Can It Really Work? Problems with Extending EINSTEIN 3 to Critical Infrastructure[1]

Steven M. Bellovin[2], Scott O. Bradner[3], Whitfield Diffie[4], Susan Landau[5], Jennifer Rexford[6]

1 Introduction

Effectiveness should be the measure of any deployed technology. Does the solution actually solve the problem? Does it do so in a cost-efficient manner? If the solution creates new problems, are these easier and less problematic to handle than the original issue? In short, is the solution effective? In the rush to protect the U.S. after the 9/11 attacks, effectiveness was not always the primary driver in determining the value of the systems being proposed. In this context we consider the potential extension to the private sector of EINSTEIN 3, a federal program to detect and prevent cyber intrusions.

Providing services to the public is a fundamental role for U.S. federal civilian agencies, and beginning in the mid 1990s, many agencies turned to the Internet. This shift was not without problems. While confidentiality, integrity, and authentication dominated early federal thinking about computer and Internet security, the threats agencies faced included phishing, IP spoofing, botnets, denials of service (DoS), distributed denials of service, and man-in-the-middle attacks. Some exploits were done purely for the publicity, but others had serious purpose behind them. By the early 2000s, the growing number of attacks on U.S. civilian agency systems could not be ignored, and in 2004 the U.S. began an active effort to protect federal civilian agencies against cyber intrusions[7]. This

[1] The authors would like to thank Matt Blaze and John Treichler for various insights and suggestions in the writing of this paper, and we would also like to acknowledge multiple useful conversations with Sandy Bacik, Tahir El Gamal, and Vern Paxson. A shorter version of this paper will appear as *As Simple as Possible---But No Simpler,* COMMUNICATIONS OF THE ACM (August 2011).

[2] Professor, Department of Computer Science, Columbia University.

[3] University Technology Security Officer, Harvard University.

[4] Vice President for Information Security, ICANN and Visiting Scholar, Center for International Security and Cooperation, Stanford University.

[5] Elizabeth S. and Richard M. Cashin Fellow, Radcliffe Institute for Advanced Study, Harvard University.

[6] Professor, Department of Computer Science, Princeton University.

[7] DEPARTMENT OF HOMELAND SECURITY, NATIONAL CYBER SECURITY DIVISION, COMPUTER EMERGENCY READINESS TEAM (US-CERT), PRIVACY IMPACT ASSESSMENT EINSTEIN PROGRAM:

classified program, EINSTEIN, sought to do real-time, or near real-time, automatic collection, correlation, and analysis of computer intrusion information as a first step in protecting federal civilian agency computer systems[8].

EINSTEIN grew into a series of programs, EINSTEIN, EINSTEIN 2, and EINSTEIN 3, all based on *intrusion-detection* and *intrusion-prevention systems* (IDS and IPS). A network IDS monitors network traffic and reports suspected malicious activity, while a network IPS goes one step further by trying to automatically stop the malicious activity (e.g., by dropping the offending traffic or actively "fighting back" against the suspected adversary).

In the original effort, EINSTEIN intrusion-detection systems were to be located at federal agency Internet access points, the intent being to gather information to protect U.S. federal government networks. If traffic appeared "anomalous," session information would be sent to US-CERT, the U. S. Computer Emergency Readiness Team, a federal government clearing house for cyber intrusion information[9]. Hard as it may be to believe, prior to EINSTEIN, information sharing between federal civilian agencies on cyber attacks was purely on an ad hoc basis[10]. The original EINSTEIN effort was not very successful. EINSTEIN information sharing did not happen in real time, and the voluntary nature of the program meant that many agencies did not participate. The next version, EINSTEIN 2, required the participation of all U.S. federal civilian agencies. Then the project appeared to take an unusual turn.

In September 2007 the *Baltimore Sun* reported the National Security Agency was developing classified plans for protecting *private* communication networks from intrusion[11]. This was more than a bit contradictory --- a classified U.S. federal government program for protecting widely used private-sector systems --- but little information was available about this "Cyber Initiative."[12] The result was that public comment was limited. Then in January 2008 the Cyber Initiative became marginally better known. The Bush administration issued National Security Presidential Directive 54 establishing the Comprehensive National Cybersecurity Initiative (CNCI), a largely classified program for protecting federal civilian agencies against cyber intrusions. EINSTEIN was one aspect of CNCI that was made public, though large portions of the

COLLECTING, ANALYZING, AND SHARING COMPUTER SECURITY INFORMATION ACROSS THE FEDERAL CIVILIAN GOVERNMENT September 2004 at 3.

[8] DEPARTMENT OF HOMELAND SECURITY, NATIONAL CYBER SECURITY DIVISION, COMPUTER EMERGENCY READINESS TEAM (US-CERT), PRIVACY IMPACT ASSESSMENT EINSTEIN PROGRAM, September 2004, at 4.

[9] US-CERT collects information from federal agencies, industry, the research community, state and local governments, and sends out alerts about known malware; see http://www.us-certs.gov/aboutus.html.

[10] DEPARTMENT OF HOMELAND SECURITY, NATIONAL CYBER SECURITY DIVISION, COMPUTER EMERGENCY READINESS TEAM (US-CERT), PRIVACY IMPACT ASSESSMENT EINSTEIN PROGRAM: COLLECTING, ANALYZING, AND SHARING COMPUTER SECURITY INFORMATION ACROSS THE FEDERAL CIVILIAN GOVERNMENT September 2004 at 3.

[11] Siobhan Gorman, *NSA to defend against hackers: Privacy fears raised as spy agency turns to system protection,* BALTIMORE SUN, September 20, 2007, at 1A.

[12] Ibid.

program remained classified. So public understanding of EINSTEIN's intent, how it worked, what risks it raised, and what it protected remained limited.

In July 2010 the *Wall Street Journal* reported Raytheon had an NSA contract to study the value of sensors in recognizing impending cyberattacks in critical infrastructure cybernetworks[13]. Public reaction was swift and highly critical[14]. NSA responded with a statement that, "PERFECT CITIZEN [sic] is purely a vulnerabilities-assessment and capabilities-development contract. This is a research and engineering effort. There is no monitoring activity involved, and no sensors are employed in this endeavor."[15] While the project may have been solely a research effort, the idea of extending EINSTEIN-type protections to the private sector is increasingly being proposed by DC policy makers[16].

Extending an EINSTEIN-like program to the private sector raises serious technical and managerial issues. While federal civilian systems directly serve two million employees, critical-infrastructure systems in the U.S. serve a population of over three hundred million Americans. Scale matters. Can a program that effectively protects the communications of federal agencies with a hundred thousand employees really do the same for the communications giants that serve a hundred million people instead? The number of communications handled by these providers is dwarfed, however, by the number of communications of the "smart grid," the power network that will use digital technology to monitor and control power generation and usage.

Nor will size be the only problem in transitioning EINSTEIN systems from federal civilian agencies to the civilian sector. While the U.S. government can mandate the types of technologies used by federal agencies, the same is not typically true about systems used in the private sector. The fact that communications technologies are in a state of constant innovation further complicates such control.

Finally, expanding EINSTEIN-type technology to critical infrastructure is complicated by the complex legal and regulatory landscapes for such systems. Putting it simply, there are fundamental differences between communication networks supporting the U.S. federal government and those supporting the private-sector critical infrastructures. These differences create serious difficulties in attempting to extend EINSTEIN-type technologies beyond the federal sector. Such issues appear to be ignored by policy pundits who are in a headlong rush to protect critical infrastructure.

While few doubt the value of IDS and IPS as part of a cybersecurity solution, can EINSTEIN really work? What attacks does EINSTEIN prevent? What does it miss?

[13] Siobhan Gorman, *U.S. Plans Cyber Shield for Utilities, Company*, WALL STREET JOURNAL, (July 8, 2010).

[14] Ryan Singel, *NSA Denies It Will Spy on Utilities*, WIRED (July 9, 2010), http://www.wired.com/threatlevel/2010/07/nsa-perfect-citizen-denial/.

[15] Ibid.

[16] See, for example, Nicholas Hoover, *Cyber Command Director: U.S. Needs to Secure Critical Infrastructure,* INFORMATION WEEK (Sept. 23, 2010), http://www.informationweek.com/news/government/security/showArticle.jhtml?articleID=227500515.

How good is EINSTEIN as a security solution? What are the privacy implications of using the interception program? This paper is an attempt to provide answers to these questions, answers that are quite important in view of efforts to expand EINSTEIN beyond its original mandate.

We begin by presenting the EINSTEIN architecture in section 2. In section 3, we discuss the technical and policy concerns raised by the use of EINSTEIN 3 by federal civilian agencies. We observe that the current EINSTEIN deployment across the federal sector raises privacy and security concerns and propose changes in policy to alleviate these. In section 4, we examine two critical infrastructures, the power grid and information and communications technologies (ICT). Our observations are that while critical infrastructure should, of course, deploy intrusion detection and intrusion prevention systems, the consolidation and real-time information sharing model central to the EINSTEIN 3 cannot effectively migrate to these private-sector systems. In section 5 we conclude by proposing ways to protect the cybernetworks of ICT and the power grid, and present guiding principles for securing critical infrastructure's cybernetworks.

2. EINSTEIN 3 Architecture

The CNCI goals were protecting against current cybersecurity threats and anticipated more sophisticated future ones attacks[17]. CNCI involved a dozen initiatives, the first being to manage the federal enterprise network as a single network. EINSTEIN was part of this, as was the Trusted Internet Connections (TIC), a program that by consolidating federal connections to the public Internet would help ensure that these were professionally protected [18].

Under the TIC program, federal civilian agencies use TIC Access Providers (TICAPs) to operate the TICs. Large federal agencies would use a few TICs (generally two to four) while small agencies would share TICs. Some agencies have been certified as capable of acting as their own TICAP but most will have to seek service from an approved TICAP[19]. The reduction in external access points, from a few thousand to around a hundred, was crucial to the EINSTEIN 2 and EINSTEIN 3 efforts.

EINSTEIN 2 uses devices located at TICs to monitor traffic coming into or exiting from government networks. Located at the agency's TICAPs[20], the EINSTEIN 2 sensors would collect communications session data; this could include packet length, protocol, source and destination IP address and port numbers, timestamp and duration information of

[17] COMPREHENSIVE NATIONAL CYBERSECURITY INITIATIVE - http://www.whitehouse.gov/cybersecurity/comprehensive-national-cybersecurity-initiative

[18] Ibid.

[19] DEPARTMENT OF HOMELAND SECURITY, US-CERT/ISS LOB, Trusted Internet Connections (TIC) Initiative --- Statement of Capability Evaluation Report, June 4, 2008.

[20] DEPARTMENT OF HOMELAND SECURITY, US-CERT/ISS LOB, Trusted Internet Connections (TIC) Initiative --- Statement of Capability Evaluation Report, June 4, 2008 at 10.

communications to/from federal civilian agencies[21]. The EINSTEIN 2 sensors would alert US-CERT whenever traffic matching *signatures*, patterns of known malware (e.g., the IP address of a server known to be hosting malware or an attachment known to include a virus), were observed in incoming packets[22]. Agency participation lagged, and EINSTEIN 2 was made mandatory for federal agencies[23].

To strengthen protections, EINSTEIN 2 is configured to do real-time detection of patterns of anomalous communications behavior. Doing so requires observing large volumes of traffic so that the anomaly detector is able to develop a model of what "normal" traffic looks like. One of the purposes of consolidation was to provide sufficient data within each Internet connection for the EINSTEIN boxes to study[24].

The third effort, EINSTEIN 3, will change the game from intrusion *detection* to intrusion *prevention*. Intrusion prevention systems devices will be located at the agency TICAPs, which will redirect traffic destined to or from the U.S. federal government network through the EINSTEIN 3 device without affecting other traffic (that is, without affecting communications not destined for U.S. federal government networks)[25]. As of this writing, EINSTEIN 3 is in preliminary stages, having been tested only on at a single medium-sized federal civilian agency[26]. Initially EINSTEIN 3 will recognize cyber threats by analyzing network traffic to see if it matches known signatures[27]. Commercial IPSs will develop signatures to be used in their devices, and it is reasonable to expect that the government will work out a way to use these. Commercial IPSs respond to threats in two ways: by discarding suspect traffic before it reaches its destination and by sending carefully crafted messages to the source of the threat.

The aim of EINSTEIN 3 is "to automatically detect and respond appropriately to cyber threats before harm is done"[28]; EINSTEIN 3 devices will perform *deep packet inspection*, examining not only transactional information but also packet content. A communications-

[21] DEPARTMENT OF HOMELAND SECURITY, NATIONAL CYBER SECURITY DIVISION, COMPUTER EMERGENCY READINESS TEAM (US-CERT) Department of Homeland Security, National Cyber Security Division, Computer Emergency Readiness Team (US-CERT), PRIVACY IMPACT ASSESSMENT EINSTEIN PROGRAM, September 2004, at 6-7.

[22] COMPREHENSIVE NATIONAL CYBERSECURITY INITIATIVE - http://www.whitehouse.gov/cybersecurity/comprehensive-national-cybersecurity-initiative

[23] DEPARTMENT OF HOMELAND SECURITY, COMPUTER EMERGENCY READINESS TEAM (US-CERT), PRIVACY IMPACT ASSESSMENT FOR EINSTEIN 2, May 19, 2008, at 3.

[24] CLAY JOHNSON III, "Memorandum for the Heads of Executive Departments and Agencies," M-08-05, Office of Mgmt. & Budget, Exec. Office of the President (November 20, 2007).

[25] DEPARTMENT OF HOMELAND SECURITY, COMPUTER EMERGENCY READINESS TEAM (US-CERT), PRIVACY IMPACT ASSESSMENT FOR THE INITIATIVE THREE EXERCISE, March 18, 2010.

[26] Personal communication with Susan Landau (September 1, 2010).

[27] DEPARTMENT OF HOMELAND SECURITY, COMPUTER EMERGENCY READINESS TEAM (US-CERT), PRIVACY IMPACT ASSESSMENT FOR THE INITIATIVE THREE EXERCISE, March 18, 2010 at 5.

[28] COMPREHENSIVE NATIONAL CYBERSECURITY INITIATIVE - http://www.whitehouse.gov/cybersecurity/comprehensive-national-cybersecurity-initiative

interception analogy would be that EINSTEIN 2 behaves somewhat like a trap-and-trace device[29], while by collecting content of the communications, EINSTEIN 3 functions somewhat like a wiretap[30]. The analogy is not perfect, however, since EINSTEIN 3 will disrupt communications believed be carrying malware (in contrast, wiretaps simply record).

By limiting the number of access points, the TICs concentrate the data, enabling a better search for "clues" about anomalous behavior. This improves the likelihood of discovering new threats. The limited number of access points makes it potentially feasible to establish a program of monitoring and intervention for *all* federal civilian agency access to the public Internet, and also limits the cost of the EINSTEIN effort both in terms of capital cost (e.g., fewer EINSTEIN boxes) and in operational expenditures (fewer people required to manage the system).

Initial concerns about the EINSTEIN effort focused on privacy threats raised by the project. Because EINSTEIN IDSs and IPSs would operate on all traffic destined for federal networks, the system would intercept private communications of federal employees (e.g., if a federal employee used an agency computer to check a private email account during lunch). In this, a federal employee is not different from employees at regulated industries using company-supplied equipment for personal communications; they, and the people with whom they communicate, are also subject to company monitoring.

We now turn to examining technical and policy issues raised by the EINSTEIN architecture.

3. Technical and Policy Concerns Raised by the Einstein 3 Architecture

To understand EINSTEIN's effectiveness, one needs to look at the architecture and the numbers. But the public EINSTEIN documents very limited details, so instead we start by doing some thought experiments---not inappropriate for a technology named EINSTEIN. Consider the technical complexities of a centralized IDS/IPS system serving multiple federal civilian agencies with two million users, which include:

- Scale: Denial-of-Service (DoS) attacks can be daunting; they have been measured at 100 Gb/s. It is unlikely that the current generation of any network device would

[29] A trap-and-trace device captures the transactional information of an incoming communication; in the case of a phone call, this would be the phone number. A trap-and-trace device does not capture content.

[30] These analogies are not exact. For one thing, EINSTEIN 2 and EINSTEIN 3 devices scan only a subset of communications. Minimization consists of singling out communications matching previously determined patterns or exhibiting anomalous behavior. More significantly, wiretaps do not prevent the occurrence of communications in which there is evidence of criminal activity, but the EINSTEIN 3 devices will do so. We note that because both EINSTEIN 2 and 3 are used only on communications to/from federal civilian agencies, these interceptions are not considered electronic surveillance from a legal point of view.

be able to resist the DoS attacks at this rate --- let alone new attack rates likely in the near future. Indeed, it is likely that new DoS attacks will be developed using the IDS/IPS monitoring functionality to disrupt otherwise legitimate traffic.

- Ability to do Correlation: Correlation is about discovering previously unknown threats in real time as they appear. But this is impossible to do in all but very small networks. The crux of the issue is that no one knows how to use a percentage of the traffic---whether compressed, diarized[31], or sampled---to characterize arbitrary new threats. If one is hoping to deter all threats (and not just previously known ones), *all* incoming data must be correlated and analyzed.

  Think of potential correlation solutions as having two variables: architectures can range from highly "centralized" to fully "decentralized" and sensors can be "smart" or "dumb," that is, having the ability to do lots of computation locally, or not.

  If analysis is done locally at the data collection point, then the need to see all incoming data requires that *all* raw signals be sent to *all* sensors. This quickly becomes unmanageable. If there are n sensors, then each sensor must look at the data from (n-1) other sensors, and there are n(n-1)/2 pairs of data traversing the network. This is simply unmanageable when n is at all large (EINSTEIN is designed to have between one and two hundred). Note that this solution also introduces a new problem: protecting the sensors, which would carry security-sensitive information.

  An alternative approach would be to centralize the data to perform the correlation. Because summarizing the data cannot solve the problem, all the data must travel through the system to the centralized detector. (We note that in an IP-based environment, the packet summary information is 20-30% of the data, which means that summarizing does not provide savings in the same scale that it would for telephone communications.) This is both enormously costly for a network of any scale, as well as unable to provide the millisecond response needed in a serious attack.

  (Of course, one could try a middling solution: neither fully decentralized nor fully sharing signals. Depending on where one puts the set points, the problems above will still occur.

[31] "Diarize" is used by the trade to mean making a diary of the data; in the case of a telephone call, this might be the to/from, time, and length of the call, while for IP communications, this would be the metadata of source and destination IP addresses, TCP source and destination ports, and perhaps length of packet.

The two alternative solutions---dumb sensors and decentralized architectures or smart sensors and centralized architectures---have the worst of both worlds: they would either miss the problems, or involve enormous investment. Neither are viable.)

In short, correlation at the scale and speed at which a system serving two million users is expected to operate is not achievable using common production technology.

- Device Management: The devices will need periodic updates. Protecting IDS/IPS control mechanisms and pathways against intrusion, disruption, modification and monitoring will be very challenging.

- Signature Management: Signatures are likely to be a mix of classified signatures developed by the government and unclassified signatures from commercial IDS and IPS vendors. These will have to be protected from those operating the IDS/IPS systems as well as from Internet-based attackers.

- Data security: Network communications are increasingly encrypted through company VPNs, etc.; in some cases federal regulations require the use of encryption (e.g., in sharing medical records). In order for the IDS/IPS systems to prevent malware from reaching end users, communications transiting the IDS/IPS must be decrypted. This introduces a new vulnerability, and makes the IDS/IPS a particularly ripe place for attack.

The above are issues for *any* IDS/IPS system centralizing monitoring and protection function through few pipes.

Now consider EINSTEIN. The Trusted Internet Connections initiative, which supports EINSTEIN, will ensure that all communications between federal civilian agencies and the Internet---both those by generated by people and those by services---occur via managed connections. Since some government agencies exchange very large quantities of research data with their partners in the private sector---data sets on the order of terabytes---some connections involve quite high bandwidth. With that in mind, we return to our thought experiment, this time supplying some numbers.

- Scaling is a Problem: Although the actual performance of the EINSTEIN 3 device is not public, the cost impact of requiring a significant amount of real-time monitoring of Internet streams can be illustrated by examining a “typical” case based on the speed of products publicly available. We begin by noting that a fully realized TIC program to minimize the number of interconnect points, the number will be more than a hundred and maybe in the low hundreds.

Consider a single shelf Cisco CRS-1 router of the type used in both Internet backbones and to aggregate traffic from local networks before sending it to the Internet. According to Cisco's press releases, more than 5,000 of these routers have been sold and deployed. When fully loaded, the CRS-1 will accept 64 10 Gb/s duplex communications links, operating at a total bit rate of 1.28 terabits/second. While some routing nodes are smaller, some are much larger, using a number of CRS-1s connected together to handle the required load.

While neither the exact nature of the algorithms planned for EINSTEIN 3 nor the equipment configuration planned for it have been disclosed, it is reasonable to assume as a model that the computation required for performing the IDS/IPS function at a federal civilian agency will be similar to that in commercial network defense products built and sold by Narus, Cloudshield, and others. It seems highly unlikely that a single EINSTEIN 3 device can run sufficiently fast to monitor the high-speed connections between some of the federal civilian agencies and the Internet or private sector agency partners. There are obviously differences in the details of the various industry products, but a review of their specifications reveals that a unit capable of examining, in real time, 20 Gb/s of Internet traffic costs about $80K and consumes about 2 kW (and another 2 kW for cooling). Scaling this up to handle the full rate entering and leaving a single half-rack CRS-1 would therefore require 64 such units, leading a cost of roughly $5M, roughly 250 kW of power consumption, and roughly 32 equipment racks.

This has two important implications: (i) each router used for directing traffic will require 64 times as much equipment to perform EINSTEIN-type security---clearly a losing battle---and (ii) the EINSTEIN program, at least the instantiation of EINSTEIN 3, would be roughly one billion dollars solely for equipment costs.

- Device Management: Installed in TICAPs, many of the EINSTEIN devices will be in non-government facilities, but will need to be remotely controlled by US-CERT. Ensuring that the control mechanisms and pathways are protected against intrusion, disruption, modification and monitoring will be challenging. Ensuring that such control paths are isolated from the Internet is likely to be a minimum requirement, but history has shown that isolated systems sometimes do not stay isolated[32]. And, as the Stuxnet case so vividly demonstrates, even isolated systems can be vulnerable to attacks[33].

EINSTEIN 3 devices are not designed to work autonomously. They are designed to be managed by, and report to, one or more control systems. A number of large Internet service providers (ISPs) and large enterprise networks have developed procedures and control systems to provide secure management of multiple network devices, such as routers or firewalls. Due to the requirement of being able to quickly determine an attack is underway, and to react to that attack by reconfiguring other Einstein devices, the management requirements for EINSTEIN devices is likely to be far more dynamic than what is required for current ISP or enterprise network devices. Developing the tools needed to manage the EINSTEIN 3 devices may turn out to be a significant technical challenge.

- Introducing New Points for Attack: Indeed, once it becomes known that EINSTEIN 3 devices are being deployed, it is likely that new types of DoS attacks will be quickly developed specifically to interfere with EINSTEIN monitoring functionality. It is unlikely that the current generation of any network device, including the EINSTEIN devices, would be able to resist DoS attacks at the rate that have already been seen. One concern is that such DoS attacks could be mounted to divert attention from, or to mask, low speed but carefully targeted penetration attacks on the protected infrastructure. Overwhelming the EINSTEIN device or overwhelming the analysis systems supporting such devices may become an effective approach---and one that may be very hard to thwart.

  Intrusion prevention systems that automatically decide when they should block or interfere with communications are quite problematic. If an adversary knew that such a system was in place, it would be an obvious attack point. The adversary could craft traffic that was designed to trigger the protection functions. This could cause the device to block traffic when it did not need to, making the protection device an operational part of a denial-of-service attack. Complex systems such as EINSTEIN 3 provide more attack points, and the analysis and control systems required for the EINSTEIN 3 installation may be more susceptible to this type of attack than current simpler IDSs and IPSs.

[32] According to Bruce Schneier, it took the Melissa virus just twenty-four hours to spread from the Internet to classified Department of Defense networks (Bruce Schneier, *Govnet*, CRYPTO-GRAM NEWSLETTER (November 15, 2001), https://www.schneier.com/crypto-gram-0111.html.

[33] John Borland, *A Four-Day Dive into Stuxnet's Heart*, WIRED (December 27, 2010).

To avoid the vulnerability created by automatically interfering with traffic, it may be better to have such decisions reviewed by a human operator. The advantage is that, assuming the human operator has the proper level of expertise, she is less likely to incorrectly interfere with non-threatening traffic; the disadvantage is that the response time is far slower. And there also may not be enough trained experts available to perform such reviews for the very large system of EINSTEIN boxes (and the resulting high data rate).

- Is correlation feasible? As we have already noted, correlation, particularly at the scale and speed at which EINSTEIN 3 is expected to operate, is simply not achievable using common production technology.

- Complexity of combining classified and non-classified signatures: Both classified and unclassified signatures will be used for intrusion detection[34]. As already noted, some signatures that EINSTEIN 3 will use will be developed by the government and will be classified while others are likely to come from commercial IDS and IPS vendors. The protection of classified signatures and the protection of any captured network traffic will be a challenge for the EINSTEIN devices located in the TICAPs, particularly for the commercial providers. The signatures will have to be protected from the TICAP operator and from Internet-based attackers. The latter is particularly important since knowing what the EINSTEIN device is looking for would simplify an attacker's approach.

- Security implications of EINSTEIN within the federal system: Internet traffic is increasingly encrypted[35], and many government web sites also offer encrypted services (e.g., the IRS). Furthermore government employees will be accessing non-government encrypted services on a regular basis (e.g., banking sites). The current set of public EINSTEIN 3 documents do not discuss how EINSTEIN 3 will handle encrypted traffic, however. Policies will need to be developed. One option would be for the EINSTEIN devices to ignore the contents of encrypted traffic, but that would provide an unmonitored attack pathway. Devices that are in the communications path, such as EINSTEIN 3, can be designed to mimic cooperating web sites (by using those websites' identities and credentials) to both expose the encrypted traffic to EINSTEIN 3 and permit that traffic to be stored.

[34] Ibid.

[35] For example, Google recently made encrypted access the default for many of its applications (Sam Schillace, *Default https access for Gmail*, Gmail Blog, January 12, 2010, http://gmailblog.blogspot.com/2010/01/default-https-access-for-gmail.html; *Search more securely with encrypted Google web search,* The Official Google Blog, May 21, 2010, http://googleblog.blogspot.com/2010/05/search-more-securely-with-encrypted.html).

> These policies should be openly developed to ensure that the public understands the implications of the EINSTEIN 3 system.

These technical complexities make it highly unlikely that EINSTEIN can accomplish the purposes for which it is being designed.

EINSTEIN 3 also raises policy concerns. Any IDS looking for long-term subtle attacks must store large amounts of traffic for non real-time analysis. This data could also be useful in tracking down wrongdoing by government employees or people with whom they communicate. Even if current EINSTSEIN 3 software is not designed for such analysis, the system is likely to store data that the government agencies might like to use---and there is also danger of misuse. Thus it is imperative that a detailed log is generated for all functions that the EINSTEIN 3 device has been configured to do.

Policies will have to be developed to detail legitimate uses of the EINSTEIN 3 devices. The only way to ensure, however, that such policies are followed is to produce detailed logs that cannot be altered. Logs must be out of reach of people who might misuse the EINSTEIN 3 devices, and these must be regularly and automatically scanned to reveal unexpected activities. Given the technology's potential for tracking individuals, policies should be developed to enable access to the logs if questions arise about how the EINSTEIN 3 devices are being used. There should be regular scrutiny of these logs by agency Inspector Generals.

Extending EINSTEIN 3 to non-government critical infrastructure would require similar policy development. We now turn to discuss this, as well as the technical fit of EINSTEIN 3 in critical infrastructure.

4. Expanding Einstein Capabilities to Critical Infrastructure

Certain critical infrastructures such as ICT and the electric power grid are essential not only to the running of society, but also to the functioning of the U.S. government, and thus the federal government has a direct vested interest in the security of the computer networks supporting these infrastructures. But direct vested interest does not mean that the federal government can force its solution onto the private sector. The fact that private industry controls 85% of critical infrastructure[36] means that the situation is not straightforward. In fact, it is far from straightforward.

We begin by discussing the general issues involved in performing real-time intrusion detection and intrusion prevention on a nation-wide scale. We then consider two critical

[36] U. S. GOV'T ACCOUNTABILITY OFFICE, GAO-07-39, CRITICAL INFRASTRUCTURE PROTECTION: PROGRESS COORDINATING GOVERNMENT AND PRIVATE SECTOR EFFORTS VARIES BY SECTORS' CHARACTERISTICS (2007) at 2.

infrastructures --- ICT and the power grid --- in some detail. In this discussion, we are assuming the approach to be the full EINSTEIN architecture, that is: TICAPs with cross-site correlation and an automatic reaction to anomalous events. Our critiques follow from there.

The Complexities of Information Collection

The EINSTEIN architecture forces a limited number of federal civilian agency access points to the Internet. In the federal sector this reachability to a limited number of access points was not particularly difficult to achieve or enforce. However much various federal agencies might clash with one another for responsibilities and resources, ultimately these agencies serve the same customer. Even if agencies A and B compete in some spheres, it is perfectly reasonable to expect they would cooperate in enabling real-time correlation of transactional information to find that U.S. sites are being attacked.

To provide EINSTEIN-type protection in the private sector would require similar coalescing of connections to the public Internet. It is far more difficult to imagine a collaboration model here. Many suppliers of critical infrastructure are genuine competitors. The manager of an EINSTEIN device has control over the communications that run through the device. Who would run the EINSTEIN devices for competing companies? Putting company A in the control seat of connections to the public Internet makes it very powerful. Would its competitor B willingly use the services of a TICAP osted at A? Even though B should encrypt its communications end-to-end, there are any number of nefarious activities that A might employ to impede its competitors, including using the IDS/IPS to throttle the communications of company B? Even short communications delays can have massive impacts for companies[37]. Would B have to pay A for its services?

A related issue is device management. Because EINSTEIN 3 devices store classified signatures, control of the private-sector systems should be handled under the aegis of the federal government (and specifically by the agency supplying the signatures). Such a solution presents myriad complexities, and the history of real-time data sharing between the private and public sector has not been a positive one.

In 1998 Presidential Decision Directive 63 (PDD-63) made protection of critical infrastructure a national objective. Since then public-private partnerships have been recommended, created, and failed, only to be re-recommended, re-created and fail again. The 1998 PDD-63 created Information Sharing and Analysis Centers (ISACs)[38], but was

[37] Peter Svensson, *Comcast Blocks Some Internet Traffic*, Internet on msnbc.com (October 19, 2007 at 09:36:11 ET) http://www.msnbc.msn.com/id/21376597/.

[38] For example, the IT-ISAC was created by the IT industry for systematic sharing and exchange of information on "electronic incidents, threats, attacks, vulnerabilities, solutions and countermeasures, best security practices and other protective measures" and includes such industry leaders as CA, Computer

superseded in 2003 by Homeland Security Presidential Directive 7, which made DHS responsible for coordinating plans for protecting critical infrastructure. This included developing plans for coordinating public-private partnerships. In 2006, DHS issued a National Infrastructure Protection Plan with public/private partnerships with two councils for each sector, a government one and a private-sector one, to handle planning and coordination. The issue of public-private partnerships arose again in 2009 with the 60-day *Cybersecurity Review*[39] conducted at the behest of President Obama.

In 2010, the Government Accountability Office reviewed public-private partnerships, and concluded, "Federal partners are not consistently meeting private sector expectations, including providing timely and actionable cyber threat information and alerts, according to private sector stakeholders"[40]. Problems included a lack of timely information, a lack of access to secure settings in which to exchange private information, and a lack of "one-stop" shopping---one federal office from which to find out information[41]. This does not bode well for private-sector use of EINSTEIN-type systems.

One example of the type of issues that would arise is signature collection. How would signatures amassed by private parties, e.g., the critical infrastructures themselves---or the companies with which they contract---be added to the EINSTEIN devices? Concerns run from mundane issues of whether signature formats will be public to knotty management concerns. Because private parties would not control the EINSTEIN devices, presumably they would not be able to directly add signatures to the IDS and IPS. This would have the counterproductive effect of removing private companies from the process of protecting their own customers. Such lack of direct control will create various problems, and would, at a minimum, create delay in adding signatures found by the private companies onto the EINSTEIN devices.

The issue of control runs deeper. Most private sector systems currently already run IPS and IDS on their networks. If EINSTEIN-type systems were deployed on their communications networks, what would happen to the systems currently in use? A possible solution would have communications relayed through two IDS/IPS systems, one supplied by the federal government, one by the company involved. The problems with this "solution" are clear.

Another issue arises from deployment. U.S. telecommunications infrastructure extends outside U.S. territorial limits. Using EINSTEIN boxes at foreign endpoints creates serious security problems for the technology---how would classified signatures be

Sciences Corporation, IBM, Intel, Juniper Networks, Microsoft, Oracle, Symatec, and Verisign (see https://www.it-isac.org/aboutitisac.php).

[39] *Cyberspace Policy Review: Assuring a Trusted and Resilient Information and Communications Infrastructure,* May 2009.

[40] United States Government Accountability Office, CRITICAL INFRASTRUCTURE PROTECTION: KEY PRIVATE AND PUBLIC CYBER EXPECTATIONS NEED TO BE CONSISTENTLY ADDRESSED, Report to Congressional Requesters, July 2010 at 13.

[41] United States Government Accountability Office, CRITICAL INFRASTRUCTURE PROTECTION: KEY PRIVATE AND PUBLIC CYBER EXPECTATIONS NEED TO BE CONSISTENTLY ADDRESSED, Report to Congressional Requesters, July 2010 at 14.

protected in such an environment?---while placing the boxes where cables enter the U.S. is simply not viable; a single modern cable carries about two or more terabits/second[42] and each incoming cablehead hosts several cables. EINSTEIN cannot cope with such numbers.

The distributed control between government and the private sector also raises legal concerns. Who bears fiscal responsibility for attacks that occur from known problems---ones that the private entities had uncovered---but that had not yet been added to the system? Distributed control leaves gaps, including the issue of who would bear responsibility for attacks that neither the U.S. federal government nor the private entities had yet uncovered. In mandating an EINSTEIN-like system be used on a private network, would the federal government indemnify the owners if cyberattacks occurred?

Privacy becomes a much greater concern were EINSTEIN technology to be extended from federal systems to the private sector. EINSTEIN 2 collects and retains transactional information in order to check for anomalous patterns. The collection includes packet length, protocol, source and destination IP address and port numbers, information already shared with Internet routers. In *Smith*[43], the Supreme Court ruled that information such as dialed numbers shared with third parties do not require government investigators to obtain a warrant. Thus extending EINSTEIN 2-type technology to the private sector might not invoke Fourth Amendment protections[44].

EINSTEIN 3 is another matter. This technology would scan and analyze not just metadata, but content. Information would be stored on suspicion of being malware, not on the knowledge that it is so. Harvard law professor Jack Goldsmith has argued that using EINSTEIN-type technologies to monitor communications for malware is akin to conducting "non-law-enforcement searches without individualized suspicion in numerous contexts," and cited highway checkpoints and inspections of regulated businesses as precedent for such monitoring sans warrants[45].

Communications form a special class however. Wiretap warrants require a higher standard of proof than standard search warrants. Goldsmith proposes handling potential invasiveness of an EINSTEIN-type system with "significant use restrictions" on the communications stored through EINSTEIN, limiting the set of crimes for which a sender

[42] Since most video is on national networks, this is almost entirely voice and data. There is very little video in cross-border or undersea cables.

[43] *Smith v. Maryland,* 442 U.S. 735 (1979).

[44] Whether some of this data (e.g., IP address) is subject to Fourth Amendment protection is currently under litigation; see, e.g., Brief for Jacob Applebaum, Birgitta Jonsdittor and Rap Gonggrijp in the matter of §2703(d) order relating to Twitter Accounts: Wikileaks, Rop_G, IOERRO, and Birgittaj as Amici Curi in Support of Objections of Real Parties in Interest Jacob Applebaum, Birgitta Jonsdottir and Rop Gonggrijp to March 11, 2011 Order Denying Motion to Vacate Misc U.S. District Court, Eastern District of Virginia, Alexandria Division (March 31 2011), No. 10-4 10GJ3703.

[45] Jack Goldsmith, *The Cyberthreat, Government Network Operations, and the Fourth Amendment,* The Future of the Constitution, GOVERNANCE STUDIES AT BROOKINGS (December 8, 2010), at 12. See, in particular, note 32.

could be prosecuted to computer-related and national-security ones[46]. This proposal sounds somewhat better in theory than it is likely to be in practice. Wiretap law is replete with instances where an initially restrictive collection is substantially expanded over time. Consider, for example, the 1967 Omnibus Crime Control and Safe Streets Act[47]. Title III of the act delineated the requirements for obtaining a wiretap warrant. Because of a history of law-enforcement abuse of wiretaps[48], Congress sharply limited the circumstances under which law-enforcement investigators could obtain a wiretap for a criminal investigation. The law listed twenty-five serious crimes for which a wiretap order could be obtained, and these were the only crimes for which a wiretap order for a criminal investigation could be issued. With time, that list was amended, and the number of crimes for which a wiretap warrant can be obtained now stands at slightly under one hundred[49].

A similar situation occurred for the Foreign Intelligence Surveillance Act, which puts forth the requirements for a foreign-intelligence wiretap order. While some expansions have been due to changes in technology (e.g., the shift to fiber optic cable that partially precipitated the FISA Amendments Act), other expansions of the law, most notably lowering the need for foreign intelligence from being a "primary" purpose of the order to simply being a "significant" one[50] have substantively changed the original law. Goldsmith's proposed limitation may not actually work very well in practice. An IDS/IPS mechanism that scanned private-sector communications networks for malware, but which used the gathered information for criminal investigations is highly problematic from a Fourth Amendment point of view and would be unlikely to gain public support---at least if the technology's import is clear.

Data retention raises concerns on another dimension. Given that competing firms run critical infrastructure, how would information be shared? Privacy and competition issues severely complicate such data sharing. There may be legal restrictions on disclosing personally identifiable information. New policy provisions and new laws would be needed in order to handle the information sharing that an EINSTEIN system would require in the broad private-sector environment (as opposed to the federal civilian agency sector).

As a result of the liberalization of U.S. cryptography export regulations in 2000[51] , encrypting communications has become much more common. The peer-to-peer VoIP

[46] Jack Goldsmith, supra note 38 at 15-16.

[47] P.L. 93-351 82 Stat. 197.

[48] United States Congress, Senate, Select Committee to Study Governmental Operations with Respect to Intelligence Activities, Final Report of the Select Committee to Study Governmental Operations with Respect to Intelligence Activities: Supplementary Detailed Staff Reports on Intelligence Activities and the Rights of Americans: Book II, Report 94-755, 1976.

[49] Title 18, Part I, Chapter 19, § 2518.

[50] This change is a result of the USA PATRIOT Act (P.L. 107-56) Title II, § 218.

[51] Department of Commerce, Bureau of Export Administration, Revisions to Encryption Items. 15 CFR Parts 734, 740, 742, 770, 772, and 774, Docket No. RIN: 0694-AC11, Effective January 14, 2000. at § 4g.

system Skype uses end-to-end encryption[52], which ensures only the sender and recipient may understand the conversation. Many large enterprises employ virtual private networks (VPNs), where communications are encrypted on a server within the corporate network, then travel the public communications network and are decrypted once the communication is again within the corporate network, for communication. Indeed, while private carriers transport the confidential communications of the U.S. government, these are often encrypted end-to-end. (If federal government communications are to be secured---say if such communications from a San Francisco switching office to a federal agency on the East Coast---then the communications architecture is likely to enter the leased fiber-borne "T1 line" to the destination. Communications would first be encrypted according to NSA-approved or NIST-approved methods[53], then enter the T1 link. Fully protected against being read, the communication travels the "public highway" to the East Coast, where it would be decrypted after it reaches its endpoint. This method of communications security has advantages --- and disadvantages. While the architecture secures the communication during its transit, it does not ensure reliability and the arrival of the communication[54].) But while the use of end-to-end encryption protects an intercepted communication from being understood, it completely disrupts EINSTEIN technology; EINSTEIN-type devices operating on encrypted communications would not be able to examine the content of the communications.

EINSTEIN devices would be able to examine transactional information, but only if the communications were not traveling through a VPN or were otherwise encrypted---in which case, the only information revealed during interception would be that the communications' destination is within the corporate network[55] (information about the ultimate endpoints of the communication would become available once the traffic is within the corporate network).

[52] Skype, P2P Telephony Explained --- For Geeks Only, http://www.skype.com/intl/en-us/support/user-guides/p2pexplained/ (last visited February 1, 2011).

[53] Which system was used depends on whether the communication was classified.

[54] Consider, for example, the events of July 2001. Several cars on a 60-car CSX train going through the Howard Street Tunnel in Baltimore derailed, and a fire broke out. The high-temperature fire took five days to put out. During that time large amounts of road traffic in Baltimore was disrupted. There were also other disruptions, one of the more notable ones being that of communications traffic along the East Coast. The problem was that seven of the largest U.S. ISPs used a fiber optic cable that ran through the Howard Street Tunnel and the fire burnt through the pipe housing the cable (SAIC, U. S. Department of Transportation, EFFECTS OF CATASTROPHIC EVENTS ON TRANSPORTATION SYSTEM MANAGEMENT AND OPERATIONS: HOWARD STREET TUNNEL FIRE, BALTIMORE CITY, MARYLAND, *Effects of Catastrophic Events on Transportation System Management and Operations: Howard Street Tunnel Fire, Baltimore City, Maryland,* July 18, 2001). The moral: unless the U.S. government owns the entire physical infrastructure of the communications network, U.S. government communications will always be subject to the "backhoe problem." That said, the communications security described above is sufficient for federal civilian agencies for all practical purposes.

[55] This is true even if a VPN user were sending a mail to someone outside the corporation. The communication would travel from the user to the corporate VPN server, where it would be decrypted and then sent to the mail server. At that point, it would travel as mail. From the point of view of an interceptor, the communication's destination is the corporate mail server.

Extending EINSTEIN-type technology to private-sector critical infrastructure would mean that *anyone* using the infrastructure would be subject to an invasion of their privacy. Effectively all Americans would be subject to inspection of their communications. At the same time, because enterprise communications would likely be using VPNs, if EINSTEIN-type surveillance were to become de rigeur for ICT, we might find ourselves in the odd situation in which corporate communications were routinely afforded privacy from surveillance, while private communications of private citizens were not. One can imagine "solutions" to this, solutions likely to complicate law-enforcement wiretapping.

It is by now clear that an extension of EINSTEIN-type technology to the private sector would be remarkably complicated both from a policy and technical viewpoint. The most basic issue, however, is how to process the massive amounts of data that may traverse an Einstein-type system. As is usually the case in such situations, complexity lies in the details. We turn to the potential role of EINSTEIN-type technology in two specific examples of critical infrastructure.

The Complexities Posed by ICT

By interposing an eavesdropper on all communications traveling over the network, an EINSTEIN-type system on a public communications network would be disruptive in many ways.

Whether an EINSTEIN-type system can work in the public communications sector is completely based on the numbers: how many packets flow through an EINSTEIN device per second, how long it takes to examine these, and how many can be stored for later examination. In the 1990s the rate of communications transmission was sufficiently slow that the communications bits could be effectively examined and stored --- at least if one did sampling. Fiber optics changed the equation; the technology of fiber optic transmission and packet routing has outstripped that of computation for the past twenty years, and that trend is likely to continue in the forseeable future. Computation-based monitoring of a significant portion of the Internet is likely to be very costly and impractical in all but very special cases.

The cost of storage is now dropping even faster than the rate of transmission is increasing, and instead there might be a temptation to store *all* questionable communication to be examined later. Recall the Cisco router described in section 3.What if, instead of examining all inputs to the CRS-1 in real time, we recorded the traffic for later examination if a threat signature were detected elsewhere. The combined input and output rate of a fully loaded single-shelf CRS-1 is 1.28 Tb/s, which translates to 160 GBytes/sec. Thus, to store all the comings and goings for a single high-end router for a day would require storage equal to about 14 petaBytes per day. This amount of storage is equivalent to 86,400 high-end iPods. Clearly the long-term storage of a router's traffic flow for later consideration is not practical. The numbers preclude EINSTEIN technology from sharing all the packets that pass through, though sharing abstracts, summaries, or snippets might work (depending on size and form of comparison being done).

Sharing transactional information would be one way to do this. Despite current limited legal protections given to transactional information, communication transactional information is itself a rich source of private information. Golle and Partridge have observed, for example, that if one can determine home and work location of a user (easily done, for example, from determining the cell location of communications made between the hours of 11 pm and 7 am and between 9 am and 5 pm respectively), can allow re-identification of a previously "anonymous" user[56]. Long-term storage of transactional data for later study creates a new security risk, while centralizing the data would create an even bigger one. The latter argues for providing privacy protections to the data. How well will this work in practice? Such techniques may destroy much of the value of the data for the IDS/IPS.

Finally --- and perhaps the most important issue --- arises from ICT's role in society. It is appropriate for an IDS and IPS to act conservatively, and thus to prohibit those types of communications that are not explicitly allowed. So an IPS should naturally disallow a new form of communications technology, whether Instant Messaging, Skype, twitter, Facebook, or some new application, until it is determined by the IDS/IPS designers that the new communications forms are not malware. Although there may be costs to the public if the Veterans Administration or the Department of Health and Human Services does not immediately implement the newest communications technologies such as Facebook or twitter, such a conservative design makes sense for a federal system IDS/IPS.

This approach does not make sense for an EINSTEIN-type system protecting the public ICT. Unless the EINSTEIN technology only uses blacklisting ("prohibit communications with these signatures"), EINSTEIN-type technologies at ICTs will prevent early deployment and testing of innovative communications technologies. That would be an enormous mistake. ICT's function is large and diverse, and it needs to encourage new forms of technology. These are exactly the characteristics that make TICs and IPSs difficult to employ except in constrained situations, such as the U.S. government sector.

Simply put, the model of few TICs cannot apply to ICT. Underlying EINSTEIN's inapplicability is that ICT infrastructure has few commonalities with the U.S. federal government. ICT has many, many pipes and many of those are big (10 gigabits/second---and greater). There are many, many more communications providers than departments of the federal government. Absent U.S. federal government requirements---which would be very hard to achieve---ICT players have no incentive to cooperate; indeed, because they are commercial competitors, they have a strong disincentive to do so. Meanwhile, EINSTEIN itself creates risks. Concentrating ICT traffic anywhere---central to the

[56] Philippe Golle and Kurt Partridge, *On the Anonymity of Home/Work Location Pairs,* PERVASIVE COMPUTING, SEVENTH INTERNATIONAL CONFERENCE, NARA JAPAN, May 11-14, 2009,

EINSTEIN 3 concept of discovery---creates its own vulnerabilities[57]. Various commonly used technologies for information protection, such as VPNs, will thwart the EINSTEIN model for detecting "bad" behavior. And finally, unlike the federal employees communicating the federal government computers, the ICT customer---the public---has Fourth Amendment and statutory rights that are greatly threatened by this technology.

The Complexities Posed by the Power Grid

On a first glance, it seems that the EINSTEIN technology would be an extremely good match for the power grid. The grid is heavily reliant upon computer networks, both at the consumer level, where such networks are used to bill customers, and at the grid management level, where computer networks coordinate power generation and transmission. The industry is moving towards "smart grid," a two-way digital communication and control system in which the utilities will send messages to devices in the home and office about energy prices at that moment (e.g., on a hot summer day when the temperature is causing high demand for air conditioning), and users' systems will respond accordingly (e.g., by shutting down until prices are lower)[58].

We already have ample demonstration of security problems. In 2007 researchers at the Idaho National Laboratory showed how to access a power plant's control system through the Internet. Running an emulator, the researchers destroyed a 27-ton power generator by power cycling at very short intervals[59]. In 2009 there were news reports that the power grid had been penetrated by spies who might have left rogue code behind[60]. In 2010 the Stuxnet worm targeted Supervisory Control And Data Acquisition (SCADA) systems used to monitor and control industrial processes---specifically those controlling Iranian nuclear centrifuges[61]---amply demonstrating proof of concept[62].

Increasing amount of electronic communications from the smart grid means there will be need to directly protect customers (e.g., from attackers who snoop on the communication with smart meters or, worse yet, send forged messages about electricity usage). Meanwhile the fact that the power industry is heavily regulated should help with lowering barriers to sharing cyberattack data among the energy providers. It would seem the cyber networks of the power grid would be ripe for EINSTEIN.

[57] See *supra*, note 54.

[58] Litos Strategic Communication for the Department of Energy, THE SMART GRID: AN INTRODUCTION (2008), at 11.

[59] CNN, Sources: Staged cyber attack reveals vulnerability in power grid (September 26, 2007).

[60] Siobhan Gorman, *Electricity Grid in U.S. Penetrated by Spies*, WALL STREET JOURNAL (April 8, 2009).

[61] William J. Broad and David E. Sanger, *Worm Was Perfect for Sabotaging Centrifuges*, NEW YORK TIMES (November 18, 2010).

[62] The worm was apparently introduced through an infected USB flash drive (Derek S. Reveron, Cyberattacks After Stuxnet, NEW ATLANTICIST, October 4, 2010), but could both update itself and spread through the Internet (Symantec, *How Stuxnet Spreads,* NEW YORK TIMES, January 16, 2011, at A1).

On closer examination, the fit is less clear. The power grid cybernetwork is actually four networks with different users, different levels of protection, and different protection needs. We begin by enumerating these:

- Providing customers with data about electricity usage: Consumers often have web access to account information, such as their latest bill and summaries of electricity usage. This communication takes place over the Internet and relies on the customer's own Internet connection.

- Providing utilities with information about electricity usage: Utilities increasingly rely on computer networks to remotely read customer electricity meters. Many utilities build and deploy their own networks over many kinds of low-bandwidth “last mile” technologies; these include microwave, power line, radio, cellular, and wireless mesh networks. User privacy is important to avoid revealing sensitive information, such as whether and when customers are at home[63].

- Controlling the customers' smart devices: With the move toward a smart grid, utilities will increasingly communicate directly with devices such as refrigerators, dish washers, or air conditioners at the customer sites, in order to adapt electricity usage to current demands. The technologies for smart devices are still in an early stage. Rather than the utilities supporting a diverse array of communication media, devices are likely to rely on customers' Internet connections for communication with the utilities.

- Managing the power grid: Communication networks play an important role in managing power generation and distribution, including coordination between various electricity providers, operations, economic markets, and transmission systems. While this communication could take place over private networks, in practice many companies rely on the public Internet in one form or another. Some utility companies may also rely on the "cloud"---servers hosted in data centers---to run their management systems and share data with third parties.

The first and third cases---customers and devices communicating with the utilities over the Internet---is an ICT issue, and one we have already discussed with respect to EINSTEIN’s applicability. We focus instead on the networks for reading and controlling customer usage and for managing the grid. Deploying EINSTEIN 3 would face many difficult challenges, and the first is complexity.

There are a large (and growing) number of energy providers communicating in complex ways over a mix of public and private networks. According to Lockheed Martin, by 2015

[63] See, for example, Mikhail A. Lisovich, Deirdre K. Mulligan, and Stephen B. Wicker, Inferring Personal Information from Demand-Response Systems, IEEE SECURITY AND PRIVACY (January/February 2010), at 11-20.

the smart grid will offer up to 440 million potential points of attack[64]. Not only is power highly distributed to millions of customers, power generation is increasingly distributed, with a large number of small providers, including individual households, contributing energy to the grid. These "last mile" networks are an important part of the cyber-security problem facing the power grid, but they are hard to protect without a very large-scale deployment of security infrastructure.

At the same time, the grid involves many independent (sometimes competing) parties with complex trust relationships. The grid is, at best, a loosely coupled federation[65], making it difficult to consolidate into a small number of network attachment points as the U.S. federal government is achieving through TIC. Even if consolidation were possible, the requirements for real-time data and high reliability make it undesirable to circuitously direct data through few consolidated access points. Yet any practical deployment of EINSTEIN 3 would have to occur at locations where these small, heterogeneous networks aggregate. For example, a provider could place an EINSTEIN 3 device at a site that aggregates the connectivity to all of its customers, or at "peering" locations that connect the provider to other parts of the grid. As such, any deployment of EINSTEIN 3 in the power grid would likely involve a large number of locations, which may be logistically and financially unwieldy, and make any ability to do correlation of anomalous behavior much less likely.

The second major problem is function mismatch. The IDS/IPS solutions useful for protecting U.S. federal government computer networks may not be a fit for the power grid and may in fact have to be completely redesigned for use in the power grid. Just as in the telecommunications sector, many parties in the energy grid already have their own IDS/IPS and firewall solutions from a variety of vendors, making the EINSTEIN 3 equipment at least partially redundant. A more complex issue is reporting. Energy providers must generate SCADA[66] reports as part of Critical Infrastructure Protection (CIP) requirements for the North American and Federal Energy Regulatory Commission (NERC/FERC)[67]. Existing IDS/IPS solutions are often integrated with other important functionality such as quality-of-service, compression, SCADA-specific reporting, and integration with existing management tools that are not naturally part of EINSTEIN 3-type devices.

SCADA presents a particular problem. SCADA systems are typically not used in Internet applications, and thus parsing the messages sent and received by these protocols would

[64] Darlene Storm, *440 Million New Hackable Smart Grid Points*, COMPUTERWORLD (October 27, 2010, 3:11 PM), http://smartgrid.ieee.org/news-ieee-smart-grid-news/1663-440-million-new-hackable-smart-grid-points .

[65] Larry Karisny, *Smart Grid Security: Ground Zero for Cyber Security*, MUNIWIRELESS (June 2, 2010 at 12:51 pm), http://www.muniwireless.com/2010/06/02/smart-grid-security-ground-zero-for-cyber-security/ (last viewed February 8, 2011).

[66] SCADA (Supervisory Control And Data Acquisition) systems are used to monitor and control industrial processes.

[67] Juniper Networks, Smart Grid Security Solution: Comprehensive Network-Based Security for Smart Grid (September 2010), www.juniper.net/us/en/local/pdf/solutionbriefs/3510346-en.pdf

require custom extensions to EINSTEIN 3. Perhaps more importantly, these systems have vulnerabilities subject to unique attacks, such as the Stuxnet worm that attacked Siemens SCADA systems in several countries in the summer of 2010. The EINSTEIN 3 system in the power grid would need to create and continually extend a library of signatures for these, increasing the cost and effort in running the EINSTEIN 3 program. These requirements mean that EINSTEIN 3 equipment cannot be extended to subsume all of this functionality without a major redesign---at great expense and uncertain outcome. Future trends further complicate the problem.

Certain grid communications, particularly in the back-end systems that control electricity generation and distribution, are highly sensitive to delay, which forcing traffic through a small number of EINSTEIN 3 locations would only increase. At this time, the grid does not have hard requirements on communication delay, but this could easily change with a move toward finer-grain control of electricity generation and distribution.

Meanwhile fundamental to any security solution for power grid communication is encryption[68]. Systems like EINSTEIN 3 can, at best, **detect** attacks while they are happening. Encryption of the critical communication in the grid can help **prevent** many of these attacks in the first place. Supporting encryption is challenging, as it requires support from the many customer meters and smart devices, as well as having secure ways to exchange keys between customers and the utilities. We discuss encryption in the next section, but note that whatever encryption solutions are chosen will have a significant influence on whether and how systems like EINSTEIN 3 should be deployed. This strongly implies that the basic security architecture for the grid should be resolved before significant effort is made to deploy EINSTEIN 3 within the power grid.

It is now time to turn to security solutions.

5. Securing the Cybernetworks of ICT and the Power Grid: What Ought be Done

In 2005 Governor of Arizona Janet Napolitano said, “You show me a 50-foot wall, and I’ll show you a 51-foot ladder.”[69] She was discussing the physical fence being built between Mexico and the United States. Over time, the wall became a virtual one, in which electronic sensors, radar, and cameras were used to alert border guards about illegal crossings. In 2011, as Secretary of the Department of Homeland Security, Napolitano canceled the project[70], which had cost one billion dollars over its five-year

[68] Currently encryption is not required. When it is implemented, the implementation is often very poorly done. See Joshua Pennell*, Securing the Smart Grid: The Road Ahead,* NETWORK SECURITY EDGE (February 5, 2010), http://www.networksecurityedge.com/content/securing-smart-grid-road-ahead?page=2.

[69] Linda Greenhouse, Legacy of a Fence, NEW YORK TIMES (January 22, 2011, 5:07 PM), http://opinionator.blogs.nytimes.com/2011/01/22/legacy-of-a-fence/ (last visited February 2, 2011).

[70] Julia Preston, Homeland Security Cancels ‘Virtual Fence’ After $1 Billion is Spent, NEW YORK TIMES (January 14, 2011).

effort. The secretary concluded the project was not viable. It would have been better, of course, to have realized this earlier[71].

Had the "virtual fence" been evaluated for effectiveness from the start, it might never have gotten off the ground. The savings in time would have been quite valuable; even more important were the lost opportunities to pursue alternative solutions, opportunities lost because of diverted resources. Effectiveness matters, and should be measured at all points along the development cycle of a project.

EINSTEIN 3 is an electronic fence. In section 4 we argued that EINSTEIN 3 protections are inappropriate and infeasible for the commercial telecommunications infrastructure and the power grid. What might be done as a practical alternative?

Beginning with telecommunications infrastructure, it is instructive to consider how U.S. telephony infrastructure was protected when AT&T was essentially the sole provider of telecommunications services in the U.S. At the time the company owned and operated the vast majority of the country's long-haul transmission systems (AT&T Long Lines). It operated two basic types of services over these: retail switched long-distance service, and the long-term lease of "private lines" to both private companies (e.g., the New York Stock Exchange) and governmental organizations (e.g., the U.S. Department of Defense).

The combination of legal requirements and good engineering practice led the design of a network that was secured from a large variety of threats by three basic methods:

- Physical security: the carriage of U.S. government traffic on the AT&T network led to the requirement to physically secure and monitor all AT&T transmission and switching facilities.
- Transmission security: at least to a reasonable degree, the signals carried over AT&T's transmission facilities were protected from intercept. While only a few signals were encrypted, all were carried by means physically or technologically resistant to interception (e.g., on buried coaxial cable, or on multiplexed microwave signals)
- Separation of control and content: for a variety of reasons, AT&T embarked in the middle 1970s on an aggressive effort to separate the control information used to set up phone calls, and control the network in general, from the circuits used to actually carry the call[72]. This approach, termed "out-of-band signaling," and now usually known as Signaling System #7, is now the rule in telephone systems (but not in data networks like the internet). With the "signaling" separated from the content it was possible to make the network more robust in many ways, to improve its operating efficiency, to introduce new services such as 800 calls, and,

[71] This is not a comment on Secretary Napolitano, who had inherited the program.

[72] G. E. Schindler, Jr. (ed.), A HISTORY OF ENGINEERING AND SCIENCE IN THE BELL SYSTEM --- SWITCHING TECHNOLOGY (1925-1975), Bell Telephone Laboratories, Inc. (1982).

of importance here, to dramatically reduce an adversary's ability to intercept calls or to manipulate the telephone network itself.

There are two obvious differences between the modern telecommunications infrastructure in the present compared to that of the U.S. of thirty years ago: (i) AT&T is not the only long-distance provider any more, and (ii) much more data is being transmitted than voice. A more nuanced comparison reveals the following differences, all of which lead to increased security compared to the present day:

- Physical security: for a variety of reasons, but mostly owing to the financial cost involved, the plethora of modern North American telecommunications providers, many of them small and undercapitalized, provide little practical physical security for their transmission and routing equipment.
- Transmission security: even though the wholesale conversation to digital transmission from the old analog methods would appear to permit equally wholesale use of encryption-based transmission security, it is still rarely used.
- Separation of the control and content "planes": for reasons of cost in the early days of the ARPAnet, and now locked into decades of legacy practice, the Internet operates on the principle of passing both the control and content information for an application over the same "pipe". This practice permits a wide variety of attacks on both the users of the network and the network itself.

In an interesting case of "back to the future," rather than proposing EINSTEIN 3 protections for telecommunications infrastructure, perhaps we should consider reintroducing telecommunications design principles that were in place in three decades ago. While requiring it of all network operators might be neither desirable nor practical, it would not be unreasonable to consider that only "certified" network operators be considered when procuring communication services supporting critical civil or military activities. This certification should include, following the formula (i) physical security, (ii) transmission security via encryption or arguably equivalent protection, and (iii) the use of techniques that isolate the control of the network itself from the content being carried by it. Such a separation would secure that which needed securing without the disruption provided by an IDS/IPS that would prevent the innovative telecommunications services the dynamic ICT sector keeps providing.

The cyberinfrastructures of the power grid, although vulnerable to cyber attacks, present a very different case. While critical infrastructure could (and perhaps should) not be accessible via the Internet, the system should be able to prevent malicious behavior---whether the attack is launched remotely or not. The controlling computer, aware of the generator's limitations, should refuse to initiate commands that would damage the equipment. Still, this solution merely introduces another problem---ensuring the controller software itself is reliable. But in this problem lies the key to protecting power grid infrastructure.

Unlike ICT, the cybernetworks of the power grid do not provide, or need to use, hot-from-the-developers communication technologies. This, and the fact that changes in power grid technology happen slowly---at least when measured by Internet years---greatly simplify the problem of protecting the cyberinfrastructure of the power grid. Compared to operations that control the generator, software changes in power grid cyberinfrastructure occur relatively infrequently. Software updates could be delivered via a trusted courier instead of over the network.

The broader solution to many of the security problems facing the power grid is cryptographic. No instruction to change behavior and or replace software should be accepted unless it is digitally signed. Once appropriate cryptographic measures are in place, the physical origins of the commands is no longer a concern; these commands can come in person, by telephone, the Internet, or satellite radio. The essential mechanism is guaranteeing that the agent with the authority to give a command possesses the correct authorizing key and is the only possessor of that key. The scale and diversity of authority can raise challenges in distributing and managing keys. Fortunately, the power grid consists of just a few thousand power companies in the U.S., and not all of these companies run generators. This is not a particularly large number of users for a key-management system.

Cryptography also offers a way of controlling smart devices and providing data about electricity usage. For example, encrypting communication from the electricity meter to the power company prevents rogue parties from passively snooping on the transmissions. Authenticating the messages from the power company to smart devices prevents unauthorized parties from remotely controlling these devices. Ensuring that electricity meters and smart devices have keys and the necessary cryptographic machinery is no trivial matter. Yet, grappling with these issues is crucial to ensuring the security of the power grid, whether or not a system like EINSTEIN 3 is ever deployed.

These arguments do not mean EINSTEIN-type solutions have no value. Rather they mean that the effectiveness of such solutions should be weighed against alternatives before they are developed, and development should proceed with the technologies must likely to provide the needed security.

There are a number of problems to be solved first in order for EINSTEIN-type solutions to succeed. For example, in the ICT domain, the issue of de-identified data sharing is one worth exploring. Recent research on “privacy-preserving” algorithms identifies ways to compute answers to data-analysis questions without revealing the raw input data. The classic example is the "millionaire problem," where two people want to know who is richer without revealing the precise amount of their wealth to each other[73]. In the context of IDS/IPS systems, multiple sites---each run by a different company---may want to identify malicious users that send excessive traffic, without divulging the total traffic

[73] Andrew Yao, *Protocols for Secure Computations*, PROCEEDINGS IEEE SYMPOSIUM ON FOUNDATIONS OF COMPUTER SCIENCE (1982) at 160-164.

received at each site or revealing the access patterns of the well-behaved users[74]. Promising solutions already exist for many of these kinds of data-analysis tasks, and further innovations in this area could lower the barrier for collaborative security solutions to protect critical infrastructure.

Another direction to pursue is opening up the EINSTEIN architecture to public view. While using classified signatures on a private-sector IDS/IPS creates a complicated control mechanism, the decision to have some signatures classified is not itself unreasonable. That is in contrast to the decision to classify the architecture, which is not a sensible choice. A fundamental principle in cryptography, Kerchoffs' Law, is that a cryptosystem's security should depend not on the secrecy of the algorithm but solely on the secrecy of the key. Similarly, an IDS/IPS security solution should depend solely on the secrecy of the signatures being use.

Public examination of the architecture allows a full appraisal and will establish greater confidence and trust in the system. The lack of a public vetting of the EINSTEIN 3 architecture being used in protecting federal civilian agencies means that there has been virtually no informed public discussion on the efficacy of using EINSTEIN-type technologies in protecting critical infrastructure. Consider the virtual fence at the border, the project that Secretary Napolitano canceled. "The problem with the [virtual fence] was that it is the wrong kind of technology to be deployed across the entire U.S.-Mexico border," Napolitano said. "It was too expensive, it was too elaborate and it was not flexible enough to meet the fact that immigration patterns change."[75] In the absence of a public vetting of EINSTEIN 3 technology, it too is likely to be too expensive, too elaborate, and not sufficiently flexible as attacks vectors change. In order to consider such a heavy-weight security solution, the architecture should be made public. This should happen early in the life of the program.

The publicly available documentation on EINSTEIN does little to clarify the technology's limitations. While experts understand that signature-based schemes can only protect against known attacks, the publicly available documentation on the EINSTEIN technology does not state this. U.S. Deputy Secretary of Defense William Lynn has characterized the cyberexploitations of U.S. business and government sites as what "may be the most significant cyberthreat that the United States will face over the long term."[76] The technically unsophisticated reader would have no idea from reading the EINSTEIN documentation that the technology provides essentially no protection

[74] Benny Applebaum, Michael Freedman, Haakon Ringberg, Matthew Caesar, and Jennifer Rexford, *Collaborative, privacy-preserving data aggregation at scale,* PROCEEDINGS PRIVACY ENHANCING TECHNOLOGIES SYMPOSIUM (July 2010).

[75] Lauren Gambino, *Failed virtual border fence has politicians pointing to success in Yuma area*, CRONKITE NEWS (January 31, 2010), http://cronkitenewsonline.com/2011/01/failure-of-border-fence-has-politicians-pointing-to-success-around-yuma/

[76] William Lynn III, *Defending a New Domain*, FOREIGN AFFAIRS (September/October 2010) at 3.

against such attacks[77]. This should be made clear to policymakers. The inflated implications of what EINSTEIN can handle---phishing[78], IP spoofing, man-in-the-middle attacks[79]---noted in section 2 is likely to lead to unrealistic expectations regarding the problems EINSTEIN-type solutions can solve, and are not unlike the claims made for the virtual border fence.

After examining the complications of applying EINSTEIN 3-type solutions to ICT and the power grid, it should be clear that the current architecture of EINSTEIN 3---concentrated Internet access points cooperating to perform intrusion detection/prevention---does not provide a viable model for protecting the cybernetworks of critical infrastructure. EINSTEIN 3 is a virtual fence that has the potential to work when you can funnel all comers through your gates---that is EINSTEIN 3 applied to the federal civilian agency sector---but not when architecture and control are highly distributed. Private infrastructure is likely to remain inherently more distributed and less trusting of partners than U.S. federal government services. To be viable, what is needed for protecting critical infrastructure's cybernetworks are new IDS/IPS solutions that scale to a large number of vantage points and analyze traffic without divulging private user data or proprietary business data. That should be the direction pursued in protecting these networks, not that of molding them into centralized systems more akin to the public switched telephone network. Sometimes hammers are just not appropriate solutions. So it is in this case.

[77] We say "essentially," since by eliminating some malware, the exploitations launched by the highly targeted attacks may stand out more. That is, however, a second-order effect, and one that cannot be counted upon.

[78] EINSTEIN should be able to prevent phishing and spear phishing attacks that use known malware. Highly-targeted spear phishing exploitations using zero-day attacks are unlikely to be stopped.

[79] DEPARTMENT OF HOMELAND SECURITY, COMPUTER EMERGENCY READINESS TEAM (US-CERT), PRIVACY IMPACT ASSESSMENT FOR THE INITIATIVE THREE EXERCISE, March 18, 2010 at 3.