis

In the third step of the baseline development process, an investment analysis was conducted using information about programmatic investments gathered from the CSIA IWG agencies. The investment information was categorized according to the taxonomy of technical topics. This produced a set of investment priorities, which could then be compared to the interagency technical priorities to identify topics in which there might be investment gaps relative to these priorities. Differences between investment priorities and interagency technical priorities are not unexpected and should not be viewed as problem indicators, since individual agencies may be investing in mission-driven priorities that are not considered to be interagency technical priorities. The objective of this comparison was not to identify and characterize such agency-specific priorities as unnecessary. Rather, the goal was to identify topics that are interagency technical priorities and in which there might be underinvestment.

It should be noted that the agency funding information gathered in this process was pre-decisional and of varying granularity; it was collected only to indicate Federal agency spending emphases in cyber security and information assurance R&D. Thus, the baseline derived from this information should be viewed as useful in the aggregate but not a comprehensive source of detailed investment data.

R&D Technical Topic Perspectives

In the fourth and final step, agency representatives with expertise in specific technical topics of cyber security and information assurance R&D provided perspectives on the status of R&D in the topic, characterizing the topic's technical importance, the current state of the art, and gaps in current capabilities that will require R&D advances to close. The technical perspectives are provided in Part II of this report, which begins on page 31.

R&D Technical and Funding Priorities

Table 1 on pages 18-19 shows the top interagency technical and funding priorities that were identified by the prioritization process, under the associated broader categories. The table is intended to highlight areas where funding emphasis is needed, but this does not mean that funding is not needed in other technical topics as well. Nor do the top technical priorities identified represent all possible cyber security and information assurance R&D topics that are important to the Federal government. The list was developed to identify the 10 topics deemed most pressing within a larger group of priorities, though more than 10 are listed due to ties in the rankings.

Commentary on Analysis of Priorities

The cyber security and information assurance R&D prioritization activity generated findings that will be useful in agency and multi-agency discussions and coordinated planning activities to implement this Plan. Not surprisingly, the analysis shows alignment between R&D priorities and spending on near-term efforts focused on improving the security of existing Federal IT infrastructure in the face of existing threats and seeking ways to add security features for new capabilities.

The technical topics that were both identified as interagency technical priorities and ranked as investment priorities are authentication, authorization, and trust management; access control and privilege management; attack protection, prevention, and preemption; wireless security; and software testing and assessment tools. All CSIA IWG agencies reported multiple programs in attack protection, prevention, and preemption and many are supporting work in access and authentication technologies and wireless security. Several agencies have programs in software testing and assessment tools. A closely related topic

– automated attack detection, warning, and response – was among the top-funded priorities although it was not rated as a top technical priority.

The following topics are ranked as interagency technical priorities but are not among the top funding priorities: large-scale cyber situational awareness; secure process control systems; security of converged networks and heterogeneous traffic; detection of vulnerabilities and malicious code; IT system modeling, simulation, and visualization; inherently secure, high-assurance, and provably secure systems and architectures; composable and scalable secure systems; architectures for next-generation Internet infrastructure; and privacy issues.

Several reasons may explain why a topic identified as an interagency technical priority is not a funding priority. Agencies may generally agree that a topic is important but perceive that such work is not within their funding scope. They may view work in certain topics as within the purview of other agencies, or more appropriately addressed by the private sector rather than government investments. Alternatively, a priority and funding disparity could reflect a time lag in funding response to a topic that only recently emerged as an interagency technical priority. In the Federal budget cycle, agency budgets for a fiscal year are the result of technical and budget planning two years earlier, so it takes time for a new interagency technical priority to appear as an agency funding priority. Or a disparity could simply indicate a lack of broad recognition of a given technical topic's importance.

Thus, the fact that a topic is a top technical priority and not a top funding priority does not make the root cause for this incongruity evident. Understanding the issues associated with such disparities as well as identifying steps to remedy them are tasks most effectively managed through close interagency coordination. Further examination of these cases by the CSIA IWG is warranted as part of its activities to implement this Plan.

None of the topics in the Foundations of Cyber Security and Information Assurance category, which includes many topics focused on achieving fundamental advances in the engineering of more secure IT systems, ranked as top technical priorities. While some agencies support such long-term research, the analysis shows that many agencies currently are emphasizing technical topics associated with current threats and vulnerabilities. However, that emphasis does not explain why none of the Foundations topics rose to the level of a top technical priority. These topics are generally important because of their role in supporting the development of other technologies.

Additional analysis is needed to ascertain whether these topics are simply not as important despite their foundational implications, or whether they are more valuable than these results suggest but are unrecognized as priorities. As the coordinated, interagency roadmapping process moves ahead, agencies will need to evaluate such baseline findings in light of this Plan's objectives and recommendations. The Plan's first recommendation, for example, calls for a Federal focus on strategic and longer-term R&D needs, including technological foundations for next-generation IT infrastructure.

A related observation based on the analysis is that agencies in general are supporting a large number of discrete cyber security and information assurance R&D activities, but these efforts are broadly distributed across technical topics and in many cases are limited in scale and scope. Improved coordination and interagency information sharing are needed to begin to leverage agency investments and the associated expertise embodied in these activities. Coordination and planning enable agencies to forge interagency goals that maximize each agency's R&D contributions to results that no single agency could attain on its own.

TABLE 1

Top Technical and Funding Priorities

Federal Cyber Security and Information Assurance R&D

CSIA R&D AREAS Categories and Technical Topics	TOP PRIORITIES Technical	 Funding
1. Functional Cyber Security and Information Assurance		
1.1 Authentication, authorization, and trust management	✓	✓
1.2 Access control and privilege management	✓	✓
1.3 Attack protection, prevention, and preemption	✓	✓
1.4 Large-scale cyber situational awareness	✓	
1.5 Automated attack detection, warning, and response		✓
1.6 Insider threat detection and mitigation		
1.7 Detection of hidden information and covert information flows		
1.8 Recovery and reconstitution		
1.9 Forensics, traceback, and attribution		
2. Securing the Infrastructure		
2.1 Secure Domain Name System		
2.2 Secure routing protocols		
2.3 IPv6, IPsec, and other Internet protocols		
2.4 Secure process control systems	✓	
3. Domain-Specific Security		
3.1 Wireless security	✓	✓
3.2 Secure radio frequency identification		
3.3 Security of converged networks and heterogeneous traffic	✓	
3.4 Next-generation priority services		
4. Cyber Security and Information Assurance Characterization and Assessment		
4.1 Software quality assessment and fault characterization		✓
4.2 Detection of vulnerabilities and malicious code	✓	
4.3 Standards		
4.4 Metrics		
4.5 Software testing and assessment tools	✓	✓
4.6 Risk-based decision making		
4.7 Critical infrastructure dependencies and interdependencies		

Top Technical and Funding Priorities (continued)

CSIA R&D AREAS Categories and Technical Topics	TOP PRIORITIES Technical	 Funding
5. Foundations for Cyber Security and Information Assurance		
5.1 Hardware and firmware security		
5.2 Secure operating systems		
5.3 Security-centric programming languages		
5.4 Security technology and policy management methods and policy specification languages		
5.5 Information provenance		
5.6 Information integrity		
5.7 Cryptography		✓
5.8 Multi-level security		
5.9 Secure software engineering		✓
5.10 Fault-tolerant and resilient systems		
5.11 Integrated, enterprise-wide security monitoring and management		
5.12 Analytical techniques for security across the IT systems engineering life cycle		✓
6. Enabling Technologies for Cyber Security and Information Assurance R&D		
6.1 Cyber security and information assurance R&D testbeds		✓
6.2 IT system modeling, simulation, and visualization	✓	
6.3 Internet modeling, simulation, and visualization		
6.4 Network mapping		
6.5 Red teaming		
7. Advanced and Next-Generation Systems and Architectures		
7.1 Trusted computing base architectures		✓
7.2 Inherently secure, high-assurance, and provably secure systems and architectures	✓	
7.3 Composable and scalable secure systems	✓	
7.4 Autonomic systems		✓
7.5 Architectures for next-generation Internet infrastructure	✓	
7.6 Quantum cryptography		
8. Social Dimensions of Cyber Security and Information Assurance		
8.1 Trust in the Internet		
8.2 Privacy	✓	

Cyber Security and Information Assurance R&D Priorities: Comparison with PITAC and IRC

Because of their direct relevance to Federal interagency cyber security and information assurance R&D priorities, the research topics identified as priorities in the PITAC cyber security report and the IRC *Hard Problem List* provide useful points of comparison with the technical and funding priorities presented in this Plan (Table 1).

The 2005 IRC list includes the following eight hard problems:

- Global-scale identity management
- Insider threat
- Availability of time-critical systems
- Building scalable secure systems
- Situational understanding and attack attribution
- Information provenance
- Security with privacy
- Enterprise-level security metrics

Although they represent differing levels of granularity and categorization, the IRC list and the PITAC research priorities (cited on pages 13-14) are substantially aligned with the interagency technical priorities in this Plan. Specifically:

- The PITAC priority of authentication technologies is both a top technical and a top funding priority for the CSIA IWG agencies. This priority also maps to the IRC hard problem of global-scale identity management.
- The PITAC priority of secure software engineering and software assurance maps directly to the technical topic of secure software engineering, identified as a top funding priority by CSIA IWG agencies. Other CSIA technical and funding priorities such as software testing and assessment tools and detection of vulnerabilities and malicious code also contribute to software assurance. This area corresponds closely to the IRC hard problem of building scalable secure systems, which includes elements of software engineering that include design, construction, verification, and validation.
- The PITAC priority of holistic system security is broad in scope and does not map directly to a single CSIA topic area, but is relevant to the CSIA topic of analytical techniques for security across the IT systems engineering life cycle, which is a funding priority for CSIA IWG agencies. This PITAC priority also can be linked to other top technical CSIA R&D priorities such as inherently secure, high-assurance, and provably secure systems and architectures, and composable and scalable secure systems. These priorities also map to the IRC's building scalable secure systems hard problem.
- PITAC's monitoring and detection priority maps to two CSIA R&D priorities: large-scale cyber situational awareness (both a top technical priority and a top funding priority), and automated attack detection, warning, and response (a top funding priority). This corresponds to the IRC's hard problem of situational understanding and attack attribution.
- The PITAC priority of modeling and testbeds for new technologies maps to multiple CSIA R&D priorities: cyber security and information assurance R&D testbeds (a top funding priority) and IT system modeling, simulation, and visualization (a top technical priority).
- The PITAC's priority of metrics, benchmarks, and best practices maps directly to the IRC's hard problem of enterprise-wide security metrics. Although this area was not ranked as a top CSIA R&D priority via the information-gathering and analysis process that led to the technical and funding priorities identified in Table 1, the CSIA IWG recognizes the importance of and need for metrics. A recommendation in this Plan calls for the development and use of metrics to improve cyber security and information assurance.

- Although privacy was not called out as a single technical area among the PITAC priorities, it was mentioned as a subtopic within three of its priorities (authentication technologies, holistic system security, and non-technology issues that can compromise cyber security). In contrast, the IRC did focus specifically on privacy, having identified security with privacy as one of the IRC's hard problems. Similarly, privacy was identified as one of the CSIA IWG's top technical priorities.
- Other PITAC research priorities and IRC hard problems not identified by the CSIA IWG as interagency R&D priorities are clearly mission-related priorities that are receiving emphasis within individual agencies. For example, the DHS focus on infrastructure protection is represented in a program aimed at securing fundamental Internet communication protocols, including the Domain Name System and routing protocols – squarely within the scope of the PITAC priority of secure fundamental protocols. Both DoD and DHS are funding work in recovery and reconstitution, which corresponds to the PITAC research priority of mitigation and recovery methodologies. DoD, DHS, and intelligence community work in forensics, traceback, and attribution corresponds to the PITAC priority of cyber forensics.
- Of the 10 research priorities identified in the PITAC report, only non-technology issues that can compromise cyber security were not considered top interagency priorities by the CSIA IWG. CSIA IWG representatives agreed that these non-technology issues are important, but did not view them as rising to the level of other topics in the interagency technical and funding rankings.
- The two areas identified as hard problems by the IRC that were not viewed as interagency R&D priorities by the CSIA IWG are priorities within certain agencies. DoD and the intelligence community are both funding R&D in insider threat detection, which addresses the IRC hard problem of insider threat. DoD and the intelligence community also have an interest in the IRC hard problem area of information provenance, because of its direct relevance to management of classified information.

It might be expected that the interagency R&D priorities identified in this Federal Plan would align closely with the IRC list, as both lists have emerged from the Federal R&D community and are derived from the perspectives of Federal agencies about R&D challenges associated with carrying out their missions. The priorities identified by PITAC, however, were developed by a non-government group of subject-matter experts and therefore represent the perspectives of a different community. Thus, the degree of correspondence and alignment among the results of R&D prioritization activities conducted independently by the CSIA IWG, the PITAC, and the IRC is particularly noteworthy.

EXECUTIVE OFFICE OF THE PRESIDENT OF THE UNITED STATES
NATIONAL SCIENCE AND TECHNOLOGY COUNCIL

Findings and Recommendations

The technology trends outlined in this report make clear that the U.S. faces a long-term engagement with a new type of challenge to its security and economic stability. Cyber threats are asymmetrical, surreptitious, global, and constantly evolving. Moreover, the pervasive interconnectivity of the IT infrastructure on which all sectors of society now rely makes cyber attacks an increasingly attractive prospect for adversaries that include terrorists as well as malicious hackers and criminals.

This Plan outlines a Federal R&D strategy for strengthening the security and assurance of the IT infrastructure. The specifics of this strategy are articulated through the following findings and recommendations:

1. Target Federal R&D investments to strategic cyber security and information assurance needs

Finding: The private-sector marketplace for cyber security and information assurance technologies is thriving, but new products and advances are focused mainly on areas for which large and profitable customer bases currently exist – principally preventing, protecting, defending against, and responding to today's cyber threats.

Recommendation: Federal cyber security and information assurance R&D managers should reassess the Nation's strategic and longer-term cyber security and information assurance needs to ensure that Federal R&D focuses on those needs and complements areas in which the private sector is productively engaged. In general, agencies should ensure that resources are available to support work in the top technical priorities, address technical and funding gaps among the priority areas as well as in the broader collection of technical topics, and help develop the technological foundations for next-generation IT infrastructure, as described in this Plan.

2. Focus on threats with the greatest potential impact

Finding: Today's most prevalent cyber threats are not the most significant threats to the Nation's critical and IT infrastructures or to the economy, nor will they necessarily remain the most prevalent threats in the future. However, the constant hacker attacks – often closer to nuisances than true threats – that consume IT managers' daily attention and security budgets pervasively skew R&D efforts toward defenses against routine low-level attacks. Because of the lower relative probability of severe and highest-impact attacks, such strategic threats are not adequately being addressed at the research, development, or deployment levels.

Recommendation: Although cyber security and information assurance technologies developed by the private sector will undoubtedly evolve along with threats, this evolution can be significantly accelerated by laying sound technological foundations through R&D efforts. Federal agencies should focus cyber security and information assurance R&D investments on high-impact threats as well as on investigation of innovative approaches to increasing the overall security of IT systems.

3. Make cyber security and information assurance R&D both an individual agency and an interagency budget priority

Finding: As budgets become constrained, it is important to focus on recognized priorities in order to maximize the impact of existing funding resources. Such a focus is particularly valuable in

R&D in information technologies, where overall advances require gains in many scientific disciplines and component technologies.

Recommendation: Agencies should consider cyber security and information assurance R&D policy guidance (e.g., the joint memorandum from OMB and OSTP [discussed on page 13] that identifies cyber security as an interagency R&D priority) as they address their mission-related R&D. Agencies should also be aware of the interagency cyber security and information assurance R&D priorities identified in this report, and should give appropriate weight to these areas in budget formulation and technical program planning.

Recommendation: To achieve the greatest possible benefit from investments throughout the Federal government, cyber security and information assurance R&D should have high priority for individual agencies as well as for coordinated interagency efforts.

4. Support sustained interagency coordination and collaboration on cyber security and information assurance R&D

Finding: Cooperative interagency activities through the CSIA IWG enabled the development of this Plan. Sustained coordination and collaboration among agencies will be required to accomplish the goals identified in the Plan. Ongoing coordination can expand communication about shared cyber security and information assurance issues across disparate agencies, ensure that there is minimal duplication of R&D efforts across agencies, and help leverage agencies' expertise and strengths to achieve common goals. Collaborative activities such as testbeds and demonstrations can maximize the gains from R&D efforts. For example, several agencies can develop cooperative R&D plans to address complementary parts of the research agenda, and joint funding may make it possible to address common needs for which no single agency has sufficient resources.

Recommendation: Agencies should designate representatives to participate in development of the interagency R&D roadmap proposed in Recommendation 7 and other interagency cyber security and information assurance R&D activities. Agencies should participate in interagency R&D coordination and collaboration on an ongoing basis. Agency leadership at high levels also has an important role to play and should formally and/or informally support cooperative activities that involve multiple agencies. Such cooperation is particularly desirable in the cyber security and information assurance domain, where the goal is improved security procedures, tools, and techniques that can have broad impact.

5. Build security in from the beginning

Finding: Many of today's IT infrastructure vulnerabilities are the result of bugs and flaws in IT systems' software and hardware. In addition, much of the current infrastructure was not built with security as a core requirement. It was initially developed in a trusted community where today's threats did not apply. Now it is being used in ways that were not originally envisioned, but that require a greater level of trust than can be provided in the absence of security. The current standard approach to security relies on patching vulnerabilities and deploying a large assortment of security countermeasures aimed at known types of attacks. While this approach functions with varying degrees of effectiveness as a reactive mitigation strategy, it is not an effective long-term path to a fundamentally more secure infrastructure.

Recommendation: The Federal cyber security and information assurance R&D portfolio should support fundamental R&D exploring inherently more secure next-generation technologies that will replace today's patching of the current insecure IT infrastructure.

6. Assess security implications of emerging information technologies

Finding: Both new information technologies and emerging research areas can be expected to introduce novel security issues and vulnerabilities in the IT infrastructure. Moreover, it is likely that as new capabilities are added to the existing IT infrastructure, the difficulty of fixing some vulnerabilities will be exacerbated.

Recommendation: The Federal government should assess the security implications and the potential impact of research results in new information technologies as they emerge, including in such fields as optical computing, quantum computing, and pervasively embedded computing. Given the pace of technological change in the IT domain, this analytical capability should be an integral component of Federal cyber security and information assurance R&D planning and coordination activities.

7. Develop a roadmap for Federal cyber security and information assurance R&D

Finding: While scientific advances can occur unexpectedly and serendipitously, progress in areas of strategic importance must be accelerated through concerted attention and planned and coordinated efforts. Accelerating development of new cyber security and information assurance technologies for the Nation's IT infrastructure will require agreement among Federal R&D agencies on interagency technical priorities and coordinated activities to address them. This Plan provides baseline information about current Federal cyber security and information assurance R&D to serve as the starting point for the necessary multi-agency coordination.

Recommendation: Federal agencies – working together and in collaboration with the private sector – should use this Plan's technical priorities and investment analyses to develop a roadmap of cyber security and information assurance R&D priorities. This effort should emphasize coordinated agency activities that address technology and investment gaps and should accelerate development of strategic capabilities. Agencies should adopt the collaborative roadmapping process in an ongoing way as a means to strengthen Federal research in cyber security and information assurance, to intensify the R&D focus on high-priority areas, and to leverage agency investments more effectively in support of strategic goals.

8. Develop and apply new metrics to assess cyber security and information assurance

Finding: It is widely acknowledged in the IT industry and the national research community that a major research challenge is posed by the lack of effective methods, technologies, and tools to assess and evaluate the level of component, system, and network security. The baseline analysis of Federal investments found that, while the technical topic of software testing and assessment tools is both funded and ranked as a top R&D priority, the topic of metrics is not in either the top funding or top priority rankings.

Recommendation: As part of roadmapping, Federal agencies should develop and implement a multi-agency plan to support the R&D for a new generation of methods and technologies for cost-effectively measuring IT component, system, and network security. As more exacting cyber security and information assurance metrics, assessment tools, and best practices are developed through R&D, these should be adopted by agencies and applied in evaluating the security of Federal systems, and should evolve with time.

9. Institute more effective coordination with the private sector

Finding: Much of the Nation's IT infrastructure and interconnected critical infrastructures is owned and operated by the private sector. Furthermore, both private as well as public (i.e., government) sectors rely broadly on mainstream commercial-off-the-shelf technologies to build out and secure their

respective parts of the IT infrastructure, making the effective transition of technologies from R&D into widely available products a key issue. Addressing these needs will require ongoing communication and coordination between the public and private sectors to maximize the gains from each sector's activities.

Recommendation: The Federal government should review private-sector cyber security and information assurance practices and counter measures to help identify capability gaps in existing technologies, and should engage the private sector in efforts to better understand private-sector views on cyber security and information assurance R&D needs and priorities. Improved awareness in the Federal government and the private sector of the other's views on R&D needs, priorities, and investments will enable both research communities to develop and pursue complementary R&D efforts that meet strategic national needs, while at the same time making the best use of limited funding resources for cyber security and information assurance R&D.

Recommendation: Federal agencies supporting cyber security and information assurance R&D should improve communication and coordination with operators of both Federal and private-sector critical infrastructures with shared interests.

Recommendation: Federal coordination efforts should encompass development of information exchanges and outreach activities that accelerate technology transition as an integral part of Federal cyber security and information assurance R&D activities. Because widespread use of effective security technologies is in the national interest, obstacles to adoption and deployment of the results of R&D activities should be addressed.

10. Strengthen R&D partnerships, including those with international partners

Finding: From its origins nearly 40 years ago in Federally funded R&D programs to meet Federal needs, the Internet has grown into a remarkable global infrastructure used by nearly a billion people. As this Plan emphasizes, however, the Internet also is interconnected with some of the Nation's most sensitive physical and IT infrastructures.

Recommendation: Given the scale, complexity, and diversity of this multifaceted fabric of connectivity, the Federal government should foster a broad partnership of government, the IT industry, researchers, and private-sector users to develop, test, and deploy a more secure next-generation IT infrastructure. The Federal government should initiate this partnership by holding a national workshop to solicit views and guidance on cyber security and information assurance R&D needs from stakeholders outside of the Federal research community. In addition, impediments to collaborative international R&D should be identified and addressed in order to facilitate joint activities that support the common interests of the United States and international partners.

Conclusions

The IT infrastructure of the United States today is essential to the functioning of government, private enterprise, and civil society, including its critical systems for water, energy, transportation, and public safety. Federal leadership is both warranted and needed to encourage development of long-term goals and technical strategies for improving the overall security of this vital national interest.

The need for Federal leadership is underscored by the cyber security conditions described in this report. In summary:

- The increasing availability of techniques to attack the economy through the IT infrastructure provide an asymmetric, low-cost advantage to adversaries of all kinds around the globe.
- Ubiquitous vulnerabilities in today's IT infrastructure and a rapidly evolving spectrum of threats tie up available resources in a recurring cycle of defensive patching that does not strengthen the infrastructure as a whole.
- The threats and vulnerabilities of tomorrow may be substantially different from today's.
- The Nation's critical physical infrastructures are connected to and rely upon the IT infrastructure, and thus are also liable to suffer impacts from cyber attacks.
- Integration of emerging technologies into the IT infrastructure increases the breadth of and access to vulnerabilities.
- The degree of interconnectivity with the IT infrastructure and among critical infrastructures will continue to rise in the years ahead.
- The Federal government cannot unilaterally deploy countermeasures across the existing IT infrastructure, nor can it unilaterally develop and deploy a more secure infrastructure. Effective solutions will come only through a combination of R&D breakthroughs and cooperation among all stakeholders.

This Plan outlines a Federal role in cyber security and information assurance R&D that:

- Ensures that cyber security and information assurance R&D is a strategic Federal priority
- Recognizes that R&D in defensive measures for the current IT infrastructure, while a short-term necessity, is not a substitute for strengthening the infrastructure in more fundamental ways
- Supports longer-term R&D to develop the next-generation technologies needed to build in, rather than bolt on, security throughout IT infrastructure architectures
- Sustains interagency collaboration to maximize the gains from Federally funded R&D
- Fosters partnerships between the Federal government and private-sector stakeholders in IT infrastructure security.

The Nation needs next-generation IT infrastructure R&D to produce the breakthroughs from which new cyber security paradigms will take shape. With this Federal Plan in place, the next step is to undertake the multi-agency efforts it recommends.


EXECUTIVE OFFICE OF THE PRESIDENT OF THE UNITED STATES
NATIONAL SCIENCE AND TECHNOLOGY COUNCIL

Part II:

Technical Perspectives on Cyber Security and Information Assurance R&D

EXECUTIVE OFFICE OF THE PRESIDENT OF THE UNITED STATES
NATIONAL SCIENCE AND TECHNOLOGY COUNCIL

Technical Perspectives on Cyber Security and Information Assurance R&D

Part II provides technical perspectives on the cyber security and information assurance R&D topics identified in Part I. The R&D topics are grouped into eight broad categories. Each technical perspective, prepared and reviewed by agency officials with expertise in the topic, describes the topic and its importance, the state of the art, and gaps in current capabilities.

1. Functional Cyber Security and Information Assurance

The R&D topics in this category address technologies and capabilities that minimize the impact of compromises or potential compromises of data, networks, and systems, or that enable them to prevent, detect, resist, or respond to attacks. Topics in this category are:

- Authentication, authorization, and trust management
- Access control and privilege management
- Attack protection, prevention, and preemption
- Large-scale cyber situational awareness
- Automated attack detection, warning, and response
- Insider threat detection and mitigation
- Detection of hidden information and covert information flows
- Recovery and reconstitution
- Forensics, traceback, and attribution

1.1 Authentication, Authorization, and Trust Management

Definition

Authentication is the process of verifying the identity or authority of a network or system user (which can be a human user or a computer-based process or device) through a secure means such as digital signatures, passwords, tokens, or biometric features. Authorization, which takes place after authentication, refers to the privileges granted to an authenticated user who has requested access to services or resources. (Section 1.2 discusses access control in greater detail.) Authentication and authorization are interdependent; authorization to use a network or system resource frequently includes establishing the identity of the user requesting access (e.g., identity-based authentication) or verifying that a trusted third party has certified that the user is entitled to the access requested (e.g., credential-based authentication). Privilege is a security attribute shared by users whose identities have been authenticated. Cross-domain credentialing allows distinct systems, connected across a network, to provide access based on the secure identification procedure performed by one of the other networked systems. Trust management consists of making assessments of sets of credentials to determine whether they constitute adequate evidence for authorization.

Importance

Authentication is fundamental to all information security because it connects the actions performed on a computer to an identified user that can be held accountable for those actions. The expanding means available for accessing networks make security breaches and uncontrolled user access a growing concern. As enterprise IT systems continue to grow in complexity and number of users, authorization technologies that enable authenticated users to be assigned varying levels of system access privileges will play an increasingly critical role in security management.

State of the Art

Authentication of a user is based on one or more of three factors: a physical attribute (e.g., fingerprint or biometric data), an artifact (e.g., an automatic teller machine [ATM] card or cryptographic token), and/or a data key (e.g., a password). Each has advantages and disadvantages. The best-known and most common authenticators are conventional static passwords. Compromised static passwords, however, are a common vulnerability because users are careless about keeping their passwords secret, password security policies (such as mandatory format rules and periodic changes) are difficult to enforce, and malicious attackers have technological and social tools for discovering and accessing passwords. The use of multi-factor authentication methods may increase assurance. For example, an ATM might require both an ATM card and a password or personal identification number to provide a higher level of assurance than is provided by either factor alone.

Biometric technologies for authentication use measurements for identifying people – for example, their fingerprints, voice, retinal scans, or even handwriting – that can be used in IT authentication. But biometric data raise privacy issues that may in some instances limit their usage. Moreover, while biometric authentication can be used to provide stronger assurance of identity beyond that achievable with static passwords, biometrics are also susceptible to compromise. For example, recent experiments with artificial fingers have shown that fingerprint recognition devices can be fooled.

Capability Gaps

The current technologies described above all have limitations that frustrate efforts of system security managers to increase overall security levels for networks, systems, and information. Next-generation concepts that both streamline and harden authentication, authorization, and trust management technologies and tools are needed to help mitigate vulnerabilities associated with changing network dynamics and increased security threats. Specific R&D needs include:

Device authentication: Device authentication requires equipping devices with characteristics that can be reliably recognized. For devices and associated processes that generate requests, authentication using cryptographic protocols may be required. Some of these protocols have been developed, but there has been little experience with deploying them and building systems that make good use of them.

Scalable authentication: Federated identities are a capability that enables organizations to share trusted identities across the boundaries of their networks – with business partners, autonomous units, and remote offices. These technologies offer the prospect of scalable authentication needed for scalable trust management. However, there are continuing challenges in defining common authentication identities and, more important, in the forms of authorization that the inter-domain authentication will support. This problem has been partially addressed in some of the most common application areas such as the use of credit cards for electronic commerce on the Internet. However, scalable authentication, or global-scale identity management, remains a challenge (e.g., see the *Hard Problem List,* INFOSEC Research Council, November 2005, for elaboration).

1.2 Access Control and Privilege Management

Definition

Access control and privilege management begin with the administrative and mechanical process of defining, enabling, and limiting the operations that users can perform on specific system resources. The permission or limitation of operations is based on the business rules or access policies of the organization.

Access control policies are enforced through a mechanism consisting of a fixed system of functions and a collection of access control data reflecting the configuration of the mechanism. Together, these map a user's access request to the decision of whether to grant or deny access. The access control data include a set of permissions, each indicating a user's authorization to perform an operation (e.g., access, read, write) on an object or resource. Permissions are not individually specified. They are organized in terms of, and mapped through administrative operations or a predefined set of rules on to, a set of user, subject (process), and resource attributes associated with a specific type or class of policy.

For example, under an access control management approach called Role-Based Access Control (RBAC), permissions are defined in terms of roles that are assigned to users and privileges that are assigned to roles. Other approaches include label-based access control mechanisms that are defined in terms of labels applied to users, processes, and objects, and discretionary access control mechanisms that are defined in terms of user identifiers, user groups, and access control lists.

Importance

Although access control is often specified in terms of limitations or protections, the ability of an organization to enforce access control policy is what ultimately enables the sharing of greater volumes of data and resources to a larger and more diverse user community.

State of the Art

Various security mechanisms now exist for enforcing secure access within host operating systems and across heterogeneous bodies of data. In an attempt to streamline the management of access control, RBAC models and more recently an RBAC standard have been developed. RBAC offers administrative efficiency and the capability to intuitively administer and enforce a wide range of access control policies.

In RBAC, permissions are associated with roles and roles are assigned to users in order to grant user permissions corresponding to those roles. The implementation of this basic concept greatly simplifies access control management. Roles are centrally created for the various job functions in an organization, and users are assigned roles based on criteria such as their positions and job responsibilities. Users can be easily reassigned roles. Roles can be granted new permissions as new applications and systems are incorporated, and permissions can be revoked from roles as needed. For example, if a user moves to a new function within the organization, the user can be assigned to the new role and removed from the old one with associated privileges updated automatically. In the absence of RBAC, the user's old privileges would have to be individually identified and revoked, and new privileges would have to be granted.

Although RBAC represents a clear improvement over simple table lookup models of the access control matrix (data structures such as access control lists), the RBAC model does not solve all access control and privilege management problems. Discovering and defining roles and mapping roles to enterprise resources and applications, commonly referred to as role engineering, are costly and difficult. Although the development of best practices and tools to ease the transition to RBAC would be helpful, these capabilities provide only an interim solution to the research objectives described below. Ultimately, access control should be redefined and re-engineered from the ground up to reflect the increasing scale and complexity of networks and systems of systems. The goal should be a redefinition that preserves access control advancements while providing a generalized context to accommodate well-known and ad hoc access control policies, is easy to deploy and manage, and is safe in its configuration.

Capability Gaps

To move toward the next generation in access control and privilege management technologies, advances in three separate but related R&D areas are needed: 1) scalable access control data management methods and tools; 2) flexible access control mechanisms capable of enforcing a wide variety of access control policies; and 3) methods and techniques for defining safe and secure access control configurations.

Scalable access control data management: Many organizations have hundreds or even thousands of systems, hundreds to hundreds of thousands of users, and thousands to millions of resources that must be protected. Managing access control data across these systems, users, and resources is a monumental task and perhaps the most expensive and error-prone of all security disciplines.

Identity-based access control models work well for small workgroups. But as the number of groups and users and the number and variety of resources they need to access grows to an enterprise- and cross-enterprise scale, access control information stored in applications, databases, and file systems grows so large that managing and controlling access changes can overwhelm even the most knowledgeable administrators. Visualizing and reasoning about a virtual ocean of access control data become impossible. For example, many enterprises are unable to make even the simplest queries, such as what system accounts exist for a given user. Consequently, organizations have resorted to implementing poor administrative practices such as account sharing and cloning of permissions, resulting in permissions becoming over-distributed and difficult to manage.

Flexible access control mechanisms: One size does not fit all access control policies. Access control mechanisms are as diverse as the types of business practices and applications that need to enforce them. An access control mechanism that meets the policy requirements within one market domain may be inappropriate in another.

Effective access control mechanisms provide a context for policy configuration, embodiment, and enforcement. Policy configuration refers to the administrative operations of creating and managing access control data. Embodiment refers to the storage of access control data that reflect the policy. Enforcement applies access control data so that users and their processes adhere to the access control policy. Since the mid 1970s, security researchers have sought to develop access control models as abstractions of access control systems. When implemented, the models provide a generalized context that supports a wide collection of policies, while adhering to an agreed-upon set of security principles such as least privilege (restricting a user to the minimum privileges needed to complete authorized tasks) and separation of duty (assigning roles and privileges such that no single user can perform multiple sensitive tasks). Revocation (removing privileges previously granted to principals) is also a key feature of these models.

The process for users to specify rich policies remains challenging. This is partly a user-interface problem and partly a problem of designing an intuitive model through which security configuration options can be conveyed to users. Some progress has been made in designing flexible mechanisms, though challenges remain (e.g., implementing least privilege or revocation on a wide-scale basis is difficult). These mechanisms have not yet been widely deployed.

Safety: In the context of access control, safety is the assurance that an access control configuration will not result in the leakage of a privilege to an unauthorized user. Safety is fundamental to ensuring that the most basic access control policies can be enforced. Unfortunately, there is a tension between the need for safety and the desire for flexibility. The safety of an access control configuration cannot be specified using a general access control model. Consequently, safety is achieved either through the use of limited access control models or the verification of safety via constraints. Currently, almost all safety-critical systems use limited access control models because constraint expression languages are too complex for easy administrative use. However, researchers have determined that most constraints belong to one of a few basic types (e.g., static, dynamic, or historical).

Therefore, a key research goal is to develop ways to formulate constraints that allow the safety of access control configurations to be ensured, while having these constraints be flexible enough to support practical applications.

1.3 Attack Protection, Prevention, and Preemption

Definition

An attack is an attempt to gain unauthorized access to a network's or a system's services, resources, or information, or to compromise a network's or a system's integrity, availability, or confidentiality. Network or system owners can adopt practices and technologies that improve resistance to attacks or that prevent attacks from disrupting communications or operations, or from compromising or corrupting information.

Importance

Attack protection, prevention, and preemption are essential functional cyber security capabilities. Their goal is to provide an enterprise-wide capability to intercept a malicious attack, thereby preventing disruption, compromise, or misappropriation of networks, systems, or information. Robust attack protection, prevention, and preemption capabilities help mitigate threats and reduce the ability of adversaries to exploit vulnerabilities.

There are two different attack protection, prevention, and preemption strategies. The proactive strategy shields healthy network or system components or services to prevent them from becoming contaminated, corrupted, or compromised. The reactive strategy temporarily isolates compromised network or system components or services to prevent them from contaminating, corrupting, or compromising healthy assets. To be effective, both the proactive and the reactive security capabilities need to be deployed at all levels of enterprise systems.

In addition, attack protection, prevention, and preemption capabilities should be governed by a flexible, adaptable concept of operations. Not all attacks have the same scope or operational impact. Accordingly, the configuration and operation of the attack protection, prevention, and preemption capability should change in accordance with attack severity and intent (i.e., the approach must be adaptable to the nature of the attack and the assets being attacked).

State of the Art

A variety of laws, regulations, and/or institutional policies require agencies and other organizations to be able to respond to security incidents, prevent disruption to normal operations, and isolate compromised networks and systems. Many current commercial offerings are primarily limited to reactive intrusion-detection tools using signature- and rule-based algorithmic techniques, which use preset identification rules to distinguish authorized from unauthorized access. These tools are labor-intensive to use, require constant updating, and provide only limited protection. Even though updates are released much more quickly today than in the past, the result is an arduous configuration control and patch management task.

For example, major vendors are constantly issuing updates and patches to operating systems or applications to fix security holes. In some instances, these updates and patches reopen existing vulnerabilities or create new ones while fixing the targeted problem. Many organizations, such as those operating safety-critical infrastructure systems, have policies that require all upgrades and patches to be thoroughly tested before being deployed to operational systems. But hackers now are also able to reverse-engineer patches to discover the vulnerabilities and rapidly launch attacks that exploit them before the patches can be widely deployed. This becomes a recurring cycle as new upgrades and patches are released more frequently and reverse engineering methods used by hackers improve.

Capability Gaps

Amid these security conditions, reactive capabilities and manual responses are inadequate. Automated responses that operate in milliseconds and emphasize preemption and prevention are needed, along with

next-generation systems that are fundamentally more robust and resilient. Furthermore, organizations need to abandon the view that any single product can secure its IT infrastructure. Rather, the focus should be on developing an integrated set of tools and techniques that provide a comprehensive, layered, enterprise-wide attack protection, prevention, and preemption solution.

Proactive behavior-based systems may offer the best option for developing the next generation of attack protection, prevention, and preemption capabilities. These systems will not depend on signatures or rules to identify attacks. Proactive behavior-based tools identify precursor events early in the attack timeline. These systems, when the technologies mature, will provide the capability to identify and preempt unknown and novel attacks. Some research has been done in this area, and early attempts at behavior-based responses are starting to emerge in commercial products. This work should be continued with the goal of making robust products available, and it should be expanded to include the capabilities highlighted below.

Protection is needed at all layers of a protocol stack, such as the seven-layer International Standards Organization (ISO)/Open Systems Interconnect (OSI) Technical Reference Model (TRM) (see box below). Current attack preemption R&D primarily addresses Layer 3 (network layer) attacks generated by outsiders. Additional protection, prevention, and preemption features and functionality that are needed include a host-based intrusion prevention capability that is independent of the platform, operating system, and applications.

Related research is needed to increase and verify the robustness and resilience of networks, systems, and components to withstand attacks, especially unknown or novel attacks. Work is also needed to improve the ability of networks, systems, and components to

ISO/OSI Technical Reference Model Layers

Layer 1 – Physical
This layer conveys the bit stream – electrical impulse, light or radio signal – through the network at the electrical and mechanical level. It provides the hardware means of sending and receiving data on a carrier, including defining cables, cards, and other physical aspects.

Layer 2 – Data Link
At this layer, data packets are encoded and decoded into bits. It furnishes transmission protocol knowledge and management and handles errors in the physical layer, flow control, and frame synchronization. The data link layer is divided into two sublayers: the Media Access Control layer and the Logical Link Control layer.

Layer 3 – Network
This layer provides switching and routing technologies, creating logical paths, known as virtual circuits, for transmitting data from node to node. Routing and forwarding are functions of this layer, as well as addressing, internetworking, error handling, congestion control, and packet sequencing.

Layer 4 – Transport
This layer provides transparent transfer of data between end systems, or hosts, and is responsible for end-to-end error recovery and flow control. It ensures complete data transfer.

Layer 5 – Session
This layer establishes, manages, and terminates connections between applications. It sets up, coordinates, and terminates conversations, exchanges, and dialogues between the applications at each end. It deals with session and connection coordination.

Layer 6 – Presentation
This layer provides independence from differences in data representation (e.g., encryption) by translating from application to network format, and vice versa. It works to transform data into the form that the application layer can accept, and it formats and encrypts data to be sent across a network, providing freedom from compatibility problems. It is sometimes called the syntax layer.

Layer 7 – Application
This layer supports application and end-user processes. Communication partners are identified, quality of service is identified, user authentication and privacy are considered, and any constraints on data syntax are identified. Everything at this layer is application-specific. This layer provides application services for file transfers, e-mail, and other network software services. Tiered application architectures are part of this layer.

Source: Cisco Systems, Inc.

dynamically reconfigure themselves in order to preempt or minimize the damage from an attack.

1.4 Large-Scale Cyber Situational Awareness

Definition

Cyber situational awareness can be defined as the capability that helps security analysts and decision makers:

- Visualize and understand the current state of the IT infrastructure, as well as the defensive posture of the IT environment
- Identify what infrastructure components are important to complete key functions
- Understand the possible actions an adversary could undertake to damage critical IT infrastructure components
- Determine where to look for key indicators of malicious activity

Cyber situational awareness involves the normalization, deconfliction, and correlation of disparate sensor data, and the ability to analyze data and display the results of these analyses. Situational awareness (SA) is an integral part of an information assurance (IA) common operational picture. Such a picture provides a graphical, statistical, and analytical view of the status of computer networks and the defensive posture.

Importance

Situational awareness is the key to effective computer network defense. A robust situational awareness capability is necessitated by the highly interconnected nature of information systems and computer networks, the degree to which they share risk, and the coordination and synchronization requirements of response efforts.

Analysts and decision makers must have tools enabling timely assessment and understanding of the status of the networks and systems that make up the IT infrastructure. This situational understanding must be presented at multiple levels of resolution: 1) a top-level, global indication of system health; 2) exploration of various unfolding threat scenarios against various components of the system; and 3) more local-level details of recognizable or previously unseen anomalous activities.

State of the Art

Most current SA technologies perform limited correlation and fusion of primarily low-level network and node-based sensor data. These technologies do not provide a picture of the broader state of health of larger network systems. Status information tends to be localized and provides limited insight into the impact on business and mission processes or interactions among the various components of the larger IT infrastructure. As a result, attempts at corrective action are difficult to coordinate to assure that response actions do not propagate undesirable effects onto critical elements of the infrastructure.

Current visualization schemes focus on presenting large amounts of sensor data in formats that support attempts by human analysts to often manually perform the necessary fusion. They do not adequately enable effective visualization and understanding of potentially unfolding malicious or anomalous activity in the broader context of the IT infrastructure.

Capability Gaps

Despite their strategic and operational significance, current SA capabilities are immature. Security analysts must analyze large volumes of data from sensors and network management systems. In the absence of improved capabilities, emerging technologies will present more information than can reasonably be analyzed with existing capabilities. A particularly difficult problem is finding trends and patterns in attacks or probes (i.e., electronic attempts to circumvent network and system protections or to identify weak points in an information system) that may be occurring. New technologies are needed to help security analysts deconflict, correlate, and understand large volumes of data to support informed decision making.

Research is also needed to determine how best to design the human-computer interface to portray many aspects of SA to fit analysts' cognitive processes. In particular, visualizations need to go beyond current

efforts that are focused on understanding large volumes of low-level sensor data. Methods are needed to model and present to decision makers multiple, possibly competing, scenarios and hypotheses of unfolding potential attacks, in some cases with sufficient warning to preempt these attacks if possible, or at least minimize damage and support rapid, effective response and restoration.

Generation of situational awareness and understanding must be based on fusion of a broad range of cyber sensor data and traditional information sources, including open source data. The large amounts of such multi-source data must be filtered, transformed, fused, and correlated to provide insightful and actionable information to analysts and decision makers. External and internal contextual information about the system also is needed to enable understanding of observed abnormalities, whether malicious or otherwise. Current capabilities must be expanded to capture the broader context in which malicious or anomalous activities may be occurring.

A key proving ground for a robust SA capability is in sustaining the overall functionality of the Nation's IT infrastructure. The IT infrastructure can be expected to increasingly support various ongoing and planned Federal and private-sector activities around the globe. These network-centric processes will depend on assured availability, integrity, and confidentiality of the infrastructure's networks, systems, and information. A sophisticated SA capability will be needed to determine where and when the required operational reliability has been, or may be, adversely affected.

Systems that provide SA capability must themselves be designed to resist subversion or manipulation. Such protection is essential to keep adversaries from using an SA system to directly or indirectly trigger inappropriate, and perhaps harmful, responses to detected anomalous activities, or to hamper recovery activities following attack.

1.5 Automated Attack Detection, Warning, and Response

Definition

Automated attack detection, warning, and response capabilities enable systems and networks to recognize that they are under attack, respond defensively, and alert human operators. Today's static signature- and rule-based technologies can detect certain types of network disturbances and can respond by alerting human operators. But these technologies generally cannot recognize novel forms of attack, and they have limited abilities to automatically act to defend the system and make repairs to keep it functioning. Automated attack detection requires next-generation tools based not only on predefined signatures but also on technologies based on dynamic learning techniques. These techniques must be integrated and sensors distributed at the host and network layers in order to provide coverage of both outsider and insider threats. Automated responses should include not only warnings but defensive actions that occur within the propagation time of an attack in order to mitigate it.

Importance

The effects of a wide-scale cyber attack could be devastating, especially if coupled with a physical attack. Static intrusion detection and prevention mechanisms that reside at network boundaries may not always be capable of stopping malicious code and worm attacks that can gain a foothold from within the network and spread rapidly. Organizations should adopt security strategies that deploy mechanisms at all levels of a network or system. But next-generation tools that can deal dynamically with ever-more sophisticated attack technologies are also needed. The need for new automated attack recognition and warning technologies spans threats from scripted attacks through ad hoc hacking to Trojan horses, viruses, self-replicating code, and blended threats.

Today, the spread of a new Internet worm results in several tiers of reaction: 1) knowledgeable network operators try to block the worm by configuring switches, routers, and firewalls; 2) an updated signature is created to stop the worm via antivirus and intrusion prevention systems; and 3) a patch is

created to fix the underlying vulnerability. Effective response times range from hours to weeks today but will need to be under a second in order to deal with attacks such as flash worms. Only sophisticated automated response techniques can provide such speedy protection.

State of the Art

A large percentage of the network security market still relies on signature- and rule-based systems. The migration from those systems to anomaly-detection and dynamic self-learning systems is just beginning. One new capability involves high-level security event managers, technologies that correlate alerts from multiple sensors and network logs with the goal of providing network alerts with fewer false positives and minimizing the amount of data the network operator must examine. While the commercial sector has been slow to target automated attack response, Federal research is developing the capability to automatically detect and respond to worm-based attacks against networks, provide advanced warning to enterprise networks, study and determine the worm's propagation and epidemiology, and provide off-line rapid response forensic analysis of malicious code.

Capability Gaps

Security-event management systems, which provide IT operators with a synthesized real-time overview of network and system activity, mark a step forward in situational awareness. But they represent the high end of the current state of the art and few address the requirement for automated response. Significant R&D will be needed to move automated attack detection and warning technologies beyond rules and signatures. Future generations of these technologies need to be preemptive rather than reactive, but automating system defense behaviors will require advances in machine learning technologies. These advances must be accompanied by new techniques and tools that help network operators understand and interact with the automated decision making process.

For example, a defensive computer system that can block or quarantine a suspected attack will have to be a highly trusted system, and even then, strong methods must also be put in place to enable operators to quickly reverse any actions taken by a system that they deem inappropriate. The combination of human and machine learning through making correct (true positive and true negative) decisions will reinforce the underlying intelligence. Weighting mechanisms are likely to be required so that the defense system will evaluate multiple factors, including the likelihood and potential impact of attack, and will be more or less likely to isolate systems depending on their criticality and the degree of certainty of an attack.

Defining and characterizing defensive courses of action is another capability gap. When a system is attacked, there is very little time for humans to react. To minimize the potential damage and stop or quarantine the cyber attack, there is a need for decision support systems that can rapidly provide defensive courses of action and alternatives to network defenders.

1.6 Insider Threat Detection and Mitigation

Definition

An insider threat can be defined as the potential damage to the interests of an organization by a person who is regarded as loyally working for or on behalf of the organization. Within the IT environment of networks, systems, and information, an organization's interests can be embodied in both implicit and explicit security policies; in this environment, an insider threat can be more narrowly defined as the potential violation of system security policy by an authorized user. Although policy violations can be the result of carelessness or accident, the core concern is deliberate and intended actions such as malicious exploitation, theft, or destruction of data, or the compromise of networks, communications, or other IT resources. Detection involves differentiating suspected malicious behavior from normal as well as unusual yet acceptable behavior. Mitigation of the insider threat involves a combination of deterrence, prevention, and detection.

Importance

The intelligence community (IC) and DoD communities are environments in which access to classified information is available to appropriately

cleared members. One of the most harmful and difficult to detect threats to information security is the trusted insider who uses privileges in a malicious manner to disrupt operations, corrupt data, exfiltrate sensitive information, or compromise IT systems. Loss of intelligence operations and/or information will ultimately compromise the Nation's ability to protect and defend itself against future attacks and to safeguard military and intelligence assets working abroad. In fact, some of the most damaging cyber attacks against the IC have been launched by trusted insiders. Such attacks will become an increasingly serious threat as increased information sharing results in greater access to and distribution of sensitive information. The private sector, where corporations maintain valuable and highly sensitive proprietary information, and where banking institutions manage the flow of and access to electronic funds, share similar concerns over insider activity.

State of the Art

Techniques to mitigate the insider threat focus on monitoring systems to identify unauthorized access, establish accountability, filter malicious code, and track data pedigree and integrity. While an array of partial measures exists for countering the insider threat, these measure are limited in scope and capabilities. Among the challenges that add to the difficulty of this problem are:

- The scale and diversity of the computing infrastructure, in terms of numbers and types of platforms, missions supported, infrastructure architectures and configurations, and worldwide geographic distribution
- The size, variety, and fluidity of the workforce in general and, in the case of military missions, the need to interface with allied and ad hoc coalition partners
- The variety of highly complex computer security environments that range from unclassified systems to classified networks, and from private sector systems and networks that support business and electronic commerce to critical infrastructure process control systems
- Policy discovery, which is the process by which the kinds of access permitted to insiders is determined. Such policies are difficult to formulate.

The trusted insider operates within this large interconnected world of information systems relatively unchecked and unmonitored beyond the basic security mechanisms used primarily to detect untrusted outsiders and prevent them from penetrating and exploiting information assets. These factors make insider threat a complex problem that is beyond the scope of commercially available tools.

Capability Gaps

Both prevention and detection of malicious insider activity can be made more effective through use of models that capture and predict the knowledge, behavior, and intent of insiders. The subjectivity of user behavior makes it difficult to distinguish acceptable, or authorized, behavior from unauthorized behavior regardless of whether the user is considered trusted or untrusted. The analysis becomes even more difficult in the complex environment described above. Thus, a capability that can reliably model and differentiate between normal, unusual but acceptable, and unacceptable user behavior in a complex setting, and that can detect and react to the identified malicious insider activity, will provide a solid foundation for successful mitigation of the insider threat.

Accurately and quickly identifying malicious insider activity within large volumes of electronic records (e.g., network access and audit logs) also requires effective behavioral models. In an illustrative scenario, a network of sensors calibrated based on such models would monitor and record user activity throughout the complex enterprise. Analytical tools could then be used to sift through the various data stores, correlating the information and presenting it in a meaningful format. The correlation techniques could use the models to sort through and map the myriad pieces of information into a complete picture of the insider activity, which could be presented via data visualization to interested parties.

For addressing insider threats, more advanced methods of document control and management – the

ability to provide an unalterable accounting of document access and dissemination – constitute a major Federal capability gap, given the sensitivity of many types of Federal information and the increasing demands for information sharing. The intelligence and DoD communities, which routinely handle many levels of data and document sensitivity, have particularly acute concerns in this area. Document control and integrity must be applied to the entire document as well as to labeled portions in a manner that is resistant to tampering and circumvention, and must be managed in accordance with data sensitivity and originator controls. The document management capability must ensure that appropriate control requirements are translated into an implementable security policy that is applied to the document and preserved from cradle to grave.

A capability is also needed to prevent document tampering and to maintain the integrity of the information, the associated sensitivity labels, and any dissemination controls contained within the document. One approach to dissemination controls is called labeled paths, in which knowledge of the path of access from point of creation to destination can be used to determine if all accesses along the path are permitted. An enhanced digital rights management regime should be developed that enforces control and accountability over information in accordance with a specified security policy whenever an individual attempts to read, write, modify, print, copy, distribute, or destroy information or the associated sensitivity labels and dissemination controls. This capability would allow fine-grained customization of a user's control over all or part of the document in accordance with a specific security policy.

This general concept of limiting exposure to insider threats applies more generally as well. Policies that embody the principle of least privilege would make it less likely that insiders would have privileges that enable them to conduct activities they should not be permitted to conduct. However, applying this concept requires technical means for translating the principle of least privilege into practice via configuration and enforcement of access control and other security policies at the implementation level.

1.7 Detection of Hidden Information and Covert Information Flows

Definition

Steganography, derived from the ancient Greek words for "covered writing," is the art and science of writing hidden messages in such a way that no one apart from the intended recipient knows of the existence of the message. Detection of covert information flows relies on the ability to detect information hidden within a stream of information that is transmitted from one system to another.

Steganographic data are hidden or embedded through the use of mathematical techniques that add information content to digital objects – usually images, video, and audio, but also other digital objects such as executable code. When sophisticated techniques are used, little or no degradation in quality or increase in data size in the resulting object is perceptible. A steganographically embedded message may also be encrypted. There is no universally applicable methodology for detecting steganographic embeddings, and the few general principles that exist tend to be ad hoc. In cyberspace, steganography provides a capability for transmitting information undetected.

Steganalysis is the examination of an object to determine whether steganographic content is present, and potentially to characterize or extract such embedded information. Watermarking refers to embedding content that conveys some information about the cover object (e.g., copyright information). In this context, digital data forensics is the use of science and technology to determine information about how a digital object was formed or modified, even in the absence of the original object.

Importance

International interest in R&D for steganographic technologies and their commercialization and application has exploded in recent years. These technologies pose a potential threat to U.S. national security. Because steganography secretly embeds additional, and nearly undetectable, information

content in digital products, the potential for covert dissemination of malicious software, mobile code, or information is great. Hundreds of steganographic software tools are readily available either commercially or as freeware and shareware downloads from the Internet. The affordability and widespread availability of these tools makes steganography an enabling technology for U.S. adversaries. The threat posed by steganography has been documented in numerous intelligence reports.

State of the Art

R&D advances in steganalytic solutions have led to operational prototypes for evaluation and use by the DoD and intelligence communities. Vulnerabilities in steganographic tools and techniques have been made public. Techniques that detect certain classes of embedders have been developed, which detect the presence of and estimate the amount of embedded content. Blind universal steganalyzers have begun to detect and classify cover images in which content has been embedded using one of several embedding techniques. The effectiveness of steganographic key searches has been demonstrated for a few embedding methods, and limited parallelization of key search capabilities has begun. Once a steganographic key has been found, the embedded content can be extracted from the cover object and exploited. Many of these advances have addressed digital image cover objects.

Covert information flows can be achieved by hiding information within a legitimate flow of information, or even by manipulating attributes (e.g., timing) of an information flow. For example, a security researcher has demonstrated the ability to use DNS requests to create covert channels for information flows. Covert information flows can be extremely difficult to detect, when information flows are being monitored for covert channels. More importantly, many legitimate information flows (including DNS requests) typically are not monitored at all for the embedding of covert information.

Capability Gaps

Targeted steganalysis techniques are useful only for a single embedding algorithm or class of algorithms. A full spectrum of such techniques is required to effectively combat the use of steganography by adversaries. More effective steganalysis is needed to counter the embedding algorithms that continue to be developed and improved, such as matrix, model-based, and wet paper code embedding. Blind universal steganalyzers need to be fully evaluated, shown to be scalable, and deployed. Key search algorithms for realistic key lengths require substantial computational power. To be more widely useful, steganalysis capabilities must be made more efficient through the use of specially programmed high-performance computing platforms. Steganalysis is less advanced for audio, video, documents, and other forms of data than for static digital images.

Advanced methods for detecting covert channels for information flows need to be developed and, when risks warrant a high level of security, these capabilities need to be deployed to monitor a variety of different types of legitimate information flows to identify covert communications. Resources to evaluate, integrate, and deploy the numerous basic research advances are limited and should be enhanced.

1.8 Recovery and Reconstitution

Definition

Recovery and reconstitution refer to the capabilities needed in the wake of a cyber attack to restore the functionality and availability of networks, systems, and data. Recovery and reconstitution methods must be adequate to cope with the consequences of cyber attacks that are carried out quickly, cause extensive damage, and propagate in uncontrolled ways.

Importance

Recovery and reconstitution must be addressed and implemented in all aspects of a system – networks, operating systems, middleware, applications, and data. Capabilities for timely recovery and reconstitution are especially important in mission-critical systems, which must be able to degrade gracefully, meaning that they must be able to survive a cyber attack even if damaged and recover to an operable state that sustains mission-critical functions. Systems must be made self-healing and self-restoring to as great a degree as possible. Self-restoring means that, as portions of a system fail, a new system is dynamically formed by rerouting

critical traffic and migrating critical data to undamaged nodes. The recovery and reconstitution aspects of dynamic response depend on accurate, timely detection of cyber attacks. The spreading of malicious code across a network needs to be stopped, for example, and damaged nodes need to be recovered while residual malicious code is eradicated. This technical area is closely linked to large-scale cyber situational awareness, which provides the information required to perform recovery and reconstitution.

State of the Art

Current technologies for recovery and reconstitution are limited. The most common recovery and reconstitution techniques are redundant processing, physical backups, and the use of special service providers to implement recovery capabilities for organizations. These procedures focus on system faults, failures, and accidents, not purposeful, malicious cyber attack. Technologies in use today are able to return databases, applications, and data to an operational state after non-malicious faults or failures. Research in self-regenerating systems is investigating technologies to enable systems that have been exploited to restore themselves autonomously, but the techniques are still in their infancy.

Capability Gaps

Recovery techniques tend to be aimed at data recovery rather than at reconstituting large-scale systems or networks. Research is needed to better understand the extent to which today's Internet would recover from attacks, particularly wide-scale attacks, and how such recovery can be accomplished.

To be effective, recovery and reconstitution must be rapid and must be guided by accurate and timely information. Damage-assessment technologies are needed that can quickly provide network defenders with an accurate snapshot of the overall enterprise, what has been attacked, where the damage is, what type of damage has been incurred (whether the attack is against the confidentiality, the integrity, or the availability of the system), and what parts of the system have been affected. Defenders also need robust decision-support systems that can rapidly present possible defensive courses of action. In some cases, autonomic (self-managing) system responses directed at recovery and reconstitution potentially could maintain a basic level of operation while further analysis of the cyber attack is being conducted.

While there might be enough network redundancy to allow continued communications, an attack might reach deep into a networked system. Techniques are needed to assess whether data are damaged and, if so, the extent and impact of that damage. Damaged data require recovery to an earlier undamaged state followed by reconstitution. Applications may need to be reloaded to ensure that malicious code has been eradicated. Remediation may also be needed to eliminate vulnerabilities that enabled the attack in the first place. Rapid post-attack reconstitution of IT systems requires the ability to create checkpoints that capture the state of a large-scale system and to not only retrieve undamaged data (as mentioned above) but also to roll back damaged systems to earlier functional (uncompromised) states. Such rapid reconstitution is an alternative to rebuilding the entire system from scratch and is vital for mission-critical systems.

1.9 Forensics, Traceback, and Attribution

Definition

Forensics, traceback, and attribution are functions performed in the process of investigating cyber anomalies, violations, and attacks. They help answer such basic investigative questions as: what happened to the computer, system, or network; where an attack originated; how it propagated; and what computer(s) and person(s) were responsible. Cyber forensics can be defined as the application of scientifically proven methods to gather, process, interpret, and use evidence to provide a conclusive description of a cyber attack; this evidence enables operators to restore systems, networks, and information after an attack. Forensic analyses help operators correlate, interpret, understand, and predict adversarial actions and their impact on system, network, and IT infrastructure operations; and provide evidence for a criminal investigation.

The goal of traceback capabilities is to determine the path from a victimized network or system through any intermediate systems and communication pathways, back to the point of attack origination. In some cases, the computers launching an attack may themselves be compromised hosts being controlled remotely from a system one or more levels farther removed from the system under attack. Attribution is the process of determining the identity of the source of a cyber attack. Types of attribution can include both digital identity (computer, user account, IP address, or enabling software) and physical identity (John Doe was the hacker using the computer from which an attack originated). Attribution can also support a new model of authorization using accountability as a basis for deciding which operations or resources to trust.

Importance

The prospect of accountability under law is a key deterrent of common crime. Unfortunately, in cyberspace accountability is nearly nonexistent. Public access points often allow anonymous access to the Internet. Even when some form of identification is used at the initial point of access, users can then proceed to move about with relative anonymity through the use of a variety of tools that include anonymization services or data and communication path obfuscators. Due to the lack of authentication capability associated with Internet communication protocols (or the frequent lack of implemented authentication when the capability exists), it is often possible to spoof (i.e., forge or manipulate) the apparent source of hostile communications, allowing malicious actors to keep their location and identity hidden. Moreover, hackers often hide their tracks by hacking and communicating through numerous compromised machines, making it difficult to determine the host from which certain Internet traffic originates.

This issue is exacerbated by the multinational nature of the Internet, which allows these network hops to be routed through compromised hosts located in countries that may not have strong relationships with U.S. law enforcement, and may not cooperate with investigative efforts. In addition, it remains difficult to associate a digital identity with a specific human being. Forensics, traceback, and attribution can help mitigate the shortcomings of the Internet's design by denying attackers anonymity and safe haven.

State of the Art

Current commercial investment in computer forensic tools is focused on the needs, practices, and procedures of law enforcement. Typically, law enforcement is incremental and conservative in adopting new IT processes and capabilities. Officials require vetted processes, repeatable procedures, and high assurance of the integrity of any collected data. Such law enforcement processes and procedures are not easily adaptable to new technologies. Cyber investigations are often conducted after the fact on systems that have been turned off, so that crucial information still in system memory (such as malware, running processes, and active network sessions) may be lost. Because evidence preservation and analysis can take 90 days or more, digital trails on the Internet revealed by the analysis may already be cold. For the most part, the current generation of investigative tools substantially lags behind the capabilities that the law enforcement community would like to have.

Today, many investigative procedures in computer network defense begin with ad hoc inquiries into computer or network states for the purposes of rapid situational awareness and development of courses of action to contain ongoing cyber attacks. Most analysis is performed by human operators, including timelining and event reconstruction from network- and host-based intrusion detection systems (IDS). IDS reports about possible intrusion activity and sources, however, are not automatically validated. Rather, human investigation is needed to distinguish legitimate events from intrusions. Many forensic IDS and enterprise security tools are not forensic in the scientific sense but support a human cognitive investigative process.

Traceback capability is limited by the ability of attackers to spoof source IP addresses. Some standard network information sources (such as traceroute and

DNS registries) can often trace a path back to a host Internet service provider (ISP). Router netflow (a metering technology for network measurements) information, when available, can also be useful. Geographic location information may be accurate at the country or state level but may not be practical with satellite-based ISPs.

Dynamic IP address assignment and spoofing make attribution a significant technical challenge. Generally, only with cooperation from the attacker's ISP might the attacker be identified. Many times, however, the evidence to attribute the attack to an individual remains inconclusive. Open wireless access points, Internet cafes, and similar venues that allow Internet access without positive identification and authentication further exacerbate this problem.

Capability Gaps

The following IT capabilities, needed for forensics, traceback, and attribution for both law enforcement and network defense, do not yet exist and require R&D:

- The ability to track individuals or computers as they access the Internet from various ISPs and IP addresses, and particularly over non-cooperative networks, to address cases in which network owners or service providers are unwilling to cooperate or where networks themselves have been compromised
- Live investigation and preservation of digital evidence on a target computer performed remotely over the Internet. This would enable investigators to react to attacks in real time and preserve potential evidence still in memory.
- Network forensics, especially in the discovery, investigation of, and response to stepping-stone attacks. Substantial network information may need to be available for traceback to be reliable.
- Techniques for sampling evidence in ways that provide confidence in the results. These techniques would constitute a valuable decision support tool for computer network defense analysts.
- The ability to determine analytical confidence factors that can predict the error rate associated with having processed only a portion of available evidence (for example, the finding that some subset of selected evidence yields a particular percent confidence level in the analysis). This would be a valuable decision support tool for computer network defense analysts.

2. Securing the Infrastructure

The R&D topics in this category are focused on improving the inherent security of the information infrastructure, ranging from protocols on which the Internet relies to critical infrastructure systems that depend on a secure information infrastructure to operate. Topics in this category are:

- Secure Domain Name System
- Secure routing protocols
- IPv6, IPsec, and other Internet protocols
- Secure process control systems

2.1 Secure Domain Name System

Definition

The Domain Name System (DNS) is a globally distributed database that provides two-way mappings between domain or host names (for example, www.whitehouse.gov) and IP addresses (for example, 63.161.169.137). Nearly all Internet communications are initiated with a DNS request to resolve a name to an IP address. Although it is arguably one of the most critical components of the Internet's architecture, the current DNS is not secure and lacks authentication and data integrity checks. Protocol exchanges in the system (e.g., resolver-to-server and server-to-server) are subject to malicious attacks. An example of such an attack is "zone hijacking" in which third parties impersonate entire DNS zones and redirect network traffic to their own machines for malicious purposes.

Importance

Because the DNS is at the core of most Internet communications, developing technologies that make it more difficult to undermine the DNS infrastructure could mitigate some existing vulnerabilities. The Domain Name System Security Extensions (DNSSEC) provide origin authentication and data integrity checks for DNS lookups. This is accomplished by adding digital signatures and public keys to the DNS. When an Internet application sends a DNS query for a host name, it can request that DNSSEC security information be returned with the response.

A secure DNS would comprise a signed DNS tree and one or more "trust anchors." Trust anchors are public keys associated with DNS zones high up in the DNS hierarchy such as .gov, .mil, .com, or the root "." that serve to create chains of trust at lower levels of the hierarchy by digitally signing keys for domains under them: The public key for the domain whitehouse.gov would be signed by the key from .gov. If a host trusts the .gov key, it can verify and trust a newly learned whitehouse.gov key, and consequently, the whitehouse.gov domain information.

Beyond providing name-to-address resolution, the DNS infrastructure is being expanded to address security and robustness issues. Examples of these uses include adding mail authentication information as part of current anti-spam proposals and adding public keys and digital certificates in support of other secure communication protocols. With this expanding role comes an increased need to assure the authenticity of the DNS responses and an increased possibility that the DNS itself will be targeted for attacks. As end systems become more secure, attackers wishing to disrupt Internet-based communications may discover that it is easier to achieve their goal by subverting the underlying protocols such as DNS on which the end systems rely. In addition, as the DNS is expanded to serve other security-related requirements (e.g., anti-spam techniques), it may become even more of a target for malicious attacks by those who seek to subvert these efforts.

State of the Art

Under the auspices of the Internet Engineering Task Force (IETF), the latest versions of the DNSSEC specifications are being completed and early implementations of DNSSEC-capable servers are emerging. Many DNSSEC technical challenges are associated with deployment and operational issues. Key organizations – including the Internet Corporation for Assigned Names and Numbers for the root for the DNS tree, large registries for domains such as .com and .net, and some country code registries – are examining the potential for DNSSEC deployment in major sub-trees of the DNS. These

initial evaluations of DNSSEC's viability for large-scale deployment have resulted in the identification of several areas for further R&D.

Capability Gaps

Zone enumeration and privacy: DNS information can be, and is, used for malicious purposes. Spammers use DNS registry information to compile lists of e-mail addresses and attackers can try to map an enterprise's network resources by looking at the DNS. Because it is public, DNS information is inherently difficult to protect from misuse, but several implementations can block certain types of potentially malicious or probing lookups. Unfortunately, the initial design of DNSSEC mechanisms to support authenticated negative responses (i.e., a reply stating that the requested name does not exist) also makes it easy to circumvent these safeguards and to enumerate all records in a given DNS domain. Concerns over this side effect range from the possibility that it makes it even easier to find information that is already generally accessible, to the legal worry that the DNSSEC mechanisms may make privacy protection more difficult. The authenticated negative response parts of the DNSSEC specification will require a new round of requirements analysis, design, standardization, and implementation.

"Last-hop" issues: While significant effort has been devoted to developing DNSSEC protocols and procedures for use among DNS servers, how applications and host DNS clients (i.e., resolvers) interface with, control, and respond to a DNSSEC-enabled infrastructure is largely undetermined. The relative lack of involvement and inputs from the host and application development communities represents a challenge to overall adoption, deployment, and use of DNSSEC. To address these "last-hop" issues, industry must draft requirements, specifications, and basic control processes for application interfaces, as well as certain protocol extensions and policy and management mechanisms to define the ways in which applications and hosts interact with DNSSEC.

Administrator tools and guidance: Even when the specifications are completed, work will remain to deploy DNSSEC and foster its adoption. Many network operators do not perform DNS administration tasks full time and may not have sufficient expertise to deploy DNSSEC correctly. Tools and guidance documents are needed to assist network administrators in the new tasks of deploying and maintaining a DNSSEC domain. The issues that must be addressed include: public key management, DNS domain signing, and new and increased registrant-registry communications.

Performance analysis: The DNSSEC extensions will make DNS transactions more complex. DNS servers and clients will have to perform cryptographic operations and judge the validity of responses in addition to performing DNS name lookups. Some network operators argue that DNSSEC will slow DNS to the point of making it impractical to deploy on a large scale. The validity of this claim has not yet been determined. To understand and resolve such concerns, a measurement basis must be designed and implemented for DNS and DNSSEC technologies. This measurement basis would be made up of models of both the contents of the DNS tree and the traffic seen at various levels in the tree, and test and measurement tools capable of exercising and evaluating specific implementations or partial deployments using such models. Developers and administrators could then use these tools and reference data sets to test various DNS configurations and to gauge the relative performance impact of DNSSEC technologies.

2.2 Secure Routing Protocols

Definition

Routing infrastructures – made up of the equipment, protocols, data, and algorithms that compute paths through interconnected network devices – organize disparate collections of connected devices into viable end-to-end paths over which all network data flow. They reside in multiple layers of network architectures – from Layer 2 systems that interconnect local area network switches, to Layer 3 systems that perform IP routing, to content- and context-based systems at Layer 4 and above.

Importance

IP routing (Layer 3) protocols interconnect public and private networks. The IP routing infrastructure comprises tens of thousands of individual routing domains employing numerous protocols and technologies operating as a hierarchical interdependent global distributed system. Routing infrastructures are among the least protected components of the overall IT infrastructure.

Many large-scale routing systems are not highly robust because there are inherent trade-offs between responsiveness and stability. To date there have been few focused attacks on routing infrastructures. However, as hosts and applications are hardened in response to common attacks on networks, network attackers may increasingly focus their attention on the underlying routing control systems.

Currently deployed Internet routing protocols are vulnerable to several classes of malicious attack. The conceptually simplest attack is the compromise and control of routers. The routing protocols and the associated router resources are also vulnerable to attack from remote nodes. These attacks focus on the resources of the router's control plane, the peering relationships between connected routers, and/or the data that the routing protocol exchanges between router peers. Attacks can also focus on the lower-layer resources (e.g., physical links, and lower-layer protocols) associated with the routing infrastructure.

Successful attacks on routing protocols can result in loss of connectivity (e.g., black holes, partitions), eavesdropping and theft of data, sub-optimal routing, or routing system disruption. All common currently deployed routing protocols are vulnerable to these attacks. As additional routing system services such as traffic engineering and QoS-sensitive routing are implemented, the potential for disruptions by attacks on the routing infrastructure increases.

State of the Art

IP routing technologies and protocols vary greatly depending upon their application. The most widely used is unicast routing among fixed, non-mobile hosts, which employs a two-level hierarchy of protocols. Interior Gateway Protocols (IGPs) are used within a single administrative or management domain (called an autonomous system) that typically has sparse connectivity but little control of topologies. IGPs typically exploit all possible paths for optimal responsiveness with little concern for policy and trust. Inter-domain protocols, known as Exterior Gateway Protocols (EGPs), route traffic between autonomous systems. EGPs typically enforce policy (i.e., using only policy-feasible paths) and emphasize global stability. The primary EGP deployed today is the Border Gateway Protocol (BGP).

Other routing protocols such as multicast or broadcast protocols for one-to-many and one-to-all communications are emerging. While their deployment in the commercial Internet is limited, they play important roles in various private and special-purpose networks. Protocols for wireless, mobile, and ad hoc networks are also rapidly growing in importance and deployment; this class of protocols makes different assumptions about the composition of networks and the trust relationships between components. Ad hoc routing assumes that there are no fixed infrastructure services to rely on, that all routing relationships are ephemeral, and that the composition of the network is constantly changing. Mobile routing is based on the same assumptions, with the addition that the nodes, or entire networks, are constantly moving.

Capability Gaps

Securing the routing infrastructure is a difficult technical problem. Complete, viable security solutions for most routing technologies have not yet been

designed and standardized. Difficulties arise from the attributes and assumptions of routing systems themselves. The solutions must address protocols that vary widely in their design and operation. A single protocol typically must address both peer-to-peer and multi-party communications, single-hop and multi-hop messages, as well as mutable and immutable data components.

Adding cryptographic protections to routing protocols also poses difficult technical issues; the topology of a given network is not always known and clock synchronization is difficult; multiple trust relationship graphs exist (e.g., customer to service provider, address administration, intra-domain vs. inter-domain); routing services must be able to bootstrap themselves and thus typically cannot depend upon other components of the infrastructure for their most basic start-up operations.

Additional security constraints are imposed by dynamic performance requirements and the need to address the trade-off between Internet stability and scalability. Global convergence and stability properties should not be compromised by security mechanisms; however, these properties are poorly understood for the largest routing systems (e.g., global BGP). There are also constraints on the platforms on which routing protocols operate. Specifically, at the core of the Internet there are orders of magnitude difference in the processing capabilities of the control and data planes, while at the mobile edge there may be constraints on processing power and battery life. In addition, security mechanisms must include viable means for incremental and/or partial deployment, day-to-day operations and management, as well as favorable risk and cost/benefit models.

There are no widely deployed secure routing protocols in use today. The current state of the art in protecting routing infrastructures uses basic techniques (e.g., passwords, TCP authentication, route filters, private addressing) that mitigate only rudimentary vulnerabilities and threats. The R&D community has been pursuing more complete solutions both at a theoretical level and through specific extensions to commonly used protocols. To date, these proposed extensions have not achieved widespread commercial implementation or deployment, in part because they are perceived as optimizing for security concerns at the cost of not adequately meeting scalability and performance requirements and constraints.

Renewed interest in routing security has begun to develop in the IETF and R&D communities. New proposals for secure variants of BGP are attempting to provide a better balance between security and performance. Security of ad hoc routing protocols continues to be a major practical concern.

To expedite development, adoption, and use of secure routing technologies, several key R&D areas need to be addressed, including:

- Risk analysis – understanding the potential risks associated with security vulnerabilities and other forms of focused, large-scale disruptions to the routing systems
- Secure protocol architectures – new designs for the decoupling of routing and security functions that address separation of control and data planes and incorporation of programmable technologies in the data plane (e.g., along the lines of DARPA's active networks efforts)
- Flexible and survivable secure routing – flexible designs that address security as one component of overall viability and survivability of the routing infrastructure. The designs should also address environments in which reputation management is a continuum rather than a binary decision and in which security systems selectively and dynamically adapt mechanisms to trade off threat mitigation for performance, scalability, and cost
- Efficient security mechanisms for routing – new cryptographic techniques to ensure the authenticity, integrity, and freshness of routing information, and that perform more effectively and efficiently than those previously proposed
- Secure routing systems – system-level designs that integrate other security technologies (e.g., intrusion and anomaly detection, firewalls) as part of the secure routing system

- Secure self-organizing networks – classes of routing technologies that support wireless ad hoc and sensor networks as well as large-scale, peer-to-peer, and grid-like distributed systems. Self-organizing networks may not necessarily assume the existence of any fixed infrastructure, and they pose security challenges associated with secure group formation, membership management, and trust management between dynamic groups.

2.3 IPv6, IPsec, and Other Internet Protocols

A number of IT infrastructure-related protocols are being developed under the auspices of the IETF and the Internet Research Task Force. These protocols address not only security but the growth of networking and the diversification of network uses. The following discusses the most significant of these emerging protocols.

Internet Protocol version 6 (IPv6)

Definition and Importance of IPv6

IPv6 was developed to enhance the capability of IPv4 by providing a vastly increased address space, to provide header space to meet security and other requirements, and to provide additional capability enhancements. The additional address space is needed to support expected large increases in the number of networked devices due to Internet growth, sensors and sensornets, and mobile network devices.

State of the Art of IPv6

IPv6 is being implemented in testbeds by several Federal agency networks, including DARPA's Advanced Technology Demonstration Network (ATDnet), DoD's Defense Research and Engineering Network (DREN), DOE's Energy Sciences network (ESnet), and NASA's NASA Research and Education Network (NREN) and Integrated Services Network (NISN). In addition, numerous Federal research network exchange points, including Americas Pathway (AMPATH), Next Generation Internet Exchange (NGIX)-East, NGIX-West, StarLight, Pacific Wave, and Manhattan Landing (MANLAN), currently support IPv6. These testbeds and exchanges implement IPv6 in dual-stack mode and also support IPv4. Currently, IPv6 traffic on these networks and exchanges is for testing IPv6 services and capabilities.

Substantial IPv6 operational and applications traffic is not expected on mainstream networks until a significant proportion of nodes and applications are IPv6-capable. IPv4 and IPv6 are expected to coexist within the Internet for some time, through mechanisms that include dual-stack routers, hosts, and other devices, as well as tunneled communications through pockets that are exclusively IPv4 or IPv6. IPv6 will not in itself fully eliminate some of the existing obstacles to end-to-end security protection. For example, because Network Address Translation/Translator (NAT) boxes, which hinder the use of Internet Protocol Security (IPsec), may continue to be used under IPv6, new firewall and other security technologies will need to be developed for operation in IPv6 environments.

A dual-protocol Internet presents numerous security pitfalls. Even enterprises that are running solely IPv4 or IPv6 need to be aware of all possible combinations because nodes within the network can individually enable one or both protocols; unexpected tunneled traffic can travel through a firewall if the firewall rules are not sufficiently robust and comprehensive.

OMB has directed that by June 2008, all Federal agency backbone networks must implement IPv6 (at least in a dual-stack mode) and agency networks must interface with this infrastructure.

IPv6 Capability Gaps

The immaturity of current IPv6 security tools results in high levels of risk for breaches of IPv6 security. Research is needed to provide a full suite of security tools and support to make IPv6 as secure as current implementations of IPv4. Robust DNS security for IPv6 needs to be developed and implemented. Specific research needs include:

- Security threat and vulnerability models to assess the security implications of widespread IPv6 implementation

- New scalable technologies to provide end-to-end security
- Techniques and tools capable of managing the proliferation of addresses and devices facilitated by IPv6
- Scalable routing to handle the demands of the IPv6 address space
- Packet filtering for IPv6 at speeds comparable to IPv4 (e.g., line rate access control list processing)
- Tools and infrastructure for managing and testing IPv6, including a rigorous conformance and interoperability testing infrastructure. Current industry standards would not support government procurement requirements or regulations.
- Business cases, detailed timelines, and scenarios for deployment and use of IPv6

A business case needs to be developed in support of deploying IPv6 functionalities on a scheduled basis. A May 2005 GAO report on IPv6 lists some benefits and risks as a start toward making that case. Such a case should address current functionality of IPv4 (including IPsec functionality) and the availability of tools to support IPv6.

Internet Protocol Security (IPsec)

Definition and Importance of IPsec

IPsec is a suite of protocols standardized through the IETF to provide security at the network layer (Layer 3). IPsec can provide confidentiality, integrity protection, peer authentication, traffic analysis protection, and replay protection. Its companion protocol, Internet Key Exchange (IKE), negotiates and manages the details of the IPsec protections and the secret keys used to provide these protections.

State of the Art of IPsec

IPsec and IKE are most frequently used to create virtual private networks and protect mobile users who need to access protected business networks and resources from outside the protected network. Many firewalls and routers incorporate IPsec/IKE functionality and many operating systems have built-in IPsec/IKE clients. The protocols are currently being updated to version 3 for IPsec and version 2 for IKE. IPsec is a mandatory component of IPv6. Although it is optional for IPv4, it has been added to many operating systems and to many gateways and routers. Numerous add-on IPsec clients are also available.

Capability Gaps of IPsec

Security of the emerging IPsec/IKE standards: The current version of IKE (IKEv1) has undergone formal protocol analysis and the current versions of IPsec and IKE have been subjected to considerable security and functional analysis. Because even minor changes to security protocols can introduce security holes due to unexpected feature interactions or other unforeseen problems, new versions of these protocols should also be rigorously tested prior to implementation and deployment.

Use of certificates and smartcards within IPsec and IKE: Public key certificates are the recommended mechanism for peer authentication within IPsec/IKE. However, they have been a source of numerous problems, including lack of interoperability among disparate domains, failure of IKE negotiations as a result of message fragmentation (due to the size of certificates that are sent as part of IKE messages), and time-outs related to the certificate revocation list checking process. For IPsec/IKE to be applied in a widespread, scalable, and secure manner, certificate problems must be addressed. Further testing and research are needed to ensure that PKI can be used with IPsec in a scalable, secure, and interoperable manner.

Host Identity Protocol (HIP)

Definition and Importance of HIP

IP addresses perform two functions: unique endpoint identifier and routing locator. Functional overloading causes problems in such diverse areas as route aggregation, host multi-homing, and network renumbering. During the development of IPv6, attempts were made to split the IP address into two parts, each performing one of these functions, but no satisfactory solution was found. The Host Identity Protocol (HIP) is another attempt to separate these functions. It introduces a new Host Identity (HI)

name space, based on public keys, in the TCP/IP stack. This cryptographic identity can also be used to provide authenticated, secure communications.

Capability Gaps

Currently, the HIP base protocol works well with any pair of cooperating end hosts. However, to be more useful and more widely deployable, HIP needs support from the existing infrastructure, including the DNS, and a new piece of infrastructure, called the HIP rendezvous server, which facilitates the use of HIP for mobile hosts. HIP is considered to be sufficiently promising that an Internet Research Task Force research group has been chartered to explore its global ramifications within the Internet architecture.

Network Address Translation (NAT)

NAT is employed in private networks to keep the hosts' addresses secret for security and privacy purposes. It is also used in networks that have exhausted their allocation of address space to implement private addresses that may duplicate addresses used elsewhere on the Internet. In this case, a pool of public, globally unique addresses is used for communications with destinations outside the private network. When such messages cross the NAT box, the private address of an outbound communication is converted to a public address and the public destination address of an inbound communication is converted to the corresponding private address. While effective at addressing these issues, NAT complicates IPsec security: it is compatible with some types of IPsec functionality but incompatible with others, some of which have work-arounds. The existence of NAT must be considered when implementing IPsec or any other type of end-to-end security.

Mobile Internet Protocol (MIPv4, MIPv6)

MIP allows transparent routing of IP data to mobile nodes on the Internet and includes separate specifications for IPv4 and IPv6. Each mobile node is identified by its home address, regardless of its current point of attachment to the Internet. While away from its home, a mobile node is also associated with a "care-of" address that provides information about its current attachment point. The protocol provides for registering the care-of address with a home agent. The home agent sends data destined for the mobile node through a tunnel to the care-of address, and the data are delivered to the mobile node at the end of the tunnel. MIPv4 is currently deployed on a wide basis such as in cdma2000 networks.

For MIP to function correctly, several types of security protection are essential: home agents and mobile nodes must perform mutual authentication; replay protection is necessary to ensure the freshness of update messages; and data may be encrypted to provide confidentiality. MIPv6 mandates IPsec for the protection of binding update messages, which direct the home node to forward data to the care-of address, between mobile nodes and home agents.

Multicast Communications Protocols

Multicast communications carry traffic from a single source host to multiple destination hosts. They are used for applications as diverse as video broadcasts, teleconferencing, distance learning, multi-player video games, and news, stock market, and weather updates. While a multicast message is to be delivered to multiple destinations, only one copy of the message is transmitted along a given network segment on its path to these destinations. The processing and traffic levels are less than if each recipient's message were transmitted individually.

Multicast traffic can require security protection, whose nature and strength vary based on the multicast group's purpose, characteristics, and membership. Numerous secure multicast protocols have been proposed. Some are applicable to any multicast group but have restricted computational feasibility and scalability; others are optimized for the characteristics of a particular group. Some have been tested under wide-scale deployment; others are still experimental or theoretical. A single secure multicast protocol that is computationally feasible and scalable for all groups, all senders, and all receivers remains a research goal. Additional work is needed to determine the requirements, applicability, scalability, and security characteristics of the various approaches.

2.4 Secure Process Control Systems

Definition

Industrial process control systems (PCSs) perform monitoring and control functions in such diverse critical infrastructures as electrical power generation, transmission, and distribution; oil and gas transport; and water pumping, purification, and supply. Some of these systems, such as Supervisory Control and Data Acquisition (SCADA) systems, typically span large geographic areas and rely on a variety of communication systems, compounding the difficulty of making them secure.

Importance

Although some technologies to secure SCADA systems and other PCSs exist, many organizations are not using these technologies to effectively secure their operational systems. Until recently, there was little perceived threat to these systems, in part because their proprietary nature was viewed as making them difficult to attack. The evolution of critical infrastructure system architectures from isolated stand-alone proprietary systems to distributed networked systems – coupled with the deregulation and market competition that have opened access to third parties and increased integration of PCS infrastructure with business networks – has led to increased security exposure of PCSs.

In this environment, new security approaches are needed that take into account the unique operating requirements of some PCSs such as sensitivity to communication latency, low bandwidth, small sizes of end devices, real-time information flows, divergent message formats based on monitored events, dynamic message routing, reliability, fault tolerance, and survivability. Needed security capabilities include methods, technologies, and tools for deriving security requirements and metrics, performing security analysis, designing security controls, and testing and evaluating the effectiveness of implemented controls. Because industry will bear most of the cost of improved security for PCSs, security solutions need to be economically viable.

State of the Art

Security solutions for PCSs historically have been minimal and ad hoc. Even now, some systems have limited security architectures with little or no adherence to computer security principles such as least privilege or separation of duties. In many environments, security has been added piecemeal rather than designed in from the start, and no widely accepted metrics exist for measuring the security levels of these systems. Although PCSs are often mission-critical, their security has often not kept pace with that of e-commerce systems, in which commercial necessity has driven rapidly improving methods and tools for processing transactions securely.

Capability Gaps

Today, R&D in security for PCSs is fragmented, with researchers scattered across academia, research labs, and industry. A developmental effort will be necessary to increase attention to this topic and integrate research skills spanning a number of technical R&D areas to address the specialized security requirements of PCSs. Capabilities are needed in:

Novel security properties: Research is needed to develop understanding of the security properties and classes of vulnerabilities that may be unique to PCSs and their implications for developing appropriate security policies and enforcement mechanisms.

Security metrics: A prerequisite for improving the security of PCSs is a comprehensive, accepted set of methods for measuring and comparing their security and safety properties. Appropriate metrics for these systems should be developed cooperatively with industry.

Testing and assurance: The benefits from security solutions can be realized only if they are implemented correctly and the resulting system is tested as a whole. System and software testing is expensive, typically consuming half of system development budgets. Improved, cost-effective methods are needed. By taking advantage of the specialized characteristics of PCSs, it should be possible to develop methods and tools that are more cost-effective than generalized software testing approaches.

National testbed and testing program: Development of test methods based on test and evaluation criteria can form the basis for security evaluations of PCSs by commercial laboratories. Mission-critical PCSs will require more thorough testing such as the FAA's regimen for certifying aviation software. A widely endorsed infrastructure for validating security evaluations and issuing security assurance certificates for PCSs would be beneficial.

Developing these capabilities will require expertise in diverse areas such as process control, computer hardware logic, network topology analysis, security vulnerability assessment, security metrics, and security testing and evaluation methods. Academic institutions may have researchers in a few of these areas but may not be able to put together teams with the combination of skills needed over sustained multi-year periods. Federal leadership may be needed to foster this R&D, given that high-assurance PCSs have not diffused into broad use in the commercial marketplace.

3. Domain-Specific Security

The R&D topics in this category focus on specialized security needs associated with particular classes of technologies within specific IT domains. Topics in this category are:

- Wireless security (traditional wireless Internet as well as mobile ad hoc networks)
- Secure radio frequency identification (RFID)
- Security of converged networks and heterogeneous traffic (data, voice, video, etc.)
- Next-generation priority services

3.1 Wireless Security

Definition

This area involves measures to provide cyber security and information assurance to users, devices, and networks that use radio frequency (RF) or infrared (IR) physical layers for communication. These measures include wireless network protection, intrusion detection, and analysis of and response to threats. Wireless security technologies typically operate at Layer 1 (physical) and/or Layer 2 (data link) of the OSI model (see box on page 36) but are tightly integrated with higher-layer security mechanisms to contribute to a holistic security solution.

Importance

Security architectures for wired networks rely at least to some degree on physical security to deny would-be intruders access to local networks and data. Within the walls of many organizations, wired network traffic is unencrypted and nodes may not be individually firewalled because the physical security provided by door locks, cable shielding, guards, fences, and the like is viewed as sufficient. However, RF and IR signals pass through and across many of the physical boundaries of wired networks, rendering physical security ineffective.

Wireless networks enable network topologies not even considered in the wired networking world. For example, mobile ad hoc networks have no network boundary or gateway in the traditional sense. Instead, each node can access and be accessed by many or all other network nodes, and also from outside networks. Thus both traditional wireless networks and mobile ad hoc networks have characteristics that render traditional perimeter-based security architectures such as corporate firewalls and wired intrusion detection sensors ineffective. The lack of physical security and network boundaries in wireless networks pose challenges to computer and network security.

State of the Art

Current wireless security technologies focus mainly on data confidentiality and frequently do not provide robust availability, integrity, authentication, non-repudiation, and access control. These technologies are more suitable for benign environments in which jamming and interference are not problems, some physical security is present, and exposure of network management and topology information does not pose a high risk. Data security in wireless networks is usually provided by Layer 3 (network) and above techniques, such as virtual private networks and secure tunnels (e.g., secure shell and secure socket layer). While these techniques provide strong encryption for higher-layer data streams, they do not address vulnerabilities at the physical and data-link layers. This allows wireless network attacks such as wireless-specific intercept, DoS, man-in-the-middle, jamming, and spoofing. Wireless networks demand cross-layer situational awareness.

Capability Gaps

Commercial interests have fostered improved data security in wireless networks but other wireless security capabilities are immature.

Additional protection mechanisms at the physical and data-link layers are needed, including adaptive antennas and coding techniques that are jam-resistant and can respond to threats in real time. Protected Layer 2 management protocols are needed to eliminate spoofing and DoS attacks. Wireless-specific intrusion detection capabilities using RF sensors are needed to supply network monitoring systems with data unique to the wireless network. In addition, wireless protection, detection, and response technologies should be integrated with higher-layer mechanisms across both wired and wireless network

domains. Network situational awareness tools need to include these features to provide a comprehensive understanding of the entire network, including existing and anticipated threats.

3.2 Secure Radio Frequency Identification

Definition

Radio frequency identification (RFID) tags, also called smart tags, are poised to replace the bar code as a mechanism for rapid object identification. The most common application of bar code technology is the Universal Product Code (UPC), which has been used for several decades on consumer product labels. An RFID tag consists of a microchip and a metallic antenna used to receive signals from RFID readers and emit responses. Active RFID tags include self-contained power sources and, as a result, generally have greater range, processing power, and data storage. Although passive tags, which are powered solely by the energy in signals from readers, are less capable, they are expected to become far more plentiful due to their lower cost.

Importance

Smart tags can be scanned by RFID readers at a rate of several hundred tags per second, with neither line of sight nor close proximity required. While RFID has numerous advantages over bar code technology, it also raises privacy and security concerns. Privacy issues arise from the ability of RFID to track individuals and inanimate objects, scan personal belongings from a distance, or aggregate and correlate data from multiple tag-reader locations. Types of attacks include eavesdropping and unauthorized scanning, traffic tracking and analysis through predictable tag responses, spoofing, DoS disruption of supply chains, and corporate espionage due to lack of reader access control.

RFID technologies are expected to be increasingly used in applications such as inventory control, logistics and real-time location systems, manufacturing, baggage handling, retailing, supply chain management, and transportation. Potential government applications range from RFID-enabled passports to supply chains for military logistics and commerce.

State of the Art

The small size, power constraints, computational processing constraints, and limited chip count associated with RFID technologies preclude complicated data processing and use of sophisticated cryptographic algorithms. Per-unit cost is a barrier to widespread use, but trends of increasing processing capabilities and decreasing costs are expected over time.

Two types of risks are associated with the security of RFID tags. The first is the possibility of DoS attacks against RFID tag readers that would render them incapable of tracking assets and inventory or reading product prices in point-of-sale applications. Criminals might use such an attack to make readers inoperable in order to hide criminal activity. The second and more serious type of risk involves the basic security functions associated with RFID tags and readers, such as encryption of information and authentication of RFID communication signals. Inadequate RFID security could result in unauthorized eavesdropping on communication signals, unauthorized tracking of assets, or spoofing of readers by intentionally misleading tags. This could lead to unauthorized access to sensitive information about individuals or supply chains, price tampering, counterfeiting, theft, and other illegal activity.

Capability Gaps

Although technology developers are beginning to address RFID security requirements, additional work is necessary. Research is needed on lightweight cryptography in the context of power and processing resource constraints under which RFID tags operate. Given the small gate count and limited memory and computational capabilities in RFID tags, the cryptographic techniques available to RFID designers and developers are limited. Cryptographic standards and reference implementations of RFID cryptographic algorithms are needed to enable the development of interoperable technologies and a competitive marketplace. Beyond simply encrypting transmitted information, needed capabilities extend to authentication of tags and readers to avoid

unauthorized scanning of tags, tracking of individuals or assets, or spoofing.

Research is needed on end-to-end security for the complete RFID life cycle since the privacy and security issues raised by RFID technologies are present from the manufacturing stage until the tag is destroyed. These issues are associated not only with the tags themselves but also with the readers and database management systems that store and process RFID information. Industry and government need to work together to develop technologies and policies to securely use RFID while maintaining confidentiality of sensitive data, protecting citizens' privacy, addressing the needs of law enforcement, and complying with government rules and regulations. Work is also needed to quantify the benefits of RFID technologies in various application domains such as cargo tracking and passport control.

3.3 Security of Converged Networks and Heterogeneous Traffic

Definition

The telecommunications sector is undergoing a transition from traditional voice and voice-band (fax and modem) data communication over a public switched telephone network (PSTN) to a next-generation network (NGN) reflecting the convergence of traditional telecommunications with IP-based communications. As existing and new services become available on the NGN, it will be necessary to provide at a minimum the same level of security as the current PSTN. As the NGN evolves toward a packet-based network, it will be necessary to provide security for multiple broadband, QoS-enabled transport technologies across different service providers, independent of any specific access or transport technology.

Importance

Next-generation network services require security policies that ensure consistent application of security measures across a range of network types and access technologies and across service provider networks. The foundation for research on NGN security should include comprehensive NGN models that provide a structured framework for identifying needed security services, including gaps in current security standards, determining security services that need to be deployed, and assessing the risks and benefits of deploying specific security technologies by systematically evaluating security deployments.

State of the Art

Although the evolution of the PSTN and IP-based communication networks has already begun, the converged networks that are expected to form the next generation of telecommunication networks do not currently exist as they are envisioned. Capabilities available in today's PSTN that will also be required as part of the NGN security architecture include: access control, authorization, non-repudiation, confidentiality, communications security, data integrity, availability, and privacy. Existing approaches to providing these capabilities in IP-based networks include encryption and virtual private networks. Redundant communication paths are likely to continue to be used in the future, but are not sufficient to meet all of the NGN security requirements.

Capability Gaps

Research is required in these NGN security issues:

- Large-scale identity management technologies for use in addressing, rather than for location as in the traditional PSTN
- Highly scalable authentication architectures and techniques that make use of multiple authentication factors (e.g., name or ID, password, subscriber identity module [SIM] card containing user authentication information, smart card, physical or software token)
- Techniques to enable non-repudiation on a user-to-user basis, unlike existing capabilities that are focused at the network rather than the user level
- Technologies that enable data integrity, confidentiality, and availability across control and media planes of the network and across all security layers
- Technologies that enable the above security requirements to be met while at the same time assuring some degree of protection of privacy-sensitive information

NGNs will rely on existing technologies and new approaches for providing services and applications, network management, signaling and control, and transport. In addition to the initial development of new technologies, these new capabilities will need to be transitioned to appropriate standards development communities in order to assure evolution toward globally scalable and interoperable NGN network architectures.

In addition to these generic requirements, security requirements associated with important application domains must be addressed in NGNs. Such requirements include those associated with next-generation priority services in support of national security/emergency preparedness (NS/EP) telecommunications (section 3.4, below) and the real-time requirements associated with security of process control systems (section 2.4, page 53).

3.4 Next-Generation Priority Services

Definition

Telecommunication priority services provide priority access to telecommunications capabilities in support of national security and emergency preparedness (NS/EP). Priority services in traditional wireline and wireless telecommunications enable users of these services to communicate in crises when the public switched telephone network (PSTN) may experience high network congestion or diminished network capacity. Existing priority service architectures cover call origination, PSTN network access, transport, network egress, and call termination. As the telecommunications industry increases the use of IP-based telephony, and as the use of other IP-based communications (e.g., e-mail, IP-based video conferencing, and other Internet-based information exchange) become increasingly prominent in the NS/EP community, the need arises to support priority service in the IP domain.

Importance

Stakeholders ranging from local first responders to Federal decision makers rely on NS/EP telecommunications to communicate during crises such as emergencies, attacks, and natural or manmade disasters, as well as during subsequent recovery and reconstitution efforts. The purpose of the NS/EP telecommunications infrastructure is to:

- Respond to the NS/EP needs of the President and the Federal departments, agencies, and other entities, including telecommunications to support national security leadership and continuity of government
- Satisfy priority telecommunications requirements under all circumstances through use of commercial, government, and privately owned telecommunications resources
- Incorporate the necessary combination of hardness, redundancy, mobility, connectivity, interoperability, restorability, and security to obtain, to the maximum extent possible, the survivability of NS/EP telecommunications

State of the Art

The Government Emergency Telecommunications Service (GETS), established by the National Communication System (NCS) in 1995, provides a nationwide ubiquitous voice and voice-band data service that interoperates with and uses the resources of selected government and private facilities, systems, and networks through the application of standards, and provides access to and egress from international service. GETS provides priority access and specialized processing in local and long-distance networks. It is maintained in a constant state of readiness to make maximum use of all available communications resources should outages or congestion occur during a crisis. GETS is survivable under a broad range of circumstances, ranging from local to widespread damage, and provides routing, signaling, and network management enhancements that result in a higher probability of call completion in congested networks. GETS augments and improves the public switched network with capabilities that include enhanced routing schemes and priority use of call-by-call priorities over the PSTN.

To complement GETS wireline services, in 2002 the NCS deployed Wireless Priority Service (WPS), a

subscription-based priority service through commercial wireless service providers that ensures NS/EP communications availability when wireless communications users experience high levels of blocking and congestion. WPS allows authorized personnel to gain access to the next available wireless channel to initiate NS/EP calls.

Capability Gaps

The Internet plays an increasingly important role in communication and exchange of information generally, and consequently in the NS/EP community in particular. Furthermore, the previously independent infrastructures for traditional circuit switched telecommunications and IP-based communications are in the process of evolving into converged next-generation networks. Because the technological infrastructures for providing existing priority services do not extend into the IP domain, this area of potentially essential telecommunications capability and capacity lacks the ability to prioritize critical communications traffic to support NS/EP missions in times of crisis. Instead, communications over IP-based networks are based on best-effort delivery, an approach that may be problematic in conditions of high congestion or reduced capacity.

To achieve assured delivery of NS/EP voice, video, and data traffic over the Internet, research is needed to determine what types of routing overlay and meta-application models are required to authenticate NS/EP users, prioritize packet traffic, detect network congestion, reallocate and queue resources, and re-route Internet traffic accordingly. Models that should be investigated include out-of-band network flow management and the use of virtual channels to carry management control information.

Numerous approaches using the IP model are possible. However, prioritized delivery of individual packets at lower layers of the OSI model (see box on page 36) does not guarantee that transactions will receive priority processing on end systems and servers. Since any single protocol is likely to be insufficient to guarantee priority, several approaches may need to be combined to form an operational system. Different types of IP-based data (e.g., voice-over-IP, streaming video, and e-mail) may be treated differently due to varying degrees of sensitivity to network characteristics such as latency, jitter, packet loss, throughput, and availability.

Next-generation priority services should be resilient in the face of large-scale outages of the Internet infrastructure and Internet support infrastructures such as electric power and telecommunications. They should also be resilient to cyber attacks originating within the Internet such as DoS and worms. The services should have ubiquitous coverage so that they apply to various physical and link layer technologies, locations, applications, and network topologies. Furthermore, they must work within single-provider networks as well as in cross-provider environments. In addition, next-generation priority services will need to satisfy the more generic NS/EP functional requirements discussed above in the context of existing priority services (e.g., availability, reliability, survivability, scalability, affordability).

4. Cyber Security and Information Assurance Characterization and Assessment

The R&D topics in this category address approaches, methods, technologies, and tools for evaluating, testing, and measuring security and risk in IT infrastructure components and systems, and in the infrastructure as a whole. Topics in this category are:

- Software quality assessment and fault characterization
- Detection of vulnerabilities and malicious code
- Standards
- Metrics
- Software testing and assessment tools
- Risk-based decision making
- Critical infrastructure dependencies and interdependencies

4.1 Software Quality Assessment and Fault Characterization

Definition

This area includes methods, technologies, and tools to assess overall software quality and to characterize defects according to their impact on the most significant quality attributes of software, such as security, safety, and reliability. This area also includes efforts to enable external assessment of these characteristics.

Importance

Overall software quality must be assured as a foundation for other security efforts. A key consequence of today's lack of validated methods and metrics to evaluate software quality is that economic mechanisms (such as risk analysis, insurance, and informed purchasing decisions) that help advance other disciplines either do not exist or have little effect on software development. Software quality assessment would help the R&D community understand what defects lead most directly to security problems and focus R&D on those problems.

State of the Art

The gap between the state of the art and the state of the practice in developing near-defect-free, secure software is large. Methods such as the Software Engineering Institute's Team Software Process can produce 20-fold to 100-fold reductions in the number of defects, yet these methods are not widely used. Research in tools to assure that code has not been modified without authorization is promising. There are tools that can assure that software is free from common security defects such as the vast majority of buffer overflows. However, these tools also are not always used. The cyber security and information assurance community currently does not have a good understanding of what would motivate purchasers to insist upon such measures and developers to use them.

Capability Gaps

More robust assessment processes and automated quality assurance: Tools are needed that enable software developers to assess the quality and security of the software they are designing throughout the development process. Automated techniques for analyzing software would reduce the time and effort required to assess software quality, thereby enabling developers to evaluate their designs more frequently. These methods should include mechanisms for assuring that code has not been tampered with by a third party (for example, proof-carrying code). COTS evaluation methods and resources should be augmented to take advantage of these capabilities, and software purchasers should have access to assessment methods and results. R&D is needed to improve the effectiveness, accessibility, and adoptability of such mechanisms.

Connect defects to attributes of highest concern: Research is needed to help categorize defects and to understand their relative severity. In-depth analysis should be performed to identify the most common vulnerabilities in various contexts (e.g., by operating

system, application type, and programming language). The results should be detailed enough that processes and tools can be designed to specifically counter the most common vulnerabilities, and that education efforts can be used to target the most common problems. Some rudimentary high-level statistics are available (e.g., statistics on buffer overflows or race conditions), but these statistics are not detailed enough to shed light on which processes or tools can effectively be used to counter a given vulnerability.

4.2 Detection of Vulnerabilities and Malicious Code

Definition

This research area focuses on methods, technologies, and tools to detect vulnerabilities and malicious code and to provide assurances about software security and other quality attributes. These capabilities are targeted at source code during development or re-factoring as well as in-use object or binary code.

Importance

Because software vulnerabilities are increasingly being exploited not by "recreational" hackers but by criminals or other adversaries with more malicious intent, there is a need for improved capabilities for detecting, mitigating, or eliminating vulnerabilities before they are exploited. In addition, IT developers need to be able to make assurances about the characteristics of their software. Today, developers do not have adequate feedback to make security-critical decisions and IT consumers cannot objectively evaluate the security of the software products they are purchasing.

State of the Art

Recent developments in source code analysis address some of the scalability, usability, and sustainability issues that date from the early days of software programming. However, information is lacking about what analyses are most effective, how best to use them, their applicability to legacy code, and how to overcome obstacles to adoption and sustained use. Many other promising targets for code analysis have yet to be explored.

Capability Gaps

Research is needed to establish new methods, technologies, and tools for vulnerability and malicious code detection. Approaches should be evaluated for effectiveness and the most promising ones identified for further R&D. Specific thrusts should include:

Improve source, object, and binary code scanning tools: Software developers benefit from working in an environment in which they have immediate feedback about the characteristics of the software they are developing. Creating such an environment will require improved source, object, and binary code scanning tools (static analysis tools) and automated execution testing tools (dynamic analysis tools) that search for security vulnerabilities. Many such tools and services exist. However, their capabilities need to be expanded to a broader range of problems, and obstacles to their adoption such as high false-positive rates, inadequate scalability, and inapplicability to legacy systems should be overcome. In many cases, these tools and services need to "understand" the design intent behind the code. Promising research that automatically extracts such intent with minimal guidance from the developer should be encouraged. Tools should take advantage of source code where available, but improved tools are also needed for detecting vulnerabilities in object and binary code when the original source code is not available.

Develop malicious code detectors: Because it is easy to make an intentional injection of malicious code look like a simple mistake, such code is difficult to detect reliably. Promising developments in simulation, code scanning, function extraction, covert channel detection, and backtracking would benefit from further R&D.

Improve the interoperability of analysis tools: Many individual tools are available to aid in code analysis (such as decompilers, debuggers, slicers), but they often do not work well together. Research is needed to ascertain how analysts use and want to use tools in combination (including identifying information flows and common processes), and to determine how to make tools more interoperable, including defining standard interchange formats to support such flows.

4.3 Standards

Definition

The goal of cyber security standards is to improve the security of IT systems, networks, and critical infrastructures by increasing the security, integrity, and reliability of commercial products. A cyber security standard defines both functional and assurance requirements within a product or technology area. Well-developed cyber security standards enable consistency among product developers and serve as a reliable metric for purchasing security products. Cyber security standards cover a broad range of granularity, from the mathematical definition of a cryptographic algorithm to the specification of security features in a Web browser, and are typically implementation-independent.

The best results emerge from standards development processes that are consensus-based, in which all stakeholders – users, producers, and researchers – participate. In such approaches, each group of stakeholders contributes unique perspectives to the process: users understand their needs, producers understand the current state of technology, and researchers understand future trends. A standard must address user needs but must also be practical, since cost and technological limitations must be considered in building products to meet the standard. Additionally, a standard's requirements must be verifiable; otherwise, users cannot assess security even when products are tested against the standard.

Importance

The security of IT systems begins with the security of their hardware and software components, and includes both security technologies and non-security hardware and software. Federal agencies, industry, and the public rely on commercial security products to protect information, communications, and IT systems. Adequate product testing against well-established cyber security standards facilitates technical improvements and helps give users confidence that the products meet their security needs. Both Federal agencies and the public benefit from the use of tested and validated products. In the absence of adequate testing, product weaknesses can render systems and critical infrastructures vulnerable to common attacks and other malicious activities.

State of the Art

As cyber security issues grow in strategic importance to the Nation, the need for methods and tools to assess and verify the security properties of computer-based systems and components is becoming more apparent to both developers and consumers. Security standards such as the Common Criteria Protection Profiles are being developed to address government national security systems. Security standards are also being developed and applied to systems throughout the Federal government.

At the same time, the requirement for security standards is new enough that many current IT products do not provide the level of security needed to protect sensitive information, electronic commerce, and critical infrastructures. In general, widely accepted cyber security standards that meet the needs for IT security in unclassified and/or sensitive civil government and commercial environments have not yet been developed. Moreover, the cyber security standards used in the national security arena are seldom built by consensus and often specify requirements beyond existing commercial technological capabilities.

Today, evaluation processes to determine compliance with security standards also are limited. Standard development processes frequently do not consider how this testing and validation will be done. Even if a standard has been well thought out from the standpoint of verifiability, specific test methods are seldom developed upon completion of the standard.

Capability Gaps

Developing cyber security standards is time-consuming, and developing associated test methods can be even more demanding without proper planning. Both require efforts from well organized stakeholders. Today's standards and test methods community is small. Much of the expertise lies in the commercial sector, with little in the user and research communities. Often, the commercial sector does not initially perceive the return on investment from

developing or using cyber security standards. Strategies are needed to encourage earlier buy-in to the process by all stakeholders.

Compliance testing helps assure that a product meets a cyber security standard and also helps isolate and correct security problems before the product enters the marketplace. Expeditious, cost-effective, and detailed technical test methods need to be developed for each component of an IT system. R&D that couples standards development with the creation of associated test methods in ways that reduce the time and cost of product validations can have the most immediate benefit to stakeholders. Testing methods, which detail the tests and documentation necessary to determine compliance with each requirement of a security standard, must be science-based and able to generate consistent, measurable, repeatable, timely, and cost-effective product validation results. Some automated techniques for generating tests exist today, but automating test generation for generalized cyber security standards is beyond current capabilities. The application of formal methods to these problems warrants additional investigation.

Looking farther into the future, strategies for identifying emerging technologies that will require standardization need to be developed. Given the pace of technological change, research in management of technology life cycles, including the effects of new technologies, is essential.

4.4 Metrics

Definition

Metrics can be defined as tools designed to facilitate decision making and improve performance and accountability, such as through the collection, analysis, and reporting of performance data. Operators can use such quantifiable, observable, and measurable data to apply corrective actions and improve performance. Regulatory, financial, and organizational factors drive the requirement to measure IT security performance. A number of laws, rules, and regulations require IT performance measurement in general and IT security assessment in particular. These laws include the Information Technology Management Reform Act (also known as the Clinger-Cohen Act), the Government Performance and Results Act, the Government Paperwork Elimination Act, the Federal Information Security Management Act, and the Healthcare Insurance Portability and Accountability Act. Other drivers are the national and homeland security implications of IT infrastructure vulnerabilities.

Potential security metrics cover a broad range of measurable features, from security audit logs of individual systems to the number of systems within an organization that were tested over the course of a year. Security metrics measure diversified multi-dimensional data collected in real time and analyzed. Effective security metrics should be used to identify security weaknesses, determine trends to better utilize security resources, and measure the success or failure of implemented security solutions. Ultimately, the metrics should help characterize an organization's overall security posture from risk/threat/vulnerability, budgetary, and regulatory standpoints.

Importance

Although numerous products and best practices have been developed to provide security solutions, determining and measuring their effectiveness is difficult in the absence of validated metrics. Organizations can improve security accountability by deploying IT security metrics. The process of data collection and reporting enables security managers to pinpoint specific technical, operational, or management controls that are not being implemented or are implemented incorrectly.

Ideally, metrics should be available that can measure different aspects of an organization's IT security policies and mechanisms. For example, the results of risk assessments, penetration testing, and security testing and evaluation can be quantified and used as data sources for metrics. Security managers and system owners can use the results of the metrics-based analysis to isolate problems, justify budget requests, and target investments to areas in need of improvement, thereby obtaining the most value from available resources. Security metrics assist with determining the effectiveness of implemented security products, processes, procedures, and controls by

relating results of security issues (e.g., cyber security incident data, revenue lost to cyber attacks) to organizational requirements and security investments. Departments and agencies can demonstrate compliance with applicable laws, rules, and regulations by implementing and maintaining security metrics programs.

State of the Art

Today, developing comprehensive security metrics for an organization's networks, systems, and information is hampered by two key issues: 1) the sheer volume of potential sources of security information and the fact that much of the information must be gathered, collated, and analyzed by hand; and 2) the reality that researchers do not yet adequately understand how to quantify, measure, and evaluate cyber security to inform decision making. Possible information sources include incident handling reports, test results, network management logs and records, audit logs, network and system billing records, configuration management and contingency planning information, and training, certification, and accreditation records. However, without metrics and tools for automating the collection of significant data and streamlining their analysis, evaluating security effectiveness is speculative at best, with hackers and attackers possibly having better awareness of an organization's security standing and weaknesses than the organization itself.

Capability Gaps

In the absence of sound methods and valid, persuasive evidence, the private sector's ability to make well-informed, risk-based IT security investments is limited and overall levels of cyber security in the IT infrastructure remain low.

It is difficult for organizations to justify allocating resources for a security metrics program, particularly in the context of constrained budgets. Rather than expending the time and resources to gather and analyze security data, organizations too often limit their cyber security activities to simply purchasing commercially available security products. Many organizations have not taken even the first step in building a security metrics system, which is to establish a baseline or framework of the key types of data that will go into measuring security effectiveness. Improved identification of key types of metrics information, more intelligent tools, and automation of metrics data collection and analysis are needed. But the security metrics field is relatively new, with a limited number of experts.

4.5 Software Testing and Assessment Tools

Definition

A test is an execution of a software program or system to determine one or more of its characteristics. Software assessment makes that determination through a static examination of the software. Static examination may focus either directly on the software or indirectly on such related aspects as specifications or development records.

Software testing and assessment tools assist in, and often automate, the exacting testing and assessment tasks. Software testing and assessment usually presuppose a specified plan. Testing, assessment, and interoperability provide evidence that a software implementation satisfies requirements such as functionality, compliance, security, reliability, usability, efficiency, and portability. Testing and assessment can – and should – occur at every phase of the software development process, including requirements analysis, design, coding, and acceptance.

Importance

Security vulnerabilities can surface almost anywhere in software that is ubiquitous in the Nation's IT infrastructure. According to a May 2002 NIST report entitled *Economic Impacts of Inadequate Infrastructure for Software Testing,* the cost to the Nation of inadequate software quality testing is estimated at $59.5 billion annually. In addition, the report states that increased software complexity and decreased average market life expectancy heighten concerns about software quality.

According to the NIST report, only 3.5 percent of errors are found during software requirements and design phases. This statistic suggests why a release-and-patch approach to software is increasingly

untenable. The report concludes: "The path to higher software quality is significantly improved requirements, specifications, and software testing early in the life cycle." It argues that reference implementations, metrics, and suites of standardized testing tools could help address software inadequacies. A 2003 NSF-funded Computing Research Association report, *Grand Research Challenges in Information Systems,* argues that to "conquer system complexity," R&D advances are needed that make complex systems easier to design.

State of the Art

The following types of tools can be used in software testing and assessment:

- Design tools assist in functional design, internal design, and code design. They also analyze requirements and designs for ambiguities, inconsistencies, security vulnerabilities, and omissions.
- Development environments assist in writing, compiling, and debugging code.
- Inspection tools assist in requirement reviews and code walk-throughs.
- Test design and development tools are used to develop test plans and abstract test cases, manage test data, and generate automated tests from specifications.
- Execution and evaluation tools develop concrete test cases and test harnesses, set up testing environments, perform selected tests, record test executions, log and analyze failures, and measure testing effectiveness and coverage.
- Artifact examination tools scan code and executables for bugs or vulnerabilities and reverse-engineer control flow.
- Support tools assist in project management and documentation, and control and track configuration.

State-of-the-art testing and assessment tools are not widely used. Commercial software often is inadequately specified and tested, resulting in products with many bugs. The lack of formal specifications can result in the introduction of vulnerabilities during the design or implementation stages. Poorly structured software may allow a fault anywhere in thousands or millions of lines of code that results in security vulnerabilities. Faults may in some cases be deliberately added "back doors" rather than inadvertent mistakes. Existing software test and assessment tools could help developers avoid or catch many systematic errors. The most advanced tools generate automated tests from rigorous specifications, help analysts understand code, and monitor execution. These tools coupled with the best of today's software development methods can improve software quality.

However, even these tools are often difficult to use, limited in scope and effectiveness, unable to work with other tools, and may lack clear demonstrations of effectiveness. While exploratory work suggests that more powerful tools are possible and that existing capabilities can be packaged in easier-to-use formats, developing these improvements is expensive. Advanced tools require large software subsystems as test infrastructures in which to analyze code, make inferences, and track and correlate information. Because of competitive pressures, there are barriers to collaboration among stakeholders who could benefit from cooperation on test methods, testing suites, and development environments. The situation is analogous to a hypothetical state of air travel in which each airline has to design, build, and test its own airplanes, airports, reservation networks, and air traffic control systems.

Capability Gaps

As the NIST report states, improving software quality will require better testing throughout all phases of development. Improved testing throughout software development will in turn require tools that can provide more comprehensive analysis, increased automation, and ease of use to produce more thorough testing at a lower cost. Because bugs are significantly more expensive to fix when they are discovered in later phases of the development process, developing higher-quality, lower-cost software will require more testing early in the development process.

The primary capability gap is at the very beginning of the software cycle, in the requirements, specifications, and top-level design phases. Current tools do not provide the precision and functionality to capture specifications in sophisticated modeling languages. Such languages could be used to generate measures of system complexity and completeness, identify design inconsistencies and ambiguities, minimize the likelihood of security vulnerabilities in an artifact that has been built to specification, and generate code and automated tests. The development of computational tools to support the requirements analysis and design phases of system development would be an improvement over what has traditionally been a manual process.

Gaps also exist during the software implementation phase. Avoiding errors and vulnerabilities in the design phase, or avoiding them in the development phases, does not eliminate the need for later testing because errors may be introduced as design decisions are made and details are added. Testing and assessment at the unit and subsystem level are more effective than waiting until the final system is built.

The lack of software assessment and testing tools for those who work with the final product is another capability gap. A comprehensive analyst's workbench is needed for post-development security analyses such as by the administrator checking the suitability of a COTS package for a given purpose or the contracting officer determining whether to accept a delivered software product. Some functions of the workbench would be to find control flow, remove obfuscations, structurally edit blocks of code, and maintain information about design intent. Such as system can also incorporate and enhance existing tools that scan for known vulnerabilities, test for anomalous behavior, and present functional slices of binary or source code.

Calibrated, validated tests and assessments are needed across all phases of software development. In the absence of advances, developers are unlikely to use a tool if it is not clear how much assurance is derived from the results. They are unlikely to choose one tool over another if increased security cannot be predicted and is not objectively evident afterward. Finding a security flaw is significant. Today, little can be said with certainty about the code if it is tested and no flaws are found. Although testing will never guarantee the absence of bugs, better testing capabilities will provide higher confidence in the security of systems than exists today.

4.6 Risk-Based Decision Making

Definition

Risk-based decision making assists managers in making more informed decisions through qualitative and quantitative mechanisms that account for desirable and undesirable outcomes. The development of investment strategies and resource allocation models – funding cyber security or other efforts – relies on information from risk management processes that facilitate identification and evaluation of threats, vulnerabilities, and impacts (economic and otherwise) of attacks relative to costs.

Importance

Today, cyber security is often a secondary issue for managers whose primary considerations are shareholder or stakeholder value, return on investment, and earnings. Research suggests that the impact of a cyber attack can range from the inconsequential, to a brief interruption of regular operations, to an incapacitating blow to the ability to conduct business. However, what is more important about the threat of cyber attack from a risk-management perspective is that the past is not necessarily a good predictor of future events. The threat, vulnerability, and risk space is dynamic. Attackers are constantly developing new approaches, and the increasing interdependence of critical infrastructures and the IT infrastructure heightens the risk of economic consequences from successful future attacks. Because cyber security will become a primary consideration only when it is factored into management's ability to earn profits, R&D is needed to develop sophisticated risk-based models for evaluating total return on investment.

State of the Art

In the business community, risk-based decision making methods have traditionally focused on risks of business interruption, project failure, natural hazard, and financial impact. These risks are so well understood that commercial insurance is available to cover them. However, analyses of business interruption and reconstitution rarely consider cyber attacks, and those that do so generally do not consider low-probability, high-impact events. This may be because there are not yet any universally accepted tools for measuring the costs and benefits of expending resources to reduce cyber security risk.

Capability Gaps

This challenge provides an opportunity for researchers to develop accurate models for risk-based decision making in cyber security. An assessment of business risk from possible cyber attacks must identify threats, vulnerabilities, and consequences. The risk or probable level of loss can be calculated as a function of threats, vulnerabilities, and consequences. Each of these three risk factors can be reduced by various counter measures. The risk factors cannot readily be estimated based on the frequency of occurrences in the past, because the most prevalent types of attacks in the past may not be the most common ones in the future, and because some types of potentially devastating attacks have not yet occurred but might in the future.

While it is, in principle, possible to estimate how much a given countermeasure will reduce the corresponding risk, research in this area has been minimal. If the total reduction in risk due to a given countermeasure can be quantified, then the countermeasure can be given an expected value, from which a justifiable price can be derived based on risk. The cost of each countermeasure can be compared with the reduction in probable loss over a relevant time period. Once an adequate assessment of overall risk is made, the development, implementation, and deployment of cyber security measures can be carried out with some confidence that the best investments are being chosen, given the available information.

R&D in a number of areas is desirable to establish the knowledge base for risk-based decision making in cyber security. Analyses of organizations' risk exposure factors (e.g., industry, location, size, and security countermeasures) and analyses of potential financial losses to organizations from attacks of varying intensity, precision, and recurrence would provide useful baseline information and tools for institutional planning that do not exist today. Such information is often viewed as proprietary or sensitive, and organizations are reluctant to share it. Empirical studies of the deployment of risk-based decision making methods in improving cyber security would also be useful in evaluating the applicability of traditional models and developing new models tailored to organizational requirements. Studies of the insurance industry's posture toward cyber security, including the current accessibility, quality, and scope of policy underwriting related to cyber security, should also be undertaken.

4.7 Critical Infrastructure Dependencies and Interdependencies

Definition

Today, the infrastructures of U.S. sectors deemed to be critical to the national interest are increasingly dependent on the IT infrastructure. Research in this topic aims to develop a thorough scientific understanding of the ways in which critical infrastructure (CI) sectors depend on the IT infrastructure, and the extent to which the CI sectors are interconnected through components of the IT infrastructure and are interdependent. This will enable analyses of the potential impacts of cyber attacks on the operations of the CI sectors, including cascading consequences resulting from CI interdependencies, and assessments of the effectiveness of possible protective and mitigation measures for CI sectors.

Importance

When consequences of a cyber attack on CI systems (including process control systems) are significant, consequence analyses of interdependencies of systems within an infrastructure and potential cascading effects across infrastructures can enable decision makers to

understand possible outcomes of their decisions and assess trade-offs between alternative actions.

Such understanding is of particular concern to critical infrastructure sectors that rely heavily on IT systems to operate, such as the banking and finance sector. Linking developing infrastructure interdependency consequence models with IT system operations and security models provides an opportunity to manage the risks from cyber and physical threats in a holistic way within and across critical infrastructures.

State of the Art

The general principles of control theory and control systems are relatively well understood. In addition, the physical and virtual commodity flows that IT systems control are generally well understood at an industry or company level. However, failures and consequences are often situation-dependent, and neither the interdependencies between the infrastructures nor the relationship between failures in control systems and large-scale consequences are well understood. Understanding these potential impacts requires understanding the function of the IT and control systems, the operation of the infrastructure, and interdependencies with other infrastructures.

Network engineering models are useful for understanding system impacts of disruptions at one or more network nodes. Modeling network performance is a fundamental part of infrastructure analysis that is used for graphical representation of the infrastructure, for estimating system performance under adverse operating conditions, and for verification. Highly aggregated systems-level models are useful for assessing potential dynamic responses and propagating effects across infrastructures, but additional model development is needed.

Capability Gaps

Coupling critical infrastructure consequence analysis to cyber security faces two major challenges: 1) the lack of information about the behavior of control systems in abnormal and malevolent environments; and 2) the lack of modeling, simulation, and decision-support tools that capture the coupling of control and IT system operations to infrastructure operations.

Modeling and simulation: A key challenge in developing better understanding of the dependence of infrastructure operations on IT and control systems is that little real-world data about control system behavior in abnormal and malevolent environments is available. Abnormal events are sufficiently uncommon that data records documenting such events are minimal. Malicious events are uncommon as well. CI modeling and simulation that combine the control systems and the controlled processes can increase understanding of the connections to event consequences. However, commercial CI sector organizations are reluctant to share the knowledge needed to improve model fidelity because they fear that competitors may gain some advantage from that knowledge, or for liability reasons. The alternative is modeling and simulation using available data and validations of the simulations against publicly accessible information about abnormal and malicious events.

Improved modeling and simulation methods will make it easier to predict the behavior of generic networks in various scenarios such as by performing "what if" analyses that are equivalent to virtual experiments. Integration of such models into larger infrastructure models will contribute to understanding the CI sectors' interdependencies. As this capability matures, coupled network engineering infrastructure and IT infrastructure models could be used to predict impending failures and visualize threatened outages based on loss of critical components.

Robust decision making: Robust metrics are needed for optimizing risk-based crisis response when multiple infrastructures are mutually dependent. Additional communication protocols and data aggregation techniques are needed for real-time visualization and forecasting to aid in responding to intrusions or loss of functionality or confidence in highly trusted systems. Modeling and simulation tools, coupled with metrics, visualization, and forecasting tools, can be used in crisis response as well as in analysis and assessment to provide decision makers with a better basis to make prudent, risk-based strategic investments and policy decisions to improve the security of critical infrastructures.

5. Foundations for Cyber Security and Information Assurance

The topics in this category focus on fundamental technological elements that serve as building blocks for developing and engineering a more secure IT infrastructure. Topics in this category are:

- Hardware and firmware security
- Secure operating systems
- Security-centric programming languages
- Security technology and policy management methods and policy specification languages
- Information provenance
- Information integrity
- Cryptography
- Multi-level security
- Secure software engineering
- Fault-tolerant and resilient systems
- Integrated, enterprise-wide security monitoring and management
- Analytical techniques for security across the IT systems engineering life cycle

5.1 Hardware and Firmware Security

Definition

Hardware includes not only computers but also externally connected components – cables, connectors, power supplies, and peripheral devices such as a keyboard, mouse, and printer – that enable the system to execute functions.

Firmware is the low-level software or sequences of instructions that are written onto programmable read-only memory (ROM) in a computer or peripheral device, enabling the device to determine its capabilities, render them functional, and coordinate operations. Some firmware may be part of the operating system (OS) kernel (i.e., the core of a computer OS that provides basic services for all other parts of the OS) and may execute in privileged mode. In some cases, firmware provides an interface to the rest of the OS so that the system can operate the device. In other instances, firmware is executed during the computer's boot process (i.e., when the OS is loaded into the computer's main memory or random access memory) – for example, the Basic Input/Output System (BIOS), which executes before the OS is loaded. Other firmware resides on peripheral devices, allowing the OS to use the devices effectively.

Importance

Hardware or firmware attacks can undermine even the most sophisticated application-level controls or security mechanisms. Malicious firmware that has unrestricted access to system components (e.g., if it is part of the OS kernel) has considerable potential to cause harm, introduce backdoor access (an undocumented way of gaining access to a computer, program, or service), install new software, or modify existing software. If the underlying hardware and firmware cannot be trusted, then the OS and application security mechanisms also cannot be trusted.

State of the Art

Hardware and firmware, including points of interconnection, are subject to attack. One notable example of an attack against firmware is the Chernobyl virus (also referred to as the CIH virus, after the author's initials); first discovered in Taiwan in June 1998, it destroys a system's flash BIOS, resulting in lost data. PC users trying to overclock their processors often distribute reverse-engineered and "improved" motherboard BIOSes. Rootkits and other software attacks can execute code on secondary processors (e.g., graphics processing units) or hide malicious code in flash or electrically erasable, programmable ROMs. Wireless access points can be altered to deliver more power and broadcast range, thereby making eavesdropping easier. Keystrokes can be tracked by small hardware devices such as keyboard dongles that can capture and record every keystroke that is typed. To mitigate such vulnerabilities and reduce risks, several hardware and language-based approaches have been proposed.

Capability Gaps

Trusted computing platforms, and the corresponding OS modifications to leverage them fully, have the

potential to improve some key areas of information security, especially the level of trust in platform hardware. Research is needed to understand weaknesses and covert channels open to hardware and firmware attacks. New approaches and rigorous methods for certifying hardware and firmware are particularly needed for an environment in which the IT infrastructure's hardware and firmware are increasingly developed and manufactured offshore. Areas in which R&D advances are needed include:

Hardware support for security: Efforts are underway to protect hardware and firmware and enable secure, trusted computing platforms. Cryptographic accelerators speed the processing of cryptographic algorithms. Smart cards can be used to protect authentication keys and for multi-factor authentication. Although work is ongoing, more R&D is needed on integrating more secure components into a trusted computing platform.

Authentication-based firmware security: The authentication-based approach (sometimes referred to as "secure bootstrap") seeks to ensure firmware integrity by using digital signatures to authenticate the origin of the device and its transmitted data, chain of custody, and physical protection. This approach ensures that the firmware has not been changed since it was approved. It is a means for preserving an existing relationship of trust but cannot establish trust. Authentication alone cannot ensure that untrusted code is safe to run. The authentication-based approach is currently the preferred strategy because the technology is better developed and its implementation is more straightforward than language-based approaches. Increased emphasis on development and deployment is needed.

Language-based firmware security: Language-based security is an approach to address security vulnerabilities in a variety of domains, including firmware security as well as others (see section 5.3 on page 72). It leverages programming, program analysis, and program rewriting to enforce security policies. The approach promises efficient enforcement of fine-grained access control policies and depends on a trusted computing base of only modest size. Unfortunately, these techniques are not as well established or advanced as authentication-based approaches, and are more difficult to implement. However, language-based security offers some promising advantages worth investigating further. The main advantage of the language-based approach is the ability to establish a basis for trust, regardless of the source.

In the language-based approach, each time a firmware module is loaded, it is verified against a standard security policy. The verification step prevents the compiler from being bypassed, spoofed (i.e., forged to make it appear to have come from somewhere or someone other than the actual source), or counterfeited. Confidence in verified device drivers requires trust only in the verifier, not in the compiler and the code it produces. The security policy is designed to rule out the most obvious forms of attack by combining type safety (all behaviors are fully specified in the programming language semantics) and various architectural constraints.

5.2 Secure Operating Systems

Definition

An operating system (OS) manages and protects a computer's hardware and software resources. Other software that is subsequently loaded depends on the OS's core services such as disk access, memory management, task scheduling, and user interfaces. The portion of OS code that performs these core services is called the kernel. The OS provides a stable, consistent way for applications to request services from the hardware without having to know details about the hardware (e.g., the underlying processor, communications mechanisms, or peripherals). In addition, the OS creates system abstractions (e.g., the abstract data types, privileges, and hierarchical domains used by applications). The OS enforces critical elements of the enterprise security policy, including information confidentiality and integrity and mandatory or discretionary access control mechanisms.

Importance

Combined with hardware and firmware, OSs are the foundation for all computing, forming the core

software for both general purpose and specialized computers, including process control systems, network routers, smart switches, servers, and PCs. An OS is loaded into a computer by a boot program and then manages all the other programs (i.e., applications, services, or application-enabling programs such as middleware) running on the computer. All computers, from small embedded processors to large servers supporting tens of thousands of users, require an OS. Most OSs have been designed and implemented to provide a wide range of features and services, but security has often not been a fundamental requirement. An OS that can be subverted, penetrated, or bypassed cannot provide a sound base for critical applications. Without OS security, even a carefully constructed application is vulnerable to subversion.

State of the Art

Computers that connect to the Internet enter a cyberspace filled with untrusted networks and systems, which allow malicious attacks on unprotected machines. The large and growing number of new attacks that exploit OS vulnerabilities constitute a continuing threat that was not present or even anticipated two decades ago. A secure OS should provide users with the assurance that implemented policies are enforced, enforcement mechanisms cannot be bypassed, and attempts to tamper with the enforcement mechanisms are detected. The current state of the art has not achieved these capabilities.

Capability Gaps

Needed capabilities include tools to facilitate high-assurance OS development, abstraction layers that allow for the portability and upgrading of secure OSs, criteria to assess OS vendors' security claims, including those for distributed and multi-layered trust architectures, and models and tools to support user interfaces and assist administrators in policy configuration and management tasks. OS prototypes that address these issues need to demonstrate protection mechanisms. Ongoing efforts conducted in collaboration with hardware developers should continue to work toward interoperability across platforms to permit the rapid adaptation of hardware functions to specific applications. The following are major areas requiring R&D advances:

Tools and resources: Operating systems today do not adequately serve as secure, high-assurance servers able to separate proprietary, public, classified, and unclassified information and to enforce the applicable separation policies. In addition, modern OSs do not enforce discretionary (i.e., user-controlled) security policies with high assurance. To address these needs, efforts to design secure OSs should build on concepts that have emerged from past and ongoing work aimed at secure and trusted platform development. This mechanism could be supported by hardware, firmware, and software components.

The development of a secure, common, low-level BIOS-like mechanism is one approach that can be used to serve as the base for future trusted OS development. This mechanism may be supported by a combination of hardware, firmware, and software components. The goal is to create well-specified security abstractions and interfaces that would persist across upgrades to the hardware, firmware, and software that an OS supports. Constructs that must be enabled by such a mechanism need to be identified and prototype implementations developed.

Existing specification tools do not provide adequate formal methods for mapping policies to implemented controls and countermeasures and then for evaluating how effectively the policies are implemented. Formal verification tools tailored to secure system development also are needed. These tools must include formal specification and analysis methods.

A weakness of some past generations of secure OSs was an emphasis on maintaining a secure state once the OS had achieved its initial runtime state, while not giving sufficient attention to the start-up processes occurring prior to the runtime state. Future secure OS development efforts should make use of the concept of secure bootstrapping and secure system initialization techniques to ensure that the process of reaching a runtime state is also effectively secured. New mechanisms that support secure distributed communications between secure operating systems

can be utilized to replace the reliance of current OSs on low-assurance and untrusted components.

High-assurance OS requirements: Many OSs require configuration tools to input policy- and security-critical information. Research is needed to develop the requirements and assurance metrics for these tools and to implement prototypes. In addition, standardized high-assurance OS requirements and well-defined construction requirements must be established to effectively support development and evaluation of high-assurance OS security. The high-assurance requirements must distinguish between state-of-the-art and research-level security engineering. In addition, advanced methods should be developed to support and facilitate secure systems documentation and evaluation processes.

Trusted paths: A reliable OS should enable trusted paths (i.e., protected, unspoofable communications channels between trusted parts of the system and the user) for transmitting sensitive information. Current mainstream commercial OSs typically lack trusted paths and are so complex that security cannot be verified. A trusted session built on trusted paths should be extensible across a distributed enterprise. In addition, trusted paths and sessions need to support business processes in a manner that ensures both system security and overall quality of service.

OSs need to provide for dynamic secure reconfiguration to support new applications without substantial security modifications. To meet this objective, mechanisms are needed to control and enforce privileges before, during, and after the reconfiguration. These mechanisms should incorporate enterprise business models and should be easy to use. Several major OS file systems lack discretionary control mechanisms that allow more than "owner control" of files and directories.

Virtual machines: Support is needed to ensure that virtual machines (multiple instances of OSs operating simultaneously on a single computer) operate securely. Partitioning a machine into virtual machines to support concurrent execution of multiple OSs poses several challenges. For example, varied OSs must be accommodated and the performance overhead introduced by virtualization must be small. The virtual machine manager should also enforce OS isolation and ensure that necessary inter-OS communications do not compromise security.

5.3 Security-Centric Programming Languages

Definition

Languages are used throughout software development, from low-level assembly languages and machine code, to conventional programming languages used for application development, to high-level modeling and specification. Security requirements are also expressed in languages. Security-centric programming languages address security as part of the language and incorporate features (or the absence of features) to increase the assuredness of code written in the language.

Importance

The rising number of reported cyber security vulnerabilities due to errors in software requirements, design, and implementation necessitates research to develop tools that can better specify security requirements and programming languages that produce inherently more secure code. Such tools and languages could also be used in developing certifiably secure systems.

State of the Art and Capability Gaps

Key R&D challenges in security-centric languages include better support for expressing security attributes in high-level and low-level languages, methods to transform the higher-level security requirements and descriptions of secure components into implementations, more accurate and scalable high-level analyses for cyber security, and efficient runtime enforcement of trust management policies. R&D is also needed to seamlessly integrate the results of this work into current software development processes. R&D topics include:

Secure language design: The design of security-centric languages focuses on the explicit specification of security features. The emphasis is on clarity and ease of use. Research is needed to enable modeling and high-level programming languages such as rule-based languages to explicitly express security features. Security can often be integrated into a language by extending existing features and attributes or removing others. However, when changing the language is impractical, libraries and development environments could provide enhancements or restrictions that help produce more secure software.

Secure implementation methods: Secure implementation methods should map system descriptions and security components in higher-level languages to system descriptions and security components in lower-level languages. Because lower-level languages support the functions that a system can perform, secure implementation methods must be designed to ensure that security flaws are avoided and that vulnerabilities are not introduced. In addition, automated methods for generating well-defined components of complex secure systems could reduce the cost of developing, validating, and maintaining those components.

Security analysis: Security analyses check whether security requirements are satisfied, support secure implementation methods, and detect instances of known vulnerabilities. An example of the use of analysis to verify that security requirements are satisfied comes from the model-carrying code paradigm: Using this principle, mobile code is accompanied by a model of its security-relevant behavior, allowing the recipient to analyze the received code to verify that the model's behavior is consistent with the given security policy. An example of low-level code analysis is checking whether a C or assembly-language program has attempted to illegally access memory. Proof-carrying code and typed assembly languages provide other approaches to such checking. Information-flow analyses that produce fewer false positives and effectively handle common languages are needed, as are more scalable verification techniques.

Analyses that guarantee conformance to security requirements can eliminate or reduce runtime compliance checking. In contrast, analyses that detect instances of known vulnerabilities can ensure only that specific common defects are absent, not that the overall security objectives are met. Examples include program analyzers that detect potential buffer overflows and format-string vulnerabilities (which enable users to initiate potentially harmful code manipulations) or vulnerability scanners that detect common weaknesses in operating systems and network configurations.

Secure execution support: Secure execution support relies on efficient runtime techniques to help achieve security goals. One example is language-based techniques that have led to more flexible and efficient mechanisms for enforcing access control policies. However, as traditional access control is superseded by trust management in enterprise systems, efficient enforcement of trust management policies is required. The distributed nature of the problem (i.e., the policy itself is dispersed) must be addressed and languages for expressing the policies need to be developed.

Development frameworks: Improved development frameworks for software assurance could dramatically reduce the effort required to specify and implement the process of security analysis. Frameworks are needed for efficiently achieving these results using declarative rules and transformation of higher-level languages into lower-level code specified in an aspect-oriented style.

5.4 Security Technology and Policy Management Methods and Policy Specification Languages

Definition

An organization is governed by security policies that describe the rights and responsibilities for accessing IT systems and information at various organizational levels, establishing priorities for use of the organization's resources, and releasing information. Policy specification languages are the means by which these policies are expressed and implemented across an organization's systems, networks, and information.

Importance

Security policies derive from various sources, including an organization's business rules, regulations established by Federal regulatory agencies, or public law. Federal agency policies are also governed by Presidential directives as well as government-wide and local policy making. In addition, state and local governments, universities, and companies establish their own policies to manage their activities.

Policy languages enable organizations to specify and promulgate policies and customize them to specific needs. IT system security policies are important to organizations not only to protect resources, including intellectual property, but also to promote business activities. As organizations become more complex, policy languages may be the primary means by which an enterprise interacts with its IT systems.

State of the Art

Cyber security must be managed across an entire organization, including systems and networks connected to the IT infrastructure. Formulating and implementing organization-wide security policies is difficult. The policies may not completely cover all relevant situations, can be mutually incompatible, or may exceed the scope of organizational control. They can also be ambiguous and difficult to evaluate for consistency and compatibility. Moreover, even if policies are expressed unambiguously, identifying and assessing their collective impact and understanding how to implement them can be challenges. Administration of systems and policies, especially those relating to security, becomes even more difficult when an organization's IT system infrastructure is largely outsourced.

Taxonomy of policies: Taxonomies provide principles and practices for the classification of systems. Taxonomies for integrated organization-wide security monitoring and management are slowly emerging. Aspects of a taxonomy may include both the source of a policy (e.g., security policies may be derived from public law) and technologies to implement policies. A generally accepted taxonomy or ontology for describing and organizing organization-wide security concepts would be a significant step forward.

Languages for enterprise security policy: Policy languages for security, and particularly for trust management and rights management, should provide high-level controls for architecting process execution and information sharing. The languages that have been developed to express the controls and constraints on organizational operations do not easily support the specification of detailed controls. These high-level policy languages must enable system administrators and decision makers to select specific controls, assess their potential impact, identify potential adverse interactions, and adjust the controls without undue complication.

Controls over different resources may interact with unanticipated effects. For example, firewall policies at different layers and locations may result in awkward firewall management or even new vulnerabilities. Certification of system security is complicated by the large number of events and potential interactions that must be considered. Currently, no languages have sufficient generality to express organization-wide security policies, but some languages can be used in limited contexts.

Legacy systems: Established systems operate based on policies that are implicit in software but not always explicitly known to the user. Legacy systems present the challenge of extracting and reverse-engineering accurate, meaningful policies from code. Some organizations outsource the management of their information systems to control costs, benefit from

economies of scale, and gain centralized control over legacy software applications that may have been developed in isolation. However, this business strategy carries risks. Legacy systems may not be fully integrated because of unresolved conflicts among security policies and organizational responsibilities, and thus may present vulnerable targets for attack.

Capability Gaps

Organizations and their operations will be hampered without the necessary security and security policies for accessing IT systems, networks, and information. Each organization must determine what constitutes sufficient security and how to manage an integrated organization-wide IT infrastructure. Unfortunately, no ready-made solutions yet exist. The major gaps in the technical foundations for improved security regimes include the need for a rigorous semantic basis for policy specification languages and the need for assured consistency and acceptability of security policies between organizations.

5.5 Information Provenance

Definition

Information provenance can be defined as the accurate historical record of an information object such as a digital text, an image, or an audio file. Provenance begins with identification of the original form and authorship of an object or its constituent components and continues with identification of each subsequent alteration to the object. Provenance information can include not only what was changed but also who or what produced the change, when the change was made, and other attributes of the object. As reliance on networked information and transactional processes grows, the need for technical means of establishing information provenance becomes increasingly important. The goal of information provenance capabilities is to track the pedigree of a digital object from its origin through all transformations leading to the current state.

Importance

Today, most information is aggregated from many sources and, even when the sources are sound, the aggregation processes can create weak points in information management mechanisms. Many new and emerging applications can store information in different formats for static, stream, and interactive multimedia use. Some products transform data from one format to another, thus making many products interoperable but at the same time compounding security challenges. Separating releasable data from sensitive data becomes more difficult with these aggregation and transformation processes. Furthermore, the vulnerabilities in transformation processes add a new dimension to concerns about reliability. Information provenance techniques are necessary to provide reliable histories of information that is received from, and delivered by, IT systems.

Provenance of information can help a user determine whether to trust it and how to interpret it. Information provenance techniques are also needed to control information sharing. Partners (e.g., allies, collaborators, or corporations engaged in a joint project) generally want to share information, but only to a limited extent. For example, there may be a constraint that information of certain types can be shared only among certain partners. Enforcing such constraints is complicated by alternate data formats and by transformations that combine data and deliver derived results. For example, the classification level of information derived from otherwise unclassified sources may prevent its public release. In addition, as diverse datasets are combined, accurate information may be interspersed with inaccurate information. To verify the provenance of the information, data about its source and derivation (or aggregation) must be propagated with the information itself.

State of the Art

Information provenance combines key concepts from operating system security, access control, and authentication. R&D advances have enabled application of some information provenance techniques in such areas as security software and management of large-scale collections of digital materials, often referred to as digital libraries. But next-generation versions of related technologies, such as metadata processing, taxonomies and ontologies,

and digital rights management need to be integrated into more sophisticated information provenance capabilities.

Capability Gaps

Information provenance methods, technologies, and tools are needed to help people and systems take better advantage of digital information by understanding its history and credibility. Taxonomies and ontologies for provenance metadata that go beyond those for security classifications and procedures for controlling secure documents are needed. Fully developed and integrated information provenance standards and digital rights management capabilities are needed to realize the benefit of information provenance for the IT infrastructure, and particularly for cyber security and information assurance. R&D areas include:

Metadata: Information provenance depends on tracking and maintaining data about information, known as metadata. Metadata might include the source of the information, transformations producing the derived information, and even procedures to extract content. Although it is not practical to maintain all details about information provenance, metadata can provide a conservative subset. Metadata can be maintained at varying degrees of granularity, depending on the sensitivity and criticality of the information.

Metadata labeling and granularity: A fine-grained approach to controlling data in repositories requires extensive labeling. For example, the label on the output of a transformation may reflect the type of transformation and its level of sensitivity. When a single object contains components constructed from different sources, OS-level mandatory access control is poorly suited to separate the components. Refining the granularity of provenance will be needed for documents that may have components constructed from different sources, whose provenance may need to be tracked separately. Manageable frameworks for fine-granularity provenance will require research drawing on ideas from secure OSs and language-based security, as well as interoperability and human usability.

Taxonomies, ontologies, and standards: Currently, there are no accepted information provenance standards, such as standard taxonomies and ontologies for classifying types of metadata. Metadata that support information provenance concepts and principles could have a powerful effect on the development of information systems. Assumptions about data used by these systems would be explicit and could be used in determining the amount of trust associated with the information. Reliable provenance information and tools for analyzing provenance will require ontologies for provenance metadata and corresponding operations.

Access controls: At the OS level, provenance may be based on mandatory access control, where the system controls the metadata labels applied to each field of data. Mandatory access control documents the data's sensitivity and classification and assures that transformation and aggregation results are labeled at the highest level among the input sources. Multi-level security remains a coarse-grained approximation for provenance and ensures the separation of classification levels. For example, multi-level security guarantees that data labeled at one level of classification do not contain information marked at a higher level of classification.

Memory management: A small portion of allocated memory may contain sensitive information that may expand during processing. The balance of the allocated memory may be of lesser sensitivity. The provenance of these different portions of memory must be tracked separately to accurately label different outputs.

Code management: Provenance could also be used to determine the portions of a process that can be permitted to execute as code. For example, code that is downloaded from the Internet might not be permitted to execute in certain IT environments. Currently, mobile code systems use a variant of this approach to distinguish locally resident libraries that are trusted to execute with high privilege from downloaded code, which can be executed only with low privilege.

Digital rights management: Rights management languages such as eXtensible rights Markup Language may help to ensure that metadata are used uniformly and consistently across a range of systems. Digital libraries of documents with detailed provenance information could benefit from this approach, if it proves successful.

5.6 Information Integrity

Definition

Integrity is the attribute of information that addresses its authenticity, correctness, and reliability. Protecting and monitoring information integrity are the goals of technologies and tools that prevent tampering and detect unauthorized modification or destruction of information.

Importance

Information integrity is a prerequisite for trust throughout the IT infrastructure. Without integrity, data, information, messages, and systems cannot be trusted. Without trust in the underlying information, higher-level functionalities, including measures to protect and safeguard the system itself, cannot be relied upon.

Data integrity assures that unauthorized modification of a system's data resources is detected and that messages or data in transit, including headers and content, are unchanged between the data's source and destination. Data resources include system configurations, data structures, the code controlling the behavior of the operating system, and other system or application software. Integrity controls also provide non-repudiation – that is, proof of the origin and integrity of data that can be verified by a third party, which prevents an entity from successfully denying involvement in a previous action. Integrity is necessary for a system to provide reliable services, including security services. Many attacks begin by undermining system integrity.

State of the Art

Information integrity is an essential foundation for most security services. An authentication service, for example, cannot succeed if the information it relies on is inaccurate or unreliable. But systems and information have differing integrity requirements. As a result, security managers must ensure that information is appropriately protected, that the appropriate levels of security are met, and that resources are applied in proportion to the integrity requirements and their cost-effectiveness. Information integrity requirements are not static, however, and integrity must be reassessed throughout a system's life cycle.

Information integrity can be compromised through accidental or intentional action by system developers, system administrators, operations and maintenance staff, end users, routine equipment failures, or malicious actors. To ensure effectiveness, information integrity planning and analysis must be coordinated with other concurrent engineering activities. To achieve and maintain information integrity, a variety of engineering methods and techniques may be implemented. For instance, using a key-hashed message authentication code, information integrity is achieved by hashing the contents of each message with any header fields that also require protection. When the peer receives the message and the hash, the peer re-computes the hash value and checks that it equals the hash received.

Assuring system integrity may require a broad range of tools, including hardware and software solutions. R&D in system integrity is needed to reinforce this holistic approach to integrity through new engineering techniques and methodologies.

Capability Gaps

Advances are needed in information integrity technologies to enable managers to continuously evaluate integrity throughout a system's life cycle and to ensure that integrity is maintained regardless of system mode or state (e.g., start-up, shutdown, normal operations, preventive maintenance, emergency shutdown). Developing these capabilities will require new software engineering and analysis techniques for integrity. Key technical areas in which R&D is needed include:

Integrity across formats: Electronic information now includes not only text but audio, video, signals, and other forms of data, and information generated in one

form can be transformed into others. Methods, techniques, and tools need to be developed for maintaining the integrity of information when it undergoes one or more format transformations.

Scalability: Because the volume of data processed in real time across the IT infrastructure will continue to increase, new highly scalable techniques extending from small networks to global information systems are needed for assuring information integrity.

Granularity: Some types of information (e.g., financial transactions, intelligence data, medical records) may require fine-grained scrutiny, while others (e.g., Internet search results) need less rigorous assessment. Techniques for assuring information integrity need to include capabilities at varying appropriate levels of granularity.

Integrity across security layers: Capabilities for checking information integrity should include techniques for every layer of system security.

5.7 Cryptography

Definition

Cryptography is the study of secret writing. In practice, cryptography applies mathematics, engineering, and information theory to provide data confidentiality, integrity, and authentication protections. Cryptanalysis – the other side of the same coin – is the art of defeating these protections. Designers of network and system security features employing cryptography must understand cryptanalysis to do their jobs effectively.

Importance

Cryptography is a foundation science for assuring information security. It is now widely used to protect both stored and transmitted information and to implement digital signatures for a variety of applications, such as electronic commerce. Strong cryptographic algorithms and protocols are built into even ordinary Web browsers and are widely used to secure communications and transactions. However, vulnerabilities in cryptographic protocols often lead to their being broken, and weak or badly implemented cryptography still occurs.

State of the Art

Several decades ago, cryptographic research in the U.S. was largely the purview of Federal intelligence agencies. Today, a substantial open cryptographic research community is well established, with academic and industrial as well as Federal participants.

Designing information systems that resist attack often consists largely of applying techniques that make it difficult for an attacker to break a cryptographic algorithm. The security of a well-designed cryptographic system rests only on maintaining the secrecy of its keys; if adversaries know all details of the system except the keys, a successful attack should still be unlikely. Emerging approaches for cryptography include multi-party computation (computing on encrypted data) and partitioning computations across separate hosts. A new thread of work is developing on engineering production-quality realizations of new algorithms to help ensure that the state of practice benefits from advances in cryptography research in a timely fashion.

The difficulty of vetting algorithms and designing robust modes with good, provable security properties argues for the development and use of carefully vetted standards rather than a more ad hoc approach to cryptography. Symmetric key (block ciphers and stream ciphers) and asymmetric key (or public key) are today's main types of cryptographic algorithms. Symmetric key algorithms, in which both parties use the same key, are usually much faster than public key algorithms, in which there is a private key and a related public key, and are used for most bulk data encryption. Public key algorithms are typically used for key management, authentication, and digital signatures. Hash functions, which are also widely used, have no key but produce cryptographic checksums of messages.

In general, the security of cryptographic algorithms cannot be proven, so extensive cryptanalysis is required to vet them against attack. These heavily analyzed algorithms are then used as primitives in carefully constructed protocols or modes of operation for particular functions. It has now become standard practice to require analysis of the mathematical techniques to ensure that breaking these modes requires

a large amount of computation or is equivalent to breaking the underlying cryptographic primitive.

Block ciphers encrypt fixed-size blocks of plaintext into fixed blocks of ciphertext. Stream ciphers generate a "keystream" of pseudo-random bits, which are logically combined with data to produce the ciphertext and can potentially be applied more quickly (i.e., with less latency) than block ciphers. Secure symmetric key block cipher algorithms are now available along with asymmetric key algorithms for key transport, key agreement, and digital signatures.

Capability Gaps

The algorithms described above perform well on today's powerful desktop computers. Custom semiconductor devices provide even higher performance. Performance issues arise mainly at the extremes, such as in ultra-high-bandwidth communications trunks, and in computation-constrained environments, such as in RFID chips or other inexpensive, low-power devices. Both types of uses require more efficient algorithms than those currently in use.

Although high-performance (e.g., high-bandwidth) applications would benefit from the availability of parallelizable algorithms, existing standardized authenticating encrypting modes are not readily parallelizable. More efficient, parallelizable constructs are known, but they rely on block cipher primitives and have intellectual property conflicts. Stream ciphers, which may offer significantly better performance than block ciphers, can be easily compromised; therefore, they are most effective when tightly bound to a mode with authentication. One promising cryptographic research area is combined stream cipher/authentication algorithms. These algorithms stress speed and lightweight implementation in order to produce a secure, efficient, authenticated stream cipher that is robust against inadvertent misuse or misapplication.

Cryptographic hash functions, which have been called the Swiss army knives of cryptography, have been the subject of little published research. Recently, however, one widely used algorithm and several less widely used ones have been broken. Continued research is needed toward understanding these functions at a level comparable to that of block ciphers and improving their performance.

Identity-based public key cryptography is a promising and rapidly developing area. Although the technique was proposed decades ago, the first practical identity-based algorithm was developed in the last few years. In this scheme, a unique identifier associated with a person (for example, an e-mail address) is used as a public key, which can then be used to encrypt a message to that individual. An authority generates public encryption parameters to be used with the identifier to encrypt messages to that identifier and provides the private decryption key to the person with that identifier.

Quantum computing is potentially the greatest long-term threat to existing cryptography. Practical quantum computers today are only a research goal, and opinions differ over whether they can ever be built. While symmetric key cryptography effectively resists known quantum computing attacks if key sizes are large, quantum computing algorithms are capable of breaking known public key algorithms.

Quantum encryption is theoretically unbreakable even by quantum computers, but such encryption is today somewhat cumbersome. Moreover, quantum encryption does not replace conventional public key cryptography for digital signatures. Some functions are provably as difficult to compute in a quantum computer as on a conventional computer. However, there is some question as to whether a practical public key cryptosystem can be developed that is as hard to break on a quantum computer as on a conventional computer. If not, the arrival of mainstream quantum computing – which some believe will occur within a few decades – would mean the end of public key cryptography and digital signatures. The cryptographic community would then have to return to the use of more complicated symmetric key management methods, or perhaps to even more cumbersome quantum encryption key management methods. Alternatively, the community may turn to new classes of algorithms that have yet to be developed.

5.8 Multi-Level Security

Definition

Multi-level security (MLS) addresses the protection and processing of information at varied levels of security classification. MLS systems must permit simultaneous access to information by authorized users and processes while denying access to unauthorized ones.

Importance

Today, information-sharing requirements are global for both the Federal government and the private sector. Many elements of critical infrastructures rely on communications over untrusted networks with varying levels of security policies. Sharing information and maintaining distributed control among MLS systems using such networks are significant challenges.

State of the Art

Sensitive information systems that were previously isolated are now linked to other networks that, in many cases, have minimal levels of trust. Unfortunately, few multi-level or cross-domain secure systems support data sharing or transfers between networks with disparate security levels. Many MLS systems support one-way data transfer, filter or parse only specified, fixed-format data types, or require human intervention. Since the inception of net-centric warfare, the need for automated, higher-trust devices that permit transfer of additional data types has grown.

Capability Gaps

Designs for multi-level and cross-domain security systems must allow direct and efficient interaction between various users operating at differing security classification levels, while still enforcing the applicable security policies. These systems must be able to detect and prevent cyber attacks while ensuring availability to authorized users. The systems must support required functionality, operate in near-real time, and provide adequate mechanisms to ensure confidentiality, integrity, and availability of information. In addition, the systems must provide sufficient assurance to support system security certification for use in a target operational environment.

The challenge in MLS is to balance security against risk, cost, timeliness, and usability. In particular, research advances are needed in engineering tools and techniques for MLS application development, inherited trust, strong separation principles, and internationally recognized cross-domain solutions.

Information flow: In multi-level and cross-domain systems, procedures to control information flow assure that only individuals with appropriate authority are allowed to access data or affect associated processes. Security policies specify the rules governing these procedures and processes. Such systems require the analysis of data access policies and programs to protect against the opening of covert communication channels that can allow uncontrolled information flow.

The Internet-mediated interconnectivity of multiple networks operating at varying security classification levels increases the exposure of vulnerabilities to attack. Higher-trust networks must ensure data confidentiality, preventing the intentional or inadvertent disclosure of sensitive information while protecting against adversaries on lower-trust networks who may attempt to establish an outflow path for sensitive data. In addition, the integrity and authenticity of data used in higher-trust networks must be assured to mitigate potential compromises by malicious users.

Policy enforcement: Since there are no mechanisms to identify malicious intent and networks at all levels are susceptible to insider threats, strong security policies must be implemented. To address the security challenges of shared network-centric environments, data and resource protection policies must be enforced. Data protection services include non-discretionary policies that dictate how data are shared, discretionary access mechanisms that allow users to specify data access rights, and transfer policies. Resource protections include secure communication to prevent unauthorized modification or surveillance, mechanisms to ensure resistance to attacks, malicious content detection, and the use of trusted processes to manage separation between networks and data.

Trusted operating systems base: Several trusted operating systems provide mechanisms sufficient to

support a strong security policy. However, much of today's security is based on application-level user functionality and often does not use security mechanisms present in an OS. For example, a system can enforce strong network separation but may continue passing sensitive data if transfer policy enforcement is weak. A trusted OS can be the foundation for building a secure system, but an overall system design also needs to integrate strong security policies across all components of the system.

Validation and verification: A secure system needs an effective security policy, according to which information flow and system behavior are governed. Under the current certification process, new multi-level and cross-domain systems must undergo penetration testing. Unfortunately, certification testing is time-consuming and does not guarantee that all security features are correctly implemented. Tools to automate this testing process are needed. Validation and verification software tools based on formal methods can provide evidence of the strengths and weaknesses of system security policies. Benefits resulting from the use of such tools may include reduced time from system design to implementation and improved system security.

Data labeling: Applying and correlating data labels among applications, OSs, and trusted database schemas are difficult because commercial products define and use labels in different ways. For example, mandatory access control is based on clearance levels associated with users and classification levels associated with data. In contrast, a single classification level – for example, sensitive – in the OS may correspond to a range of sensitivity labels in a trusted database. Mandatory access control labels are defined in Federal statute, while sensitivity (discretionary) labels are defined by the organization. Finally, untrusted applications cannot be relied on to support trusted labels. Models need to be developed to support interrelationships among mandatory and discretionary labels.

Coalition issues: Data integrity and authenticity are achievable only in certain environments. In many instances, the data source and whether the data have been modified in transit cannot be determined. Data communicated over the Internet are often not authenticated or provided integrity protection because of the expense of implementing these security policies. In some instances, non-scalable, system-unique authentication mechanisms are implemented or, alternatively, a higher level of risk is accepted. New, internationally recognizable solutions for communities of interest are needed. Token standards, certificate issuance, cross-certification, and inherited trust are just a few of the issues to be addressed.

5.9 Secure Software Engineering

Definition

Software engineering is the application of engineering methods, technologies, tools, and practices to the systematic development of computer software. Security is a component of software engineering, and security requirements must be met just as the requirements for functional correctness must be. The primary goal of secure software engineering is the design, development, verification, and validation of secure and correct software. Secure software engineering is required throughout the software engineering life cycle.

Importance

Cyberspace increases the importance of security in software engineering because the infrastructure that makes systems highly interconnected also makes them highly accessible to adversaries and, if not adequately secured, vulnerable to attacks. Exploitable vulnerabilities in the IT infrastructure frequently are traceable to failures in software development and engineering. Secure software engineering will become increasingly important as service-oriented architectures are more widely adopted. Service-oriented architectures are loosely coupled, interoperable application services, developed using small sets of standardized components that enable these services (e.g., Web services) to be broadly provided and incorporated.

State of the Art

Secure software engineering is a concern across all phases of a system's life cycle: initiation, acquisition or development, implementation and assessment,

operations and maintenance, and sunset or disposition. During the initiation phase, the confidentiality, integrity, and availability objectives are specified. Assets that need to be protected are specified and a preliminary risk assessment is performed. In the development phase, the security control baseline is selected and modified, as required, and the security controls are designed. Security imposes additional overhead on a system's development and performance that must be balanced against the cost and risks of a system compromise.

To ensure the integration of security controls into a system during the implementation and assessment phase, programming languages, compilers, and libraries that are certified against specific security criteria should be used. Unfortunately, there are few widely used certified compilers and libraries. In addition, code should be verified against the applicable specifications to ensure that the confidentiality, integrity, and availability objectives have been met and that security has not been compromised.

Verification using formal methods still remains elusive despite significant advances. Testing software components and assessing a system require a combination of static methods (e.g., reviewing the software/firmware code and documentation to identify flaws and potential vulnerabilities) and dynamic methods (e.g., testing the software and firmware against suitably chosen scenarios). Frequently, unit testing and testing of distributed software are key development activities. Assessments may lead to refined security requirements and controls that necessitate revision of a system; this revision process continues during the operations and maintenance phase.

Capability Gaps

Research results in software engineering techniques that address security issues in areas such as programming languages, software development, and code analysis have proven effective in some phases of the system development life cycle. Further R&D should focus on incorporating the best security development practices, such as from language-based security, access control and trust management, and protocols, into software engineering methods. Advances in secure software engineering principles, methods, tools, techniques, and semantics of languages will be required. Ideally, secure software engineering practices should include the following elements:

Security policy models: Potential threats and vulnerabilities and their associated controls and countermeasures define the security environment in which security objectives must be met by the software under development. Security controls can be configured to achieve these objectives. Security specifications and architectures provide choices for an overall strategy for configuring security mechanisms. As software components are integrated and applications and services interoperate, the consistency and consolidation of security policies must be validated to assure that no new vulnerabilities are introduced and systems are not unduly constrained. Few, if any, development tools currently meet all these needs.

Certification: Secure software engineering should enable stakeholders to assess software security and thus should provide methods to demonstrate trust at each phase of the software life cycle. This requires testing and verification methods that are thorough, practical, scalable, and relevant to security, with static and dynamic code analyses tailored to exposing vulnerabilities in code and weaknesses in security models. Security certification becomes more complicated as programs grow larger, architectures become more adaptive, and software involves more polymorphisms.

Human-computer interfaces/human-system interactions (HCI/HSI): While security functions in software are designed to be transparent to users, both users and system administrators need ways to understand the nature of security policies. System administrators need sophisticated graphical interfaces that ease the difficulties of configuring system security. Presenting an overview of cyber security for a distributed system is an HCI/HSI design challenge. Currently, HCI/HSI design is more art than science

and there are few broadly applicable guidelines for best practices in the context of IT system security.

Updates, upgrades, and system migration: The difficulty of facilitating change in an operational environment is frequently underestimated and is always complicated by the breadth and depth of understanding required to plan, design, build, and systematically introduce change. For example, frequent software updates and occasional software upgrades present a software engineering challenge. These updates and upgrades are often required to operate correctly and effectively across different platforms and applications and across the IT infrastructure without introducing new vulnerabilities. There are no tools and techniques currently available to facilitate this aspect of software engineering with sufficient assurance.

Interoperations and trust: Composability of assurances must be considered when assembling software components during development and when combining services dynamically at runtime. Service-oriented architectures present particular challenges to traditional validation, verification, certification, and security assurance because of their dynamic nature. Trustworthiness of a system cannot be automatically assumed but must be verified.

Trusted libraries: Code libraries have long been used in software development. Libraries facilitate the re-use of components, tools, techniques, and best practices, and should be extensible to adjust to new insights and developments. While today the more successful a library is, the more valuable and trusted it is perceived to be, security, trust, and information provenance should also become explicitly evaluated aspects of libraries.

Ontologies and taxonomies: Libraries require structure and organization, which often come from ontologies and taxonomies. Taxonomy concerns the practice, principles, and rules of classification, while ontology refers to a model of components and their interactions in a given knowledge domain. The purpose of any taxonomy or ontology is to facilitate sharing of knowledge through a commonly accepted language of concepts and terminology. Such sharing helps realize the full benefits of the IT infrastructure, especially for service-oriented architectures and for data exchange. However, coordination of cyber security taxonomies and ontologies is a fundamental challenge in the development of standards, guidelines, and best practices. Broadly accepted standards could have a strong positive effect on the development of IT systems. For example, there are currently no accepted standards for metadata for software re-use. Code available for re-use could have metadata describing the degree of trust that can be attributed to the code.

5.10 Fault-Tolerant and Resilient Systems

Definition

Fault-tolerant systems are systems designed to continue to function with predefined performance measures despite the malfunction or failure of one or more components. A resilient system will easily recover from disruptions within an acceptable period of time.

Importance

Fault tolerance is generally focused on mitigating the impacts of non-malicious events such as accidents and random failures. New principles need to be added to the concept in order to develop systems that are resilient in the face of malicious activity and hostile attacks. In a highly distributed system environment such as the Internet, component and node failures are common. Resilient systems (also referred to as "fail-secure" systems in the context of IT security) that retain their security properties amid component failures could mitigate potential risks that may arise as a result of such failures. Systems designed to maintain predictable timeliness properties must also be resilient against denial of service attacks and disruption of system resources.

In some mission- and safety-critical systems, such as the national power grid, an attacker who can manipulate control variables faster than the system can respond could potentially produce catastrophic results. "Resource-secure" systems are constructed to

preclude such abuses. Resiliency and real-time fault-tolerance are crosscutting requirements in many mission- and safety-critical systems. Both these requirements as well as security requirements need to be effectively addressed without adversely affecting system functionality and performance. For example, a malicious attack should not cause service disruptions because the system failed to maintain timeliness of response or to recover from unexpected behavior.

State of the Art

System dependability can be achieved only by understanding the relationships between security and the properties of fault tolerance and resiliency. The relationships among these properties have been investigated primarily in studies of various types of synchrony (i.e., timeliness) assumptions to resolve fault-tolerance problems. However, the fundamental relationships between security and fault tolerance and between security and timeliness are largely unexplored, especially in Internet-enabled systems.

Capability Gaps

Building secure, dependable systems that simultaneously exhibit predictable timing behavior, withstand failures, and rapidly recover will require extending fault-tolerance computing techniques to system security properties. R&D challenges include:

Transforming techniques: Simply transferring – rather than transforming – existing fault-tolerance techniques to cyber security may introduce vulnerabilities that an attacker can exploit. The reason for this is that fault-tolerance techniques are developed based on the assumption that random failures are inevitable and can be mitigated, but this assumption does not necessarily hold for malicious attacks. The transformation approach, currently an active research area, requires protecting from malicious attacks each fault-tolerant computing technique placed in a cyber security scenario.

Safety and liveness must be exhibited at many levels of system abstraction. (A safety property stipulates that "nothing bad" will happen during the execution of a system. A liveness property stipulates that "something good" will happen, eventually, during the execution of a system.) A human is frequently involved at some level of an assured computing system to coordinate post-attack recovery. Unfortunately, human activity (whether an accidental error or an intentional action) can result in impairments to assured computing. Non-random and directed faults introduced by a malicious attacker challenge the transformation of fault tolerance to cyber security, because faults introduced by legitimate users may be indistinguishable from malicious attacks.

Performance guarantees: Research in real-time systems has traditionally focused on dedicated systems, using a task-scheduling model to capture the application workload and to guarantee absolute performance. In integrating real-time and security enforcement, real-time system technologies must be expanded to handle more flexible and dynamic workloads in order to guarantee differing levels of performance. Absolute guarantees must be provided to safeguard the responsiveness of critical control functions in cyberspace so that, even under attack, essential services can continue to operate. Other less critical infrastructure services can be designed with weaker levels of guarantees, such as an occasional timing failure whose frequency of occurrence is bounded. Research on typing, formal theorem proving, automated runtime monitoring, and other approaches is addressing resource-bounded safety, but more work is needed.

End-to-end security policies: Many future real-time multimedia applications will depend on IT infrastructure support for end-to-end security. The Internet currently provides building blocks, such as Internet Protocol security (IPsec) virtual private networking gateways, to support secure network operations. Mobile users can access mission-critical network resources by establishing secure connections to an office network's IPSec-compliant gateway and firewalls. However, to protect an Internet telephony session, the security policies of different network security devices along the end-to-end route must be consistent and interoperable. Inconsistencies in a set of individually valid security and routing policies might introduce undesirable side effects such as

unexpected violations of end-to-end security properties.

Dependability: Dependability is a necessity in both application and infrastructure software. Research is needed to devise new fault-tolerant techniques to localize and minimize the damage caused by untrusted malicious software, such as by using virtual channels or other methods. This research can benefit from advances in real-time systems research, including open-systems resource management and real-time queuing theory.

5.11 Integrated, Enterprise-Wide Security Monitoring and Management

Definition

An enterprise consists of one or more organizations cooperatively engaged in achieving a common goal. An enterprise is often large and may include multiple supply chain partners. Integrated security monitoring and management provide real-time situational awareness, decision support, and command and control over the security of the enterprise's systems, networks, and information.

Importance

The absence of integrated, enterprise-wide security monitoring and management may impede an enterprise's ability to respond rapidly and intelligently to cyber attacks and may leave its information systems more vulnerable, less trustworthy, and less usable.

State of the Art

Security monitoring and management for many enterprises are based on a layered defense. At a minimum, this may include the use of passwords, firewalls, anti-virus protection, and cryptographic protocols. Management of network security may or may not be organized hierarchically. But for an enterprise, any local information needs to be aggregated into a common operational security picture across all systems, networks, and information. Among the technical hurdles in building this capability are system complexity and interactivity. For example, network behaviors in large, highly interdependent infrastructures often contribute to the emergence of anomalous behaviors affecting the connected systems. Geographic dispersion of the components of an enterprise can also increase dependence on the IT infrastructure and can increase the complexity of administration. In addition, change control and configuration management may introduce system vulnerabilities if not implemented correctly.

Capability Gaps

Defining, understanding, quantifying, using, and managing the IT security of a large-scale enterprise are substantial challenges. The goal is a command and control capability for enterprise-wide information security that integrates monitoring and management systems to provide sophisticated decision support for the enterprise as a whole. Needed capabilities include:

Macro-models of network activity: Improving the understanding of activities taking place on enterprise networks is an open area for investigation. Varied models are available for monitoring and diagnosing global network behavior. For example, models of the Internet incorporate graphs, influence diagrams for causal analyses, fluid flow and other differential equations, and stochastic networks. Fluid flow analogies have been useful in identifying "sources" and "sinks" (dark or unused IP addresses) within the Internet. A recent research trend is toward more realistic scale-free Internet models that depict the Internet as consisting of both intensively connected hubs of nodes and sparsely connected nodes.

Analyses of these models have led to better understanding of Internet robustness against random node failures and fragility with respect to hub failures. To accompany the general Internet models, more specialized models focus exclusively on specific Internet problems. In one example, statistical mechanics and queuing theory are used to model Internet traffic congestion. Epidemiological models of computer virus and worm propagation are inspired by biological models of the spread of natural viruses; these include graphs, hidden Markov models, and

human virus propagation mechanism analogies. These emerging capabilities need to be integrated into new models of system complexity and security applications in large-scale enterprises.

Enterprise situational awareness: An accurate, real-time view of policy, enforcement, and exceptions is necessary for administration of networked systems of any size, and the larger the organization, the more difficult the problem. Situational awareness of the application of enterprise policy helps ensure that malformed policies can be identified and corrected both as the enterprise grows and when anomalous behavior appears. Visual and graphical presentations used for situational awareness provide a limited picture when an enterprise is large, but there is no agreement on the best approach for integrating and synthesizing heterogeneous data in a real-time decision-support capability.

Enterprise-wide information infrastructure: The fact that organizational components of a distributed enterprise are often controlled and administered independently introduces a variety of challenges associated with operating an enterprise-wide information infrastructure. Enterprises will coordinate and in some cases consolidate IT systems and networks to gain better control over information resources. Technologies to facilitate integration of enterprise-wide security monitoring and management will be required to achieve the interoperability necessary to support network-centric operations.

Global identity management: Problems arise when enterprises need to interoperate between distinct administrative domains. Agreed-upon approaches to managing identity information are necessary to allow systems and people to interact with systems in multiple domains. Development of standardized identity formats and associated software would enhance both security and functionality. This effort should begin with investigations to develop a detailed understanding of various organizations' requirements.

5.12 Analytical Techniques for Security Across the IT Systems Engineering Life Cycle

Definition

Security across the IT systems engineering life cycle must be defined, measured, and evaluated. Analytical techniques facilitate detecting, quantifying, measuring, visualizing, and understanding system security.

Importance

Security is a core development concern throughout the systems engineering life cycle, from initial conceptual design to retirement and disposal. Practical analytical techniques are needed to certify system security, estimate security costs, and evaluate the tradeoffs of various security controls and countermeasures at every stage of the life cycle.

State of the Art

Security is difficult not only to analyze thoroughly but also to effect. Assuring system security requires verifying that the system and its information cannot be compromised. Few IT systems are formally verified or rigorously tested from top to bottom. No system has been tested for every possible combination of configurations and events. No deployed system can be certified as completely secure. Furthermore, there is no assurance that a system composed of secure components is itself secure. New and unexpected vulnerabilities often emerge simply through the process of integrating a system's components, and minor changes to a program may have unintended effects on security.

Research has been directed at understanding software as mathematical objects at an appropriate level of abstraction. Mathematical theory provides a framework in which to analyze software and security. Techniques come from logic, automated theorem proving, model checking, and operations research. Many programming languages still lack rigorous mathematical semantics for all their features. Game theory, with its foundations in operations research,

logic, modeling, economics, and even sociology, views security as a game played between a system and an adversary, and offers new means to understand interactions between attackers and defenders, and to design security systems accordingly.

Formal methods for software engineering are mathematical techniques for the specification, development, and verification of systems. These methods are used to prove that a program meets its specification, including security requirements. Despite considerable effort, formal methods have been unable to prove properties of large complex programs because of the many combinations of events that must be analyzed – a time-consuming and computationally intensive exercise. Still, formal methods in general remain a promising approach to assuring correctness of code or the security properties of software systems.

Empirical approaches have often been used to understand and predict properties associated with software and its development. Such approaches have led to useful insights into software processes. Estimates based on empirical studies of large projects are useful in planning and managing software projects, although estimates can be subject to considerable variability and uncertainty.

Capability Gaps

Security is an important but often overlooked component of software engineering. Analytical techniques for understanding security requirements are being developed. Fundamental principles for software development and software engineering are difficult to identify and describe. Interdisciplinary tools for modeling, analysis, mitigation, and remediation that focus on the dynamic aspects of systems and networks are needed.

System models need to include multi-scale and multi-resolution capabilities with varying levels of abstraction and scale-free properties. These capabilities will enable more computationally efficient analyses that focus on only one level of abstraction. More accurate and efficient models are needed to understand the behavior and security of complex IT systems.

Current formal methods are not sufficiently robust for large-scale software development. Yet advances in these methods are needed to enable developers to address network environments and system design simultaneously. Formal methods must adapt to service-oriented architectures, incorporate social factors such as economics, and enable predictions of, for example, macro-level properties of system behaviors. In addition, analytical techniques and formal methods should lead to better approaches for certification of software and particularly of assurance properties of software systems. Analytical techniques addressing security must be founded upon fundamental principles and must be composable, scalable, usable, and widely applicable. Such capabilities will augment the tools and techniques for secure software engineering across the life cycle of software systems.

6. ENABLING TECHNOLOGIES FOR CYBER SECURITY AND INFORMATION ASSURANCE R&D

The topics in this category address more generic enabling technologies rather than those specific to IT system security. Such enabling technologies can be applied to the design, construction, and evaluation of IT systems in general and the Internet in particular in order to improve cyber security and information assurance. The topics in this category are:

- CSIA R&D testbeds
- IT system modeling, simulation, and visualization
- Internet modeling, simulation, and visualization
- Network mapping
- Red teaming

6.1 Cyber Security and Information Assurance R&D Testbeds

Definition

A testbed can be defined as a framework for experimentation for large development projects or as an infrastructural platform to support testing and deployment of technologies at a smaller scale. Unlike theoretical models and simulations, testbeds are made up of the physical hardware and software components of a real-world operational environment. In testbed environments, researchers can deploy experimental tools, generate execution scenarios involving multiple component interactions in real time, and collect and analyze the results. Because they allow researchers to investigate what can "break" a system, testbeds provide a uniquely rigorous way to test scientific theories and new technologies.

Importance

Given the scale and complexity of networks and enterprise systems and the constantly evolving variety of cyber threats and system vulnerabilities, testbeds have a particularly important role to play in cyber security and information assurance R&D. Disruption, denial of service, and denial of access attacks, for example, involve making networked resources unavailable to the people and systems that use them. These attacks include insertion of tasks into a process stream to divert attention, saturate resources, displace capacity, or disrupt communications across both the cyber and physical infrastructures.

A testbed focusing on these threats can enable investigations of reconfiguration, redundancy, or re-routing options, and self-healing and self-sustaining capabilities to rapidly restore services or to provide a minimum level of service until full recovery actions can be implemented. Such a testbed would also enable researchers to generate and test models to help mitigate vulnerabilities by identifying optimal configurations in an emulated real-world environment.

An R&D testbed can also be used to help researchers develop methods to protect against infiltration or theft, modification, and destruction of data. Testbeds can provide environments for testing the effectiveness of tools and technologies for protection against cyber attacks and attack detection, mitigation, and recovery, both for the purpose of improving technologies still in R&D stages and for testing the effectiveness of existing commercial technologies.

Because of the difficulty in performing tests of malicious attacks at the largest scales of the Internet, researchers often turn to models and simulations for gaining understanding of such events, such as the propagation of a rapidly spreading worm across the Internet. In addition to providing realistic environments for testing new protective and defensive technologies, testbeds also provide the means for validating models and simulations of IT systems and infrastructure at large scale, providing researchers with greater confidence in the results provided by the models and simulations they use.

State of the Art

Disruption, DoS, and denial of access attacks, which are now commonplace across the Internet, were not generally considered in the planning and design of critical infrastructures and computer-based systems.

Today, methods for protecting against or mitigating these kinds of attacks are not universally effective. However, because such attacks generally employ known techniques, the current state of the art in prevention and mitigation, which includes the emergence of self-healing networks and systems, is narrowing the gap. Several Federally funded networking and cyber security testbeds are aiding research in this area.

The issues of data infiltration, tampering, destruction, and monitoring have been addressed in sectors such as banking and finance, but less effort has gone into identifying and communicating information about specific threats to control systems, facilities, and other critical infrastructure assets. Under ideal circumstances, only trusted personnel handle critical system information. But such behavior cannot be assured in the face of increased network connectivity, added exposure of system vulnerabilities, and the surreptitious gathering of information that exposure can make possible. System vulnerabilities can be exploited to open a door for an intruder to monitor, remove, change, or destroy data.

Currently, there is no widely used infrastructure for experimenting with a wide variety of threats to IT systems or for testing and validating the effectiveness of new security products or novel next-generation approaches to providing cyber security and information assurance that are still at R&D stages.

Capability Gaps

Several elements of an experimental infrastructure are needed in order to support research in and evaluation of cyber security and information assurance technologies, including approaches to creating, expanding, validating, and effectively using testbeds for evaluating R&D results.

Although a testbed is itself an inherently physical infrastructure, both application of existing technologies and development of new technologies are required to support the construction of testbeds. This includes the use of virtualization technologies to create virtual nodes, whereby one machine can appear as multiple machines on a network, thereby allowing a network with a given number of machines to perform and behave as a much larger, heterogeneous network would. Also needed are rapid reconfiguration techniques that allow the physical and/or logical connectivity of a testbed, as well as the state (e.g., of operating systems or disk images) associated with testbed nodes to be easily and quickly modified. Geographically distributed networks need to be more easily integrated so that they can readily be combined to serve as a single, larger-scale testbed network.

Software infrastructure is also needed to enable effective use of cyber security and information assurance R&D testbeds, while ensuring that the testbeds themselves do not put operational networks at risk. This includes the use of software frameworks for supporting and automating experimentation; collection, analysis, and visualization of data; and software for characterizing behavior and performance. In addition, effective measures are needed to validate testbeds and verify their scalability, in order to help ensure that their behavior is realistic and representative of much larger networks.

6.2 IT System Modeling, Simulation, and Visualization

Definition

The model of an IT system abstractly describes what each component does, alone and in concert with other components. Models can capture behavior at varying levels of detail and abstraction. The properties of most IT system components are dynamic; they must respond to and interact with users, and can have behavior that varies with time (e.g., their performance may degrade over time). This makes it difficult for system modelers to identify which details to capture and which to ignore when developing a model. An effective model captures the essential properties of the components while keeping complexity manageable by hiding or omitting some details.

The purpose of modeling IT systems is to reproduce relevant properties and predict the behavior of individual and networked components. Researchers, analysts, and managers need to succinctly describe

how an IT component works under both normal and atypical conditions. Simulation and visualization are used to determine how the systems behave over time and under a variety of conditions, and to convey that information to humans. Few components of IT systems are static – the systems are constantly being modified and upgraded – so the models themselves must be dynamic.

Importance

IT systems can be flexible – components perform the same function under various circumstances – and malleable – components perform functions other than those for which they were originally designed. Modeling IT systems is fundamental to answering questions about performance, management, and security, and it enables researchers and analysts to draw conclusions about capability. For example, an accurate model of system call behavior in an operating system can be used to determine when malicious code is being executed. Such a model also enables trade-off analysis, such as to determine whether the overhead (e.g., time and cost) of checking for deviations from normal behavior patterns is worth the security it provides.

A key challenge of IT system modeling lies in the scope of its components – ranging from magnetic media to optical fiber and anything in between. Each system component presents its own problems for modelers deciding what to model. For example, while modeling the behavior of random access memory may not require an understanding of the underlying physics, knowing the physics of optical fiber is essential to answering questions about optical network throughput. When components are combined into enterprise-scale systems, complex interactions and the need for matching different levels of model abstraction add to the modeling challenge.

Simulations and visualizations based on system models can be put to a wide range of uses, including:

- Predicting system performance for a given configuration, or suggesting an optimal design
- Demonstrating the functionality of a given component in an enterprise
- Providing trade-off analysis for calls on system resources
- Exercising user interfaces
- Testing system defenses
- Monitoring routine system behavior to detect anomalies

By enabling better understanding of system functionalities and component interactions, such simulation capabilities can be used to design more secure systems.

State of the Art

IT systems require models that are able to scale up with increasing size and complexity. Today, models exist for small- to moderate-size systems. Some models for simple components are available (e.g., mean time between failure for data storage systems), as are complicated models for small-scale behavior. Sophisticated mathematical packages can provide analytical and compositional capabilities. Current research includes expanding the ability of complex system models to scale to enterprise size. Research on combining these models is on the horizon. For example, correlating models across scales (such as by using heuristics or artificial intelligence) may provide the ability to predict enterprise-wide behavior in order to detect failure, anomaly, or malicious activity.

Capability Gaps

The greatest capability gaps associated with modeling of IT systems are in scaling and composition. Few if any current modeling technologies can scale to realistic numbers of system components and throughputs for large IT systems. Complex systems theory suggests that, under certain circumstances, even simple systems cannot be modeled to great fidelity for long periods of time (chaos precludes predictability). IT systems, however, are engineered systems, so scalable modeling may be achievable. Current technologies are well suited to modeling individual components, but work is needed to compose models, such as embedding a detailed model of a component (e.g., a single computer) within an enterprise model. The difference in time scales associated with different behaviors (e.g., differing

response times from electronic components, inter-network communications, and human action) increases the difficulty of composing multiple models into models of aggregate system behavior.

Human understanding of complex systems behavior is evolving. A multidisciplinary research community is currently focusing its attention on complex data networks, with some success. For example, highly optimized tolerance theory is being used to analyze the aggregate behavior of large collections of IT components. But the mathematics for describing IT systems is still being developed. Research in this area has a stochastic flavor and has yet to broadly address mobility and long-range dependency, as well as nonlinearity and non-stationarity issues.

6.3 Internet Modeling, Simulation, and Visualization

Definition

Internet modeling, simulation, and visualization enable researchers, analysts, and decision makers to model the Internet or a portion of it; simulate a mix of traffic and measure key parameters such as congestion, queuing delays, network transit time, packet loss, and jitter; and visualize network performance graphically (though not necessarily in real time).

Importance

The Internet comprises autonomous networks that are connected via network links, across which communications take place using the Border Gateway Protocol. As the Internet grows to carry more traffic that is sensitive to latency, jitter, and packet loss (e.g., voice over IP or streaming video), modeling of Internet behavior will become increasingly helpful for ensuring that critical traffic is delivered in a timely manner. For example, understanding how a router drops packets when network traffic is normal is essential to knowing how that router will behave when the network is congested. Advanced graphical tools will enable analysts to better investigate congestion under present conditions and also as the size and complexity of the Internet grow.

The development of a comprehensive Internet modeling and visualization capability will allow analysts to evaluate the effect of new protocols (such as those that handle larger packet sizes); develop and test techniques to thwart distributed network attacks such as distributed denial of service, physical disruption, and worm propagation; estimate performance of a network design under various types of load and operational conditions; study how congestion affects time-sensitive traffic; and facilitate the configuration of networks for increased robustness.

State of the Art

Mathematical modeling techniques such as queuing theory are being used in developing Internet performance models. These techniques can be readily combined with commercial visualization software to graphically portray network performance. They have been successfully run on IP networks with 1,000 nodes; the next step is the extension to a network of Internet scale and diversity. Simulations can be used to calibrate and validate models and to demonstrate their accuracy, but because they have long run times, they are not practical for many types of network performance analysis and prediction. Simulation tools also are needed for certain types of problems for which analytical capabilities are still under development. Analytical tools can be used to study private IP network performance under baseline conditions and can predict performance under several types of network attack, such as DDoS and worm attacks.

Capability Gaps

Scaling these capabilities to Internet-size networks is a research problem. Today's technologies enable only partial or targeted models of facets of Internet behaviors. A variety of capability gaps currently exist, including a need for analytical tools to address Internet-scale complexity problems such as: the nature of distribution functions for many quantities of interest that makes closed-form queuing equations unusable; estimating performance of Internet-size networks; difficulties with heavy-tailed distributions of network traffic parameters; analytical modeling of

network traffic types and protocols; modeling of various classes of distributed cyber attacks; the lack of universally accepted topology, traffic, and protocol data associated with the Internet; and software integration of analytical and visualization tools.

6.4 Network Mapping

Definition

By design, a network map shows physical and/or logical network configurations, including information about hierarchies of connectivity (e.g., a main network and its subnetworks). Network maps are constructed using information either supplied by a human or obtained through automated discovery. Topological or logical network information can be displayed alone or superimposed on a geographical map. Network mapping can also include the automated processing and analysis of network traffic and topology data for presentation using network visualization tools.

Importance

Network mapping is essential to network engineering, monitoring, maintenance, and repair. Network size and complexity require that analysts be able to quickly visualize congestion points using near-real-time traffic data. By graphically monitoring large networks and detecting when link utilization is growing faster than expected, it may be possible to quickly determine if the network is experiencing a physical disruption or a cyber attack. Network mapping is also used in research on network performance.

State of the Art

Several network mapping systems have implemented, to some degree, many of the features of interest to users, including: drawing detailed maps from user data; automated discovery or scanning of a range of IP addresses; drawing of subnet maps using standard computer-aided design drawing capabilities; automated scheduled network sweeps that automatically update a map; showing traffic on maps; listing device IP addresses; and user-initiated vulnerability scanning on networked devices. These capabilities, however, are available for the most part at the level of local networks, not at the level of large-scale networks or across autonomous systems.

Capability Gaps

Current network mapping capabilities require time to map and display small sections of networks using limited data. For real-time and near-real-time mapping of traffic flows and congestion and to support automated network management and visualization of networks across domains, these capabilities need to be expanded to enable rapid mapping of much larger sections of the network. Network traffic data need to be parsed, scrubbed, and formatted before being accepted by the mapping software. Greater flexibility is needed for importing network data from database applications.

Other mapping capabilities needed include: improved mapping speeds; the ability to map larger networks; and the ability to map networks that are “hidden” behind network address translation devices. Additional needs include: real-time or near-real-time mapping of network traffic flows and congestion; the ability to more easily incorporate network topology changes into existing maps; rapid identification of attacks or other network problems; and automated analysis of network topology and recommendations of configuration changes to improve performance.

6.5 Red Teaming

Definition

Red teaming is a technique for analyzing IT system vulnerabilities by actually putting a system under attack. In a red team exercise, skilled outside experts plan and carry out surprise adversarial cyber attacks on an enterprise's systems to find and exploit vulnerabilities and reveal flaws in security planning, policies, and defenses. Unlike role playing or tabletop exercises, the "hostile adversaries" in red teaming exercises make every effort to outthink defenders and "win" by overcoming real cyber defenses and gaining access to actual systems, networks, and information. The attack phase of the exercise is followed by a thorough analysis of what transpired. Red teaming can be combined with or used by other types of assessment such as risk, vulnerability, threat, consequence, system management, system security, accreditation, and certification.

Importance

An effective red teaming exercise should challenge security assumptions and strategies, expose operational and technical weaknesses, and stimulate fresh thinking about an enterprise's security posture. Red teaming has been applied for varied purposes, including: testing cyber defenses and response plans; improving the design and implementation of a system and its security throughout its life cycle; system calibration; generating likely adversary actions to obtain signatures and test detection capabilities; technical analysis of adversarial scenarios; observing the effects of various decisions and prioritizations on an adversary's response; demonstrating a scenario involving real systems and operational constraints; and training.

Red teaming can be an effective tool for IT system engineering or for evaluating the security of complex systems through an increased understanding of component and system function and behavior. Red teaming can encompass globally distributed systems, numerous distributed organizations, a range of technologies, and the effects of interdependencies among systems. While Federal and private-sector red teaming activities may take place independently in order to address their respective needs, the Federal government can facilitate cooperation to assess interdependencies and improve red teaming capabilities and effectiveness.

State of the Art

Red teaming is useful for identifying technical system vulnerabilities and managerial oversights. In industry it may be used to assess the security of high-consequence targets such as those in a banking or financial infrastructure. However, much information about red-teaming methods has not yet been documented. Dedicated red teams often do not share their knowledge with other teams, and temporary red teams rarely have the resources to capture their own knowledge for re-use. There is no easy way to measure a red team's capability and performance to determine its effectiveness.

Federal and industry needs for skilled red teaming exceed the capacity of available resources. Derivatives of red teaming, such as structured self-assessments, may be used to address some issues with fewer resources. However, such an approach is insufficient for the complexity, scope, and scale of IT infrastructure security issues.

Capability Gaps

The September 2003 *Final Report of the Defense Science Board Task Force on The Role and Status of DoD Red Teaming Activities* recommended that DoD red teaming be expanded to deepen understanding of U.S. adversaries in the war on terrorism, and in particular their capabilities and potential responses to U.S. initiatives. The report also recommended developing best-practice guidelines and multiple forms of red teams.

Technical capability gaps identified in the report include:

- Security assessment metrics: A comprehensive set of security metrics for red teaming and assessments, including domain-specific metrics
- Red teaming tools: Specific procedures as well as hardware and software tools to support red teaming, enabling more efficient, consistent,

measurable, and reproducible results. Needs include tools that: 1) assist and improve red teaming processes: 2) can be used for particular technical domains (e.g., network discovery); and 3) enable analysis of red teaming activities.

- Adversarial modeling: The ability to model the behavior of particular adversary classes, groups, or individuals to make the red teaming process more accurate and realistic, particularly when fine resolution is needed. Alternative approaches for unconstrained and innovative adversarial tactics also need to be developed.
- Techniques and methods: Techniques and methods are needed to generate efficient, measurable, and reproducible results; to ensure accurate composition of results from different red teams over space and time; and to red-team numerous widely distributed and complex systems.
- Qualification and certification: The capabilities and qualifications of red teams need to be better understood so that they can be more effectively selected and used for particular tasks.

7. Advanced and Next-Generation Systems and Architectures

The topics in this category focus on methods, technologies, and architectures that will enable the creation of new generations of IT infrastructure components and systems that are designed and built to be inherently more secure than those in use today. The topics in this category are:

- Trusted computing base architectures
- Inherently secure, high-assurance, and provably secure systems and architectures
- Composable and scalable secure systems
- Autonomic (self-managing) systems
- Architectures for next-generation Internet infrastructure
- Quantum cryptography

7.1 Trusted Computing Base Architectures

Definition

The trusted computing base (TCB) is the set of all system hardware, firmware, and software that is relied upon to enforce the system's security policy. The ability of a TCB to correctly enforce a security policy depends on the mechanisms within the TCB and on the correct input by system administrative personnel of parameters related to the security policy.

A TCB architecture is a description of the interrelationships among the hardware, firmware, and software that, in combination, enforce the desired security policies for the system. In principle, a TCB architecture enables analysis to determine if certain security properties hold, and it allows continuous monitoring and verification of the integrity and properties of the TCB (including the kernel, configuration files, secure memory, privileged applications, and running applications).

Importance

The TCB is critical to the secure operation of an IT system. If the security of any component of the TCB is compromised, then the security of the entire computing system is suspect and cannot be assured.

State of the Art

The TCB kernel must interact with many processes and applications, both locally and over complex networks. Increasing system code complexity makes analysis of components of the TCB as well as interactions with untrusted components increasingly difficult. For all but the simplest of computational components and systems, it can be impractical or impossible to determine whether the TCB operates as desired and enforces all desired system security policies at all times. It is equally difficult to analyze a TCB architecture to ensure that it provides the security functionalities that are desired of a system.

Capability Gaps

Currently, it is not known how to effectively test and validate the properties of a TCB architecture, how to ensure that the TCB is fault-tolerant for failures in portions of the TCB, how to ensure that the TCB degrades gracefully under adverse circumstances, or how a TCB can best be defined and assembled for future computing architectures. Although a TCB often consists of system components provided by multiple vendors, generic, broadly accepted methods for architecting and assembling a TCB in such situations do not exist.

Small high-assurance kernels and partitioning kernels are often rejected by developers due to a variety of factors (e.g., restricted functionality), but there is a lack of robust research into techniques for practical hardening of traditional commercial-off-the-shelf kernels. Research is needed to develop monolithic and distributed TCB architecture concepts as well as methods for assuring their security properties. Processes and tools are needed throughout the design, development, and deployment of computational systems to support the verification of TCB properties and the assurance that TCB properties are not compromised.

Additional work is needed to help developers identify the minimal set of components and functions needed for TCBs (either generally or for use in a given application), while providing methods for mapping desired security properties into architectural specifications. While TCBs are intended to be assured so that they can safely interact with untrusted processes, applications, and systems, architecting systems that provide secure sharing and interoperation between trusted systems (e.g., multiple independent TCBs) remains a research area. R&D is also needed in such emerging approaches as techniques to move trust from one part of a system to another and employing strong isolation (e.g., virtual machine management – see section 5.2, page 70).

7.2 Inherently Secure, High-Assurance, and Provably Secure Systems and Architectures

Definition

An inherently secure system fully integrates cyber security mechanisms with system functionality and other properties. This may involve tailoring security mechanisms to system characteristics to achieve both security and functionality. For example, partitioning can be used to both prevent corruption of data and enforce separation of access for processes that are not allowed to share data. Critical systems whose failure or corruption would cause severe consequences require high-assurance design and development. High-assurance systems should be subjected to rigorous design and implementation checking, and control steps should go beyond routine processes used in developing software. A provably secure system is one whose security mechanisms can be proven not to fail in certain modes that would allow inappropriate access or modification of data, or would otherwise disrupt the integrity of system function and data.

Importance

Critical infrastructures increasingly integrate information using hardware and software that interoperate over the Internet and depend on the IT infrastructure. Vast amounts of information are collected and shared within government and throughout the private sector using interdependent physical and IT infrastructure. Critical distributed information resources and Web services that support operations must be protected against inappropriate access and malicious attack. These resources include Federal and military systems, critical components of the banking and finance sector, agriculture, transportation, and national disease tracking and health care delivery systems. Control infrastructures (e.g., for aviation, power, and water) must operate correctly in a multi-system context and must not permit disruption by malicious attack. Disruptions of critical components of the IT infrastructure (e.g., air transport or finance) may affect citizen safety and consumer confidence, causing national economic ramifications. Thus, the ability to detect vulnerabilities and assure that systems operate as intended is vital.

State of the Art

Provably secure systems have long been sought, though some question whether the goal is attainable. But since R&D in provably secure systems enhances understanding of security properties, advances toward this objective can nevertheless serve as a foundation for improving security technologies.

IT system validation and verification are largely based on evaluation of the development process. For decades, Federal agencies have supported a product evaluation (and re-evaluation) program for critical components of the trusted computing base used in military systems. The Common Criteria (CC) standards provide an international classification system for security functionality and a process for evaluating security products. Product evaluation and re-evaluation remains a largely manual process using process-based functionality checklists. The CC framework relies on informal description and lacks formal semantics and models for capability evaluation.

Commercial tools audit systems for vulnerabilities and perform limited analyses. However, the technology base for designing and building inherently secure and assured or verified systems remains weak. Trusted

computing base architectures for managing information domains that must remain separate have been proposed.

Improvements in analysis and model-checking tools enable evaluation of increasing numbers of lines of code for simple properties such as stack or buffer overflow and storage leaks, but precision remains a problem in these tools. The private sector now routinely applies "lightweight" verification technologies such as model checking to restricted classes of components such as device drivers.

Verification technologies have been applied selectively to check critical properties (e.g., isolation of virtual channels used for managing information flows). Research progress is seen in correct-by-construction methods for specific problems, such as programmable smart cards and encryption protocols, supported by security-domain-specific programming technologies.

Capability Gaps

Research in this area is addressing the information, testing, and verification technologies required to develop inherently more secure systems. Processes and tools are needed throughout the design, development, and deployment cycles of products and systems to support evaluation, acceptance, and (in some cases) certification of critical and non-critical IT infrastructures. Research is needed to extend and advance the maturity of high-confidence design and assurance capabilities. There is a need for formal methods to be applied to exemplary systems, such as separation kernels. Development of secure and highly assured systems requires evidence-based methods that can improve today's process-oriented and largely after-the-fact evaluation. Monolithic automated theorem-proving tools, usable only by experts, must be replaced by technologies that enable skilled developers to provide specified levels of assurance. Robust research prototypes are needed to allow experimentation with new security concepts.

Examples of R&D needs include:

Security concepts: Including security models and techniques for systems of varying types and scales (e.g., enterprise and Web services, real-time embedded control, widely distributed sensor and actuator networks); models of risk and trust that address realistic and changing (e.g., dynamic and configurable) system needs; formal models of the properties of critical systems (e.g., security, fault tolerance, real-time response); design technologies to assess properties and manage their interactions; and novel security concepts that can be integrated into emerging information technologies.

System-level security: Including assurance services (e.g., reconfiguration and repair) for autonomic or self-managing systems; capabilities to certify properties of system-level components, with assumptions about the contexts in which these properties hold; capabilities to check end-to-end and emergent properties for composite and cooperating systems; and assurance in the presence of combined cyber and physical system behavior.

Design and analysis: Including environment- or context-aware design and assurance technologies; software composition and analysis technologies; composition tools that can check cross-layer properties; processes that integrate development and evidence-based assurance; and verification technologies that are automated, open, and interoperable and that can be tailored for key properties and domains.

Case studies, experimental testbeds, and prototypes: Including reference implementations of both systems and assurance technologies, together with their design evidence, assurance technology prototypes, and evaluation case studies.

7.3 Composable and Scalable Secure Systems

Definition

Composable secure systems are assembled from components and subsystems in ways that preserve the security properties of those constituent elements while satisfying system-wide and end-to-end security requirements. Systems can be composed by applying rigorous engineering measures to assure system properties, or alternatively, the systems can be federated in an ad hoc manner while ensuring that they interoperate securely, reliably, and sustainably. A highly scalable, secure critical infrastructure should be able to accommodate variation in system parameters (e.g., number of records, users, nodes, or clients, as well as the degree of geographic distribution, heterogeneity, or complexity) without failure of security or system reliability.

Importance

Composability and scalability are interlocking issues not only in cyber security and information assurance but more broadly in architectures for increasingly complex IT systems and networks that must nonetheless be far more robust and reliable than today's. An underlying concept is modularity: if secure and demonstrably reliable components can be engineered to be fitted together reliably and without loss of security, the composed system that results can continue to be predictably scaled up through the addition of other secure, reliable components. While intuitively straightforward, composability and scalability are among the most difficult technical challenges in the development of IT systems because they require rethinking and re-integration of all elements of hardware and software engineering, including first principles, requirements, architecture, and development methods and practices. However, composability and scalability are deemed essential approaches for increasing the overall security and trustworthiness of the IT infrastructure.

State of the Art

At present, understanding of security properties and system composition methods is such that it is not possible to assure that a system composed of assured components will preserve desired security properties at the system level. Similarly, it is not possible to know a priori that methods to assure security at one level of scale will function effectively and robustly at greater scales. Some system architectures in use today are based on design assumptions and practices that are decades old. Many existing designs are not adequately robust against high failure rates or correlated failure modes, such as might occur under coordinated attacks. While research in some aspects of composability and scalability addresses increased security (e.g., trusted computing base, high-assurance system architectures, secure software engineering), few efforts to date focus on drawing together all the design and engineering principles, technical approaches, and component technologies that need to be integrated to achieve composable and scalable secure systems.

Capability Gaps

R&D, including long-term research, is needed to develop new and transitional system architectures, software engineering methods, techniques and tools, programming languages, and evaluation and assessment criteria that can enable the development of composable, scalable systems that are more secure and more robust than those of today. Areas in which R&D should focus include:

New frameworks and architectures: Including trusted computing base architectures to enable controlled sharing, interoperation, and coordination; trust management architectures for open, interoperable systems; computing and systems software architectures that integrate scalable and interoperable security measures (e.g., secure computing platforms, kernel-level, OS, and middleware services); architectures for authenticated human-computer, component-to-component, and system-to-system interaction, with tailored, easy-to-use security interfaces; coordination and interoperation platforms; reference architectures and implementations, and experimental platforms.

Secure, composable, and scalable IT system technologies and development methodologies: Including guaranteed and secure cooperative global-scale dynamic resource management; distributed open

systems and middleware technologies that enable robust, dynamically reconfigurable system operation; and system composition, integration, and management technologies for analyzing components and assembling complex systems.

Composable and scalable cyber security technologies: Including core security service architectures for systems at varied scales; security models and architectures for heterogeneous systems; limitation of misuse and damage by insiders; security status notification and coordination architectures; and scalable, assured system security services (e.g., cryptographic management, authentication, and access control).

7.4 Autonomic Systems

Definition

Autonomic or self-managing (i.e., predictive, self-aware, self-diagnosing, self-configuring, self-optimizing, and self-healing) systems increase the robustness, reliability, resilience, performance, and security of complex IT systems by: 1) exploiting expert knowledge and reasoning; 2) reacting quickly and automatically to events of interest; and 3) minimizing the need for human intervention and the adverse consequences associated with the response time needed for human intervention. An adaptive, autonomic system should ensure that services remain trustworthy and meet required service levels, even across multiple application domains within a system.

Importance

Both the private sector and the Federal government (including the military with its network-centric warfare requirements) need robust systems that respond automatically and dynamically to accidental and deliberate faults, while maintaining required levels of trustworthy service. Capabilities such as fault tolerance have made computing and information systems more resilient during failures due to accidental faults and errors. However, even with these advancements, a system can exhaust all resources in the face of prolonged failures or deliberate, sustained attacks. In addition, systems that rely on manual intervention can become more fragile and susceptible to accidental faults and errors over time if manual maintenance and updating are not performed in a regular and timely fashion. Adaptive, autonomic technologies can address this brittleness by automating activities that would otherwise require human intervention for responding to effects of problems such as imperfect software, human error, accidental hardware faults, or cyber attacks.

State of the Art

Most complex systems are installed, configured, initialized, integrated, optimized, repaired, and protected by human operators. This process is challenging, time-consuming, and error-prone, even for experts. Autonomic systems can facilitate self-configuration by automatically converting policies based on business objectives into configuration settings on system components. Within an autonomic system, a new component can incorporate itself seamlessly while the rest of the system adapts to its presence. Autonomic systems can facilitate self-optimization by continually monitoring system operation to identify parameter adjustments that can improve performance.

Failures in complex computing systems can in some instances take experts weeks to diagnose and repair, while autonomic systems facilitate self-healing by rapidly detecting, diagnosing, and repairing localized problems resulting from bugs or other hardware or software failures. Ongoing research is developing predictive diagnostic techniques that can avoid failures and mitigate problems before they impede system operation.

The current state of the art in quality of service (QoS) is communication-centric and does not consider QoS in a more generic sense, to cover all of the necessary computing elements (processing, data management, and communication) at the component, application, and system level. Many existing systems either do not adapt to changes or have ad hoc hardwired mechanisms that accommodate only a small, predefined set of changes.

Autonomic systems broadly include fault-tolerant systems, intrusion-tolerant systems, and autonomic computing. While traditional fault-tolerant systems

generally focus on addressing accidental faults and errors, intrusion-tolerant systems address intentional faults caused by a malicious adversary. Autonomic computing uses automated management techniques to install software and patches, or to otherwise respond or adapt to changes in the computing environment such as failure-induced outages, changes in load characteristics, and addition of server capacity. Autonomic computing systems may not effectively respond to failures or changes in operating conditions due to malicious attacks without being deliberately designed to do so.

Capability Gaps

Today, human operators decide how to protect systems from inadvertent cascading failures and malicious attacks, which can occur even when some classes of cyber security tools are in place. Autonomic systems facilitate self-protection by: 1) monitoring systems and automatically improving system defenses (i.e., improving protective measures, configurations of IT, and cyber security systems); 2) using sensor reports to anticipate problems before they occur or identify them as they occur, and taking steps to avoid them or reduce their consequences; and 3) identifying emerging problems arising from failures or attacks that are not corrected by self-healing measures, and responding to mitigate their impact. For example, an autonomic security system might provide automated security patch identification and deployment, or might automatically correlate security information across an enterprise to facilitate management of distributed protective measures.

Autonomic system technologies are needed to better protect systems and continually improve reliability, respond to ongoing failures and attacks, isolate compromised portions of IT systems, reconfigure networks and resources, and reconstitute systems to recover from attacks. These systems must protect against accidental faults introduced by human and software errors as well as against misuse and malicious internal and external attacks.

Autonomic systems R&D and integration of key technologies are needed in several research areas, including information models, machine reasoning and learning, and autonomic platforms and infrastructure. Autonomic systems require design and implementation of models that completely and accurately capture information about the state of a system and its components. Predictive models of autonomic systems are needed in order to help guarantee desired behavior. New languages may be needed to capture the data and semantics in these models, aggregate knowledge, parse and translate knowledge to support reasoning, and enable heterogeneous components to infer and share knowledge. Platforms and infrastructures for retrieving and manipulating models for distributed heterogeneous components need to be designed, developed, and validated through simulation. This will require the development of languages and APIs for querying components, and a communication infrastructure for facilitating interaction between distributed components.

Such advances will enable greater scalability, interoperability, security, scope, and robustness of IT systems, while reducing costs associated with operating, maintaining, and protecting them. Once this infrastructure exists, domain-specific applications will be needed to capitalize on the benefits of autonomic systems. Different domains may require the establishment of different policies to allow inferences about the operational state of a system across distributed heterogeneous components. Each application needs mechanisms for automated learning about the system, automated reasoning using knowledge about the system, establishing rules based on policies, propagating new or revised policies, and manipulating components.

Approaches to adaptive, autonomic systems include R&D in a variety of research areas. For example, biologically inspired cognitive response strategies can be developed to use machine learning and proactive contingency planning to automate cyber-based analogs to immune response and system regeneration. Work aimed at reducing the time required to achieve consistency among replicated information stores after an update can help increase the effectiveness of redundancy techniques to the point where they can enable self-healing when bodies of information are damaged or compromised. This work, being applied

to centralized servers and distributed publish-and-subscribe settings, can include reasoning about the insider threat to preempt insider attacks; detecting system overrun by inferring user goals and intent; enabling anomaly detection; and combining and correlating information such as from system layers and direct user challenges.

Research is also needed on extended and refined end-to-end QoS models. Such models must provide a quantitative basis for efficient and effective resource management for adaptive, autonomic systems, enabling them to respond to changes due to overload, component failure, malicious attacks, evolving operational requirements, and/or a dynamic operational environment.

End users need the ability to define policies based on application-specific QoS metrics to control system operation in order to apply resources in the most appropriate manner. Research needs include: developing a policy specification language to capture application-specific QoS requirements; mapping high-level application QoS specifications into lower-level system metrics; predictable QoS-aware components for processing, data management, and communication; and composability for end-to-end assurance.

7.5 Architectures for Next-Generation Internet Infrastructure

Definition

The next-generation Internet infrastructure will comprise the protocols, hardware, and software for secure, reliable communication across loose federations of networks. Its design and uses will be driven by a new generation of applications. The full range of applications is impossible to predict and unanticipated forms of network usage can be expected.

Importance

The Internet has become an essential component of the Nation's critical infrastructures. It inherits traditional requirements associated with supporting these infrastructures, but also needs to address new requirements such as real-time response, reliable and secure communications, and quality of service. Internet vulnerabilities include threats to availability (e.g., denial-of-service attacks), unauthorized access to and use of information (e.g., compromised privacy or integrity of personal or corporate records and theft of intellectual property), and potential disruption of operations for essential government and public services (e.g., financial systems, transportation, and power, water, and food supply and distribution). New architectures are essential to assured, secure operation of critical and non-critical Internet-based systems and services.

State of the Art

The Internet was designed for transparency, scalability, and redundancy to enable resilience and survivability. Packet-switched services underpin end-to-end delivery of information and communications. The early adoption of the simple, robust Transmission Control Protocol/Internet Protocol (TCP/IP) suite enabled diverse implementations and services such as the addition of virtual circuits via sockets. Performance assessment focused on statistical characteristics of aggregate traffic through the network, such as average message latency. The strong separation between network and application provided by the TCP/IP architecture enabled a variety of new networking media (e.g., wireless and optical networking) and new classes of applications (e.g., streaming video, IP telephony) to be added without major changes to the underlying architecture.

Internet enhancements have included augmentation of host-based physical addressing to include dynamic IP address assignment and virtual IP addressing. Virtual Private Network (VPN) architectures use the Border Gateway Protocol to set up tunneled networks that have separate address spaces. VPNs permit groups of trusted participants to exchange information securely, supported by encryption and key management that assures isolation for each VPN, while operating over the Internet fabric.

Most Internet services depend on a core infrastructure of backbone networks and routers, with local area networks and nodes at the edge delivering services to

end users. Existing networking protocols continue to be improved and new ones are developed for specialized applications. For example, Internet Protocol version 6 (IPv6) is a variant of IP that offers an expanded address space, and protocols have been developed for mobile ad hoc networks. Specialized controller area networks for real-time control are being developed using time-triggered physical layers and time-triggered architectures.

In some cases, as with IPv6, enhancements to protocols include incorporation of security mechanisms. In other instances, new protocols are developed to provide security, as with the Wired Equivalent Privacy protocol that is intended to improve the security of certain classes of wireless communications. Security extensions to the Domain Name System are starting to transition into operational use; efforts aimed at developing security-based improvements to routing protocols are underway; and security is a key requirement in various areas of networking research such as wireless networking and optical networking.

Capability Gaps

Development of architectures for next-generation Internet infrastructures will require research in a wide variety of technical areas. These include: improved and/or new protocols that include enhancements for security and authentication; optical networking including optical circuit-based networking, optical switching, and optical computing; network management and control (e.g., virtual control plane) technologies; new networking services for naming, addressing, and identity management; scalable, robust technologies to meet high-assurance, high-reliability, real-time computing needs; testbeds to support architecture research, development, analysis, testing, and evaluation; and new applications that take advantage of next-generation Internet architectures.

Research will also be needed on deployment issues, including development of technology transition and migration paths – taking compatibility issues into account – from current to next-generation infrastructure and services, interoperability of heterogeneous (IP-based, optical, and wireless) systems, scalability issues, and enhanced understanding of business cases and economics of commercial deployment.

7.6 Quantum Cryptography

Definition

Quantum cryptography is a set of methods for implementing cryptographic functions using the properties of quantum mechanics. These methods are based on quantum mechanics, but they need not, and currently do not, make use of quantum computing. Most quantum cryptography research is directed toward generating a shared key between two parties, a process known as quantum key distribution (QKD). Shared keys may be used directly as keys for a conventional symmetric cryptographic algorithm or as a one-time pad. A variety of protocols have been developed for QKD, but they generally share two basic features: 1) the idealized version of the protocol prevents an eavesdropper from obtaining enough information to intercept or decode messages (e.g., messages are encoded by using the shared key as a one-time pad); and 2) the communicating parties can detect the presence of an eavesdropper because eavesdropping will alter the quantum properties of the particles used in key distribution in a measurable way.

Importance

Quantum cryptography offers the potential for stronger information assurance, but QKD must be designed and implemented properly to deliver promised benefits. QKD systems may be subject to a number of attacks, depending on the implementation and the protocol, and as is always the case, even the strongest of cryptographic methods are vulnerable to flaws in design and implementation.

State of the Art

Quantum cryptographic products have been offered since 1999, with research ongoing to advance the state of the art. The most common type of quantum key distribution uses a scheme known as BB84 in which polarized photons are sent between the communicating parties and used to develop the shared key. The BB84 protocol has been shown to be secure for implementations that preserve assumptions about physical properties of the system. Many

varieties of the BB84 scheme have been developed, and other forms of quantum key distribution have been proposed.

Rapid progress has led to products capable of key distribution through many kilometers of fiber-optic cable. Additional products include quantum random number generators, single photon detectors, and photon sources. However, vulnerabilities may be introduced in the physical systems, quantum protocols, application software, and operating systems used to process keys. Existing QKD systems are not able to guarantee the production and receipt of a single photon per time slice, as required by most quantum protocols. Multiple photons emitted in a single time slice may allow an attacker to obtain information on the shared key.

Capability Gaps

Existing quantum cryptographic protocols may also have weaknesses. Although BB84 is generally regarded as secure, researchers frequently introduce new protocols that differ radically from the BB84 scheme, and a number of these protocols are vulnerable to attack. Quantum cryptographic equipment must be integrated with an organization's network, potentially leaving the QKD system and its software open to conventional network attacks. Methods for evaluating and certifying QKD systems have not yet been incorporated into existing security evaluation capabilities.

8. Social Dimensions of Cyber Security and Information Assurance

Topics in this R&D category address the impacts of cyber security on people, organizations, and society, and the implications of cyber security for law, policy, and social systems. Topics in this category are:

- Trust in the Internet
- Privacy

8.1 Trust in the Internet

Definition

While the focus of CSIA R&D is necessarily on technical advances that improve cyber security and information assurance, such technical activities take place within a broader cultural context: the overall level of public confidence, or trust, in the Internet and the varied transactions and processes it makes possible. Public trust in the Internet can be defined as the degree to which individuals and organizations feel comfortable that their sensitive information will be handled securely, their privacy will be maintained, and their systems will be free from intrusion in any interactions and transactions over the Internet.

Importance

In the years ahead, economic innovation – including development of novel applications exploiting high-bandwidth connectivity – will depend heavily on a steady strengthening of the public's trust in online transactions. Current data indicate, however, that as Internet use increases, so do the levels of electronic crime and malicious attacks that users experience. Consumers worry about identity theft, theft of credit-card information, and other fraudulent activities (e.g., through phishing, spyware, keylogging). Studies have found tens of millions of different spyware products in use and the number of phishing attacks has been known to grow by double-digit rates from one month to the next.

State of the Art

Media accounts of cyber misbehavior and crime have had a positive impact, in that they have raised public awareness of security issues and useful protective measures and have spurred vendors of software and Internet services to improve their products. Practices to build public confidence now include password-protected Web sites, anti-virus software, firewalls, trustmarks (seals of approval from trusted parties posted on sites and products), certifications, digital signatures, and mechanisms for customer service and complaints. In addition, government, academic, and private-sector organizations have begun to more systematically collect and analyze data about Internet crime and transgressions of trust. This will clarify global trends and enable more informed decision making by regulators, hardware and software developers and vendors, and consumers.

Capability Gaps

Improved user awareness and sustained enforcement of cyber crime and IT-related consumer protection laws help maintain trust in the Internet. However, the development of new technologies can also contribute to maintaining this trust, such as by deterring or preventing online activities that lead to loss of trust in the Internet.

The financial services community has a long history of developing fraud-detection technologies based on pattern recognition and classification of purchasing behavior. Online fraud detection is a less mature area that requires additional R&D. Better detection of technologies used by criminals, such as keystroke loggers used to steal credit card, banking, or other sensitive information, is also needed.

Much malicious activity has its roots in social engineering, such as a Web site designed to trick individuals into installing a malicious piece of software on their computers or an e-mail designed to trick recipients into divulging sensitive or personal information. Better technologies for detecting social engineering attacks can help reduce the exposure of users to various classes of attacks that ultimately compromise trust in the Internet. In addition, public-

and private-sector studies and analyses of malicious and criminal cyber techniques, trends, and costs should be improved and refined to provide better guidance for decision making by policy makers and consumers alike.

8.2 Privacy

Definition

Privacy can be defined as freedom from surveillance, intrusion, and unauthorized access to information. In the context of the IT infrastructure, privacy can be defined as the expectation that personal and business information, communications, transactions, or other computer-mediated activities will not be viewed, intercepted, recorded, manipulated, or used without one's knowledge or permission, except in certain well understood circumstances (e.g., information legitimately subpoenaed for law enforcement purposes).

Tensions can arise between expectations of preservation of privacy and the use of authentication and need for monitoring and analysis of information and communications to enhance cyber security. Effective cyber security must not only assure information and communication privacy but also, more broadly, protect the physical functionality of IT systems to process, store, and make available information with assured confidentiality and integrity. Good information management and system design should support both privacy and security; indeed, true protection of privacy is not possible without good security.

Most cyber security involves the use of technology to protect data or systems from improper access or alteration. Privacy issues include data access and accuracy but can extend further, to such questions as how data are used and whether citizens are informed about the collection and use of their personal data, as well as about their ability to correct inaccurate data. These are often policy issues that must be addressed using mechanisms that may or may not include technology.

Importance

Addressing privacy issues is critical to ensuring public trust in the integrity of the IT infrastructure. The Administration stated in its 2003 *National Strategy to Secure Cyberspace* that protecting privacy and civil liberties is a guiding principle of cyber security and that cyber security programs must strengthen, not weaken, such protections. The report stated that the Federal government should lead by example in implementing strong privacy policies and practices in cyberspace.

State of the Art

The scope of security in cyberspace includes the internal networks of an organization and its connections with the public, the private sector, and government entities. Each sector is subject to its own privacy laws, regulations, and/or industry practices. Given this interconnected environment, cyber security research should incorporate privacy principles that underscore the requirements and practices of both the public and private sectors. Domain-specific examples of lawmaking in the privacy arena include the Health Insurance Portability and Accountability Act, which governs use of personal medical information by third parties; the Federal Trade Commission (FTC) Act, which gives the FTC power to enforce companies' privacy promises about how they collect, use, and secure consumers' personal information; and the Financial Modernization Act of 1999 (also known as the Gramm-Leach-Bliley Act), which requires the administrative, technical, and physical safeguarding of personal information by businesses.

Although this is an evolving area of law, regulation, and institutional policy, it is clear that cyber security technologies have the ability to impact privacy. New technologies resulting from R&D have the potential to continue raising privacy issues requiring resolution; consequently, these issues should be considered as part of developing cyber security technologies.

Capability Gaps

Fully integrating privacy in cyber security R&D will entail development of principles and methods in the following areas:

Integrating privacy in IT system life cycles: Including privacy practices and processes at the earliest stages of R&D helps ensure the efficient deployment of systems that require privacy protections. Privacy risks that are identified at early

stages of development are more easily mitigated with reduced impact on the development effort. Resources and tools can include privacy-impact assessments and privacy audits, which together can establish objectives and evaluate privacy throughout the life cycle of the system and its data.

Privacy principles: The private sector, government, and citizens are each subject to privacy laws, regulations, and/or practices. Cyber security R&D should enable new technologies and their implementation to be consistent with privacy laws and widely accepted privacy principles. Examples include principles for assuring data quality and integrity; limits on data collection, use, disclosure, and retention; openness and accountability; and citizen participation and impact through notifications, accessibility, and avoiding or redressing harm from inaccurate data.

Privacy environments: The technical environments for privacy in cyber security are of two primary types: 1) the maintenance environment, which involves system architecture and the storage and protection of data; and 2) the transaction environment, which concerns how data are shared and exchanged within and across organizations. Cyber security and information assurance R&D should address privacy issues raised by both types of environments.

The maintenance environment provides for the collection and storage of information. Privacy issues concern the information itself – its scope, its accuracy, and how it is used – as well as how it is managed, such as in policies for storage, retention periods, and disposal. Privacy issues in data collection involve data sources, the type and quantity of data stored, and notification. Research questions include how to limit the scope of personal information needed and avoid collection of unnecessary data; methods to improve data accuracy; impacts of inaccurate data; concerns about data sources; and methods for providing notifications when information is collected. In storage, issues include methods for preventing privacy breaches resulting from events that range from lost or stolen computers to deliberate penetration of enterprise systems and data theft. Retention and disposal issues include how to ensure that data are not retained longer than needed, and that data are properly destroyed from all media at the end of the retention period.

The transaction environment provides for information sharing and requires access controls, authentication, and technologies for sharing. Privacy issues relating to access controls include defining the categories of authorized users and the access rights and permissions appropriate for each category. Privacy issues associated with authentication methods include invasiveness, types of data required (e.g., collection and use of biometrics), and whether or how personal information can be linked across data stores or organizations. On the topic of forensics issues, one question is how IP addresses, audit trails, or other techniques can be used to ascertain whether transactions violate privacy policies or laws.

Data sharing issues include how to establish and evaluate technical and non-technical solutions to ensure that: 1) sharing practices are aligned with enterprise policies such as any confidentiality promises made in privacy policies; 2) data will be shared only with approved parties; and 3) data shared with suppliers are subject to appropriate protections.

As information technologies and cyber security technologies evolve, better understanding of how privacy features of these technologies will be assessed and what metrics will be used to measure their effectiveness also needs to be developed.

Appendices


EXECUTIVE OFFICE OF THE PRESIDENT OF THE UNITED STATES
NATIONAL SCIENCE AND TECHNOLOGY COUNCIL

Appendix A

CSIA IWG Agency Roles and Responsibilities

This appendix provides brief descriptions of the missions of the Federal agencies that participate in the Cyber Security and Information Assurance Interagency Working Group (CSIA IWG) as well as summaries of their CSIA R&D activities and interests.

Department of Commerce (DOC) and National Institute of Standards and Technology (NIST)

Building upon the Computer Security Act of 1987 (P.L. 100-35), the Paperwork Reduction Act of 1995 (P.L. 104-13), and the Information Technology Management Reform Act of 1996 (i.e., Clinger-Cohen Act, P.L. 104-106, Division E), the Federal Information Security Management Act of 2002 (FISMA) (P.L. 107-347, Title III) provides the basic statutory requirements for securing Federal computer systems. The FISMA requires each agency to inventory its major computer systems, identify and provide appropriate security protections, and develop, document, and implement an agency-wide information security program.

FISMA, the Cyber Security Research and Development Act of 2002 (P.L. 107-305), and OMB Circular A-130 authorize NIST to conduct research and develop standards and guidelines for use by Federal agencies for securing non-national security systems. NIST carries out this mission principally through technology transfer initiatives and the issuance of NIST Special Publications and Federal Information Processing Standards (FIPSs). NIST conducts research to determine the nature and extent of information security vulnerabilities, to develop techniques for providing cost-effective information security solutions, and to support its standards and guideline programs.

FISMA authorizes the Secretary of Commerce to choose which of these standards and guidelines to promulgate, and it authorizes the director of OMB to oversee the development and implementation of these security policies, principles, standards, and guidelines. FISMA authorizes the OMB director to: require agencies to follow the standards and guidelines developed by NIST and prescribed by the Secretary of Commerce; review agency security programs annually and approve or disapprove them; and take actions authorized by the Clinger-Cohen Act (including budgetary actions) to ensure compliance. These roles and responsibilities assigned to NIST and the Secretary of Commerce do not extend to computer systems identified as national security systems. The director of OMB has the authority to take budgetary actions and report to Congress on similar matters related to national security systems.

In addition to supporting Federal activities under FISMA, the security standards, guidelines, and research results developed by NIST are also frequently used by U.S. and global industries and foreign governments as sources for IT system security policies and methods.

Department of Defense (DoD)

DoD is concerned with protecting the security of all aspects of the IT infrastructure that affect critical military infrastructures, including private-sector infrastructures on which the warfighter relies. The assured exchange of information and communications is a crucial component of military activities supporting the primary mission of DoD.

As outlined in Joint Vision 2020 and the Defense Technology Area Plan, U.S. forces depend on interrelated capabilities that combine command, control, communications, and computers with

intelligence, surveillance, and reconnaissance. All of these capabilities must be supported by an underlying foundation of information assurance to facilitate "decision superiority" on the battlefield. Successful application of these capabilities will enable full-spectrum dominance for U.S. forces in the future.

The DoD Science & Technology (S&T) program advances the S&T base for protecting critical defense infrastructures and develops tools and solutions to eliminate any significant vulnerability to cyber attacks. The program includes thrusts in the areas of analysis and assessment, mission assurance, indications and warning, threats and vulnerabilities, remediation, mitigation response, and reconstitution. The program focuses on DoD requirements for protection that go well beyond what the private sector requires and commercial technologies provide.

The Director for Defense Research and Engineering (DDR&E) is responsible for DoD S&T. The DDR&E is also the Chief Technology Officer for the Secretary of Defense and the Under Secretary of Defense for Acquisition, Technology, and Logistics responsible for scientific and technical matters and technology readiness of major acquisition programs. The three technical offices within the Offices of DDR&E are: the Office of the Deputy Under Secretary of Defense (DUSD) for Laboratories and Basic Sciences (LABS), the DUSD for Science and Technology (S&T), and the DUSD for Advanced Systems and Concepts (AS&C). In addition, DDR&E oversees the Defense Advanced Research Projects Agency (DARPA).

DDR&E oversees the Technology Area Review and Assessment process, which is DoD's mechanism to coordinate and review S&T programs throughout the department. Within DDR&E, DUSD (S&T) is responsible for policy, programmatic, financial, and strategic oversight of the department's applied research and advanced technology development. DUSD (LABS) is responsible for basic research and DoD laboratory management issues, and DUSD (AS&C) is responsible for technology demonstrations and transition programs such as the Advanced Concept Technology Demonstrations.

DoD strives for a balanced R&D program across basic and applied research and advanced development in academia, DoD laboratories, and industry. DoD's internal cyber security and information assurance research programs are concentrated at the National Security Agency (NSA), Army Research Laboratory, Naval Research Laboratory, Air Force Research Laboratory, and Army's Communications and Electronics Research, Development, and Engineering Command. DARPA funds targeted research projects with three- to five-year lifetimes. The Army Research Office, Office of Naval Research, and Air Force Office of Scientific Research fund most of the DoD-sponsored university research. DoD laboratory and sponsored industrial R&D emphasize advanced defensive technologies that DoD requires but are not available in commercial systems.

DDR&E also collaborates and coordinates with the Office of the Assistant Secretary of Defense (Networks and Information Integration) Information Assurance Directorate, which is responsible for policy, oversight, and acquisition in operational information assurance to ensure linkage between operational needs and long-term S&T investments.

Department of Energy (DOE)

DOE is principally a national security agency and all of its missions flow from this core mission. The department provides the scientific foundations, technologies, policies, and institutional leadership necessary to achieve efficiency in energy use, diversity in energy sources, a more productive and competitive economy, improved environmental quality, and a secure national defense.

DOE also works to ensure energy security, maintain the safety, security, and reliability of the nuclear weapons stockpile, clean up the environment from the legacy of the Cold War, and develop innovations in science and technology. Science and technology are the department's principal tools in the pursuit of its national security mission.

DOE's research activities complement and are closely coordinated with the activities of other Federal agencies, including DARPA, EPA, NASA, NIH,

NSA, and NSF. The department also promotes the transfer of the results of its basic research to a broad set of application fields such as advanced materials, national defense, medicine, space science and exploration, and industrial processes. The department has taken a deliberate and integrated approach to its R&D portfolio, using the strengths of all programs to address its central mission. For example, environmental security and economic security underpin national security, and each is sustained by science. Within the department, the Office of Science manages fundamental research programs in basic energy sciences, biological and environmental sciences, and computational science.

The DOE Office of Science is the single largest supporter of basic research in the physical sciences in the United States. It oversees and is the principal Federal funding agency for the Nation's research programs in high-energy physics, nuclear physics, and fusion energy sciences. In addition, the Office of Science manages fundamental research programs in basic energy sciences, biological and environmental research, and advanced scientific computing research. The Office of Science also promotes workforce development by sponsoring programs that support the scientific advancement of students and educators.

Department of Homeland Security (DHS)

DHS is leading the Federal government's unified effort to secure the United States homeland. The department's organizations are focused on a variety of missions associated with preventing and deterring terrorist attacks and protecting against and responding to threats and hazards to the Nation. The responsibilities of the department's Science and Technology (S&T) Directorate include identifying priorities for, establishing, conducting, and coordinating basic and applied research, development, testing, and evaluation (RDT&E) activities that are relevant to all areas of the department mission.

DHS's cyber security R&D portfolio engages in cyber security RDT&E endeavors aimed at securing the Nation's critical infrastructures through coordinated efforts to improve the security of today's IT infrastructure and to provide a foundation for a more secure next-generation IT infrastructure. In this context, the IT infrastructure is considered to include the infrastructure that underlies Internet communications as well as the IT components that are essential to the operations of the Nation's critical infrastructure sectors.

Cyber security RDT&E activities funded by DHS S&T are carried out by a variety of organizations, including the private sector, universities, and national laboratories. The strategic approach taken by DHS cyber security portfolio and program managers emphasizes public-private partnerships and other forms of collaboration. The goal of this approach is to encourage widespread use of more secure IT systems and components through technology transfer and diffusion of Federally funded R&D into commercial products and services.

Department of Justice (DOJ)

DOJ's mission is "to enforce the law and defend the interests of the United States according to the law; to ensure public safety against threats foreign and domestic; to provide Federal leadership in preventing and controlling crime; to seek just punishment for those guilty of unlawful behavior; and to ensure fair and impartial administration of justice for all Americans."

To execute this mission, the department fields more than 110,000 employees in 39 separate component organizations led by the U.S. Attorney General. These include the U.S. Attorneys who prosecute offenders and represent the U.S. government in court; investigative agencies – the Federal Bureau of Investigation, the Drug Enforcement Administration, and the Bureau of Alcohol, Tobacco, Firearms, and Explosives – that deter and investigate crimes and arrest criminal suspects; the U.S. Marshals Service, which protects the Federal judiciary, apprehends fugitives, and detains persons in Federal custody; and the Bureau of Prisons, which confines convicted offenders.

Litigating divisions represent the interests of the American people and enforce Federal criminal and

civil laws, including civil rights, tax, antitrust, environmental, and civil justice statutes. The Office of Justice Programs and the Office of Community-Oriented Policing Services provide leadership and assistance to state, tribal, and local governments. Other major departmental components include the National Drug Intelligence Center, the United States Trustees, the Justice Management Division, the Executive Office for Immigration Review, the Community Relations Service, the Executive Office of the Attorney General, and the Office of the Inspector General. Headquartered in Washington, D.C., the department conducts most of its work in offices located throughout the country and overseas.

Department of State

The Department of State promotes international scientific and technical exchanges in the service of U.S. ideals and interests. Such international collaboration accelerates the progress of advanced research and creates global communities of like-minded researchers. International cooperation in cyber security research is especially critical, due to the global nature of networks and the ability of cyber attacks to move rapidly across borders.

While the department does not perform or sponsor its own research activities in science and technology, it works with other Federal agencies to facilitate international research collaborations and to coordinate the exchange of information about research activities with other nations. Cooperation often takes place under the auspices of umbrella agreements, negotiated by the department bilaterally, for research cooperation with other nations. Working with other agencies, the State Department coordinates meetings and exchanges between U.S. and foreign government officials overseeing R&D and seeks to develop self-sustaining dialogues among researchers.

The goals of State Department efforts in cyber security are: 1) to foster international collaboration and technical exchange in cyber security R&D; 2) to encourage standardization and the adoption of best practices around the globe; and 3) to enhance national security by facilitating international communication about cyber threats.

The State Department has coordinated exchanges between U.S. cyber security and information assurance research officials and their counterparts in other nations and has facilitated the development of international collaborations in cyber security R&D. Some of these have been stand-alone initiatives, such as the U.S.-India Cyber Security Experts Group. Others have been components of larger collaborative R&D programs for critical infrastructure protection, such as the bilateral programs with Canada, the United Kingdom, and Japan. Federal agency participants have included NIST, NSF, DoD, DHS, and DOE. Recent programs have served to strengthen collaborative cyber security research relationships between the U.S. and Japan, India, Canada, the European Union, and several individual European countries. These activities have taken place in both bilateral and multilateral settings.

The State Department participates as an observer on the INFOSEC Research Council and provides oversight for the Technical Support Working Group.

Department of Transportation (DOT) and Federal Aviation Administration (FAA)

The mission of the FAA is to provide the safest, most efficient aerospace system in the world. In securing the national airspace system, the FAA supports DHS programs in emergency preparedness, crisis management, and continuity of government planning.

The FAA is a member of the Joint Planning and Development Office (JPDO), which is chartered by Congress to develop a vision of the aviation system in the year 2025 and a Next Generation Air Transportation System (NGATS) Implementation Plan. The JPDO includes DHS, DOC, DoD, DOT, NASA, and OSTP. Close partnerships with other Federal agencies on integration of security technologies and management of over-flight programs help ensure continuous operation of the national airspace system.

FAA cyber security and information assurance research activities seek to maximize budget effectiveness and leverage developments by other

agencies. FAA's unique requirements are based on identification of security measures that provide for the safety and security of the FAA workforce, facilities, and critical infrastructure. Cyber-defense concept modeling plays a significant role in improving the security of FAA's information infrastructure. The agency's cyber security goal is mission survivability by achieving zero cyber events that disable or significantly degrade FAA services. The Director of Information Technology Research and Development (Chief Technology Officer) is responsible for developing, managing, and executing FAA's IT and information systems security R&D programs.

FAA cyber security and information assurance research includes such topics as:

- Adaptive quarantine methods to better isolate network damage
- Information detection techniques to help officials anticipate and thwart the plans of potential terrorists
- Topological vulnerability analysis to determine the topology of dependencies among network vulnerabilities and analyze possible attacks against the network
- Development of a time-phased (2005 to 2015) National Airspace System (NAS) Information Systems Security Architecture (ISSA)
- R&D to extend the software-costing model, COnstructive COst MOdel II (COCOMO II), to include security for the full FAA acquisition operation cost for secure systems

Department of the Treasury

Though the Department of the Treasury does not have an R&D budget, through an outreach to financial-sector chief information and chief technology officers, the department has documented the sector's cyber security R&D requirements and identified near-, medium-, and long-term R&D projects to meet the requirements. In addition, the Federal Reserve has agreed to participate in a pilot project to develop a system for tracking the physical diversity of telecommunications circuits.

The suggested projects in Treasury's R&D agenda come from many sources and cover many areas. These projects are designed to identify security approaches that are easily transferable to the private sector and to increase the overall resiliency of the sector. The projects are also designed to encompass: 1) all facets of the critical infrastructure protection life cycle; 2) a wide variety of technology and business practice areas; 3) short- to long-term development; and 4) low- to high-risk research. In addition, Department personnel are continuing their discussions with experts and organizations both inside and outside the financial services sector to identify R&D projects that will help make the sector more resilient against external and internal cyber threats.

Intelligence Community

The Intelligence Community provides assessments, including national intelligence estimates, of foreign threats to the U.S. IT infrastructure. These assessments consider capabilities and intent both currently and looking forward five years and beyond.

National Aeronautics and Space Administration (NASA)

NASA is at the forefront of exploration and discovery as the world's preeminent organization for space and aeronautics R&D. NASA is currently undergoing a transformation to align its core competencies with its space exploration vision.

Given NASA's focus on its core competencies, the agency's cyber security and information security R&D portfolio is limited. Currently, NASA is participating in a multi-agency R&D project that seeks to secure data transmission and IT systems used by the National Airspace System through securing onboard networks, protecting air/ground data links, and providing improved situational awareness of the onboard environment.

NASA's other areas of interest germane to cyber security and information assurance include the maintenance of availability for satellite and spacecraft transmissions, protection of NASA's space-based

assets, and security requirements for post-IP-based space communications.

National Institutes of Health (NIH)

The National Institutes of Health (NIH), a part of the U.S. Department of Health and Human Services, is the primary Federal agency that conducts and supports medical research. Helping to lead the way toward medical discoveries that improve people's health and save lives, NIH scientists investigate ways to prevent disease and to identify the causes, treatments, and even cures for common and rare diseases.

NIH's work in cyber security and information assurance R&D is aimed at supporting the mission of the Institutes, with an emphasis on continuing the development of the security infrastructure to support distributed multi-organization federated data and computation, including fine-grained access control for biomedical information.

National Science Foundation (NSF)

NSF is an independent agency established to promote the progress of science through investment in research and education. The Federal government's only agency dedicated to the support of education and fundamental research in all scientific and engineering disciplines, NSF serves as the primary source of Federal investment in long-range, innovative research. A central goal of this investment is to ensure that the United States maintains leadership in scientific discovery and the development of new technologies. In addition, NSF's mission explicitly includes support for research to advance the national health, prosperity, and welfare, and to secure the national defense.

NSF primarily supports peer-reviewed, long-range, innovative research conducted by academic institutions and not-for-profit research laboratories. This is implemented through a wide range of investments, from grants to individual investigators, to multi-university teams, to large centers. The majority of NSF funding takes the form of research grants and cooperative agreements. When appropriate, NSF cooperates with other Federal and international agencies to enable co-funding of promising peer-reviewed research.

All aspects of IT infrastructure are included in the NSF research portfolio. NSF's cyber security R&D is managed in the Computer and Information Science and Engineering (CISE) Directorate. NSF's Cyber Trust initiative provides the central focus for cyber security research investment. Research related to other aspects of the IT infrastructure is also centered in CISE. This includes research to advance: networking and Internet technologies, communications, computer systems, operating systems and middleware, databases and information management, distributed systems, and embedded sensing and control. The NSF Engineering Directorate supports research in several related areas, including sensor networks and infrastructure risk analysis. In areas of common concern, CISE investment in IT research is closely coordinated with the other NSF science and engineering directorates.

Technical Support Working Group (TSWG)

The TSWG is a multi-agency organization that identifies, prioritizes, and coordinates interagency and international R&D requirements for combating terrorism. Overseen by the Department of State and including representatives from agencies across the Federal government, the TSWG rapidly develops technologies and equipment to meet the high-priority needs of communities engaged in combating terrorism, and addresses joint international operational requirements through cooperative R&D with major allies.

Since 1986, the TSWG has pursued technologies to combat terrorism in the broad context of national security by providing a cohesive interagency forum to define user-based technical requirements spanning Federal agencies. By enlisting the participation of U.S. and foreign industry, academic institutions, and Federal and private laboratories, the TSWG ensures a robust forum for developing technical solutions to the

most pressing counterterrorism requirements. Participants in the 10 functional subgroup areas of the TSWG can come to a single table to articulate specific threats and user-defined approaches to the rapid prototyping and development of counterterrorism devices, training tools, reference materials, software, and other equipment.

The TSWG's program development efforts seek to balance investments across the four main capabilities needed for combating terrorism:

- Antiterrorism – Defensive measures to reduce vulnerability to terrorist acts
- Counterterrorism – Offensive measures to prevent, deter, and respond to terrorism
- Intelligence support – Collection and dissemination of terrorism-related information regarding the entire spectrum of terrorist threats, including the use of chemical, biological, radiological, and nuclear materials or high-yield explosive devices
- Consequence management – Preparation for, and response to, the consequences of a terrorist event

Appendix B

The Networking and Information Technology Research and Development Program

The Networking and Information Technology Research and Development (NITRD) Program is authorized by Congress under the High-Performance Computing (HPC) Act of 1991 (P.L. 102-194) and the Next Generation Internet Research Act of 1998 (P.L. 105-305). The goals of the Program are to:

- Provide research and development foundations for assuring continued U.S. technological leadership in advanced networking, computing systems, software, and associated information technologies
- Provide research and development foundations for meeting the needs of the Federal government for advanced networking, computing systems, software, and associated information technologies
- Accelerate development and deployment of these technologies in order to maintain world leadership in science and engineering; enhance national defense and national and homeland security; improve U.S. productivity and competitiveness and promote long-term economic growth; improve the health of the U.S. citizenry; protect the environment; improve education, training, and lifelong learning; and improve the quality of life

Program Structure

The Cabinet-level National Science and Technology Council (NSTC) is the principal means by which the President coordinates the diverse science and technology programs across the Federal government. The Director of the Office of Science and Technology Policy (OSTP) manages the NSTC for the President. The NITRD Subcommittee, which reports to the NSTC Committee on Technology, coordinates planning, budgeting, and assessment for the NITRD Program. The Subcommittee is composed of representatives from 13 Federal agencies that participate in the formal NITRD budget crosscut, OSTP, the Office of Management and Budget, and the NITRD National Coordination Office. (In the NITRD Program, the term "agency" may also refer to a department, a major departmental subdivision, or a research office, institute, or laboratory.) The member agencies are:

AHRQ – Agency for Healthcare Research and Quality
DARPA – Defense Advanced Research Projects Agency
DHS – Department of Homeland Security
DOE/NNSA – Department of Energy/National Nuclear Security Administration
DOE/SC – Department of Energy/Office of Science
EPA – Environmental Protection Agency
NASA – National Aeronautics and Space Administration
NIH – National Institutes of Health
NIST – National Institute of Standards and Technology
NOAA – National Oceanic and Atmospheric Administration
NSF – National Science Foundation
NSA – National Security Agency
OSD and Service research organizations – Office of the Secretary of Defense and DoD Air Force, Army, and Navy research organizations

In addition to the CSIA agencies described in Appendix A, other agencies that participate in NITRD Program activities include:

FAA – Federal Aviation Administration
FDA – Food and Drug Administration
GSA – General Services Administration
NARA – National Archives and Records Administration

The HPC Act of 1991 authorizes the establishment of an advisory committee to provide guidance to the President on the Federal role in networking and information technology R&D to assure U.S. scientific leadership and competitiveness. Presidential Executive Order 13226, dated September 30, 2005, assigns these responsibilities to the President's Council of Advisors on Science and Technology (PCAST).

The NITRD Program's broad impact derives in part from its diversified and multidisciplinary research strategy, which funds fundamental scientific investigations across Federal laboratories and centers, research universities, nonprofit organizations, and partnerships with industry. The NITRD Program is a leading source of not only fundamental technological breakthroughs but also highly skilled human resources in the advanced computing, networking, software, and information management technologies that underpin U.S. infrastructures and quality of life.

NITRD Program Component Areas, Interagency Working Groups, and Coordinating Groups

The NITRD Program is organized into eight Program Component Areas (PCAs). The work of each PCA is guided by either an Interagency Working Group (IWG) or a Coordinating Group (CG) of agency program managers. The collaboration fostered in the IWGs and CGs results in more effective use of funding resources by leveraging agency strengths, avoiding duplication, and generating interoperable results that maximize the benefits of Federal networking and IT R&D investments to both agency missions and private-sector innovation. These groups, which report to the Subcommittee, meet monthly to coordinate planning and activities in their specialized R&D areas.

The NITRD PCAs evolve in response to changing research needs. In August 2005, the new Cyber Security and Information Assurance (CSIA) PCA was established when the NSTC's Critical Information Infrastructure Protection (CIIP) IWG was renamed and rechartered to report jointly to the NSTC's Subcommittee on Infrastructure and NITRD Subcommittee. These steps were taken to facilitate better integration of CSIA R&D with NITRD activities, reflecting the broader impact of cyber security and information assurance beyond critical information infrastructure protection. The NITRD PCAs are:

High-End Computing Infrastructure and Applications (HEC I&A)

HEC I&A agencies coordinate Federal activities to provide advanced computing systems, applications software, data management, and HEC R&D infrastructure to meet agency mission needs and to keep the United States at the forefront of 21st century science, engineering, and technology. HEC capabilities enable researchers in academia, Federal laboratories, and industry to model and simulate complex processes in biology, chemistry, climate and weather, environmental sciences, materials science, nanoscale science and technology, physics, and other areas to address Federal agency mission needs.

High-End Computing Research and Development (HEC R&D)

HEC R&D agencies conduct and coordinate hardware and software R&D to enable the effective use of high-end systems to meet Federal agency mission needs, to address many of society's most challenging problems, and to strengthen the Nation's leadership in science, engineering, and technology. Research areas of interest include hardware (e.g., microarchitecture, memory subsystems, interconnect, packaging, I/O, and storage), software (e.g., operating systems, languages and compilers, development environments, algorithms), and systems technology (e.g., system architecture, programming models).

The HEC Interagency Working Group (IWG) coordinates the activities of both the HEC I&A and the HEC R&D PCAs.

Cyber Security and Information Assurance (CSIA)

CSIA focuses on research and advanced development to prevent, resist, detect, respond to, and/or recover from actions that compromise or threaten to

compromise the availability, integrity, or confidentiality of computer-based systems. These systems provide both the basic infrastructure and advanced communications in every sector of the economy, including critical infrastructures such as power grids, emergency communications systems, financial systems, and air-traffic-control networks. These systems also support national defense, national and homeland security, and other vital Federal missions, and themselves constitute critical elements of the IT infrastructure. Broad areas of concern include Internet and network security; confidentiality, availability, and integrity of information and computer-based systems; new approaches to achieving hardware and software security; testing and assessment of computer-based systems security; and reconstitution and recovery of computer-based systems and data.

The CSIA Interagency Working Group coordinates the activities of the CSIA PCA.

Human Computer Interaction and Information Management (HCI&IM)

HCI&IM R&D aims to increase the benefit of computer technologies to humans, particularly the science and engineering R&D community. To that end, HCI&IM R&D invests in technologies for mapping human knowledge into computing systems, communications networks, and information systems and back to human beings, for human analysis, understanding, and use. R&D areas include: cognitive systems, data analysis in fields such as human health and the environment, information integration, multimodal and automated language translation, robotics, and user interaction technologies.

The HCI&IM Coordinating Group coordinates the activities of the HCI&IM PCA.

Large Scale Networking (LSN)

LSN members coordinate Federal agency networking R&D in leading-edge networking technologies, services, and enhanced performance, including programs in new architectures, optical network testbeds, security, infrastructure, middleware, end-to-end performance measurement, and advanced network components; grid and collaboration networking tools and services; and engineering, management, and use of large-scale networks for scientific and applications R&D. The results of this coordinated R&D, once deployed, can assure that the next generation of the Internet will be scalable, trustworthy, and flexible.

The LSN Coordinating Group coordinates the activities of the LSN PCA.

Three teams report to the LSN Coordinating Group:

The Joint Engineering Team (JET) coordinates the network architecture, connectivity, exchange points, and cooperation among Federal agency networks and other high-performance research networks, and provides close coordination of connectivity, interoperability, and services among government, academia, and industry to improve end-to-end user performance and avoid duplication of resources and efforts. The JET also coordinates international connectivity and interoperability.

The Middleware And Grid Infrastructure Coordination (MAGIC) Team coordinates cooperation among Federal agencies, researchers, and commercial entities to research, develop, widely deploy, and use interoperable grid and middleware technologies, tools, and services and to provide a forum for international coordination.

The Networking Research Team (NRT) coordinates agency networking research programs and shares networking research information among Federal agencies. It provides outreach to end users by disseminating networking research information and coordinating activities among applications developers and end users.

High Confidence Software and Systems (HCSS)

The goal of HCSS R&D is to bolster the Nation's capability and capacity for engineering effective and efficient distributed, real-time, IT-centric systems that are certifiably and inherently dependable, reliable, safe, secure, fault-tolerant, survivable, and trustworthy. These systems, which are often embedded in larger physical and IT systems, are

essential for the operation and evolution of the country's national defense, key industrial sectors, and critical infrastructures.

The HCSS Coordinating Group coordinates the activities of the HCSS PCA.

Social, Economic, and Workforce Implications of IT and IT Workforce Development (SEW)

The activities funded under the SEW PCA focus on the nature and dynamics of IT and its implications for social, economic, and legal systems as well as the interactions between people and IT devices and capabilities; the workforce development needs arising from the growing demand for workers who are highly skilled in information technology; and the role of innovative IT applications in education and training. SEW also supports efforts to speed the transfer of networking and IT R&D results to the policy making and IT user communities at all levels in government and the private sector. A key goal of SEW research and dissemination activities is to enable individuals and society to better understand and anticipate the uses and consequences of IT, so that this knowledge can inform social policy making, IT designs, the IT user community, and broadened participation in IT education and careers.

The SEW Coordinating Group coordinates the activities of the SEW PCA.

Software Design and Productivity (SDP)

SDP R&D will lead to fundamental advances in concepts, methods, techniques, and tools for software design, development, and maintenance that can address the widening gap between the needs of Federal agencies and society for usable and dependable software-based systems and the ability to produce them in a timely, predictable, and cost-effective manner. The SDP R&D agenda spans both the engineering components of software creation (e.g., development environments, component technologies, languages, tools, system software) and the economics of software management (e.g., project management, schedule estimation and prediction, testing, document management systems) across diverse domains that include sensor networks, embedded systems, autonomous software, and highly complex, interconnected systems of systems.

The SDP Coordinating Group coordinates the activities of the SDP PCA.

Appendix C

Acronyms

ATDnet - Advanced Technology Demonstration Network

AFRL - Air Force Research Laboratory

AMPATH - AmericasPATH

API - Application programming interface

AS&C - Advanced Systems and Concepts

ATM - Automatic teller machine

BGP - Border Gateway Protocol

BIOS - Basic input/output system

CC - Common Criteria

CI - Critical infrastructure

CIH - Chernobyl virus

CIIP IWG - Critical Information Infrastructure Protection Interagency Working Group

CIO - Chief Information Officer

CISE - NSF's Computer and Information Science Engineering directorate

COCOMO II - FAA's Constructive Cost Model II

COTS - Commercial off the shelf

CSIA - Cyber Security and Information Assurance

DARPA - Defense Advanced Research Projects Agency

DDoS - Distributed denial of service

DDR&E- DoD's Director for Defense Research & Engineering

DHS - Department of Homeland Security

DNS - Domain Name System

DNSSEC - Domain Name System Security Extensions

DOC - Department of Commerce

DoD - Department of Defense

DOE - Department of Energy

DOJ - Department of Justice

DoS - Denial of service

DOT - Department of Transportation

DREN - DoD's Defense Research and Engineering Network

DUSD - Deputy Under Secretary of Defense

EGP - Exterior Gateway Protocol

ESnet - DOE's Energy Sciences network

FAA - Federal Aviation Administration

FIPS - Federal Information Processing Standard

FISMA - Federal Information Security Management Act of 2002

FTC - Federal Trade Commission

GETS - Government Emergency Telecommunications Service

HCI/HSI - Human Computer Interfaces/Human-System Interactions

HI - Host identity

HIP - Host Identity Protocol

I/O - Input/output

IA - Information assurance

IC - Intelligence community

IDS - Intrusion detection system

IETF - Internet Engineering Task Force

IGP - Interior Gateway Protocol

IKE - Internet key exchange

Infosec - Information security

IP - Internet Protocol

IPsec - Internet Protocol security

IR - Infrared

IRC - INFOSEC Research Council

IRTF - Internet Research Task Force

ISO/OSI - International Organization for Standardization/ Open System Interconnect

ISP - Internet service provider

ISSA - FAA's Information Systems Security Architecture

IT - Information technology

IWG - Interagency Working Group

JPDO - Joint Planning and Development Office

LABS - Laboratories and Basic Sciences

MANLAN - Manhattan Landing exchange point

MIP - Mobile Internet Protocol

MLS - Multi-Level Security

NAS - National Airspace System

NASA - National Aeronautics and Space Administration

NAT - Network Address Translation/Translator

NCS - National Communication System

NGATS - Next Generation Air Transportation System

NGIX - Next Generation Internet Exchange

NGN - Next-Generation Network

NIH - National Institutes of Health

NISN - NASA's Integrated Services Network

NIST - National Institute of Standards and Technology

NITRD - Networking and Information Technology Research and Development Program

NREN - NASA's Research and Education Network

NS/EP - National Security/Emergency Preparedness

NSA - National Security Agency

NSF - National Science Foundation

OMB - Office of Management and Budget

OS - Operating system

OSI - Open Systems Interconnect

OSTP - White House Office of Science and Technology Policy

PC - Personal computer

PCS - Process control system

PITAC - President's Information Technology Advisory Committee

PKI - Public key infrastructure

PSTN - Public switched telephone network

QKD - Quantum key distribution

QoS - Quality of service

R&D - Research and development

RBAC - Role-based access control

RDT&E - Research, development, test, and evaluation

RF - Radio frequency

RFID - Radio frequency identification

ROM - Read-only memory

S&T - Science and technology

SA - Situational awareness

SCADA - Supervisory control and data acquisition

SIM - Subscriber identification module

TCB - Trusted computing base

TCP/IP - Transmission Control Protocol/Internet Protocol

TRM - Technical reference model

TSWG - Technical Support Working Group

UPC - Universal Product Code

USD (AT&L) - Under Secretary of Defense for Acquisition, Technology, and Logistics

VPN - Virtual private network

WPS - Wireless priority service

XML - eXtensible Markup Language

Acknowledgements

The *Federal Plan for Cyber Security and Information Assurance Research and Development* is the product of extensive efforts by the co-chairs and members of the Interagency Working Group (IWG) on Cyber Security and Information Assurance. In addition, experts not formally affiliated with the IWG provided specialized technical information and feedback on drafts that were also essential to the completion of this Plan.

The National Coordination Office for Networking and Information Technology Research and Development played an instrumental role in the Plan's development, including research assistance and substantive technical input as well as intensive editorial review and publication of the final document.

Representatives of the following departments and agencies participated in developing the Plan and made multiple technical and editorial contributions to the content of this document:

Central Intelligence Agency
Defense Advanced Research Projects Agency
Department of Energy
Department of Homeland Security
Department of Justice
Department of State
Department of Transportation
Department of the Treasury
Disruptive Technology Office
Federal Aviation Administration
Federal Bureau of Investigation
National Aeronautics and Space Administration
National Institute of Standards and Technology
National Institutes of Health
National Science Foundation
National Security Agency
Office of the Secretary of Defense
and Department of Defense Service research organizations
Technical Support Working Group
U.S. Postal Service

Cover Design, Graphics, and Printing

The cover design and graphics are the work of Scientific Designer/Illustrator James J. Caras of NSF's Design and Publishing Section. Printing was overseen by NSF Electronic Publishing Specialist Kelly DuBose.

National Coordination Office
for Networking and Information Technology
Research and Development

Suite II - 405, 4201 Wilson Blvd., Arlington, Virginia 22230
(703) 292-4873
http://www.nitrd.gov

National Science and Technology Council
Interagency Working Group on
Cyber Security and Information Assurance