IAnewsletter

The Newsletter for Information Assurance Technology Professionals

Volume 7 Number 1 • Spring 2004

The National Institute of Standards and Technology (NIST) Computer Security Incident Handling Guide

also inside—

contents

feature

IA initiatives

in every issue

About IATAC & the *IAnewsletter*—

IAnewsletter is published quarterly by the Information Assurance Technology Analysis Center (IATAC). IATAC is a Department of Defense (DoD) sponsored Information Analysis Center, administratively managed by the Defense Technical Information Center (DTIC), Defense Information Systems Agency (DISA).

Contents of the *IAnewsletter* are not necessarily the official views of or endorsed by the U.S. Government, DoD, or DISA. The mention of commercial products and/or services does not imply endorsement by the DoD or DISA.

Inquiries about IATAC capabilities, products, and services may be addressed to—

IATAC Director: Robert J. Lamb
Deputy Director: Abraham T. Usher
Inquiry Services: Peggy O'Connor

IAnewsletter Staff—

Creative Director: Christina P. McNemar
Art Director: Ahnie Senft
Designers: Maria Candelaria, Kathy Everett, and Trang Dam
Editorial Board: Abraham T. Usher, April Perera, Jim Peña, Brad Soules

IAnewsletter Article Submissions

To submit your articles, notices, programs, or ideas for future issues, please visit http://iac.dtic.mil/iatac/IA_newsletter.html and download an "Article Instructions" packet.

IAnewsletter Address Changes/Additions/Deletions

To change, add, or delete your mailing or E-mail address (soft-copy receipt), please contact us at—

IATAC
Attn: Peggy O'Connor
3190 Fairview Park Drive
Falls Church, VA 22042

Phone: 703/289-5454
Fax: 703/289-5467

E-mail: iatac@dtic.mil
URL: http://iac.dtic.mil/iatac

Deadlines for future Issues—
Summer 2004 June 25, 2004
Cover design: Maria Candelaria
Newsletter design: Ahnie Senft

Distribution Statement A:

Approved for public release; distribution is unlimited.

IATAC Chat

Robert J. Lamb, IATAC Director

"Become a student of change. It is the only thing that will remain constant."

Anthony J. D'Angelo

I originally asked Christina McNemar, the IATAC Promotions Director to capture some thoughts for this IATAC Chat **as she reviewed articles and directed the layout of this edition.** As in all things, she did a super job capturing the essence of these articles summarizing them with…"as I read through this edition's articles, one word kept coming to the forefront of my thoughts—change. So rapidly do the trends, applications, technology, and policies change in the information assurance (IA) arena, IA is a work in progress."

How many countless articles have you read, or heard in classes, conferences, speeches—organizations are increasingly dependent on information technology (IT), and as a result, threats against IT resources have increased in number and severity. But it is true—the potential impact from a single threat can be extremely high and costly exactly because we are so dependent. And, when disaster strikes, organizations must adapt and do so quickly and that requires a formal incident response (IR) capability to quickly identify and mitigate incidents. To that end, the National Institute of Standards and Technology (NIST) has released a new publication to help organizations build an IR capability. The feature article gives a sampling of the guidance provided in the full, 150–page publication, as well as how to obtain it.

One way we have all changed is our use of the Web. From news to online purchases, workplace applications, the Web has become the foundation of much of our information exchange processes. As organizations continue to "push the envelope," to add more functionality and collaboration to their business practices, Web applications are increasingly being leveraged to solve complicated tasks. As a result, the architectures have become more complicated and more vulnerable. Mr. Usher's article on Web Application Security provides insight into threats associated with leveraging Web applications.

Add to this already complex issue the concept of "grid computing," the scenario gets even more complex. Grid computing is a collaborative infrastructure that enables direct access to software, data, and hardware components, which is managed by many organizations in widely dispersed locations. The special report presents a wealth of information on grid computing including challenges, current and future technologies, research and development, and security implications and initiatives.

The evolving threat environment has also sparked DoD to change how they view the information operations (IO) landscape. USSTRATCOM is working diligently to make this a reality. With the recently published IO Roadmap, their main mission-thrust centers around the "operationalization" of IO into core military warfighting competency.

I'd like to include one sad note in this Chat—the departure of Abe Usher as the Deputy Director of IATAC. Since Abe joined us three years ago, he has been an agent of change within IATAC. He brought new ideas on enhanced processes and procedures, bringing tools that automated many of those activities. Along the way, his calm presence, considered thoughts, and friendship touched the lives of us all. We will miss Abe and wish him the very best of luck in his new challenges.

Finally, we continue to seek authors for future editions of the newsletter. If you or a colleague would be interested in writing an article, please don't hesitate to contact IATAC. You can download an article submission packet from our Web site at http://iac.dtic.mil/iatac/IA_newsletter.html. Suggestions and/or feedback is helpful as well. ■

The National Institute of Standards and Technology (NIST) Computer Security Incident Handling Guide

by Karen Kent

Author's Note

NIST SP 800–61 defines a computer security incident as "a violation or imminent threat of violation of computer security policies, acceptable use policies, or standard security practices." Consider the components of this definition in more detail—

- **"Violation or imminent threat of violation"** reflects that computer security incidents not only include events that have already occurred, but also include situations where a serious event is likely to occur in the near future. For example, if a new worm is likely to infect the organization's systems within the next six hours, the organization can proactively declare an incident, even though it has not yet been attacked.
- **"Computer security policies"** typically state which types of security-related activities are and are not permitted within the organization. For example, the policy typically forbids unauthorized access to data, systems, and networks.
- **"Acceptable use policies"** specify appropriate and inappropriate usage of information technology resources. Typical examples of inappropriate usage include sending threats through e-mail, distributing pornography, and sharing copyrighted music. Incident response (IR) teams often handle violations of acceptable use policies because they typically involve using similar procedures and technology, such as intrusion detection systems and computer forensics software.
- **"Standard security practices"** encompass all actions that involve violations of generally accepted practices regarding information security. This provides a definition for incidents in cases where computer security policies are incomplete or nonexistent.

In 1991, the National Institute of Standards and Technology (NIST) released Special Publication (SP) 800–3: *Establishing a Computer Security Incident Response Capability (CSIRC)*. The purpose of incident response is to react quickly when an incident occurs so that the impact to the organization is minimized. When SP 800–3 was released, most organizations did not have formal incident response (IR) programs and were just starting to face computer security-related threats, such as viruses. Of course, much has changed since then. As organizations' dependence on information technology (IT) has grown, the threats against IT resources have also increased in number and severity. The potential impact to an organization from a single incident can be incredibly high. A formal incident response capability is invaluable in quickly identifying and mitigating incidents, reducing their impact. However, although incidents affect practically every organization, many still lack robust incident response capabilities. Accordingly, in January 2004, NIST released SP 800–61: *Computer Security Incident Handling Guide*.

NIST SP 800–61 is intended for use by organizations that want to establish a formal IR capability and by existing IR teams that want to improve their performance. The goal of SP 800–61 is to help all organizations handle computer security incidents more efficiently and effectively. Nearly 150 pages long, the guide provides a wealth of information regarding incident handling. Material covered by the guide includes the following—

- Advice on establishing an IR capability, including policy and procedure creation; IR team staffing, structure, responsibilities, and skills; and interactions with other internal teams and external parties (e.g., law enforcement, media, other IR teams)
- Recommendations for performing the basic incident handling steps, with an emphasis on preparation and incident detection and analysis
- Guidance on handling five specific types of incidents: denial of service (DoS), malicious code, unau-

thorized access, inappropriate usage, and multiple component incidents

- IR scenarios and questions to be used for tabletop IR exercises
- Pointers to helpful IR-related tools and resources

The remainder of this article will provide a summary of the most important principles discussed in SP 800–61.

Understanding the IR lifecycle

Over the years, several models for the incident response lifecycle have been developed. Most models are quite similar, with the same basic steps performed in sequence, and with lessons learned being applied to improve future responses. NIST SP 800–61 uses the model depicted below in Figure 1. The following items describe each step in the IR process—

- **Preparation**—In the NIST model, preparation has two major components. The first is creating a formal IR capability, which includes developing IR policy and procedures, and staffing an IR team. The second component of preparation is to prevent incidents by implementing appropriate security measures. While incident prevention is not strictly a part of an IR program, it is key to the success of any IR team. If the team is constantly being called upon to handle easily preventable incidents, the team members are likely to become burned out and frustrated, and the team will also not be able to focus adequately on preparing for and handling more serious incidents.
- **Detection and analysis**—The next step in the model is to detect potential signs of incidents. Millions of security-related events may be recorded on a daily basis across large organizations. To ensure that event data is reviewed promptly, automated methods such as event correlation (performed by security event management software) should be used to analyze the data and identify a small subset of events that should be reviewed by a trained incident analyst. The IR team should have deep expertise on incident analysis, so that it can quickly review the available information and determine the appropriate course of action. Suspected incidents often turn out to be operational issues, and vice versa. However, the handling of each incident should be prioritized based on the impact that the incident may have to the organization's mission, regardless of whether the incident turns out to be caused by a security breach or a network device malfunction.
- **Containment, eradication, and recovery**—The goals of this step are to stop the incident, mitigate any vulnerabilities that allowed the incident to

Figure 1. IR lifecycle model

occur, and recover disrupted functionality and data. Most IR models list containment, eradication, and recovery as three separate steps. In reality, these actions are intertwined in many incidents, so the NIST model groups them together. For example, running an antivirus program against an infected system might contain, eradicate, and recover from the incident all at once.

- **Post-incident activity**—Once the organization has successfully recovered from the incident, it is important to analyze how effective the response was, identify areas of concern, and recommend and implement improvements to the IR process and the organization's security posture. This analysis often occurs through a lessons learned meeting, which brings the incident participants together to discuss the incident. Another post-incident activity is the creation of a follow-up report that documents what occurred during the incident handling. Information revealed by lessons learned meetings and follow-up reports should be applied to prevent future incidents and improve the IR process.

Establishing a formal IR capability

Organizations that face significant information security threats should establish a formal incident response capability. A major component of this is the creation of an IR policy and procedures. The organization should also determine what services the IR team should provide, and how the team should be structured and staffed. Once the team has been staffed and has established good working relationships with other groups within the organization, it is important that the team members receive ongoing training. IR exercises are particularly helpful in sharpening IR skills as well as identifying shortcomings in the IR process and procedures. This will help the IR team to contain and resolve incidents more efficiently and effectively. The following items highlight the primary actions involved in establishing a formal IR capability—

- **Create an incident response policy**—Every organization should have an incident response policy. The policy should define the term "incident" and explain the impact that incidents can have on the organization. The policy should also explain the IR-related responsibilities of specific teams within the organization, as well as the user community. For example, the policy may require all users to report incidents to the organization's IR team. In addition to delineating responsibilities, the policy should also clearly state which individuals or teams have authority to make decisions, such as disconnecting systems from networks, confiscating computers, and contacting law enforcement. If the policy does not explain authority clearly, incident handling may be significantly delayed while disagreements occur about who is allowed to do what. The IR policy may also include guidelines for prioritizing incidents and definitions of performance measures.

- **Determine what the IR team's responsibilities will be**—Besides handling incidents, most IR teams provide additional services. It is important to think first about what services the team should provide, then select a team structure and staffing model to support the services. Other services sometimes provided by IR teams are listed below. Note that most of these services are proactive, attempting to prevent incidents from occurring.
  - Communicating information on new vulnerabilities, threats, and trends to the appropriate internal parties (e.g., management, IT staff, end users); in some organizations, the IR team actually performs research (e.g., deploying honeypots) to increase their understanding and awareness of new threats
  - Performing vulnerability assessments and penetration testing to identify weaknesses and recommend corrective actions
  - Maintaining and monitoring intrusion detection systems
  - Providing education and awareness services for IT staff and end users

- **Choose an IR team structure**—The IR team structure should be based on the physical and logical divisions of the organization. For example, a small organization located in a single building would probably have a single IR team. A separate IR team in each city may best serve an organization with a major presence in three cities, and an organization with six divisions may wish to have a separate IR team for each division. However, whenever an organization has multiple teams, it is very important that a centralized IR entity also exist to facilitate consistent responses and strong communication among the teams. This can greatly mitigate the impact of incidents such as worms that are likely to affect many parts of the organization.

- **Choose an IR team staffing model**—Staffing models can range from having employees perform all incident response work to fully outsourcing it. In most organizations, employees perform most IR work, but it is increasingly common for 24x7 monitoring of security devices (e.g., firewalls, intrusion detection sensors) to be performed by an offsite managed security services provider (MSSP). Other organizations use contractors only to assist with particular incidents, such as widespread malicious code infections that are extremely labor-intensive to contain and eradicate. Contractors may also be used to provide expertise in highly specialized fields.

- **Staff the team with individuals with the right skills**—IR team members should be technically adept in several fields. Besides having a broad knowledge of security and good expertise in intrusion detection, IR team members should also be knowledgeable about the major operating systems and applications in use, as well as the fundamentals of networking. Because of the nature of the work, IR team members should also have strong problem solving, teamwork, and communication skills, the ability to work well under pressure, and good attention to detail. If team members do not have the right skills, they may not handle incidents as effectively, which may negatively impact the organization, as well as damage the team's reputation. Team members can be trained on specific

technical areas, but they must already have the basic skills, aptitude, and temperament to perform incident response work.

- **Form relationships with other groups**—Effective incident response requires the participation of teams throughout the organization. NIST SP 800–61 lists some of the teams: "information security, physical security, IT technical support (e.g., help desk agents, system administrators), public affairs, legal, human resources, business continuity planning/continuity of operations staff, and management." The IR team should establish good working relationships with each of these groups, as well as certain external groups. For example, the team may want to seek advice regarding a particular incident from other IR teams. Other external entities that the IR team may need to communicate with include law enforcement agencies and the media. In many organizations, the IR team does not initiate such communications, but the team should be prepared to participate as needed. The IR team should be responsible for creating and maintaining contact information for each internal and external group that may participate in incidents. The organization should have written guidelines and procedures that clearly state who may communicate what incident-related information with what parties under what circumstances.

- **Create incident response procedures**—Having detailed procedures for handling incidents helps to ensure that incidents are handled consistently and effectively. The procedures should be based on the IR policy and the IR team's responsibilities, with input from other groups that may participate in handling incidents. A common problem in IR is developing procedures that look great on paper but are not helpful when a crisis occurs. Organizations should validate the procedures periodically to ensure that they are accurate and realistic; this is best done through simulations and exercises, rather than finding the errors during a live incident. Organizations should also ensure that the procedures support the organization's mission, such as giving the highest priority to incidents that may severely impact operations.

- **Establish an ongoing training program**—Compared to most IT-related areas, IR teams should receive substantially more training. Not only do they need to understand general IR principles and a wide variety of technologies, but also they may handle certain types of incidents only on occasion. It is strongly recommended that organizations hold incident response exercises regularly so that IR teams (as well as other teams within the organization) can practice IR and gain experience. This is particularly valuable for junior members of the team, who may also benefit from mentoring relationships with senior team members. Other ways for team members to build technical skills is to attend training courses and conferences and to work temporarily or part-time for other teams.

Responding based on mission impact

In many organizations, incident handlers do not receive much guidance regarding the prioritization of incidents. They tend to handle incidents in the order that they are reported, giving higher priority to incidents that obviously have a major impact to the organization (e.g., compromise of a firewall, defacement of a publicly accessible Web server). Incident handlers often have little information regarding the significance to the organization of particular systems, applications, and data. A handler may not know that the compromise of application XYZ has a severe negative impact on the organization's mission. Therefore, it is critical that incident handlers receive guidance on prioritizing responses to incidents. The following items provide advice for facilitating this:

- **Establish written guidelines for incident prioritization**—Each incident should be prioritized based on three criteria. First, how critical to the organization's mission are the affected resources? Second, what is the current technical effect of the incident (e.g., administrator-level compromise, data modification, data access)? Third, how is this incident likely to change if it is not stopped immediately? For example, if the incident is not contained for 24 hours, is it likely to affect other resources, and if so, which ones? Is the technical effect of the incident likely to worsen (e.g., user-level compromise escalating to an administrator-level compromise)? By documenting these considerations and establishing formal guidelines for prioritizing incident responses, incident handlers can make consistent decisions that best support the organization's mission, even under times of extreme stress.

- **Maintain situational awareness during all incidents**—This is particularly important during the most severe incidents, such as large malicious code infestations that may adversely affect many or most of the organization's systems and users. It is usually quite challenging for all parties involved in handling the incident to communicate effectively with each other, and for the organization's management and IR team to acquire and validate the necessary information so that they can rapidly make decisions that best support the organization. Besides creating written guidelines for incident prioritization, other preparatory activities that support situational awareness are as follows—
  - **Notification framework**—The organization should have an established framework for notifying all relevant individuals and groups within the organization of the current status of an incident. It is best to have multiple notification mechanisms, in case the primary method (e.g., telephone, pager, E-mail) is unavailable. The framework should be tested regularly to ensure that it works smoothly and that the points of contact are up-to-date. The framework should also take into account external parties that may need to be contacted, such as the appropriate law enforcement agencies.

continued on page 14…

Web Application Security

by Abraham T. Usher, CISSP

Web applications have become an indispensable mechanism for efficiently conducting business in the Information Age. At the same time, Web applications often introduce risks to information and information systems. This article examines the evolution of the World Wide Web (WWW, Web or W3), why Web application attacks are increasingly common, common attack vectors, as well as methods for reducing risks associated with Web applications.

web application://

a program that uses HTTP for its core communicationprotocol and delivers Web-based information to users in HTML. [1]

Evolution of the Web

In March 1989, Tim Berners Lee released a paper for the Laboratory for Particle Physics titled "Information Management: A Proposal" which introduced the concept of linking information systems and information resources with hypertext—Human-readable information linked together in an unconstrained way. [2] This concept of hypertext eventually grew into what became known as the WWW. Through the creation and use of hypertext markup language (HTML) and related media files, it is simple for most anyone with Internet access to browse or contribute to the Web. Early Web applications in the mid 1990's were made up of primarily static content (see Figure 1).

HTML files and images were stored on Web servers that served the content directly to client Web browsers such as Netscape Communicator or Internet Explorer. As the need for additional functionality became a business reality, many Web sites became 3-tiered Web applications, including a Web server, a middleware tier for enforcing business rules (application server), and a backend database for providing persistent storage of on-line transactions (see Figure 2).

As organizations attempted to leverage Web applications to solve increasingly complicated tasks, the architecture evolved to include n-tiered (many tiered) architectures (see Figure 3 on page 10).

With added functionality comes additional risk. Many organizations now expose mission critical databases and information stores indirectly through Web interfaces. The subsequent sections of this article examine these risks in greater detail.

Why application attacks are common

Government and industry have made substantial strides in securing information systems and networks. Collaboration between the Department of Homeland Security's National Cybersecurity Division (NCSD), the Computer Emergency Response Team Coordination Center (CERT/CC), the Joint Task Force for Computer Network Operations (JTF–CNO), military Service CERTS, the Defense Information Systems Agency (DISA) and many other organizations has led to a much improved security posture for U.S. computers and networks. Most organizations require that application servers and end-user workstations be protected through a process of operating system patching, secure configuration, and network segmentation.

Unfortunately, these basic security procedures may not be sufficient for Web applications. The systems that make up a Web application can be patched, properly configured, and located behind a firewall and yet still be vulnerable to attack. Why? Because attacks against Web applications exploit flaws in the processing logic of the application. Comprehensive security of Web applications

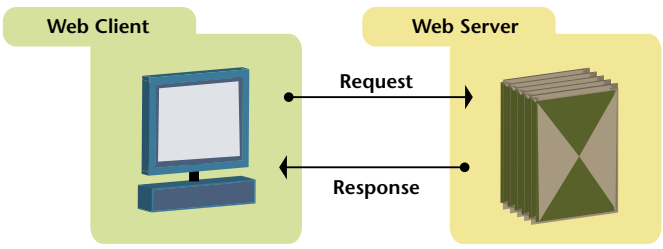


Figure 1. Static Web site

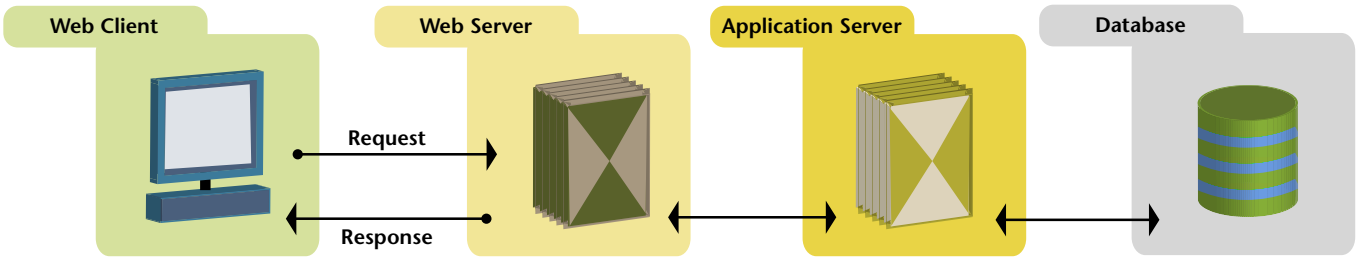


Figure 2. Three-tiered Web application

must address the multiple related security mechanisms affecting networks, system platforms, and application logic (see Figure 4 on page 11). Testing the "correctness" of the logic in a Web application's source code and configuration is non-trivial. This is because what may be appropriate in one security context (for example a public bulletin board) may not be appropriate in another security context (e.g., an e-commerce Web site).

Consider the ability to write data to the database of a Web application. This would be a beneficial capability for a public bulletin board, where end-users want to post information to be viewed by other users. However, the ability to arbitrarily write data to the system's database might be an inappropriate capability on an e-commerce site. What would happen if customers visiting their favorite on-line bookstore could modify prices in the Web application's database? While consumers might like the idea of being able to set their own price, the retailer would consider such a data modification as highly inappropriate! Because of security context ambiguity and the difficulty of determining the "intent" of source code, securing Web applications remains a complicated problem.

Threats to Web applications

There are many threats to publicly accessible network resources, especially related to Web applications. Generally these threats fall into one of the three classic categories of Information Assurance—availability, integrity, or confidentiality. These threats are described below.

Availability

- **Denial-of-Service (DoS)**—Intentional actions which prevent an information system from functioning according to its purpose or which preclude the use of an information resource to authorized users
- **Non-malicious resource consumption**—Overuse of a system or resource that prevents authorized users from accessing it

Confidentiality

- **Data interception**—External entities may monitor information transferred across an insecure medium
- **Aggregation**—Users without a valid need to know may "put together" restricted information by combining many smaller pieces of unrestricted information
- **Inference**—Users without a valid need to know may infer specific restricted information by examining general statistical information
- **Inappropriate disclosure**—Confidential information may be disclosed due to inappropriate configuration or application logic (e.g., directory traversal attack against a Microsoft IIS Web server).

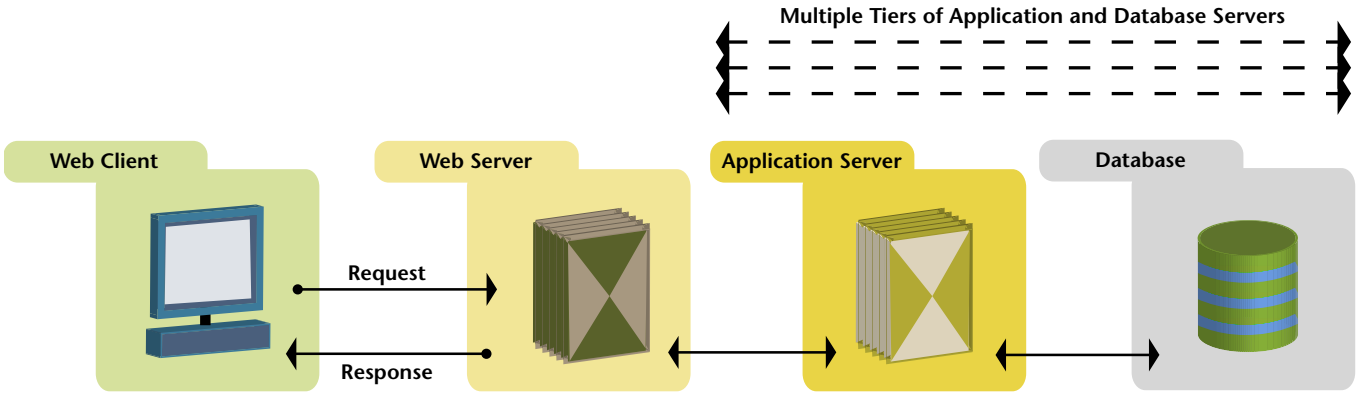


Figure 3. n-tiered Web application

Integrity

- **Web defacement**—Unauthorized modification or replacement of content on a Web application
- **SQL injection**—Obtaining unauthorized access to an on-line database by sending Structured Query Language (SQL) commands via Web forms
- **Malicious code**—Any worm, virus, or unauthorized program that violates the security policy of the system in question

Reducing risk

Quality management researcher Dr. Joseph Juran recognized a universal principle he described as the "vital few and trivial many." The principle (sometimes referred to as the 80/20 rule or the Pareto Principle) states that 20 percent of something is always responsible for 80 percent of the results. [3]. In Juran's initial research, he found that 20 percent of all defects usually create 80 percent of the problems.

The application of this concept in the IA community is obvious. With limited time and resources, IA professionals must focus on the threats that cause the highest proportion of risk. The System Administration and Network Security Institute (SANS) in conjunction with the Federal Bureau of Investigation (FBI) has created a "Top Twenty" list of vulnerabilities that correspond to the greatest amount of risk in information systems. This list provides a useful starting point for system administrators and security officers that are responsible for securing their enterprise resources.

More specific to Web security, the Open Web Application Security Project (OWASP) has released a top ten list of threats to the security of Web applications.

From the SANS Web site

The SANS Top Twenty is actually two Top Ten lists—the ten most commonly exploited vulnerable services in Windows and the ten most commonly exploited vulnerable services in UNIX and Linux. Although there are thousands of security incidents each year affecting these operating systems, the overwhelming majority of successful attacks target one or more of these twenty vulnerable services. [4]

OWASP Top Ten List

1. **Unvalidated input**—Information from Web requests is not validated before being used by a Web application
2. **Broken access control**—Restrictions on what authenticated users are allowed to do are not properly enforced
3. **Broken authentication and session management**—Account credentials and session tokens are not properly protected
4. **Cross-site scripting flaws**—The Web application can be used to transport an attack to an end user's browser
5. **Buffer overflows**—Web application components in some languages do not properly validate input and in some cases, can be used to take control of a process
6. **Injection flaws**—Insecure parameter passing between a Web server and an external system
7. **Improper error handling**—Errors not handled properly allow attackers to gain system information, deny service, or crash the server
8. **Insecure storage**—Lack of cryptographic mechanisms may increase damage from system compromise
9. **Denial of service**—Actions which prevent in information system from functioning according to its purpose
10. **Insecure configuration management**—Strong server configuration is critical to a secure Web application. More than 66 percent of on-line compromises are due to misconfiguration.

The task of manually attempting to test all risks related to the SANS Top Twenty and OWASP Top Ten items is very challenging. Fortunately there exist several products that can automate some of the process of testing the security of Web applications. Among the most popular commercial products are SPI Dynamics' Web Inspect, Sanctuum's

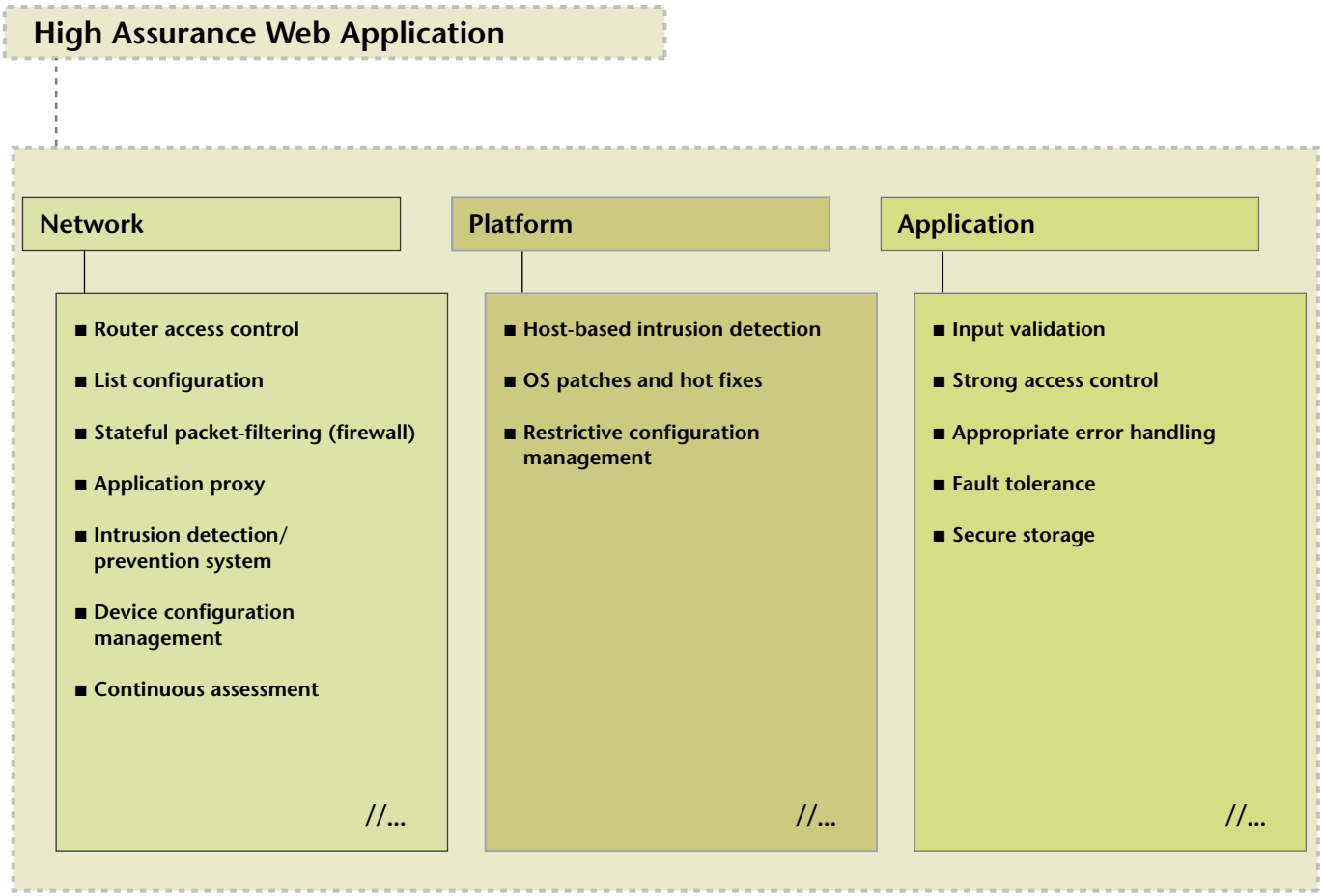


Figure 4. Security components of a high assurance Web application

AppScan, NGSSoftware's DominoScan and OraScan, Nstalker's Nstealth, and KaVaDo's Scando. Additional specific information on these products is available through IATAC's Vulnerability Analysis tools report.

Conclusion

The Web provides a wealth of opportunities for collaboration and information exchange. With its great potential for facilitating productivity, the Web also has many associated risks. IA professionals must remain vigilant in securing Web applications through a process of disciplined patching, configuration management, and application testing. ■

About the Author

Abraham T. Usher

Abraham T. Usher graduated from the U.S. Military Academy in 1996 with a B.S. double major in Modern Standard Arabic and German language studies and he also received a M.S. in Information Systems from George Mason University. Mr. Usher is a Certified Informational System Security Professional (CISSP) and may be reached at iatac@dtic.mil.

References

1. Microsoft Windows 2000 Server Documentation. Definition of "Web application."
2. Berners-Lee, Tim. Information Management: A Proposal. CERN, 1989. http://www.w3.org/History/1989/proposal.html
3. Juran, J.M. http://www.juran.com/lower_2.cfm?article_id=21
4. System Administrator and Network Security Institute (SANS). Top Twenty Vulnerabilities—The Experts Consensus. 2004. http://www.sans.org/top20/

Department of Defense (DoD) Changing Information Operations Landscape

by Tom Castellano and Paul Mays

Author's Note

The statements in this article describe work in progress within USSTRATCOM and do not represent the official policy of the JFHQ–IO.

The recently published Information Operations (IO) Roadmap from the Office of Secretary of Defense continues to advocate and direct the milestones identified in the Unified Command Plan (UCP) Change Two that establishes the United States Strategic Command (USSTRATCOM) as the Department of Defense (DoD) lead combatant command for IO. With that responsibility, USSTRATCOM is pioneering new ways of planning, supporting, and executing the information operations (IO) missions as part of a national global strike capability. The following information is provided to introduce new and old IO warriors to USSTRATCOM's vision.

USSTRATCOM, headquartered at Offutt Air Force Base, Nebraska, is currently leading IO initiatives and changes that span national, strategic, operational, combined multinational, joint, and tactical levels within the military and the United States Government. USSTRATCOM is executing two pivotal efforts that are outlined in UCP Change Two and the recently published DoD IO Roadmap. The main thrust centers around the operationalization of IO into a core military warfighting competency.

To execute these missions, USSTRATCOM established the Joint Force Headquarters–Information Operations (JFHQ–IO), specifically chartered to accomplish the IO tasks assigned in UCP Change Two and the DoD IO Roadmap. The JFHQ–IO, co-located within USSTRATCOM, is the focal point for DoD information operations. Specifically, JFHQ–IO's mission and tasks include—

- Plan, coordinate, and integrate DoD IO, support other combatant commanders for IO planning, identify IO requirements and, when directed, execute selected IO missions to support decisive national security objectives
- Provide responsive IO to achieve full spectrum military integrated operations in support of national objectives
  - IO **fully integrated** into planning and execution of DoD operations
  - Enable supporting and supported combatant commands (CCs) to affect behavior of adversaries through decisive IO

As the JFHQ–IO executes its new DoD IO role, one of their challenges is to apply the operational experience and lessons gained during the global war on terrorism to integrate and coordinate the use of IO capabilities across the DoD, intelligence community, and with allied/coalition partners. To assist in the overall IO mission, the components of JFHQ–IO (see Figure 1 on next page) include the Deputy Commander for Network Attack (Planning and Integration), Deputy Commander for Global Network Operations (name change pending approval), and the Deputy Commander, Joint Information Operations Center (JIOC).

The IO landscape

As part of the overall synchronization and deconfliction of DoD IO, the JFHQ–IO is in the midst of evaluating current and evolving IO doctrine and policy to determine how to fulfill the DoD goal of making IO a core military capability. The core IO mission areas include Computer Network Operations (CNO), Operations Security (OPSEC), Military Deception (MILDEC), Psychological Operations (PSYOP), Electronic Warfare (EW), as well as supporting activities that include Information Assurance (IA), Public Affairs (PA), Civil Affairs (CA), Counterintelligence (CI), Computer Security, Communications Security, and Physical Security. The cross-fertilization of these IO functional and reinforcing areas cover the full spectrum and involves participants ranging from national level agencies down to service unique units and capabilities.

Another realization in the IO arena is the threat landscape rapidly changes. The identification of evolving threats is increasingly difficult in the information battlespace and persistent collection compounds link and nodal analysis. This overload of information taxes even the

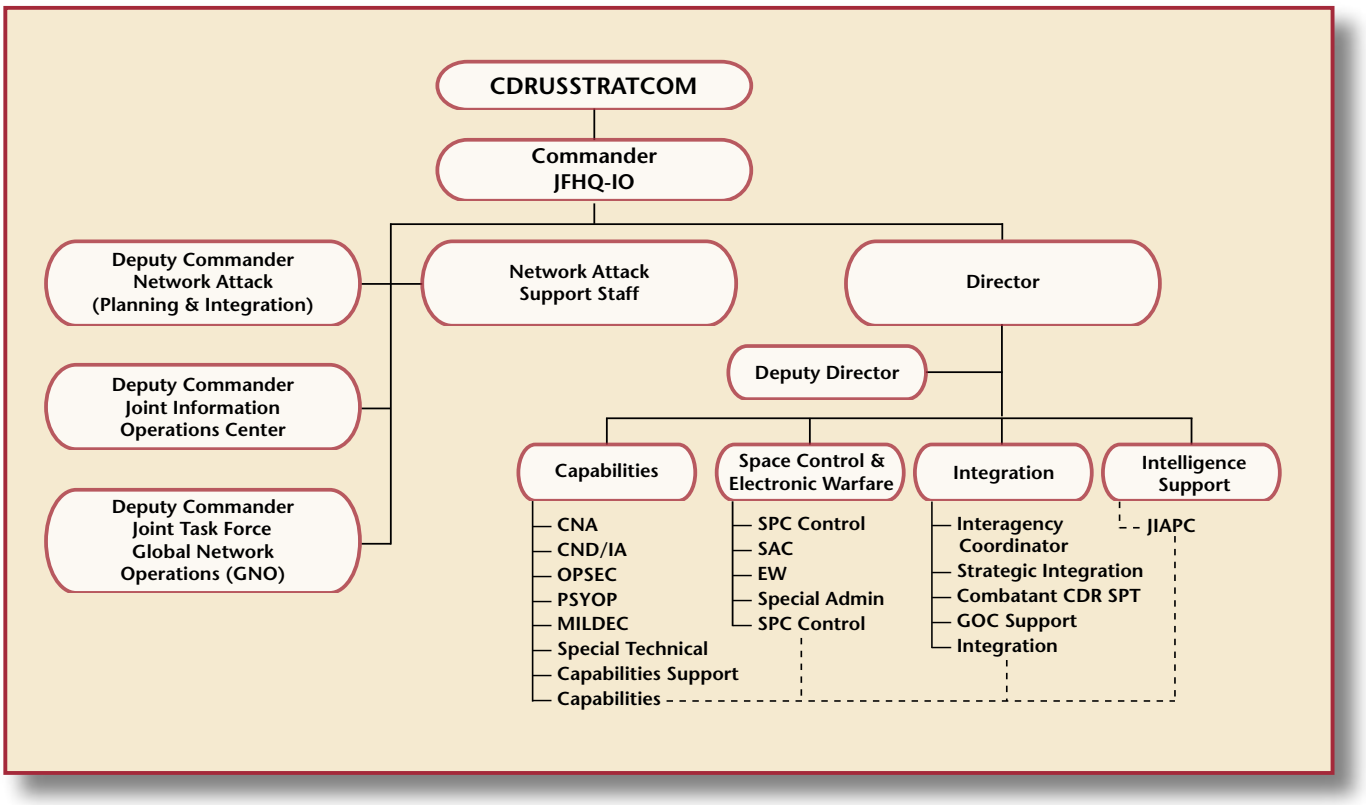


Figure 1. JFHQ-IO Headquarters staff and functional components

most sophisticated systems and cannot only sub-optimize one specific IO discipline, but rather the entire IO spectrum. The proper application and timeliness of information and knowledge is critical to the success of an information operations plan/effort.

Conclusion

USSTRATCOM and JFHQ–IO are positioned to execute their newly assigned IO tasks. JFHQ–IO, along with their components, is working toward the goal of making IO a viable warfighting capability. Although all the geographic combatant commanders have the requirement to develop IO as a core military warfighting competency, USSTRATCOM, as the DoD lead, is organized to deconflict and support IO planning and execution. The ultimate realization is to provide the combatant commander or decision maker a full suite of both kinetic and non-kinetic options that are integrated into viable courses of action. ■

continued on next page...

About the Authors

Tom Castellano

Tom Castellano is an IO analyst assigned to the IO Capabilities Division, JFHQ–IO. His background includes assignments with DIA, Joint Staff, U.S. Army and U.S. Air Force IO/IW/IA units and activities. His interests include scientific and technical research, enterprise integration and engineering, and collaboration. He can be reached at castellt@stratcom.mil, 402/294-2894.

Paul Mays

Paul Mays is an IO analyst assigned to the IO Capabilities Division, JFHQ–IO. He was previously assigned as a CNO analyst and planner with the Joint Task Force for Computer Network Operations (JTF–CNO). While in the Army, Mr. Mays received the prestigious MI Corps' Knowlton Award and participated in numerous theater level IO exercises. He is currently a Master's candidate in Management Information Systems. He can be reached at mays_paul@bah.com, 402/294-0125.

References

Cyber Security: How Can We Protect American Computer Networks From Attack?: Hearing Before the Committee on Science, U.S. House of Representatives, Oct 2001.

Military Transformation: A Strategic Approach, Fall 2003, Office of Transformation, Department of Defense, available at: http://www.oft.osd.mil/.

continued from page 7…

"The NIST Computer Security Incident Handling Guide"

- **Incident leads**—The IR team should have multiple members prepared to act as incident leads for the most severe incidents. In these cases, the incident lead will not perform any technical work; rather, the lead acts as a coordinator for the incident response. The lead stays in close contact with all parties involved in the response and ensures that relevant information is communicated effectively and consistently among all the parties.
- **Exercises and simulations**—Because major incidents do not occur very often in most organizations, IR teams typically lack experience in handling them. An effective way of improving a team's skills is to conduct exercises and simulations of major incidents regularly. This will also identify instances where procedures are not sufficient for handling major incidents; resolving these issues is likely to improve handling of future incidents.

This article has presented just a sampling of the guidance provided in NIST SP 800–61: *Computer Security Incident Handling Guide*. The full guide provides much more detailed information on establishing an IR capability, and also provides technical guidance on handling several types of incidents. The guide also contains an extensive set of IR exercise scenarios and questions that organizations can use to build staff skills and improve their IR processes. SP 800–61 is available for download from NIST's Computer Security Resource Center at http://csrc.nist.gov/publications/nistpubs/index.html. ■

About the Author

Karen Kent

Karen Kent holds a B.S. in Computer Science from the University of Wisconsin-Parkside and an M.S. in Computer Science from the University of Idaho. Ms. Kent co-authored NIST Special Publication 800–61: *Computer Security Incident Handling Guide*, and has also contributed to several books on information security and published over a dozen articles on intrusion detection. She may be reached at kent_karen@bah.com.

Information Assurance– Are You Prepared?

by Walter C. Kelley

The Defense Information Systems Agency (DISA) mission is to plan, engineer, acquire, field, and support the command, control, communications, and information systems needs of the Department of Defense (DoD) in times of war and peace. One vital aspect of that mission is to provide valuable, timely and accurate education, training, and awareness (ETA) to DoD components.

DISA's Field Security Operations (FSO) branch provides information assurance (IA) training products to the DoD community in support of this mission. The training products consist of 18 Web-based training (WBT), six computer-based training (CBT), and four VHS videos. All of these are offered and shipped to customers free of charge.

The newest addition to the product offering is Critical Infrastructure Protection (CIP). The CIP WBT provides baseline CIP awareness to enhance the knowledge of DoD personnel in the front lines of defense, DoD and other government CIP planners, infrastructure owners, managers, technicians, and users. This product provides an overview of the systems that comprise the critical infrastructure, what CIP is, the national organizational structure of CIP, how DoD fits into the national CIP organization, and DoD CIP organizational structure and responsibilities. The course goes into detail on the DoD infrastructure sectors and special function components and concludes with the six phases of the CIP lifecycle.

Another new addition and one that is quickly gaining notoriety, is the System Security Authorization Agreement (SSAA) Preparation Guide. Published in December 2003, the SSAA Preparation Guide contains guidance on completion of the SSAA while accomplishing the DoD Information Technology Security Certification and Accreditation Process (DITSCAP). After presenting an overview of the DITSCAP, this Web-based product provides detailed guidance on the contents necessary to complete a SSAA, using the outline presented in the DITSCAP Application Manual, DoD 8510.1–M. The target audience for the product is information system certification team members, information assurance managers (IAMs), information assurance officers (IAOs), system administrators (SAs), and other personnel responsible for writing, processing, or reviewing SSAAs. This product is also useful for preparation of a SSAA using the National Information Assurance Certification and Accreditation Process (NIACAP), in the National Security Telecommunications and Information Systems Security Instruction (NSTISSI) No. 1000.

Among the new Web-based training products is the very popular Information Assurance Policy and Technology (IAP&T). This product updates and replaces the older Operational Information Systems Security (OISS) CD–ROM. The IAP&T training courseware has been created so that users may successfully perform their duties as IAMs, IAOs, or SAs in accordance with DoD guidance (DoDD 8500.1 and DoDI 8500.2) pertaining to the defense of information systems. Individuals assigned to duties involved with policy and oversight, inspection and audit, or other functions supporting the IA mission (e.g., prevention, detection, and eradication of viruses; execution and evaluation of system audit records; access control; disposition of Information Systems (IS) media; and development and compliance with the risk managed approval of system operation [certification and accreditation] plans) will find this course useful and meaningful. Depending on the Command, Service, or Agency (C/S/A), the completion of this online course could help the student meet the DoD and C/S/A standards for Level 1 System Administrator certification.

In the near future, DISA's FSO plans to stand up a new server to host these WBT and CBT products on-line. If you are interested in learning more about the training products mentioned, please visit http://iase.disa.mil/eta. ■

About the Author

Walter C. Kelley

Walter C. Kelley is an Information Assurance Training Officer with DISA FSO's Projects and Plans Branch. He graduated from Valdosta State University in 1994 with a B.B.A. double major in Management and Marketing studies. Mr. Kelley may be reached at KelleyW@ncr.disa.mil.

Grid Computing

by John Killian

Author's Note

This report is the second in a series that examines new technologies that will likely have impact across DoD with implications for security professionals. This special report describes grid computing and the security implications associated with implementing, operating, and managing a grid computing environment. Technical security challenges involved with establishing a grid along with organizational, cultural, and policy issues between the disparate organizations which collaborate in a grid will be covered. A grid can be defined as flexible, secure, coordinated resource sharing across the network among dynamic collections of individuals, organizations, and resources. Grid computing introduces unique challenges related to authentication, authorization, resource access, and resource discovery. An open grid architecture is discussed in this paper in which protocols, services, application programming interfaces, and software development kits are described according to their roles in enabling resource sharing. Functional and security requirements that must be satisfied for proper implementation and use of a grid are identified and security implications are addressed. Additionally, grid technology leaders in academia, research and development, government and commercial industry are identified and examples of existing grid implementations are highlighted.

Overview

Grid computing is a hardware and software infrastructure, which consists of persistent environments that enable software applications to coordinate resource sharing and problem solving across the Internet. This collaborative infrastructure enables direct access to software, data, and hardware components managed by diverse organizations in widespread locations.

Grid computing differs from conventional distributed computing and massively parallel cluster computing by its focus on large-scale resource sharing amongst collaborative disparate organizations (see Figure 1 on page 18). As noted in *The Grid: Blueprint for a New Computing Infrastructure* [1], the original design of grid computing was intended to create an infrastructure for scientific and engineering computing environments which required large scale processing power.

Many organizations can benefit from grid computing and most share

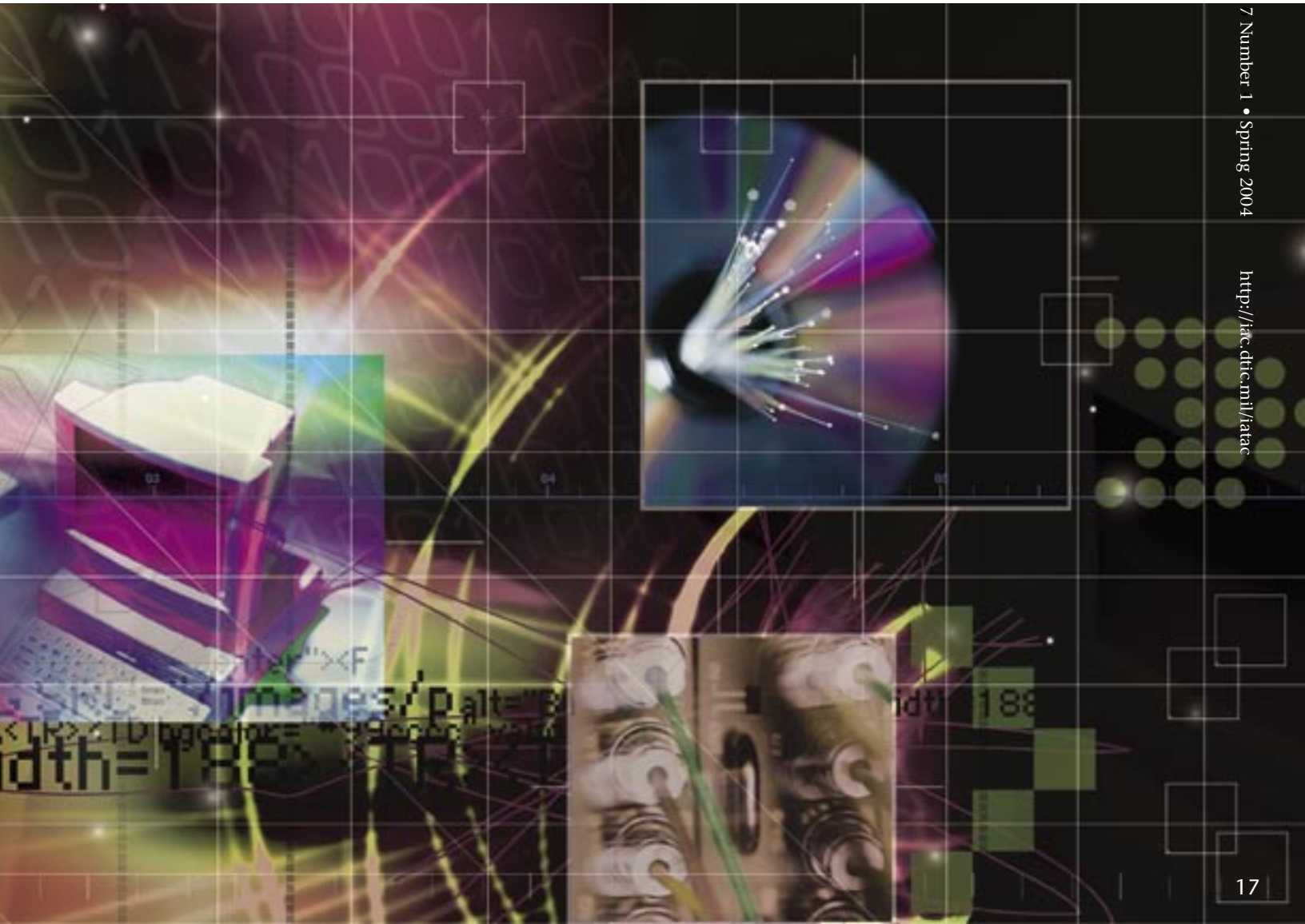
a common set of needs, although the specifics of their requirements may vary, (e.g., the number and type of participants, types of activities, duration and scale of interactions, and resources being shared). This commonality of general requirements allows some organizations to participate simultaneously in more than one grid, dynamically sharing some or all of their resources in parallel with multiple grid partners.

Resource sharing is conditional. That is, each resource owner makes resources available subject to constraints on when, where, and what operations may be carried out. Resource consumers may also place constraints on properties of the resources they are prepared to work with. The implementation of constraints requires mechanisms for expressing policies for establishing the identity of a consumer or resource (authentication) and for determining whether an operation is consistent with applicable sharing relationships (authorizations).

Sharing relationships may include both client-server resources and peer-to-peer components. Providers can be consumers and sharing relationships can exist among any subset of participants. They may be combined to coordinate use across many resources in which each is owned by a different organization. The ability to delegate authority in controlled ways becomes important in such situations, as do mechanisms for coordinating operations across multiple resources (see Figure 2 on page 21).

A key issue is the need to ensure that sharing relationships can be initiated among arbitrary parties allowing new participants to share dynamically across different platforms, languages, and programming environments. In this context, computing mechanisms serve little purpose if they are not defined and implemented to be interoperable across organizational boundaries, operational policies, and resource types. Without interoperability, grid applications and participants are forced to enter into one-on-one sharing arrangements. This introduces a problem that there is no assurance that the sharing mechanisms used between any two parties will extend to other parties. Without this assurance, dynamic grid party participation formation will be inconsistent, if not impossible. Just as the Web revolutionized information sharing by providing a universal protocol and syntax (Hyper Text Transfer Protocol [http] and Hyper Text Markup Language [HTML]) for information exchange, so Grid Computing requires standard protocols and syntaxes for general resource sharing.

To implement grid computing, developers will need to develop sophisticated applications in complex and dynamic execution environments. Users must be able to operate these applications. Application robustness, correctness, development costs, and maintenance costs are important concerns that will need to be addressed. Without standard protocols, interoperability cannot be achieved at the API level. Foster,



Kesselman, and Tuecke observe that the grid architecture emphasizes—

1. the identification and definition of protocols and services; and

2. Application Programmer Interfaces (API) and Software Development Kits (SDKs).

The grid challenge

The prevailing problem that exists with the grid concept is coordinated resource sharing and problem solving in dynamic disparate organizations. Collaborating organizations that participate in a grid are primarily focused on direct access to computers, software, data, and other resources as opposed to traditional data exchange which is the focus of conventional computer communication. This sharing is highly controlled, with resource providers and resource consumers defining clearly and carefully what is shared, who is allowed to share, and the conditions in which sharing occur. According to I. Foster, C. Kesselman, and S. Tuecke in their article "The Anatomy of the Grid" (International Journal of Supercomputer Applications, 2001), a set of individuals or organizations defined by these grid sharing rules form what is referred to as a virtual organization (VO).

For grid computing to become feasible, there must be an underlying infrastructure which provides functional and security services that satisfy common grid application concerns and requirements. VOs have a varying range of technical requirements and organizational needs. Highly flexible sharing relationships for client-server and peer-to-peer communication must be established to allow precise levels of control over how shared resources are used. Resource usages should be constrained by highly granular multi-stakeholder access control, authorization management, delegation, and local and global policies enforcement. Within the grid, the resources that are shared can range from executable code, files, data, computer components, sensors, and networks. For many grids, users have performance sensitive requirements and cost-sensitive connection needs so elements such as quality of service (QoS), scheduling, co-allocation, and accounting must be provided from within the grid environment.

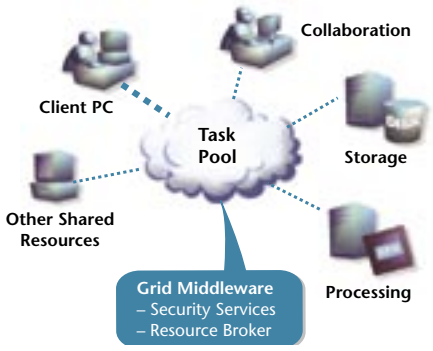


Figure 1. Grid Tasking [2]

Existing distributed computing technologies do not fully address the aforementioned concerns and requirements. Current Internet technologies such as business-to-business (B2B) environments address communication and information exchange among computers, but do not provide solutions for coordinated use of resources at multiple sites. In addition, as noted by A. Sculley and W. Woods in their book *B2B Exchanges: The Killer Application in the Business-to-Business Internet Revolution* [3], conventional transactional exchanges focus on information sharing primarily with centralized server architectures.

Traditional distributed technologies such as CORBA and J2EE allow resource sharing within a single organization. The Open Group's Distributed Computing Environment (DCE) supports secure resource sharing across sites, but most VOs would find it too burdensome and inflexible to be practical. Foster, Kesselman, and Tuecke in "The Anatomy of the Grid" describe how storage service providers (SSPs) and application service providers (ASPs) allow organizations to outsource storage and computing requirements to other parties, but only in a limited manner. Current existing technology does provide the range of resources required by grid computing, but does not provide the flexibility and control on sharing relationships required to make grid computing a reality.

Existing WWW technology–shortfalls for grid computing

The prevalence of Web technologies such as Internet Engineering Task Force (IETF) and World Wide Web Consortium (W3C) standard protocols—Transmission Control Protocol/Internet Protocol (TCP/IP), HTTP, Simple Object Access Protocol (SOAP), etc.—and languages, such as, HTML and eXtensible Markup Language (XML) makes them attractive as a platform for constructing grid computing systems and applications. However, while these technologies do an excellent job of supporting the browser-client-to-Web-server interactions that are the foundation of today's Web, they lack features required for the richer interaction models that occur in grid environments. An example of this deficiency is Web browsers today use TLS for encryption, but the browsers do not support single sign-on or delegation.

Steps can be taken to integrate grid and Web technologies. For example, the single sign-on capabilities provided in the GSI extensions to TLS would, if integrated into Web browsers, allow for single sign-on to multiple Web applications. GSI delegation capabilities would permit a browser client to delegate capabilities to a Web server so that the server could act on the client's behalf.

Application service providers (ASPs), storage service providers and similar hosting companies typically offer to outsource specific business and engineering applications (in the case of ASPs) and storage capabilities (in the case of SSPs). A customer negotiates a service level agreement that defines access to a specific combination of hardware and software. Security tends to be handled by using Virtual Private Network (VPN) technology to extend the customer's intranet to encompass resources operated by the ASP or SSP on the customer's behalf. According to Foster, Kesselman, and Tuecke, other SSPs offer file-sharing services, in which case access is provided via HTTP,

File Transfer Protocol (FTP), or Web-based Distributed Authoring and Versioning (WebDAV) with user ids, passwords, and access control lists providing identification authentication and authorization.

VPNs and static configurations make many grid sharing relationships difficult to achieve. The use of VPNs means that it is typically impossible for an ASP application to access data resources located on storage managed by a separate SSP. Similarly, dynamic reconfiguration of resources within a single ASP or SPP is very complex and is rarely attempted. The basic problem is that a VPN is not a VO. It cannot extend dynamically to encompass other resources and does not pro-

vide the remote resource provider with any control of when and how to share its resources.

Enterprise development technologies such as CORBA, Enterprise Java Beans, Java 2 Enterprise Edition, and Distributed Component Object Model (DCOM) are all systems designed to enable the construction of distributed applications. They provide standard resource interfaces, remote invocation mechanisms, and trading services for resource discovery and hence make it easy to share resources within a single organization. However, these mechanisms address none of the specific VO requirements previously mentioned. Sharing arrangements are typically static and restricted to occur within a single organization. The primary form of interaction follows the paradigm of client-server rather than the coordinated use of multiple resources.

Peer-to-peer computing (Napster, Gnutella, Kazaa, etc.) that implements file sharing systems and distributed Internet computing (SETI@home, Parabon, Entropia, etc.) are examples of more general sharing models. According to Foster, Kesselman, and Tuecke, these implementations have much in common with future grid computing architectures.

In practice, we find that the technical focus of work in the domains of enterprise development and peer-to-peer have not overlapped significantly to date. One reason for this is that peer-to-peer and Internet computing developers have been focusing on vertically integrated (stovepipe) solutions rather than seeking to define common protocols that allow for shared infrastructure and interoperability. Another reason for this is peer-to-peer file sharing often employs no access control and computational sharing usually depends on a single centralized server.

As applications become more sophisticated and the need for interoperability becomes more clear, we may see a more intense concentration of efforts between peer-to-peer, Internet, and grid computing communities. For example, analysts predict single sign-on, delegation, and authorization technologies will become important when data sharing services users demand more interoperability.

Technology leaders

There have already been some success stories in the development grid technologies and instantiation of grid implementations. One initiative that has fostered the production of useful grid tools, solutions, specifications and APIs is the Globus Project. In 1997, Globus produced the GSI. As described in previous sections of this paper, GSI uses public key protocols to aid developers in programming grid applications. Many lessons were gleaned from developing the GSI that assisted developers creating numerous production grid applications and middleware toolkits.

The Global Grid Forum (GGF) is a community-initiated forum of 5,000+ individual researchers and practitioners representing over 400 commercial, educational and governmental organizations working on grid technologies worldwide. GGF's primary objective is to promote and support the development, deployment, and implementation of grid technologies and applications via the creation and documentation of best practices, technical specifications, user experiences, and implementation guidelines.

GGF efforts are also aimed at the development of a broadly based Integrated Grid Architecture that can serve to guide the research, development, and deployment activities of emerging grid communities. Defining such architecture will advance the grid agenda through the broad deployment and adoption of fundamental basic services and by sharing code among different applications with common requirements.

Current and future technologies

Today there are a number of grid computing initiatives that are creating grids for a variety of applications and communities as well as projects developing grid technologies and technology frameworks. Several are listed here as examples.

- **DOE Science Grid**—The U.S. Department of Energy (DOE) has established the DOE Science grid. The DOE Science grid aims to provide an advanced distributed computing infrastructure based on grid middleware and tools to enable an enhanced degree of scalability in scientific computing necessary for DOE to accomplish its missions in science. The vision of this undertaking is to revolutionize the use of computing in science by making the construction and use of large-scale systems of diverse resources as easy as using today's desktop computing environments.
- **DOE DisCom2**—Another grid initiative facilitated by DOE is the DisCom2 (Distance and Distributed Computing and Communication) Grid. DisCom2 is developing technologies and infrastructure for efficient use of high-end computing platforms as geoph locations. The goals of DisCom2 include developing technologies and infrastructure for efficient use of high-end computing platforms across large distances and creating flexible distributed systems that can provide both enhanced capacity and capability computing.
- **NASA IPG**—NASA has established the Information Power Grid (IPG). This 20 year project aims to seamlessly integrate computing systems, data storage, specialized networks, and sophisticated analysis software to provide a steady, reliable source of computing power for NASA's future requirements.
- **NSF TeraGrid**—The U.S. National Science Foundation (NSF) has created the TeraGrid. Lead by Argonne National Laboratory, California Institute of Technology and the National Center for Supercomputing Applications, NSF is building a 14-TeraFlop distributed Linux cluster based grid system with a 40 Gigabit wide area network interconnect.
- **NCSA National Technology Grid**—NSF is also sponsoring the National Computational Science Alliance (NCSA) National Technology Grid. The NCSA is working to prototype a seamless, integrated computational and collaborative environment comprised of a computational grid (infrastructure-oriented) and an access grid (people-oriented).
- **EU EuroGrid**—On the international front the European Union (EU) is providing the Euro Grid. The EU sponsored this grid project to deploy a test bed among multiple European high performance computing laboratories, focusing on a suite of applications including capabilities in biomolecular simulation, weather prediction, coupled CAE simulations, structural analysis, and real-time data processing.

Research and development

As noted by Foster, Kesselman, and Tuecke, since 1996, research and development efforts within the grid community have produced protocols, services, and tools that meet

the challenges to build scalable VOs. These technologies include security solutions that support management of credentials and policies that span across organizations. Because of their focus on dynamic, cross-organizational sharing, grid technologies complement rather than compete with existing distributed computing technologies. For example, enterprise distributed computing systems can use grid technologies to achieve resource sharing across institutional boundaries. In the ASP/SSP space, grid technologies can be used to establish dynamic services for computing and storage resources which can overcome limitations of current static configurations. We discuss the relationship between grids and these technologies in more detail below.

Standards efforts

On January 20th, IBM, Akamai, The Globus Alliance, HP, IBM, Sonic Software, and TIBCO announced new Web Services specifications that will integrate grid and Web services standards. The new WS-Notification and WS-Resource Framework represent the first time a common standards-based infrastructure will be available for business applications, grid resources and systems management. These new specifications will help customers lower costs, speed deployment and, enable integration across and outside of the enterprise. These new specifications are important for key business applications and will provide customers with the ability to utilize a common Web services based infrastructure that supports grid based solutions.

Web Services standards have primarily focused on requirements for defining and managing networked services. WS-Notification proposes to also specify an agreed-upon definition for events. According to Derek Collison, vice president of Products and Technologies for TIBCO Software, "Events are what bring a service-oriented architecture to life and a standardized definition for events will accelerate delivery of real-time business to more companies at lower cost."

These new specifications provide a foundation for the Open Grid Services Architecture (OGSA). Using WS-Resource Framework and WS-Notification, grid infrastructures and applications can now be built using Web services specifications. This may facilitate customers' ability to share computing resources on demand over the Internet relying on an infrastructure that is resilient, self-managing and always available. Customers can integrate applications and share data and processing power with huge potential cost reductions and efficiency savings.

Inter-grid protocols

The grid architecture establishes standards and requirements for the protocols and APIs that enable sharing of resources, services, and code. It allows the ability to define multiple instantiations of key grid architecture elements. For instance, grid engineers can construct both Kerberos-and PKI-based protocols at the connectivity layer—and access these security mechanisms via the same API. This is enabled by use of the Generic Security Service (GSS) API. However, grids constructed with these different protocols are not interoperable and cannot share essential services.

For this reason, the long-term success of grid computing requires that one set of inter-grid protocols at the connectivity and resource layers (transport layer and link layer respectively of the IP stack) be selected and accepted. Then a widespread deployment of protocol set will need to occur. As Foster, Kesselman, and Tuecke observe, just as the Internet protocols will enable different computer networks to interoperate and exchange information, these inter-grid protocols will enable different organizations to interoperate and exchange or share resources.

The Globus Toolkit represents an approach that has had some success in developing these inter-grid protocols that need to be established to make grid computing a reality. It should be noted the development of these protocols will be a significant effort.

Security implications

The connectivity layer of the grid architecture that operates at the Internet and transport layers of the Internet Protocol (IP) stack defines the communication and authentication protocols required for grid transactions. Authentication protocols build on communication services to provide cryptographically secure mechanisms for verifying the identity of users and resources. Communication requirements for grid computing include transport, routing, and naming. Grid communications in the future will demand new protocols that take into account particular types of network dynamics.

Investigation into security engineering of grid computing has shown that there is a noticeable level of complexity in implementing security mechanisms. This grid security challenge dictates that any security solutions be based on existing standards whenever possible. As with the communication protocols, many of the security standards developed within the context of the Internet protocol suite are applicable within the grid environment.

In their article "Design and Deployment of a National-Scale Authentication Infrastructure" [4], J. Volmer, and V. Welch state that authentication solutions for grid environments need to provide several services for grid applications. One of these required services is Single Sign-On (SSO) that enables users to be able to "log on" (authenticate) just once and then have access to multiple grid resources without requiring further user intervention.

Another service is Delegation. According to M. Gasser and E. McDermott in "An Architecture for Practical Delegation in a Distributed System" [5], delegation allows a user to endow a program with the ability to run on that user's behalf so that the program is able to access the resources which the user is authorized. The program should also be able to conditionally delegate a subset of its rights to another program (sometimes referred to as restricted delegation). Each site or resource provider may employ any of a variety of local security solutions including Kerberos and Unix security. Grid security solutions must be able to interoperate with these various local solutions. They cannot actually require replacement of local security solutions, but must allow mapping into the local environment settings.

For a user to use resources from multiple providers together, the security system must not require that each of the resource providers cooperate or interact with each other in configuring the security environment. For example, if a user has the right to use sites A and B, then user should be able to use sites A and B together without requiring that A's and B's security administrators interact.

Future grid security solutions should also provide flexible support for communication protection (e.g.,

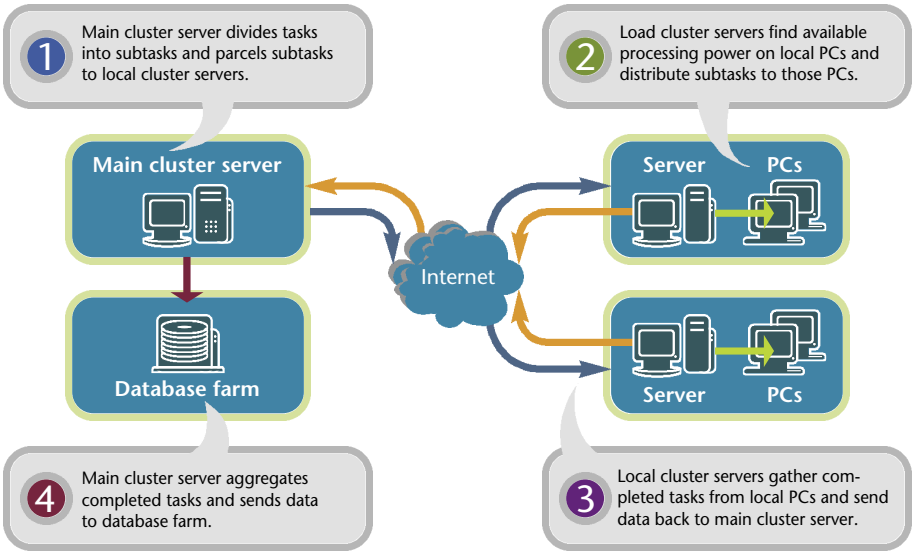


Figure 2. Grid computing: How it works. In the global grid computing scenario, unused processing power on local clusters of computers scattered across the Internet would be harnessed to address a single, complex application. [6]

control over the degree of protection, independent data unit protection for unreliable protocols, support for reliable transport protocols other than TCP) and enable stakeholder control over authorization decisions. This capability should include the ability to restrict the delegation of system privileges in various ways.

Security initiatives

Grid security infrastructure

Grid security infrastructure (GSI) uses public key cryptography as the basis for its functionality. Thus, GSI relies on the underlying authentication, encryption and data integrity services of a Public Key Infrastructure (PKI). PKI certificates suit the need to support security across organizational boundaries. Traversing across disparate domains normally prohibits employing a centrally managed security system. The problem of whom to trust limits the usefulness of PKI across organizational boundaries. PKI can integrate with SSO for users of the grid including delegation of credentials for computations that involve multiple resources and multiple sites. GSI protocols, in association with PKI services, are used for authentication, communication protection, and granting authorization.

GSI builds on and extends the Transport Layer Security (TLS) protocols to address single sign-on, delegation, integration with various local security solutions (including Kerberos), and user-based trust relationships. X.509-format identity certificates are used. Foster, Kesselman, and Tuecke describe how stakeholder control of authorization is supported via an authorization toolkit that allows resource owners to integrate local policies via a Generic Authorization and Access (GAA) control interface.

GSI uses the Secure Sockets Layer (SSL) and TLS for its mutual authentication protocol. Before mutual authentication can occur, the parties involved must first trust the CAs that signed each other's certificates. This means that they must have copies of the CAs' certificates—which contain the CAs' public keys—and that they must trust that these certificates really belong to the CAs.

GSI does not automatically establish confidential (encrypted) communication between parties. Once authentication is performed, GSI functions cease so communication can occur without the overhead of constant public key encryption and decryption. The GSI can easily be used to establish a shared key for encryption if confidentiality is required.

A related security feature of GSI is communication integrity. Integrity means that an unintended entity may be able to read communication between the two intended parties, but is not able to modify the communication. The GSI provides communication integrity by default. Communication integrity introduces some overhead in communication, but does not cause as much overhead as encryption.

GSI provides a delegation capability that is an extension of the standard SSL protocol that reduces the number of times the user must enter his passphrase. If a grid computation requires that several grid resources be used (each requiring mutual authentication), or if there is a need to have agents (local or remote) requesting services on behalf of a user, the need to re-enter the user's passphrase can be avoided by creating a proxy. A proxy consists of a new certificate (with a new public key in it) and a new private key. The new certificate contains the owner's identity, modified slightly to indicate that it is a proxy. The owner, rather than a CA signs the new certificate. The certificate also includes a time notation after which others should no longer accept the proxy.

Conclusion

There are many large-scale grid initiatives and projects currently underway. There has been development over the past few years that are making grid computing more feasible. Commercial interests are beginning to see business cases for investing into grid technologies that is adding synergy within the grid community. In the scientific and engineering sectors where collaboration with outside entities is part of the culture and indeed a necessity to survive, grid computing offers great benefits and return on investment. However, for most commercial sectors and general population, grid computing currently has some impediments to adoption and implementation.

At its most fundamental levels, grid computing involves creating or incorporating a set of additional protocols and services that build on Internet protocols and services that exist today. The current Internet alone does not easily facilitate the creation and use of computational power with shared disparate systems. Grid computing will be a mechanism for achieving greater resource utilization. It does not imply unrestricted access to resources, but it does enable controlled resource sharing. Resource owners will typically want to enforce policies locally and globally that constrain access based on group membership, ability to pay, and other characteristics. This means that an accounting capability is needed so the grid computing architecture must incorporate resource and collection protocols for exchanging usage and cost information allowing partici-

pating parties to determine if they should enable sharing.

Grid computing relies on the existence of infrastructure services such as authoritative data sources, Identity Management Systems, integrated Enterprise Directory Services, PKI, and SSO systems. Each of these infrastructure services are significantly complex, relatively expensive, can take significant time to implement and integrate with existing user systems and applications. All these interoperating infrastructure services can be difficult to manage. Coupled with the challenges involved with creating protocols and standards for the grid computing environment can present a difficult business case to justify the cost, level of effort, and risks.

Grid computing requires new programming models. Programming in grid environments introduces challenges that are not encountered in sequential or parallel computing environments. Some of these challenges include dealing with multiple administrative domains, new failure modes, and large variations in performance. However, It can be argued that these can be viewed as incidental, not central issues and that the programming Grid applications is not fundamentally different than current conventional programming.

Finally, the need to develop intergrid protocols (to make Grids work) that will enable interoperability among grid systems will most likely be a long and onerous task. Finding consensus through standards bodies and grid solution providers wanting to diverge and differentiate their product offerings can exasperate rapid development of these protocols. Determining what services should be present in a persistent fashion (rather than being duplicated by each application) to create usable grids and identifying what the key APIs and SDKs are that must be delivered to users in order to accelerate development and deployment of grid applications is also a need. This need adds more impediment to grid development and deployment. As Foster, Kesselman, and Tuecke observe, answers to these issues require further research and will take a significant amount of time and effort. ■

About the Author

John Killian

John Killian has over 24 years of IT experience including five years as an applications developer, 12 years as a network engineer, and seven years focusing in information security specializing in Public Key Infrastructure (PKI), PK-enabling (PKE) business process applications, and Web Services Security. Mr. Killian has supported the PK-enablement of production logistics application systems for the U.S. Army. In the past, he has co-facilitated the Army's PK-enabling Working Group, led the Army Digital Signature sub working group and assisted the Army Publications Directorate (APD) in enterprise digital signature requirements analysis in support of development of the Army electronic Forms Content Management Program (FCMP) system. Mr. Killian has a B.S. in Computer Studies from the University of Maryland, a certificate in Systems Engineering and currently works for Booz Allen Hamilton as a PKI and PKE Security consultant.

References

1. Foster, I and Kesselman, C., editors: The Grid: Blueprint for a New Computing Infrastructure (Elsevier/Morgan Kaufmann, 1999; ISBN: 1-55860-475-8). http://books.elsevier.com/us//mk/us/subindex.asp?maintarget=&isbn=1558604758
2. Lais, Sami: "Grid Computing" (Computerworld, 23 December 2003).
3. Sculley, A. and Woods, W., B2B Exchanges: The Killer Application in the Business-to-Business Internet Revolution; ISI Publications, 2000.
4. IEEE Computer, 33[12]:60-66; 2000, Butler, R., Engert, D., Foster, I., Kesselman, C., and Tuecke, S..
5. Proceedings of the 1990 IEEE Symposium on Research in Security and Privacy, pp. 20-30; IEEE Press, 1990.
6. Dr. Smith, Barry R., Birbeck College (London, UK): Grid Computing: What is the Grid? (http://people.cryst.bbk.ac.uk/~ubcg05s/grid/) and ALiCE Grid Computing Project, Department of Computer Science, National University of Singapore (http://www.comp.nus.edu.sg/~teoym/alice.htm)

Resources

Foster, I., Kesselman, C., and Tuecke, S.: "The Anatomy of the Grid" (International Journal of Supercomputer Applications, 2001). http://grid.lbl.gov/GPA/Anatomy.of.Grid.Foster,Kesselman,Tuecke.pdf

Foster, I.: "Internet Computing and the Emerging Grid" (Nature Web Matters, 2000). http://www.nature.com/nature/webmatters/grid/grid.html

Global Grid Forum (GGF). http://www.gridforum.org

Globus Project. http://www.globus.org

Brunett, S., Czajkowski, K., Fitzgerald, S., Foster, I., Johnson, A., Kesselman, C., Leigh, J. and Tuecke, S.: "Application Experiences with the Globus Toolkit" (Proceedings of the 7th IEEE Symposium on High Performance Distributed Computing, pp. 81-89; IEEE Press, 1998). http://www.isi.edu/~karlcz/papers/globus-apps.pdf

Tuecke, S., ANL: "Internet Draft: Grid Security Infrastructure (GSI) Roadmap" (February 2001). http://www.ggf1.nl/abstracts/SEC/draft-ggf-gsi-roadmap-02.pdf

Foster, I., Kesselman, C., Tsudik, G. and Tuecke, S.: "A Security Architecture for Computational Grids" (Proceedings of the ACM Conference on Computers and Security, pp. 83-91; 1998). http://www.csd.uch.gr/~hy555/globus/globus_security.pdf

product order form

Instructions: All IATAC **LIMITED DISTRIBUTION** reports are distributed through DTIC. If you are not a registered DTIC user, you must do so **prior** to ordering any IATAC products (unless you are DoD or Government personnel). **To register On-line:** http://www.dtic.mil/dtic/regprocess.html.
The *IAnewsletter* is **UNLIMITED DISTRIBUTION** and may be requested directly from IATAC.

Name _____ DTIC User Code _____

Organization _____ Ofc. Symbol _____

Address _____ Phone _____

_____ E-mail _____

_____ Fax _____

Please check one: ❑ USA ❑ USMC ❑ USN ❑ USAF ❑ DoD
❑ Industry ❑ Academia ❑ Gov't ❑ Other

Please list the Government Program(s)/Project(s) that the product(s) will be used to support: _____

_____

LIMITED DISTRIBUTION

IA Tools Reports (softcopy only)

❑ Firewalls ❑ Intrusion Detection ❑ Vulnerability Analysis

Critical Review and Technology Assessment (CR/TA) Reports

❑ Biometrics (soft copy only) ❑ Computer Forensics* (soft copy only) ❑ Configuration Management
❑ Defense in Depth (soft copy only) ❑ Data Mining ❑ Exploring Biotechnology
❑ IA Metrics (soft copy only) ❑ Network Centric Warfare
❑ Wireless Wide Area Network (WWAN) Security

State-of-the-Art Reports (SOARs)

❑ Data Embedding for IA (soft copy only) ❑ IO/IA Visualization Technologies
❑ Modeling & Simulation for IA ❑ Malicious Code

* You MUST supply your DTIC user code before these reports will be shipped to you.

UNLIMITED DISTRIBUTION

Hardcopy *IAnewsletters* available

Volumes 4		❑ No. 2	❑ No. 3	❑ No. 4
Volumes 5	❑ No. 1	❑ No. 2	❑ No. 3	❑ No. 4
Volumes 6	❑ No. 1	❑ No. 2	❑ No. 3	❑ No. 4
Volumes 7	❑ No. 1			

Softcopy *IAnewsletters* back issues are available for download at http://iac.dtic.mil/iatac/IA_newsletter.html

Fax completed form to IATAC at 703/289-5467

June

Transformation Technet 2004
June 8–10, 2004
Virginia Beach Pavilion Convention Center,
Virginia Beach, VA
http://afcea.org/transformation04/default.asp

Blue Force Tracking: Situational Awareness
June 29–30, 2004
Hamilton Crowne Plaza Hotel,
Washington, DC
http://www.idga.org/cgi-bin/templates/singlecell.html?topic=233&event=4816

10th Annual Gartner IT Security Summit
June 7–9, 2004
Marriott Wardman Park Hotel,
Washington, DC
http://www3.gartner.com/2_events/conferences/sec10.jsp

Annual IEEE Information Assurance Workshop
June 10–11, 2004
U.S. Military Academy, West Point, NY
http://www.itoc.usma.edu/workshop/2004/index.html

NetSec 2004
June 14–16, 2004
Hyatt Regency Embarcadero,
San Francisco, CA
http://www.gocsi.com/events/netsec.jhtml

Federal Information Security Conference (FISC) 2004
June 16–17, 2004
Antler`s Hotel, Colorado Springs, CO
http://www.fbcinc.com/eventasp?eventid=Q6UJ9A0078BN

July

2004 U.S. Department of Homeland Security Security Conference
July 28–29 2004
Sheraton Inner Harbor, Baltimore, MD
http://www.fbcinc.com/event.asp?eventid=Q6UJ9A007OXB

September

Gartner IT Security Summit 2004
September 20–21, 2004
Landmark Hotel, London, England
gg@delegate.com

InfowarCon 2003
September 30–October 3, 2004
Renaissance Washington DC Hotel,
Washington, DC
http://www.infowarcon.com/app/homepage.cfm?appname=100206&moduleid=451&campaignid=338&iUserCampaignID=718381

IATAC
Information Assurance Technology Analysis Center
3190 Fairview Park Drive
Falls Church, VA 22042

To change, add, or delete your mailing or E-mail address (soft-copy receipt), please contact us at the address above or call us at: 703/289-5454, fax us at: 703/289-5467, or send us a message at: iatac@dtic.mil