Chapter 5

Cyberspace and Infrastructure

William D. O'Neil

THIS CHAPTER addresses two related subjects: protecting cyber, electrical, pipeline, and other infrastructures against cyber attack, and protecting cyber infrastructure against all forms of attack. After a brief history of infrastructure attack and a review of the nature of networks, it outlines threats and responses, including systems engineering and dependability. Existing U.S. policy for infrastructure protection is described, along with government and industry responsibilities. The chapter next examines policy issues for protecting infrastructure against cyber attack and protecting cyber infrastructure, and then closes with specific policy recommendations.

The History of Infrastructure Attack

Infrastructure attack is a story as old as war. Time out of mind, attackers have sought to cut off their target's water supply and transportation, often with decisive results. The rise of modern infrastructure systems starting in the 19th century brought heightened concerns about vulnerability. As one widely read futurist and social critic put it in 1929:

> [S]omething on the order of one hundred key men, opening its veins of water, power, gas, sewage disposal, milk supply, [and] communication, could bring the life of a great city to an end—almost as neatly as though its every crevice had been soaked with poison gas. Even in rural areas with the growing use of electric power, the telephone, gasoline, and imported foodstuffs, the factor of dependence on an unknown technology is very great. . . . The machine has presented us with a central nervous system, protected with no spinal vertebrae, lying almost naked for the cutting. If, for one reason or another, the severance is made, we face a terrifying, perhaps mortal crisis. . . . Day by day the complexity, and hence the potential danger, accelerates; materials and structures ceaselessly and silently deteriorate. One may look for some very ugly happenings in the next ten years.[1]

Especially in the United States, early airpower enthusiasts drawing on these currents of thought became convinced that infrastructure attack held the key to victory in modern war. Infrastructures—especially the electric grid—were seen as relatively easy to take down and so critical that their slightest disruption would severely affect warmaking potential and economic activity in general.

When World War II came, however, Air Corps planners decided that electric power was not as critical as previously thought and turned their attention to other target complexes. Later analysis suggested that this was probably an error and that attacking powerplants could have been quite effective. German electric production was curtailed when attacks on the rail infrastructure cut back coal shipments severely, but concerted attack on transportation was not decided upon until late in the war. Electric production in Japan, which was largely hydroelectric, was affected less by bombing.[2]

Of the 1.5 million tons of air ordnance delivered by U.S. forces against German and

German-held targets in 1943–1945, 41 percent fell on transportation targets, largely rail, and 6 percent on oil, chemical, and synthetic rubber targets.[3] In the war against Japan, naval submarine and surface forces as well as air forces devoted much of the weight of their attack to enemy transportation, particularly at sea.[4] It was later concluded that attacks on transportation infrastructure had severely affected the enemy war effort and that even greater effort against transportation would have been worthwhile.[5]

Wars since 1945 have continued to feature attacks on transportation and often on oil as well. The 1991 Gulf War included major campaigns against both, as well as systematic attacks on Iraq's communications infrastructure. Since World War II, U.S. bombing campaigns generally have made electric power infrastructure a major target. The best documented case is the Gulf War: 88 percent of Iraq's electric grid capacity was knocked out, most in the first few days.[6] Guerrilla and terrorist forces have also frequently targeted infrastructure. Since the invasion of Iraq in 2003, there have been repeated attacks on Iraq's infrastructures, especially those for electric power and oil.[7]

So far as has been publicly revealed, however, there have not yet been military campaigns against the infrastructure of cyberspace, nor any military cyber attacks on other infrastructures. But there have been a great many attacks by hackers, whose identities and motives often are shadowy; some observers believe that some of these attacks have been state-sponsored.

Below, lessons of this history for defending infrastructures are examined, but first the nature of infrastructures themselves is explored.

Networks

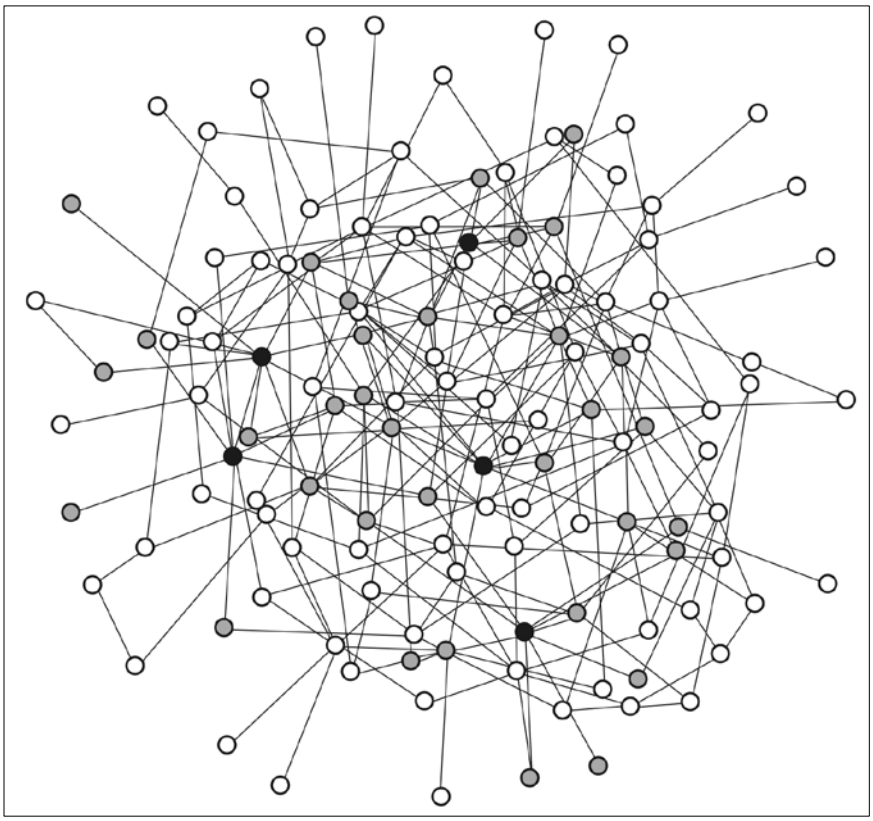
Infrastructures often depend on networks: we speak of the road network, the rail network, the telephone network, the electricity network, and more recently the Internet. The theory of networks is important to a great many fields of science and technology.[8] A network consists of the points of supply or origin, the routes of transportation or movement, and the points of destination or consumption. Nodes may function both for origin and for destination. The whole set of nodes of origin and destination, together with the linking routes, comprises a network.[99] A wide variety of manmade and natural physical, biological, and social systems can be analyzed in network terms. Each infrastructure is physically distinct, but, as the terminology suggests, they share something important at the level of abstract structure.

While infrastructure networks are not truly random, they are complex and irregular; as a result, many of the applicable tools of network theory are statistical in nature.[10] Large, irregular networks—like most infrastructures—can be seen in two broad classes. In one type, most of the nodes have roughly the same number of links. Only a few nodes have many links. Networks with this sort of egalitarian structure are often called *exponential networks* (because highly connected nodes are exponentially unlikely) or, more descriptively, *uniform-random networks* (see figure 5–1).[11]

In the other major class, most nodes are connected to nodes that also already have a great many connections (although there is still a random or random-like element). These are called *power-law networks* (for technical mathematical reasons that reflect the relative abundance of highly connected nodes) or, more commonly, *scale-free networks*, a reference to their lack of a dominant typical or average scale in the sense of number of connections per node. A more descriptive term might be *hub-and-spoke random network* (see figure 5–2).

If an accident or attack were to disable a node picked at random from the exponential network in figure 5–1, it usually would disconnect only a handful of other nodes that happen to have no connections other than to the one disabled. In the scale-free network, a random disablement would likely do even less damage, since so few nodes have any other node that connects only through them. However, the worst case in a scale-free or power-law (hub-and-spoke) network (figure 5–2) is worse than in a random-like exponential network because taking down only a few highly connected hubs does a lot of damage.

Figure 5-1 Exponential or Uniform-random Network



Cyber Networks

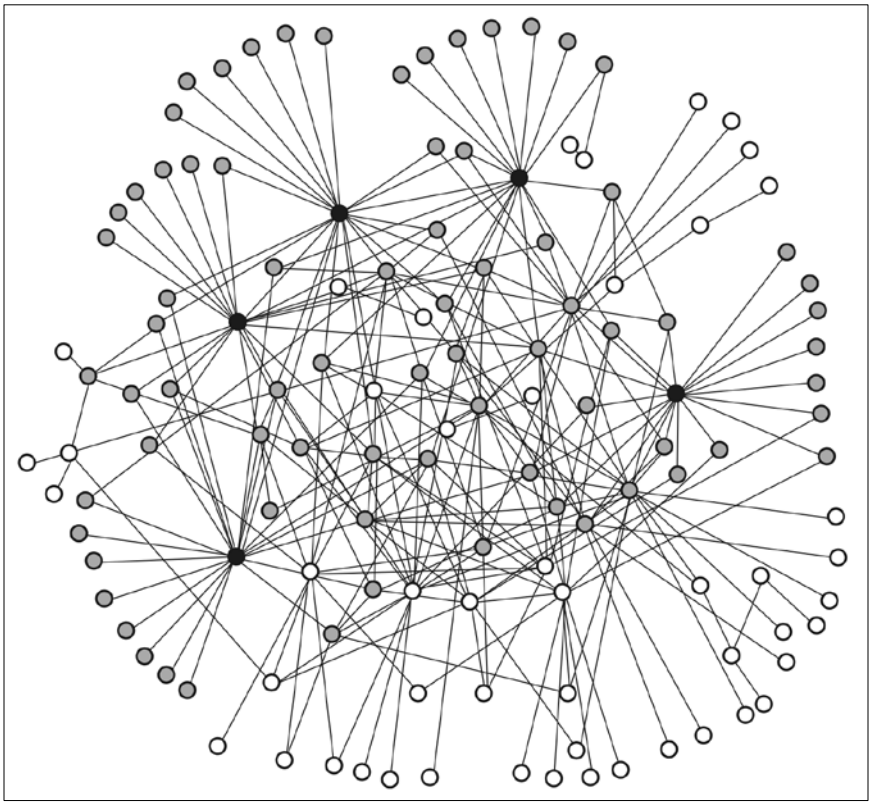
So far we have addressed the topology of the networks: the logic of how nodes and links connect. We also need to look at the physical nature and spatial location of these elements. Infrastructures are not only networks, but also networks upon networks (as outlined with respect to cyberspace in chapters 6 and 8 in this volume, "Evolutionary Trends in Cyberspace" and "The Future of the Internet and Cyberpower"). As table 5–1 shows, cyber content rests on a structure of physical elements that have physical properties and locations. Even though its topology is not identical with that of the network layers it is built upon, cyberspace itself, like other infrastructure networks, has a geography as well as a topology, and both affect its vulnerability and survivability.

The topologies of the Internet and World Wide Web are the subject of particularly intense study, for reasons including their complexity, availability for study, and practical importance.[12] No simple model can fully capture the complexity of these structures, but in broad terms, both are power-law or scale- free networks.

Scale-free networks arise typically through growth, as new nodes link preferentially to old nodes that are already highly linked, forming highly connected hubs. It is easy to see how this happens in the Web: it costs no more to link a Web page to a richly connected hub Web site such as Google or Wikipedia than to an isolated site run by a specialized organization, such as <ndu.edu>, or by an individual, such as <analysis.williamdoneil.com>.

Internet nodes consist of computers (or devices that incorporate computers). The simplest case is a single isolated computer in a home or small business. The cheapest possible connection would involve running a cable or wireless link to the computer next door. In most cases, however, this would not accomplish much, since usually only a small portion of our information needs can be supplied by our immediate neighbors.

Figure 5-2 Scale-free or Power-law Network



Of course I might be able to piggyback on the information channels available to my neighbor, but this would cut into the bandwidth available to him and so would be unattractive from his standpoint. Even though it costs more, therefore, we generally buy our service from an Internet service provider (ISP), which offers a connection to its hub or server bank (a group of high-speed computers, usually housed in a single warehouse-like building) via some miles of telephone wire, coaxial cable, fiber optic cable, wireless cellular radio link, or satellite radio link. Higher bandwidth connections that provide greater information capacity cost more, but most users find the expense worthwhile.

An ISP whose server bank services thousands of high-speed connections over an area of many acres or square miles faces similar choices. Connections to nearby ISPs would be relatively inexpensive in terms of the cost of the cable but would not meet ISP needs for a rich flow of data to meet its customers' demands. Thus, the ISP finds it worthwhile to pay for a very high bandwidth connection to a major hub with a massive server bank that handles a great deal of Internet traffic in order to tap its riches. Processes such as these, repeated at all levels, drive the Internet toward a hub-and-spoke scale-free architecture resembling that shown in figure 5–2.

Table 5-1 Simplified Schematic Overview of Levels Involved in Cyberspace

Level	Description	Examples
Cyber	Intellectual content	Data, commands, knowledge, ideas, mental models
Logical net	Services employing physical signals to carry logical messages	Telephones, broadcast radio and TV services, cable TV service, public Internet, private Internet protocol (IP)–based networks carried on common-carrier infrastructure, private-infrastructure IP- based networks, supervisory control and data acquisition networks
Hard net	Infrastructures formed from base elements that carry electrical or electromagnetic signals	Common-carrier telecommunications networks, tactical radio systems, dedicated wireline systems, community cable systems, cell phone systems
Base	Physical elements that underlie telecommunications services	Cable headworks, optical fiber, coaxial cable, radio transmitters and receivers, radio transmission paths, communications satellites, Internet routers, modems

Scale-free networks are, for reasons described above, robust in the face of random or untargeted failures, which fall most heavily on the large numbers of nodes with only a few connections. The experience of the Internet reflects this: nodes often fail or are shut down for a variety of reasons, but this has scarcely any discernible effect on overall network performance. Even more massive failures, such as those caused by widespread power outages or the 9/11 attacks, have been quite localized in their effects. (Of course, such incidents can generate a surge in traffic that may itself slow response, but this is not a vulnerability of the Internet per se.)[13]

There have been many random outages but few that preferentially target the Internet's major hubs. Nevertheless, what is true in theory would hold equally true in practice: successful

attacks on many of the biggest hubs would have severe and pervasive effects, leaving many Internet nodes isolated or able to communicate with only a small number of other nodes. Thus, protection of major Internet hubs is a cornerstone of rational policy for cyberspace infrastructure defense.

Link outages might be much less worrisome than outages of key nodes in a scale-free network like the Internet: severing links could not do as much damage to network connectivity as disabling an equal number of critical nodes. A closer look at the physical layers underlying the Internet, however, shows that this may be too sanguine a view in practice. Links that are logically and topologically separate may in fact be carried over the same physical communications infrastructure through multiplexing. Indeed, entirely separate networks, having no logical interfaces at all, may be multiplexed via one fiber optic strand. Even if they use physically separate communications lines, it is possible that those lines may share the same conduit or otherwise be vulnerable to the same damage agents. Thus, a single attack might take out thousands or tens of thousands of links at one time, potentially cutting off multiple nodes from the network. The places where this can occur must be protected to assure cyberspace infrastructure integrity. This is a particular concern for nodes located in physically isolated sites, as are many that are critical to national security. Where economy or convenience is the dominant consideration, such sites are often served by only one or two pathways for all communications links.

The Electrical Network

Loss of electricity does not ordinarily take down a major Internet hub—at least not at once, since most hubs have emergency backup power sources that can carry them for hours or days. In a larger sense, however, the Internet, like practically our entire society, is critically dependent on electric supply. If we examine specific electric grids, we find that while there are core areas representing major centers of population and industry, there are no hubs connected directly to large numbers of nodes, and most nodes have more than one link. The topology of electric grids is more like the uniform-random (exponential) network depicted in figure 5–1. Although the electric grid, like the Internet, grows and changes as nodes and links are added, modified, and sometimes deleted, its economic and technological forces are quite different from the Internet and result in a different kind of pattern. These forces are changing in important ways today, and the resulting grid will no doubt eventually look quite different, as discussed below. We must look first at historical forces to understand today's grid.

Even though it comes in different forms—alternating or direct current at any of a number of voltage levels—electricity is all of a kind.[14] With suitable conversion of form, any electrical energy will serve any electrical load.[15] It is a bulk commodity both in the sense that it is lacking in the specificity that distinguishes information and that it shows strong economies of scale. Because electricity is most economically generated at large scale, the resulting grid is dominated by a relatively small number of large central station plants, usually located at or near their energy sources. (This may well change as the costs of carbon emissions and other environmental damage are figured into the cost of generation; some generating technologies with low environmental impact, such as wind turbines and solar-electric systems, may favor smaller scale operation.) Electricity is also most economically transported in bulk, at high levels of energy and voltage.[16]

Neither bulk generation nor bulk transmission in itself dictates a uniform-random electric network. A key reason for this type of network is that in the United States, most

electrical transmission is in the form of alternating current (AC) at high voltages, and most electrical use is AC at lower voltages. A relatively simple passive device, the transformer, allows high-voltage AC (HVAC) to be tapped at lower voltage with scarcely any loss of energy. Thus, major corridors are served by a few high-capacity HVAC lines along which distribution stations are located that feed local bulk users and local retail distribution networks. The corridors themselves are determined by economic geography: they go to where the customers are. (Of course, customers may also find it economical to locate in corridors served by major transmission facilities.)

Except for those who use electric power on truly massive scales, it is more economical for customers to draw power from a nearby distribution station than to run lines directly to a distant central station. Because the distribution station draws its power from a major HVAC line, it can supply large quantities, and since all power is the same, it makes no difference where it comes from as long as there is enough. This is why, when we look at a portion of the electric grid, we see a network that more closely resembles the uniform-random pattern of figure 5–1 than the scale-free hub-and-spoke layout of figure 5–2. Thus, the electric grid is a fundamentally different kind of network from the Internet.

The earliest commercial electric utilities used direct current (DC). AC won out as the U.S. standard, in large part because it is less difficult and costly to tap HVAC transmission lines with transformers to produce lower voltage for distribution and final use than to step down from HVDC.[17] DC continued in use for specialized local applications (such as shipboard electrical systems) for some decades, but these applications too gradually died out. For moving very large flows of energy over long distances, however, HVDC lines can be more economical than HVAC. This fact has led to the use of HVDC intertie lines to connect distant "islands" of intense electric use across wide stretches with little use, where distribution stations are not needed.[18] Today, the North American electric grid (encompassing the continental United States, Canada, and a small portion of northwestern Mexico) is divided into four large regions that connect almost entirely via HVDC links. This greatly reduces the risks of a continent- wide grid failure.

When there are two or more possible routes from generator to load, electricity by its nature will flow over all of them, with the greater amount following the paths with lower resistance. If one path is cut off, the flow automatically redirects itself over the remaining links. When the flow in a transmission network is near the limits of its capacity to handle power flow, the failure of one link would throw more load on remaining links than they can carry. This would lead to a cascade of failures, as links either break down due to overheating or are shut down (by automatic switches or by human intervention) to save them from damage.

On an AC network, the current must alternate at the same frequency everywhere in what is called synchronous operation. Any failure of this frequency synchronization would produce unbalanced forces that could literally tear equipment apart. Synchronization failures also can cascade, as generation or transmission equipment drops offline to avoid catastrophic failures.

The loading on the grid varies from moment to moment, and the organizations responsible for its operation have only limited tools for managing it. Users can add loads by throwing a switch; generators and transmission equipment can go offline for a variety of reasons. Grid operators may have the ability to shed some loads (temporarily cutting off customers who have bought "interruptible power" at reduced rates), but load-shedding capacity is limited. In an emergency, a block of customers in a particular area may be blacked out to shed loads, but

many systems are not set up to allow it to be done quickly, and utilities are reluctant to do this except as a last resort. An overstressed link or node may have to be shut down, which increases the load on other components. If local overloading drags the frequency of a generator down, then it and the area it serves must immediately be disconnected from the grid. On a wide scale, this can cut the grid up into isolated islands, many or all of which might fail under local load imbalances.

Could such a failure cascade engulf the entire North American electrical grid, leaving the whole continent in the dark? Two factors make this unlikely. First, like the ripple caused by a rock thrown into a pool, the disturbance following a major fault in the grid weakens as it disperses. Second, the HVDC intertie lines that link the four major synchronous regions in North America also isolate each region from frequency disturbances in the other regions. Regardless of what may happen in any one region, the others should be able to adjust and continue normal operation without major disruption.[19]

Prior to the mid-1960s, widespread grid failures were unknown because with a few exceptions (mostly involving large hydropower systems), most electric power was generated, distributed, and delivered within a compact area served by a local power utility that enjoyed a regulated monopoly. This "cellular" rather than networked structure meant that there was little opportunity for failures to spread beyond a single utility. Moreover, regulators held utilities responsible for reliability of service and could apply effective sanctions when reliability fell short.

Although this regime led to steadily decreasing electrical prices for decades as the utilities incorporated new technology, in the 1970s and 1980s economists and political leaders argued that the monopolistic structure, even with regulation, was economically inefficient and led to added cost. At the same time, new technologies appeared to offer the potential to generate and transmit electricity economically on scales that were beyond the capacity of even large individual utility companies. Thus, starting in the mid-1970s, the Federal Government moved to deregulate electricity—in fact, to change the regulatory basis so as to encourage competition in generation and transmission. States have followed suit, although not in a uniform way, leading to fragmentation of ownership and control over generation and distribution equipment and operation.[20]

Deregulation in general has not been followed by further significant decreases in the costs of electricity, although proponents argue that it resulted in avoidance of large increases and had other benefits. It has, however, opened the way to other problems not fully anticipated and still being worked out.

Because every part of the grid influences every other part, it has been difficult to construct a deregulation regime that would allow the truly independent operation necessary for fully effective competition. In 2000 and 2001, for example, Enron Corporation and other power producers and speculators exploited the physical properties of California's electricity grid in combination with its deregulated operating rules to manipulate prices to their great advantage, at the same time causing or exacerbating electricity shortages in the state. Analyses of this event show how difficult it is to ensure the smooth running of the physically tightly coupled but economically fragmented electric market system.[21] The same limitations that permit participants to impose costs on others without inherent limits (other than those interposed by the remaining regulators) equally allow serious technical problems to develop and spread without any individual participating firm or organization having a clear interest in taking corrective action.[22]

Legislators and public officials nevertheless retain a strong commitment to deregulation, but even if restoration of the earlier regime of regulated local vertical-power monopolies were politically and economically feasible, it is not clear how it could work physically today. Many regions have come to depend on power generated far away and transmitted over long distances. Heavy long-distance power flows have become a fact of life, and any attempt to re-divide the grid into relatively small, self-sufficient cells operated by separate local firms would involve major investment costs and serious environmental concerns. But without some such structure, there is no simple way to assign responsibility for maintaining adequate and reliable power service.

The physics of electricity simply does not allow a fully laissez-faire, every-man-for-himself operating regime. Just as on the highway, there must be some consistent set of operating rules that everyone is constrained to obey if the system is to operate stably and safely. Despite warnings, this realization has been somewhat slow in emerging, perhaps in part because authorities were thinking in terms of analogies with networks that were not as tightly coupled as the electricity grid and thus less in need of highly disciplined operation. Below, we discuss what has been done to address the policy issues raised by these facts and what more may need to be done.

Lessons from a Blackout

Major outages demonstrate how tightly coupled the grid is and what this implies for its operation and protection. The most recent major outage in North America occurred on August 14, 2003, and eventually covered large areas of the northeastern United States and Canada, affecting electric service to approximately 50 million people for an extended period.[23] The extent of the blackout is illustrated in figure 5–3.

Investigation revealed a number of hardware and software failures, together with faulty operational procedures on the part of both the local utility operator and the organization responsible for ensuring the reliability of the grid in the region. Most of these did not directly contribute to the blackout, but many of them clearly could have led to major failures under slightly different circumstances. Many aspects of the operation were in violation of accepted industry standards (which were then voluntary). Even if all equipment and software had been functioning properly and fully in compliance with existing standards, however, the tools available to the operators for system awareness were critically limited. The process that led to blackout started after 3 p.m. near Cleveland, Ohio.

August 14 was hot and air conditioning loads were heavy, although not near peaks that the system had handled before. The immediate cause of the blackout was a series of instances in which high-voltage transmission lines contacted trees that had been allowed to grow too tall into the lines' rights of way. Autonomous safety systems sensed the resulting ground faults and automatically disconnected or "tripped" the lines to prevent more serious damage and fires. Over a period of 8 seconds starting at 4:10:37 p.m., automatic safety relays all over the Northeast shut down lines and generators that had violated preset acceptable operating limits; these shutdowns severed grid links and blacked out areas throughout the region.[24]

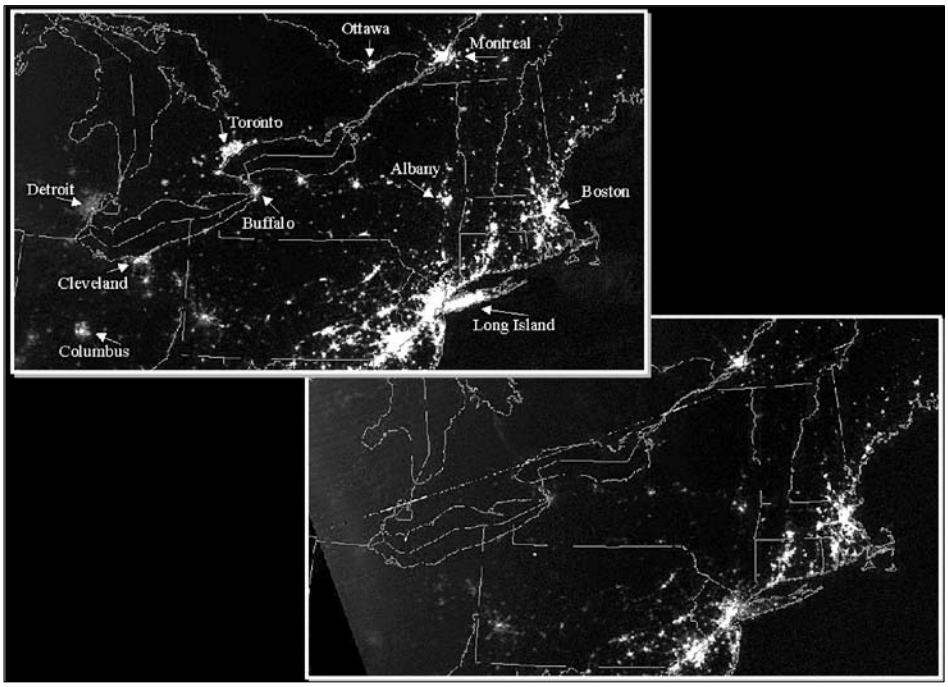
Operator response was poor, in part because critical warning and analysis systems had failed due to design defects and operator error. One by one, the faults accumulated until the point at which human intervention could no longer be effective. Even given the various hardware and software failures, the huge blackout would never have occurred as it did if the

operators had been well trained and effective in applying the existing procedures, despite their limitations. At worst, a very limited area with a few tens of thousands of customers might have been affected for a few hours. Instead, part of the problem arose from excessive and inappropriate operator reliance on limited and fallible warning and diagnosis systems.

Coming just a month before the second anniversary of September 11, the blackout caused many to wonder whether it was caused or worsened by terrorist attacks. Indeed, claims of responsibility purportedly from al Qaeda appeared within a few days of the outage. The "Blaster" Internet worm had first been seen just a few days earlier on August 11; this led to speculation that it might have been involved in the blackout. Investigation showed that neither al Qaeda nor Blaster was responsible, but it did reveal significant potential vulnerabilities.[25]

The process that set the stage for the August 2003 Northeast blackout included software failures that denied operators the information tools they were accustomed to. These failures were accidental or were intrinsic to the system design, but comparable failures could have resulted from cyberspace attacks.

Figure 5-3 August 2003 Blackout



In general, the operators of the grid rely on a variety of cyberspace services to gather operating information, communicate with other control personnel, and issue instructions. Of particular concern are supervisory control and data acquisition (SCADA) systems and energy management systems (EMS). These gather data on the operational parameters of equipment throughout a particular segment of the electric grid, report it to a central location and present it to operators, and change setpoints for equipment controllers in response to operator decisions about system configuration and operation. We address the policy and standards efforts being undertaken to meet these problems later in this chapter.

The (Secure) Grid of the Future?

Engineers and policymakers have devoted considerable attention and development effort to defining the future of the electric power grid. The visions generally involve a "smart grid" able to adapt to failures in real time with limited if any degradation.[26] In a sense, this would resemble a return to the cellular structure of the pre-deregulation grid, but with smaller cells that are regulated by adaptive software rather than governmental agencies, and with provision to take advantage of distant power sources. One key is distributed local power sources and perhaps power storage systems. Fuel cells in particular appear to promise small-scale but highly efficient generating units that could serve these purposes.[27]

Growing concern about global climate change may affect these visions in various ways. One possibility is a renewed emphasis on very large nuclear central-station generating plants. This offers zero emissions of greenhouse gases— even better than fuel cells—but would involve greater concentration of power generation.[28] Solar power is another zero-emissions option and could integrate more naturally into a cellular structure, although efficient and economical means to store the energy from solar systems for release in the hours of darkness must be found if they are to become a major electrical source.[29] Effective use of wind power at large scale depends on solutions to the challenges posed by the irregularity and unpredictability of its flow.[30] Various other advanced technologies for energy production are more speculative at this time.[31] None of the alternative sources so far conceived can obviate or substantially modify the need for a more reliable and robust grid for electrical transmission and distribution. In most cases, they would add complexity, due to their limited ability to provide steady and continuous power or to vary their output rapidly in response to load fluctuations.

Practically all proposed schemes for improved electricity delivery depend on networked "smart control," which is to say that they depend more on cyberspace. Most proposals have devoted little attention to security against cyber attack. Clearly, this must change before much more work is done along these lines in order to ensure that efforts to improve the reliability and efficiency of power distribution do not increase vulnerability to attack.

Pipeline Networks

The electrical infrastructure is unique both in its degree of coupling and in its central role, but other infrastructures—particularly two other major energy-sector infrastructures, oil and natural gas—present parallel concerns, even if overall risk levels are lower.[32] Both oil and natural gas are also networked infrastructures, with about 170,000 miles of oil pipelines and 1.4 million miles of natural gas pipelines.[33] More than 75 percent of U.S. crude oil supplies flow by pipeline, as do about 60 percent of refined petroleum products, and almost all natural

gas.[34] Notwithstanding the obvious dangers of pipes filled with flammable and potentially explosive fluids, the overall safety record of U.S. pipeline systems is good. Despite their vulnerability to sabotage, there have been few attacks or attempts on U.S. pipelines. So far, all publicly known threats have been of attack by physical rather than cyber means.

Both oil and gas pipelines make use of SCADA and operational management systems, although not at the same level as the electrical infrastructure. The issues of cybersecurity in these infrastructures are generally similar to those affecting electricity.

Infrastructure Threats

The operators of infrastructure systems of all types routinely face a spectrum of threats, from natural causes such as lightning, earthquakes, or hurricanes; from intrinsic faults such as stuck valves or circuit breakers, failing electronics, or unstable software; and from criminal action by vandals, thieves, extortionists, or hackers. Motivated by a community sense of responsibility, regulatory and legal requirements, and economic self-interest, operators take action to avert these threats, minimize their potential damage, and recover rapidly from damage when it does occur. Many national security threats resemble more intense and deliberate versions of these normal infrastructure threats. This emphasizes the need to integrate all aspects of infrastructure protection.

Our special focus here is on cyberthreats. There have been a number of attacks on the cybersystems serving infrastructure, but they have not been coordinated large-scale attacks. Damage has been limited so far. However, the Central Intelligence Agency (CIA) has warned of the threat of cyber attack, especially against electrical utilities, noting that "cyber attackers have hacked into the computer systems of utility companies outside the United States and made demands, in at least one case causing a power outage that affected multiple cities." The agency reportedly did not know "who executed these attacks or why, but all involved intrusions through the Internet." According to a CIA analyst, "We suspect, but cannot confirm, that some of the attackers had the benefit of inside knowledge."[35] There have been other attacks, some domestic, but generally they have received no publicity in an effort to avoid giving the attackers useful feedback.[36]

In some cases, cyber-extortionists clearly were behind the attacks, but in most, the identity and motivation of the attackers are unclear. Following the September 11, 2001, terrorist attacks, it was widely predicted that al Qaeda would follow up with massive cyber attacks on infrastructure targets, but these have not materialized, and the likelihood of large-scale cyber-infrastructure attacks by terrorists is disputed.[37]

Even if terrorists never find the means or motivation to do so, there is little doubt that a conventional state enemy determined to mount a military attack on the United States could launch massive coordinated cyber attacks on infrastructure as a part of an overall strategy of infrastructure attack. Thus, we need to ask how much damage could be done by cyber means and what may be done to limit it.

A crucial question is the extent to which systems can be accessed via the Internet. Wide-open access is rare, but many systems may have some Internet portals through which an attacker might reach critical functions. Active efforts are widespread to close these or at least provide highly secure protection, and to the extent these efforts succeed, it will be impossible to attack systems from the Internet. This does not rule out attacks via individuals with inside access, however.

The consequences of failures are always foremost in the minds of the engineers who design infrastructure systems and components. Well aware that complex systems in general and software in particular are prone to failure, they design to limit the consequences. Where possible, critical control functions with the potential for severe equipment damage are lodged within simple, entirely autonomous, self-contained systems such as governors and overload trips. Thus, an EMS might be able to overload a transmission line with current but could not prevent the circuit breakers from tripping and opening the circuit to forestall damage.

This strategy is not universally applicable, however. In an aircraft with a fly-by-wire control system, for instance, a runaway of the system conceivably could fly the airplane into the ground or exceed its limiting flight loads. When such vulnerabilities are unavoidable, engineers may go to extraordinary lengths to assure the reliability and integrity of the critical control system, as they do in the case of aircraft controls. Yet there may be cases where care is inadequate, leaving vulnerabilities that cyber attackers might exploit with devastating results. Systems engineering disciplines do exist to minimize the chances of this, but they require high costs and intrusive oversight and thus are unlikely to be uniformly applied. Moreover, many systems were designed before the risks of cyber attack were recognized.

There are many proponents of commercial-off-the-shelf (COTS) or open- source systems for practically all uses; such approaches can save substantial amounts of time and money, and thus suppliers and customers may be tempted to incorporate such subsystems and software modules into a system. However, they can greatly increase vulnerability to common modes of attack, perhaps catastrophically so in critical applications such as EMS or SCADA systems. Thus, their use should be guided by policies that assure management of risks and by adequate weighing of risks and costs.[38]

There is also danger in policies of commonality—that is, the widespread use of the same systems or the same family of systems in an organization or industry. No doubt this can reduce costs, but it means that exploitation of a single vulnerability could affect a wide range of operations. Here, too, care should be taken to weigh savings against risk.

Concerns about security of infrastructure systems emerged in the 1990s and have been heightened by recurrent hacker attacks. Even though these attacks have not so far gone beyond the nuisance stage (at least domestically), when the grid is operating under stress, a successful denial-of-service attack on a SCADA system or an EMS could, at least in theory, lead to a situation comparable to what occurred by accident and inattention in August 2003, when the blackout spread because operators lacked important information or could not exercise effective control. If appropriate operator action is hindered or prevented by cyberspace attack, this could set the stage for a massive failure cascade.

Worse still could be capture of a SCADA system or an EMS by attackers who could then use it to control generating and transmission equipment. An attacker who had sufficient information and understanding of the affected portion of the grid and enough access could increase stress on the system by configuring it inappropriately. An attacker who could simultaneously block or spoof system reporting and operator control functions could render the manual safeguards ineffective or even counterproductive, whether the capture was affected by physical intrusion or by remote means via cyberspace.

There are some limits, however, on how much an attacker could accomplish simply by capturing control of an EMS or a SCADA system. SCADA and energy management systems are designed to lack the capability to override self- contained automatic safety relays associated with generators, transmission lines, and transformers. Design engineers, aware of the

possibilities of SCADA and EMS failures for a variety of reasons, avoid giving them more control authority than is strictly necessary for their intended functions. Unless an attacker knew of and could exploit a serious design flaw, capture of a SCADA system or an EMS generally would not, by itself, allow the attacker to inflict major long-term damage on the electrical system, nor could the attacker initiate a major failure cascade in the absence of heavy external system loading.[39] If combined with effective physical attacks on critical equipment, however, capture of the control systems could allow an attacker to inflict much greater damage. Even without coordinated physical attacks, a cyber attacker with sufficient access and knowledge could trigger a widespread blackout comparable, within the affected area, to the Northeast blackout of August 2003. As that experience showed, it could take several days to restore full service, and there could be large economic losses, as well as risk to life and property.

SCADA systems and EMS are cyberspace systems, according to the definition used in this book, but their susceptibility to attack by cyberspace means varies. Industry standards call for them to be isolated from contact with the public Internet and other sources of possible outside cyber entry.[40] (The standards also call for them to be protected from physical intrusion and from surreptitious insider takeover.) Many instances have been found, however, in which the systems have failed to meet these standards fully. In some cases, deficiencies have been revealed through hacker attacks, but most have been discovered in the course of testing.

One potential threat that is often overlooked is that of Trojan horses introduced in the process of developing or maintaining the software. A concealed fault deliberately planted in a device or software program is a familiar fictional device,[41] but little has been done to forestall such threats, perhaps in part because no actual cases have been publicized. Although more complex to mount than a virus or denial-of-service attack, a surreptitious fifth column attack of this sort could be more damaging and more difficult to diagnose and correct. The danger is greatest in the case of open-source systems, where there is little control over who may have modified the code. But COTS and even purpose-built systems are at risk, because they might have incorporated key modules supplied by obscure low-tier subcontractors with little oversight.

While an attacker who finds and exploits a key cyber vulnerability may be able to do severe and lasting damage to a particular system, many systems will be competently and conscientiously designed and operated and will not offer such opportunities. If an attacker targets a system for which he cannot identify a catastrophic cyber vulnerability, he would have to employ physical attack to do major damage to it.

In general, moreover, it is necessary not only to do physical damage, but also to do enough of it to saturate the capacity for near-term restoration. Electrical utility companies, for instance, generally are well prepared to restore considerable numbers of downed transmission lines quickly, since natural causes such as ice storms or hurricanes can do damage of this sort. If the attacker's goals involve putting the system out of operation for more than a few days, it would do better to attack other elements. There is only very limited capacity to replace large transformers quickly, for instance, and even less to replace major generator facilities. The dangers are greater if the attacker is able to interfere with repair and restoration activities, or to mount repeated attacks with little cost, as in Iraq.[42] In any event, the importance of physical security for key facilities is clear. In recognition of the threat posed by attacks on transformers and generators, efforts are being made by utilities and coordinating groups to improve capabilities for restoring them quickly, an effort that should be wider.

Systems Engineering and Dependability

Engineers in many fields have long sought to make systems perform their intended functions dependably despite a wide spectrum of threats. They have developed a body of practice, usually referred to as systems engineering, that encompasses specification, analysis, design, and testing practices to ensure that a system will meet definite standards of dependable operation. Although it is not always fully effective, thorough application of systems engineering practice greatly improves dependability.

However, application of systems engineering has been notably weak in most areas of software development. The techniques for effective systems engineering for software are well understood and documented,[43] but the structure of the industry has not supported their application in most commercial software (making it cheaper but less dependable). However, most customers find it easier to assess and evaluate price than dependability in the abstract. Securing infrastructures against cyber attack is impossible without dependable software. Thus, any program for infrastructure protection must mandate good software systems engineering in order to be effective.

Policy and Organization

Concerns regarding protection of infrastructure are long standing, but it was in the second Clinton administration that the first steps toward a comprehensive policy were taken. A Presidential commission, convened in 1996 and reporting in 1997, emphasized government-industry cooperative efforts.[44] On May 22, 1998, President Bill Clinton signed Presidential Decision Directive (PDD)/National Security Council 63 (PDD 63), "Critical Infrastructure Protection." Although now superseded, PDD 63 was the root of most current U.S. infrastructure protection policy. The principal policy directives regarding infrastructure protection as of mid-2008 are described in this section.

The primary focus of the Executive Order on Critical Infrastructure Protection, signed by President George W. Bush on October 16, 2001, is "continuous efforts to secure information systems for critical infrastructure, including emergency preparedness communications, and the physical assets that support such systems." It assigns to the Secretary of Defense and the Director of Central Intelligence (DCI) the responsibility:

> to oversee, develop, and ensure implementation of policies, principles, standards, and guidelines for the security of information systems that support the operations under their respective control. In consultation with the Assistant to the President for National Security Affairs and the affected departments and agencies, the Secretary of Defense and the DCI shall develop policies, principles, standards, and guidelines for the security of national security information systems that support the operations of other executive branch departments and agencies with national security information.

However, the policy and oversight structure set up by this executive order has been considerably modified since its promulgation, as outlined below.

The Homeland Security Act of 2002, signed into law by President Bush on November 25, 2002, established the Department of Homeland Security (DHS) and assigned it lead responsibility for preventing terrorist attacks in the United States, reducing national vulnerability to terrorist attacks, and minimizing the damage and assisting in recovery from attacks that do occur. It gives DHS broad responsibilities for protection of critical infrastructure in

the United States against both terrorism and natural disaster. DHS, however, was not given responsibilities for protecting critical infrastructure from intrinsic or natural faults such as those involved in the Northeast blackout of August 14, 2003, or from nonterrorist attacks.

The National Strategy for the Physical Protection of Critical Infrastructures and Key Assets (NSPPCIKA) was approved by President Bush in February 2003. The focus is on protection against terrorist attack, rather than protection generally. It lays out a cooperative effort to be shared among various levels of government and the private sector without, for the most part, specifying definite responsibilities.

The National Strategy to Secure Cyberspace was approved by President Bush in February 2003. It does not focus on terrorist threats to the extent that the NSPPCIKA does, mentioning criminal threats and threats of military attacks as well. Its overall structure and approach, however, are similar to that of the NSPPCIKA. Its stated purpose is "to engage and empower Americans to secure the portions of cyberspace that they own, operate, control, or with which they interact." It identifies the private sector as "best equipped and structured to respond to an evolving cyber threat" but acknowledges that in "specific instances

. . . federal government response is most appropriate and justified," such as "where high transaction costs or legal barriers lead to significant coordination problems" and "[where] governments operate in the absence of private sector forces," and to resolve "incentive problems that lead to under provisioning of critical shared resources" as well as "raising awareness." The role of DHS is emphasized and explained in detail.

Homeland Security Presidential Directive 7, "Critical Infrastructure Identification, Prioritization, and Protection," signed by President Bush on December 17, 2003, establishes U.S. Government policy "to identify and prioritize United States critical infrastructure and key resources and to protect them from terrorist attacks." The Department of Defense (DOD) is assigned specific responsibility for protecting infrastructure relating to the defense industrial base. DHS is given responsibility for protection of most other national-level infrastructures, including many critical to DOD operations, from terrorist attacks.

National Infrastructure Protection Plan (NIPP) 2006 was agreed upon by multiple agency heads, including then–Secretary of Defense Donald Rumsfeld, in June 2006. The goals of the NIPP are "enhancing protection of the Nation's CI/KR [critical infrastructure/key resources] to prevent, deter, neutralize, or mitigate the effects of deliberate efforts by terrorists to destroy, incapacitate, or exploit them; and to strengthen national preparedness, timely response, and rapid recovery in the event of an attack, natural disaster, or other emergency." In military terms, the NIPP is analogous to a strategic plan, whereas the other directives more closely resemble broad statements of policy. Sector-specific plans under the NIPP are in the process of development and approval.

The NIPP integrates terrorist and natural-disaster threats into the infrastructure and gives brief attention to criminal threats. Like the other directives, it does not address threats of warlike attack or intrinsic failure. This is an opportunity for improvement in U.S. policy: it would be more effective and efficient to deal with such threats in an integrated and comprehensive way.

Aside from the now-superseded Clinton administration PDD 63, the focus of all these directives is strongly on terrorist threats. Criminal and military threats receive some limited treatment but practically no attention to damage that might result from intrinsic, accidental, or natural causes.

The government has policies with respect to these other threats, but, for the most part, they

are scattered among many laws, regulations, and directives relating to the responsibilities and functions of specific departments, agencies, and organizations. For that reason, we turn next to an organizational perspective.

Organizational Responsibilities

The duties of DHS are the subject of the policy documents just described. U.S. infrastructures are operated by thousands of commercial and other organizations, almost all of which take measures to ensure reliability and security of operations. The rest of this section concentrates on other organizations.

The DOD Role

Regardless of other considerations, each Federal department and agency must see to protecting its own infrastructures against all threats, cooperating and coordinating with other agencies where appropriate. In addition, DOD must defend the Nation's infrastructures against military attack and participate with allies in doing the same. It may be called upon to aid in protecting and restoring the Nation's infrastructure or those of allied or friendly nations against natural disasters. A comprehensive DOD policy regarding infrastructure defense and protection must deal with all these needs.

As a practical matter, however, DOD concerns cannot stop there. Infrastructure by its nature is pervasive and highly networked and may be largely invisible. Few clear boundaries can be drawn. Although DOD makes great efforts to be self-sufficient, it is dependent on many infrastructures not under its control. Even though defending them against some kinds of threats is not within its defined responsibilities, DOD cannot afford to neglect them.

Moreover, the distinctions among threat sources—military, terrorist, criminal, natural, or intrinsic—often are not operationally meaningful: it can be difficult or impossible to discern the actual source of a threat in time to affect operations. For instance, it may not be feasible or prudent to await definitive information about whether a specific problem is the result of military attack before taking defensive action.

Thus, the DOD policymaker confronts a dilemma regarding infrastructure protection. It is impossible for DOD to simultaneously stick solely to its own business and fulfill its responsibilities. There will inevitably be ambiguities and overlaps of responsibility and spheres of action, and consequent potential for costly conflict with other agencies and entities. There can be no bureaucratic "good fences" to make "good neighbors" with other agencies and organizations in infrastructure protection. Unless close and cooperative give-and-take relationships can be developed in advance, counterproductive friction is likely to hamper needed efforts.

None of these issues or considerations is new to DOD, which has long confronted them in various forms. But the changing nature of the threats, as well as new organizational responses in other areas of the government, has led to significant changes. These are reflected in three key policy and doctrine documents. The Department of Defense Strategy for Homeland Security and Civil Support, approved by Deputy Secretary of Defense Gordon England in June 2005, is a broad statement of policy and approach. DOD Directive 3020.40 of August 19, 2005, on Defense Critical Infrastructure Program (DCIP), also approved by Deputy Secretary England, specifically defines DOD policy with respect to protection of defense-

related critical infrastructure and assigns responsibilities. Third, Joint Publication 3–27 of July 12, 2007, *Homeland Defense*, is especially lengthy, reflecting the complexity of the issues involved, and provides commanders at all levels with authoritative guidance covering a wide range of situations and contingencies.

Only experience will tell whether DOD policy and doctrine will prove adequate, and be adequately implemented, but what has been produced is encouraging.

Other Federal Agencies

Five major Federal agencies share responsibilities relating to energy infra- structure: DHS, the Department of Energy (DOE), Department of Transportation (DOT), Federal Energy Regulatory Commission (FERC), and Nuclear Regulatory Commission (NRC). DOE and FERC are both involved in protection of all energy infrastructures (electrical, oil, and natural gas) against natural and intrinsic threats and share terrorism-protection responsibilities with DHS. NRC plays a comparable role with respect to nuclear energy infrastructure. DOT has responsibility for pipeline safety, exercised by its Office of Pipeline Safety, and coordinates with DHS regarding pipeline security.[45] DOE has several national laboratories (outgrowths of the development of nuclear weapons); its Idaho and Sandia National Laboratories are active in energy infrastructure security research and development.

The Federal Communications Commission (FCC) has responsibility for all Federal communications regulation. In the past, it commissioned a recurring series of Network Reliability and Interoperability Councils (NRICs), composed of representatives of a broad spectrum of communications industry entities as well as concerned government organizations, and chartered to provide recommendations to the FCC and to the communications industry to help "assure optimal reliability and interoperability of wireless, wireline, satellite, cable, and public data networks."[46] The commission's Public Safety and Homeland Security Bureau works with DHS on security and protection issues.

National Communications System

The National Communications System (NCS) is an outgrowth of a Cold War initiative from the 1960s intended to assure critical executive branch communications under any circumstances, including nuclear attack. An interagency group long run by DOD, it is today lodged in DHS.

Rather than build dedicated government-owned communications infra-structure, the NCS stressed close cooperation with the telecommunications industry to assure the necessary reliability. The closely related National Security Telecommunications Advisory Committee (NSTAC) provides industry-based advice to the executive branch on communications security issues.

North American Electric Reliability Organization and North American Electric Reliability Corporation

The first widespread power outage in North America was the Northeast blackout of November 1965, which affected large areas of Ontario, New York, New Jersey, and New England. In response, regional reliability councils were formed, to be coordinated by the

National Electric Reliability Council (NERC). The regional councils and NERC operated under Federal authority but were funded and staffed from the utility industry; their standards were consensual, and industry compliance was voluntary and self-policed. Experience showed that this was not adequate and that the same causes cropped up again and again in power failures. Industry leaders urged that NERC—which by then had become the North American Electric Reliability Council, including coverage of Canada and a small portion of Mexico whose grid is linked to that of California—needed to be given teeth so it could formulate and enforce mandatory standards.[47] Finally, after the August 2003 Northeast blackout, necessary legislation was passed.

Under the Energy Policy Act of 2005, FERC was given responsibility and authority for the reliability of the bulk electric power delivery system throughout the United States. It was authorized to designate an independent Electric Reliability Organization (ERO), which it was hoped would also be recognized by Canada and Mexico, to set and enforce standards throughout the North American electric grid. NERC had become the North American Electric Reliability Corporation, subsidiary to the council; it reorganized itself to comply with the requirement that it be independent and submitted its proposal. Certified by FERC as the U.S. ERO in July 2006, it is empowered to establish and enforce reliability standards, subject to FERC approval, with penalties for infraction. Among these are standards for security, including cybersecurity. Inevitably, the standards reflect a balance among security and other considerations, notably cost. They have been reviewed by concerned government and industry organizations and are widely but not universally believed to be adequate. The standards were approved by FERC in January 2008 and went into effect in March 2008.[48] The standards are oriented toward both processes and objectives and are broad enough to cover a range of situations. The ERO issues implementing instructions and assesses compliance, recommending action to the FERC to correct problems that cannot be resolved administratively.

State Agencies

Although states have embraced deregulation in various ways and degrees, state governments retain their inherent powers to regulate infrastructures operating in their territories. Most states have one or more independent agencies devoted to these functions. In addition, state and local law enforcement agencies play major roles in protecting infrastructure systems.

Information Sharing and Analysis Centers

In 1998, PDD 63 called for the establishment of an Information Sharing and Analysis Center (ISAC) as part of the Federal apparatus for critical infrastructure protection. This evolved into a series of 11 organizations: Communications ISAC, Electricity Sector ISAC, Emergency Management and Response ISAC, Financial Services ISAC, Highway ISAC, Information Technology ISAC, Multi-State ISAC, Public Transit ISAC, Surface Transportation ISAC, Supply Chain ISAC, and Water ISAC. Loosely coordinated by an overall council, the ISACs serve as conduits for government-industry and industry-industry communication about operational threats and protective measures. The Communications ISAC is the National Coordinating Center for Telecommunications, an arm of the National Communications System. The Electricity Sector ISAC is the NERC.

DHS Councils and Partnerships

DHS has established or assumed sponsorship of a number of organizations intended to advise the government on infrastructure protection matters and coordinate government and private-sector efforts. These include the Federal Inter- agency Security Committee, the State, Local, Tribal, and Territorial Government Coordinating Council, the Critical Infrastructure Partnership Advisory Council (CIPAC), and the National Infrastructure Advisory Council.[49]

Partnership for Critical Infrastructure Security

Representatives from the CIPAC Sector Coordinating Councils comprise the Partnership for Critical Infrastructure Security, a cross-sector coordinating organization established in December 1999 under the auspices of the Department of Commerce and now subsumed under the CIPAC.

Policy Issues

Since PDD 63 was issued in 1998, a great deal has been accomplished to make the Nation's critical infrastructure systems more secure and robust and to improve their protection against cyber as well as physical attack, with emphasis on defense against terrorists. In the same period, law enforcement agencies have greatly stepped up their activities in the area of cyber crime. Yet potential threats also have burgeoned. It is a race in which to stand still is to fall seriously and swiftly behind.

Throughout this period, cyber attacks have mounted rapidly. Many have been directed at the infrastructure of cyberspace itself, and a smaller but still substantial number against other infrastructures via their SCADA and management systems. In the majority of cases, it has been impossible to determine the identity or the motivations of the attackers. Vandalism, criminal gain, terrorism, intelligence-gathering, or even covert military attack are all possibilities, and usually there has been no way to tell.

These cyber attacks have been costly, but their effects, in terms of deaths, economic losses, and sheer misery and inconvenience, have been much less than those stemming from other sources of infrastructure damage. Many more Americans have been far more seriously affected by loss of electrical, communications, transportation, natural gas, and oil service resulting from stressful weather, geological disaster, accident, and intrinsic faults in design or construction. As a logical result, our society invests more attention and capital in averting and containing these more common and costly problems.

In practice, however, it often is not clear whether damage was initiated by human or natural attack. After the August 2003 Northeast blackout, it took months to determine that the cyber infrastructure failures that had an important bearing on the extent of the damage had not been caused by hostile attack. In fact, very similar damage might have been produced by a combination of physical attacks on transmission lines and cyber attacks on EMS and SCADA systems.

At the physical and engineering level, there is thus a large area of overlap in the measures needed to guard infrastructures against damage from whatever cause. To ignore this underlying unity in framing policy is to fight against nature, and it cannot fail to generate needless conflicts, gaps, and duplications of effort. Yet our survey above of the welter of policies and governing organizations reveals little evidence of unity in dealing with

infrastructure protection.

A Basis for Unified Policy

A fundamental axiom of U.S. policy in every field is that, to the greatest extent possible, responsibilities should be assigned to specific individuals or small, unified groups, and that responsibility and authority should always be closely aligned. That is the basis for our free enterprise system, for restricting the powers of government as narrowly as possible, and for assigning governmental powers to the lowest and most local level possible.

We are also wary, as a society, of the hazards of mixed motives and conflicting interests. We know all too well how difficult it is to serve two masters or pursue divergent interests.

These principles have informed America's decisions about infrastructures. Thus, for example, unlike many countries, we have never made the telecommunications, rail, or petroleum infrastructures into government departments. The Internet, created at Federal Government initiative, was divested as soon as it seemed feasible. Governmental control and operation of other infrastructures is quite limited and largely confined to local authorities.[50]

Yet we have seen how the private enterprise structure in electrical power distribution has contributed to conflicting interests and mismatches of responsibility and authority, leading to massive artificial shortages and huge blackouts. The companies involved found that they could increase profit potential by withholding electricity (as in the California energy crisis of 2000–2001) or by neglecting safeguards (as in the August 2003 blackout). They understood that these actions were undesirable from the standpoint of American society as a whole, but it is our society, after all, that mandates such powerful incentives to individual and company profit. This was predictable and even predicted, but no effective preventive measures were established. If we wish different outcomes, we must either restructure the marketplace to assure that profit motives align with society's needs, or else impose effective regulation to prevent companies from finding profit in damaging or dangerous actions.

Market Solutions

Aligning profit motives with needs for infrastructure protection against attack would be most desirable, providing maximum delegation of power and responsibility while minimizing conflict of motive and interest. The most direct approach to this is to make companies bear the costs that result from successful attacks on their facilities and services. This in principle would motivate them to do what is needed to avoid or mitigate damaging attacks, including banding together as necessary to take collective action. This would be the pure free-market solution and the one that is arguably best aligned with American principles and values.

Scarcely anyone in our society questions the efficacy of the free market in providing well-defined products and services to meet customer demands at the lowest price. But in a case such as this, experience and theory combine to raise a series of issues to be considered. First, because the incentives would take the form of threat of need to repay money already collected from customers, they would depend on the credibility of some external enforcement mechanism, inevitably governmental in character. The government would have to be prepared not only to act in a rigidly punitive fashion but also to convince the companies that it would act. It has always been difficult for democratic government to do this.

The companies most affected would be limited-liability corporations, and their inherent

limitations of liability would imply a cutoff in threat response. That is, any threat severe enough to endanger the viability of the company should evoke the same level of protection, regardless of whether its effect on society was catastrophic or only serious. Thus, a second issue is that society might be relatively underprotected against the gravest threats.

A third issue arises because any corporation is run by agents—its executives—whose own incentives may be more or less misaligned with those of the corporation's owners, its shareholders. Alignment of executive and owner interests is essential to any free-market solution to the infrastructure protection problem. But this alignment is difficult to achieve when income and cost are separated in time and the magnitude of cost is uncertain. This is a case where cost may be imposed a long way in the future and where its amount (and even incidence) is wildly uncertain. In such circumstances, it is extremely tempting for executives to focus on present-day income and neglect the highly uncertain future costs of infrastructure attack.

Finally, many of the most effective potential responses to threats of attack involve intelligence collection in foreign countries, or the exercise of police powers or military force. Broad delegation of such powers to individuals or corporations would raise issues regarding the nature of our nation and government that far transcend the bounds of this discussion.

Regulatory Solutions

Unless some way can be found to avoid problems that are inherent in a market-based approach, security for our infrastructures will have to depend on direct government control or regulation. Direct government control would raise the issues of delegation of responsibility and control and of alignment of motivations discussed earlier.

In principle, the most desirable way to regulate infrastructure security might well be by private orderings, in which industry participants, recognizing self-interest or social responsibility, evolve structures upon which society could rely. This could minimize the costs of regulation.[51] Internet governance is a prominent example of such private ordering. But there seem to have been no serious efforts yet to develop proposals for private orderings for infrastructure security. For the moment, it is an open question and a challenge to policy analysts.

In the absence of serious suggestions regarding private orderings, we turn to public orderings. In broad principle, all regulation operates by manipulation of the incentives of income and costs. There is a great difference in practice, however, between regulation that threatens to cost executives their freedom and that which merely promises to modify the firm's profit and loss calculus. Here we distinguish between incentive regulation and directive regulation.

A well-known example of incentive regulation is pollution credits (also discussed under the rubrics of emissions trading or cap and trade). The regulator creates a certain number of credits, each conferring the right to emit a defined quantity of pollutants—for instance, 10 million tons of carbon dioxide per year. The credits are allocated to firms by administrative fiat or by auction, and thereafter firms are free either to keep a credit and emit that quantity of pollutant or to sell the credit to another firm. The price of a credit acts as an incentive to the firm to invest in pollution reduction so it can sell the credit. The net effect, ideally, is to concentrate investment in pollution reduction in areas where the greatest improvements can be achieved at the least cost to society as a whole.

There are pitfalls for regulators in such schemes, and they do not always work well.

As outlined earlier, the incentive regulation regime employed to regulate electricity and natural gas distribution in California in the early 2000s offered opportunities for gaming, which were exploited by Enron and other suppliers to gain billions of dollars in extra profits without public benefit. There is wide (although by no means universal) agreement, however, that where they can be appropriately designed and well implemented, incentives provide the most efficient means of regulation.

There are, however, many areas of regulation where incentive regimes have not yet been found feasible or attractive. For instance, issuance of credits permitting a firm to cause a certain number of deaths or maimings probably would not be publicly accepted as a substitute for affirmative direct regulation of safety measures, regardless of any theoretical merits of such a scheme.

Basing regulation to decrease vulnerability to infrastructure attack on such incentives seems open to similar objections. Beyond this, however, attacks themselves are infrequent and variable enough in nature and intensity to raise severe problems in measuring vulnerability. It is a very different situation from that of carbon dioxide emissions or even workplace accidents, where it is possible to gather relatively immediate and direct data on the impact of any control measures.[52]

Thus, it appears that in many areas, effective protection of infrastructures against attack can best—and perhaps only—be assured through directive regulation of infrastructure firms. To the greatest possible extent, this regulation should take the form of performance-oriented requirements that leave to the individual firm the choice of means by which the necessary performance is to be achieved. In safety-related regulation, however, there generally are areas in which the regulators must, as a practical matter, mandate the use or avoidance of specified procedures and equipment. The problems of this sort of regulation are significant: regulators are given power to impose costs without having to answer to the firm's owners, whose only recourse is through administrative, legal, or political appeals. The regulators present the same agency problems as management and may be even less accountable to owners. The only mitigation is that, unlike managers, regulators are not able to profit personally by actions that might damage the firm.

In any event, good practice in process- and equipment-oriented regulation always dictates that firms should be given the opportunity to propose alternatives based on an evidentiary case that what they propose will produce results at least as satisfactory as those mandated in the regulation. Regulators also must be attentive to industry arguments that change in technology or circumstance results in a need for regulatory revision.

One important variant of public ordering is public ordering with private enforcement, in which the rules are publicly determined but are enforced in part or in whole by private appeal to the courts or administrative tribunals. In principle, this offers opportunities to reduce both costs and opportunities for costly bureaucratic meddling. A variant of this is now being employed in regulating electrical system reliability, and it merits attention.

The quest to improve the security of the electrical power infrastructure against natural and intrinsic threats, as well as against many forms of attack, has led recently to the establishment of a formal Electric Reliability Organization. The ERO takes the form of a private nonprofit corporation with close ties to the electric power industry, but endowed with regulatory powers under the supervision and control of a Federal agency, FERC. Earlier experience with NERC (which now runs the ERO) had demonstrated that admonition and appeals to industry-wide and national interest were not adequate, but it is expected that the

ERO will not have to regulate with a heavy hand and will not be a source of significant needless cost for the industry and its customers.

This system of regulation is only now starting to operate, and we cannot be certain how well it will fulfill expectations. Even if it operates exactly as hoped, it will not eliminate blackouts, for that is not possible with an electric grid anything like the one we have. It should be able to reduce their frequency and greatly limit their severity. If it is successful in establishing and enforcing appropriate standards, based in present knowledge, then a blackout like that of August 2003 should never happen again.

It is not possible to be as confident about ERO potential to protect the grid from deliberate attack because our experience with attack is not as comprehensive as that with natural and intrinsic casualties. But analysis and experience indicate that consistent enforcement of ERO standards should make the risk of damage from an attack significantly lower.

The ERO model offers promise as a mechanism for regulation to improve the survivability and operability of many kinds of infrastructures in the face of attacks as well as natural and intrinsic threats.

Cyberspace Infrastructure

Many U.S. infrastructures are shared to some degree with Canada and Mexico, and actions to protect them need to be coordinated closely with those taken by the governments of these nations.

The cyberspace infrastructure has far greater international connections and dependencies than other infrastructures. This raises unique problems of governance and regulation. In electrical power, the United States can and does operate with standards that differ considerably from those used in distant countries, even in such basic matters as AC distribution frequency (60 hertz here but 50 hertz in many other places). In cyberspace, however, international coordination issues are much more complex. Many details of cyberspace infrastructure protection policy must therefore be coordinated with foreign and international bodies.

How Much Is Enough?

The most fundamental of questions in defense planning is always, "How much is enough?" What level of protection is needed? There is no absolute answer, but the question must be addressed explicitly and systematically.

How are we to weigh warnings of threats to cyberspace infrastructure, and to other infrastructures, against other threats? How much of our attention and resources should be devoted to countering them? Our limited experience with such threats makes the problem much more difficult.

Examples drawn from other risk fields help to illustrate the issues. For example, in 2001, about 5,000 extra people died in U.S. traffic accidents because more than a quarter of car and truck occupants failed to wear seatbelts.[53]

In the same year, drownings killed 3,300.[54] Both tolls exceeded that of terrorist attacks and continue year after year, but public concern about terrorist attacks is far higher than concern about seatbelt use or water hazards, and far more resources are devoted to combating terrorism.

Another comparison is illuminating. The average annual toll from asteroid impacts is estimated to be lower than that from machinery accidents, but the two averages are arrived at in different ways. Machinery accidents occur frequently and relatively regularly, each involving a small number of deaths. By contrast, fatal asteroid impacts are rare, coming at intervals of 1,000 years (for relatively small incidents) to 100 million years or more (for catastrophic ones), but they could involve huge numbers of fatalities and even extinction of our species. Until recently, the asteroid threat was generally discounted, but over the past two decades, public concern has increased as evidence of historical impacts has been discovered and analyzed.[55]

Surveys and experimental studies analyzing how we evaluate and respond to perceived risks confirm, as these examples suggest, that people are not rigorously logical about such matters. The perceived dreadfulness of a threat has a lot to do with response to it, as does the form in which information regarding its probability of occurrence is received.[56] We are prone to be less concerned, relative to objective quantitative risk level, about common and familiar risks such as heart disease or motor vehicle accident than about shadowy and little- understood menaces such as cyber attack.

A further complicating factor is the concern decisionmakers often feel regarding public reactions to attacks or failures, such as fear of mass panic.[57] Less immediately, decisionmakers fear weakening of public support for necessary measures or sacrifices. For instance, almost every wartime President since Abraham Lincoln has worried that the people would be unwilling to accept casualties as the price of victory.

Social scientists find, however, that support for a conflict depends not so much on particular levels of casualties as on belief in the cause for which it is fought and the probability of success.[58] Similarly, it is found that mass panic is very rare, even in circumstances that might seem to provide ample justification.[59] For most people, emotional factors play the dominant role in determining overall response to issues such as demand for defense, but this by no means implies that the public is irrational about such subjects.[60] Our emotional apparatus evolved as it has because it aided survival in threatening environments, and it continues to serve us in this role.[61] Decisionmakers often resort to measures intended to manipulate the public's emotional responses to gain support, but while manipulation can seem effective in the short term, it often evokes a backlash over the longer term.

Ultimately, the question of how much is enough infrastructure protection can be answered only by the public through the political process. Policymakers hoping for a sound answer will do well to provide the public with clear, credible information.

Policy Recommendations

The foregoing examination of infrastructure protection issues has revealed a lack of broad and systematic policy. The following 10 recommendations to remedy this are presented for consideration at the highest levels of government.

Unify Policy Direction

It is unrealistic to expect that all of the aspects of policy relating to infrastructure protection can or should be united under a single governmental department or agency, but it is essential that a positive mechanism be put in place to assure effective interagency coordination. Because essentially the same actions are required to protect each infrastructure

against natural disasters, accidental and intrinsic failures, and threats from terrorist, military, and criminal attack, the interagency mechanism must encompass them all. The precise organization of this mechanism requires further study by Congress and the Executive.

Specialize Policy Direction

While there should be unity in overall direction, policy direction for the various infrastructures should be tailored to their specific nature and needs. Thus, for each infrastructure there should be a subordinate interagency process involving the agencies that have specialized knowledge and responsibility.

Strengthen and Unify Regulation

While directive regulation of infrastructure firms at the process level has important pitfalls, there is no evident substitute for it with regard to protection of infrastructures. Absence of effective regulation leaves firms exposed to commercial pressures that work against protection and tend to prompt a "race to the bottom." For each infrastructure there should be a single, well-informed regulator with the knowledge and incentives to strike the right balance between risk and economic benefit. The ERO represents a promising approach, which should be studied as a potential model for other infrastructures.

Define State and Local Roles

State primacy in policy and regulation for infrastructures has been undercut by the trend toward larger interstate networks, but state and local government agencies nevertheless retain a very important role. The Federal interagency mechanism for infrastructure protection policy and related regulatory apparatus must be linked closely with the relevant state agencies. How this is to be accomplished must be worked out directly with the states.

Define International Interfaces

Cyberspace infrastructure networks depend on international connections, but in this they differ in degree rather than in kind from other infrastructures. In practically all cases, it is necessary to coordinate international action to secure infrastructures most effectively. Again, the ERO appears to offer a promising model, with the United States playing a positive leadership role by offering a structure with mutual benefit and by demonstrating readiness to modify its positions to meet the legitimate interests of others.

Mandate Effective Systems Engineering for Infrastructure-related Software

Undependable software is one of the greatest vulnerabilities of infrastructure systems. The cost-driven trend to wide use of undependable COTS and open-source software is exacerbating the risks. Software dependability will not achieve the necessary standards unless effective systems engineering is mandated for infrastructure systems.

Don't Take No for an Answer

There will be some in the infrastructure industries who resist any directive regulation, regardless of justification. Their objections must not be accepted. It is instructive to look at the 40-year struggle to avert massive electrical blackouts without directive regulation; that struggle culminated in the August 2003 Northeast blackout, whose magnitude was multiplied by widespread failure to comply with existing voluntary standards. Any decisionmakers who are tempted to give in to industry pressures against regulation should consider carefully what they would say if that decision cleared the way for a successful attack.

Establish and Implement Clear Priorities

While there is no clear limit to potential threats against infrastructures, there are limits to the resources that can be used for protection. An attempt to protect everything against all possible threats will result in failure to protect the most crucial targets adequately against the most important threats. Setting and keeping priorities for the allocation of financial and management resources are essential in order to provide effective protection.

Inform the Public Clearly and Accurately

Many of the decisions regarding protection of infrastructures will be technical and should be made by those with appropriate expertise. It is a serious error, however, to imagine that the key decisions in this area can be held within any closed group. The integrity of infrastructures affects everyone in our society, and the public will demand that its views be heeded at critical junctures. A systematic ongoing effort to make full and objective information available is the best guarantee of informed and considered public input. It also is the best way to ensure that the public will feel confidence in those who direct infrastructure protection efforts and will pay appropriate attention to their recommendations.

Conduct a Continuing Program of Research

Many important questions remain unsettled and more will arise as threats, technology, and economic conditions change. The policy and regulation institutions must have the authority, resources, and responsibility to sponsor and guide broadly conceived programs of research to serve their information needs. Know- ledge can be expensive, but its absence can be much more so.

---

[1] Stuart Chase, *Men and Machines* (New York: Macmillan, 1929), 288–289, 297.

[2] Mark Clodfelter, "Pinpointing Devastation: American Air Campaign Planning before Pearl Harbor," *The Journal of Military History* 58 (January 1994), 75–101; United States Strategic Bombing Survey (hereafter USSBS), *Over-All Report (European War)* (Washington, DC: U.S. Government Printing Office, 1945); USSBS, *The Effects of Strategic Bombing on the German War Economy* (Washington, DC: U.S. Government Printing Office, October 31, 1945); USSBS, *The Effects of Strategic Bombing on Japan's War Economy* (Washington, DC: U.S. Government Printing Office, December 1946).

[3] USSBS, *Statistical Appendix to Over-All Report (European War)* (Washington, DC: U.S. Government Printing Office, February 1947).

[4] USSBS, *The War against Japanese Transportation, 1941–1945* (Washington, DC: U.S. Government Printing Office, May 1947).

[5] Attacks against Japan's oil infrastructure were judged, by contrast, to have been "almost superfluous" because the war on transportation had already largely idled the refineries for want of feedstocks. USSBS, *The Effects of Strategic Bombing on Japan's War Economy*, 46–47.

[6] Thomas A. Keaney and Eliot A. Cohen, *Gulf War Air Power Survey: Summary Report* (Washington, DC: Air Force Historical Studies Office, U.S. Government Printing Office, 1993).

[7] James Glanz, "Iraq Insurgents Starve Capital of Electricity," *The New York Times*, December 19, 2006.

[8] Network theory is an active research field, and a number of important discoveries have been made in recent years. See Albert-László Barabási, *Linked: How Everything Is Connected to Everything Else and What It Means for Business, Science, and Everyday Life* (New York: Plume Books, 2003), for an excellent nontechnical overview by a leader in modern network theory. For a more technical summary, see Réka Albert and Albert-László Barabási, "Statistical Mechanics of Complex Networks," *Reviews of Modern Physics* 74 (January 2002), 47–97. Duncan J. Watts, *Six Degrees: The Science of a Connected Age* (New York: Norton, 2003) is another sound nontechnical overview by a prominent scientist, oriented more toward social networks than infrastructures.

[9] The terms *link*, *edge*, and *line* are used interchangeably in discussing networks. Similarly, *node*, *vertex*, and *point* all refer to the same thing. Networks are also referred to as *graphs*.

[10] The most fundamental statistic describing a complex network is p_k, the proportion of the network's nodes having exactly k links connecting to other nodes. Naturally, k has to be a whole number; a node cannot have 2½ or 4.38 links. A node having k links is said to have *degree k,* and the table or set of values of p_k for all values of k is called the degree distribution of the network.

[11] It seems odd to speak of "random" networks in connection with infrastructures, where the design choices are not made by rolling dice or drawing from a hat. But infrastructures do tend to develop through a sequence of choices reflecting a variety of changing considerations, and this gives them a certain statistical or random-like character.

[12] Dmitri Krioukov et al., "The Workshop on Internet Topology (WIT) Report," *Computer Communication Review* 37, no. 1 (2007), 69–73.

[13] Computer Science and Telecommunications Board, *The Internet Under Crisis Conditions: Learning from September 11* (Washington, DC: National Academies Press, 2003).

[14] For an overview of the electricity grid, see Jack Casazza and Frank Delea, *Understanding Electric Power Systems: An Overview of the Technology and the Marketplace* (Hoboken, NJ: IEEE Press and Wiley-Interscience, 2003).

[15] In electrical terminology, any equipment or system that draws electric power is a *load*.

[16] The advent of new transmission technology such as high-temperature superconductors may reduce but not eliminate the advantage of high-capacity transmission.

[17] Transformers do not work for direct current (DC), and there is no simple DC equivalent.

[18] To distribute current from high-voltage direct current lines, it is necessary first to convert it to alternating current.

[19] There is a partial exception to this, in that one of the synchronous regions, that in northeastern Canada operated by Hydro-Québec, is a major power exporter from its large hydroelectric generating facilities. The power is almost all exported via a high-voltage direct current line that prevents frequency disturbances from spreading, but a sudden major voltage disturbance, either in this region or in the northeastern U.S. region to which it sells power, would put the other region under stress.

[20] Richard F. Hirsh, *Power Loss: The Origins of Deregulation and Restructuring in the American Electric Utility System* (Cambridge: MIT Press, 1999).

[21] Frank A. Wolak, "Diagnosing the California Electricity Crisis," *The Electricity Journal* 16, no. 7 (August 2003), 11–37.

[22] Secretary of Energy Advisory Board, *Maintaining Reliability in a Competitive U.S. Electricity Industry: Final Report of the Task Force on Electric System Reliability* (Washington, DC: Department of Energy, September 29, 1998).

[23] The comprehensive official report is illuminating about the mechanisms of failure. U.S.- Canada Power System Outage Task Force, *Final Report on the August 14, 2003, Blackout in the United States and Canada: Causes and Recommendations* (Washington, DC, and Ottawa: U.S. Department of Energy and Natural Resources Canada, April 2004). Also useful is North American Electric Reliability Council (NERC), *Technical Analysis of the August 14, 2003, Blackout: What Happened, Why, and What Did We Learn?* (Princeton: NERC, July 2004). Also see Richard Pérez-Peña, "Utility Could Have Halted '03 Blackout, Panel Says," *The New York Times*, April 6, 2004, A16. "Blackout 101," a series of tutorial presentations developed by experts to inform Congress, is available at <www.ieee.org/portal/site/pes/menuite m.2b4756efb9a16c58fb2275875bac26c8/index.jsp?&pName=pes_level1&path=pes/subpages/meetings-folder/other_meetings&file=Blackout_101.xml&xsl=generic.xsl> or <http://tinyurl.com/yur6o4>.

[24] In many cases, the lines and generators were not actually immediately threatened but only appeared to be due to the large power surges triggered by the cascade. If the safety relays had been better able to discriminate between real and apparent threats, the outage could have been much less widespread.

[25] U.S.-Canada Power System Outage Task Force, 131–137.

[26] S. Massoud Amin and Philip Schewe, "Preventing Blackouts," *Scientific American* 296, no. 5 (May 2007), 60–67; S. Massoud Amin and Bruce F. Wollenberg, "Toward a Smart Grid," *IEEE Power and Energy Magazine* 3, no. 5 (September-October 2005), 34–38; and Clark W. Gellings and Kurt E. Yeager, "Transforming the Electric Infrastructure," *Physics Today* 57, no. 12 (December 2004), 45–51. See also "Ideas Generated for Transforming the Electric Infrastructure," *Physics Today* 58, no. 5 (May 2005), 13–15.

[27] Supramaniam Srinivasan et al., "Fuel Cells: Reaching the Era of Clean and Efficient Power Generation in the Twenty-first Century," *Annual Reviews of Energy and the Environment* 24 (1999), 281–328.

[28] John M. Deutch and Ernest J. Moniz, "The Nuclear Option," *Scientific American* 295, no. 3 (September 2006), 76–83; James A. Lake, Ralph G. Bennett, and John F. Kotek, "Next- generation Nuclear Power," *Scientific American* 286, no. 1 (January 2002), 72–81.

[29] George W. Crabtree and Nathan S. Lewis, "Solar Energy Conversion," *Physics Today* 60, no. 3 (March 2007), 37–42; Ken Zweibel, James Mason, and Vasilis Fthenakis, "A Solar Grand Plan," *Scientific American* 298, no. 1 (January 2008), 64–73; Daniel M. Kammen,

"The Rise of Renewable Energy," *Scientific American* 295, no. 3 (September 2006), 84–93.

[30] Kammen; Karl Stahlkopf, "Taking Wind Mainstream," *IEEE Spectrum* (June 2006).

[31] W. Wyatt Gibbs, "Plan B for Energy," *Scientific American* 295, no. 3 (September 2006), 102–114.

[32] The United States has 4 million miles of public roads, 100,000 miles of Class I rail lines, and 26,000 miles of waterways that also represent networked infrastructures, but they are not discussed here because they are less vulnerable to cyber attack.

[33] Department of Transportation, "National Transportation Statistics, 2007," table 1–10, available at <www.bts.gov/publications/national_transportation_statistics/>.

[34] Paul W. Parfomak, "Pipeline Safety and Security: Federal Programs," CRS Report RL33347 (Washington, DC: Congressional Research Service, July 11, 2007), 1–2.

[35] Ellen Nakashima and Steven Mufson, "Hackers Have Attacked Foreign Utilities, CIA Analyst Says," *The Washington Post*, January 19, 2008, A4.

[36] Andy Greenberg, "America's Hackable Backbone," August 22, 2007, available at <www. forbes.com/2007/08/22/scada-hackers-infrastructure-tech-security-cx_ag_0822hack. html>.

[37] John Rollins and Clay Wilson, "Terrorist Capabilities for Cyberattack: Overview and Policy Issues," CRS Report RL33123 (Washington, DC: Congressional Research Service, January 22, 2007).

[38] Howard F. Lipson, Nancy R. Mead, and Andrew P. Moore, "Can We Ever Build Survivable Systems from COTS Components?" CMU/SEI–2001–TN–030 (Pittsburgh: Carnegie Mellon University, Software Engineering Institute, December 2001).

[39] It may be objected that attackers do routinely find and exploit important design flaws in the software on Internet-connected computers. However, EMS and SCADA systems are much less available for examination and probing. Moreover, the relative simplicity of SCADA systems, in particular, allows less opportunity for serious hidden flaws. Thus, devastating attacks on these systems are far less likely. This is borne out by experience, as attacks on EMS and SCADA systems by hackers have thus far been much less common and generally less serious than those directed at Internet computers and servers.

[40] Control Systems Security and Test Center, "A Comparison of Electrical Sector Cyber Security Standards and Guidelines," INEEL/EXT–04–02428, Revision 0, Idaho National Engineering and Environmental Laboratory, October 28, 2004.

[41] See, for example, Richard A. Clarke, *Breakpoint* (New York: Putnam, 2007); Will O'Neil, *The Libyan Kill* (New York: W.W. Norton, 1980).

[42] Since the March 2003 American invasion, insurgents have targeted Iraq's infrastructure, especially its oil and electricity infrastructure. There have been repeated physical attacks on oil production facilities and especially pipelines. See Iraq Pipeline Watch, available at <www.iags.org/iraqpipelinewatch.htm>. Electric grid attacks have been aimed at high-voltage transmission lines. Many people have been killed. There are no reports of cyber attacks, but neither the oil nor electrical system has much in the way of SCADA or operational management systems as potential cyber targets. The identities of the attackers and their strategies, goals, and incentives are unclear; their motives may be profit as much as politics. They have not destroyed the oil or electrical systems, but it is not clear whether that is their intention; they certainly have imposed major problems and costs. The lack of adequate and reliable electrical power has been a factor in demoralizing and angering the population and undercutting support for U.S. objectives. Loss of oil revenues has significantly weakened the Iraqi government and forced the United States to subsidize it. Substantial U.S. forces have had to be devoted to protection of infrastructure. The limitations of domestic production of petroleum products and electricity have forced large-scale trucking of fuels from Iran, Kuwait, and Turkey, and substantial effort is required to protect the fuel convoys. The problems of fuels logistics have been greatly exacerbated by an ill-considered American decision at an early stage to boost Iraqi electrical generating capacity with combustion turbines (turbogenerators driven by the exhaust from aircraft-type jet engines), which were ill suited to Iraqi needs and conditions and which required fuel that must be trucked in because it is not available in Iraq. Glenn Zorpette, "Re-engineering Iraq," *IEEE Spectrum* 43, no. 2 (February 2006), 22–35. Advance examination of Iraq's real needs on a total-system basis would have paid significant dividends.

[43] Daniel Jackson, Martyn Thomas, and Lynette I. Millett, eds., *Software for Dependable Systems: Sufficient Evidence?* (Washington, DC: National Academies Press, 2007).

[44] President's Commission on Critical Infrastructure Protection, *Critical Foundations: Protecting America's Infrastructures: The Report of the President's Commission on Critical Infrastructure Protection*, October 1997, available at <www.fas.org/sgp/library/pccip.pdf>.

[45] Parfomak.

[46] The latest Network Reliability and Interoperability Council concluded its work in 2005. See <www.nric.org/>.

[47] U.S. Department of Energy, Secretary of Energy Advisory Board, *Maintaining Reliability in a Competitive U.S. Electricity Industry*, available at <www.seab.energy.gov/publications/ esrfinal.pdf>.

[48] *Mandatory Reliability Standards for Critical Infrastructure Protection*, FERC Docket No. RM06– 22–000, Order No. 706, issued January 18, 2008.

[49] See <www.dhs.gov/xprevprot/committees/>.

[50] The Tennessee Valley Authority is the most prominent exception.

[51] Steven L. Schwarcz, "Private Ordering," *Northwestern University Law Review* 97, no. 1 (January 2002), 319–349.

[52] It is in some way similar to the problem of regulating hedge funds and like entities, which can have very infrequent but extremely costly failures. See Dean P. Foster and H. Peyton Young, "The Hedge Fund Game," Brookings Institution Center on Social and

Economic Dynamics Working Paper No. 53, November 14, 2007.

[53] Based on data in "National Transportation Statistics," available at <www.bts.gov/publications/national_transportation_statistics/>.

[54] Centers for Disease Control and Prevention, "Web-based Injury Statistics Query and Reporting System (WISQARS)," available at <www.cdc.gov/ncipc/wisqars/>.

[55] Clark R. Chapman, "The Hazard of Near-Earth Asteroid Impacts on Earth," *Earth and Planetary Science Letters* 222, no. 1 (May 2004), 1–15.

[56] M. Granger Morgan, "Risk Analysis and Management," *Scientific American* 269, no. 1 (July 1993), 32–41; Paul Slovic et al., "Risk as Analysis and Risk as Feelings," *Risk Analysis* 24, no. 2 (April 2004), 311–322.

[57] "What would we do if the United States were attacked and New York menaced? . . . A deafening roar—another and another. . . . There is another blast—and the rush to the streets begins. . . . The streets are tightly filled before a third of the office workers have poured out. Tardy ones claw and clutch and scramble, clambering on top of those who have fallen. Before long there is a yelling, bloody, fighting mass of humanity." William Mitchell, "When the Air Raiders come," *Collier's,* May 1, 1926.

[58] Christopher Gelpi, Peter D. Feaver, and Jason Reifler, "Success Matters: Casualty Sensitivity and the War in Iraq," *International Security* 30, no. 3 (Winter 2005/2006), 47–86, provides a guide to earlier literature. See also Louis J. Klarevas, Christopher Gelpi, and Jason Reifler, "Casualties, Polls, and the Iraq War," *International Security* 31, no. 2 (Fall 2006), 186–198, for a critique and response. For a survey of data on support for wars between 1942 and 1993, see Eric V. Larson, *Casualties and Consensus: The Historical Role of Casualties in Domestic Support for U.S. Military Operations*, MR–726–RC (Santa Monica, CA: RAND, 1996), 105–120.

[59] Lee Clarke, "Panic: Myth or Reality?" *Contexts* 1, no. 3 (Fall 2002), 21–26.

[60] "Risk as Analysis and Risk as Feelings: Some Thoughts about Affect, Reason, Risk, and Rationality," *Risk Analysis* 24, no. 2 (April 2004), 311–322.

[61] Antonio R. Damasio, *The Feeling of What Happens: Body and Emotion in the Making of Consciousness* (New York: Harcourt Brace, 1999).