Standards Today

A Journal of News, Ideas and Analysis

A publication of CONSORTIUM INFO.ORG

GesmerUpdegrove LLP

January–April 2011 Vol. X, No. 1

FEATURE ARTICLE:

Cyber Security and the Vulnerability of Networks: Why we Need to Rethink our Cyber Defenses Now

Andrew Updegrove

Abstract: *While network effects provide great benefits, they can also lead to great risks as dependencies evolve, especially where alternative networks do not exist. The Internet is increasingly becoming the unique host for essential capabilities and services, as government, finance, energy management, supply chains, telephony and much more abandon traditional channels and migrate to the Web. Now, the cloud computing model is being heralded as the information technology (IT) architecture of the future, raising the prospect that rapidly growing concentrations of crucial data and software (of all types) will exist in fewer and fewer data centers, despite the fact that measures are unlikely to be deployed that can provide complete security to these centers against cyber and terrorist attacks (as well as actual war). It is therefore vitally important that virtual and physical security standards be developed and implemented to guard against the very real prospect that an asymmetric attack by any of a variety of potential enemies could bring a modern society to the verge of collapse. In this article, I outline the growing risks, and suggest the types of physical and virtual security standards frameworks that should be developed and implemented to minimize vulnerability to catastrophic attacks, and to maximize the likelihood that a society could rapidly recover from a massive cyber or physical attack against its IT infrastructure. I conclude by evaluating the degree to which the Obama administration's new cyber security proposal will be sufficient to protect the United States from systemic cyber risks.*

The old Internet is dying off and a new Internet is emerging

> *Salesforce.com CEO Marc Benioff, describing the move to cloud computing.*

Throughout the long decades of the Cold War, the spectre of the end of life as we know it hung over the United States, the Soviet Union and many of their allies. All too credible images of societies reduced to lawless, pre-industrial era conditions permeated literature, television and the cinema. With the fall of the Soviet Union,

that grim prospect receded into memory, seemingly never to return unless nuclear tensions should rise again.

While the end of the nuclear madness of the Cold War is to be celebrated, we should not assume that the risk of a societal collapse caused by hostile action has disappeared entirely. Indeed, a sobering case can easily be made that just as we stockpiled our way into a nuclear stalemate we could not control, we are now building out an insecure information technology (IT) infrastructure that leaves us vulnerable to equally dire and uncontrollable consequences. Only this time there will be many enemies (some unknown) with the motive and ability to strike rather than a single enemy that can be monitored and approached in negotiations.

This new area of vulnerability goes by names such as "cyber threat" and "cyber war," and its existence is now well recognized. The severity of that threat has also been propounded, perhaps most thoroughly by former Special Advisor to the President for Cybersecurity Richard A. Clarke, in his recent book, *Cyber War: The Next Threat to National Security and What We Need to Do About It* (2010). But Clarke speaks mostly of current vulnerabilities, and does not extensively explore may expand in the future. In fact, these risks can be expected to increase exponentially in a remarkably short time if certain trends in IT deployment persist. Most significantly, the increasing popularity of the "cloud computing" IT hosting model (described in greater detail below) will result in more and more data and software being hosted in fewer and fewer locations. When combined with our increasing dependence on IT and the Internet to host all important public and private services and functions, this means that progressively more damage could be inflicted by a physical or cyber attack upon fewer data centers.

The increasing popularity of the "cloud computing" IT hosting model will result in more and more data and software being hosted in fewer and fewer locations

In this light, it is disturbing to acknowledge that our ability to defend ourselves against even the cyber threats of today is inadequate, and that our will to deploy the defenses that we do have is all too often lacking. As we look into the future, we will need to take cyber security much more seriously. If we fail to do so, our risk will increase dramatically, while our means to manage it will even more challenging.

In this article, I will make the case that realistically conceivable cyber attacks could bring a modern society to its knees if current trends in the centralization of computing resources and essential data continue unabated without the development and deployment of adequate safeguards. I will also suggest the type of practices and standards that could provide increased security, as well as minimize the impact of such attacks when they nonetheless inevitably occur.

As this article was being finalized, the Obama Administration released its own plan for increasing national cyber security. At a high level, the details of the relevant portions of those plans track my recommendations very closely, and I will close by summarizing what the Obama plan would and would not accomplish.

The nature and magnitude of the risk identified can best be explained through a hypothetical example.

A look into the near future: As the third decade of the 21st century dawned, it seemed that humanity had finally embraced the need to solve its energy and global warming crises. The watershed moment had arrived in 2013, when virtually all developed and developing nations at last signed a global treaty imposing aggressive penalties for failures to dramatically reduce greenhouse gas emissions.

That accord would have been impossible absent the multiple disasters of 2011, the year in which oil first reached $200 a barrel and stayed there, as one Arab nation after another was wracked by civil unrest. Pipelines, refineries and port facilities were sabotaged, and even destroyed, as one side or the other gained or lost ground.

The other motivating force was the months-long cascade of disasters at the Fukishima Dai-Ichi nuclear reactor complex in Japan. Perhaps the result would have been different if the enormous earthquake and tsunami had affected only a single reactor instead of six. Emergency staffs performed heroically, but just as the situation seemed at last to have been brought under control, a devastating fire erupted that once more exposed the fuel rods in three of the reactors. For months thereafter, whenever the situation in one reactor was brought under control, a new aftershock, explosion, fire, or radiation leak would set the process back in another, keeping the story in the news for most of the rest of the year.

In Europe and the United States, the possibility of starting new, or of renewing the licenses of existing, nuclear reactors evaporated. The power of the Green Party in many EU countries surged, and in the United States, the specter of an "oil shock" that would tip the nation back into recession led to a once in a lifetime bipartisan commitment to accelerate the development of renewable energy and the reduction of energy consumption.

Concurrently, a new business model called "cloud computing" was coming into its own, and millions of new servers were brought on line as a growing flood of public and private sector entities moved their software and data to third-party hosts that maintained them for a fee. Predictably, the IT industry rapidly consolidated as data hosting became commoditized and profit margins narrowed. After ten years, a handful of companies was hosting the data and software of most private and public enterprises.

With energy costs at all time highs and the percentage of national energy demands allocable to IT rising rapidly, the new cloud provider giants built enormous server farms adjacent to lowest-cost generation facilities. Government subsidies, tax breaks and social pressure all ensured that these new facilities would be the most "green" enterprises ever built. By 2020, astronauts circling the globe could easily pick out server farms the size of small cities, each situated next to a renewable energy source: hydroelectric dams in Canada, solar installations in the southwest, wind farms in Texas and off the shores of New England. The same revolution transformed the IT and power infrastructures of Europe and the Pacific Rim as well.

Governments, universities and public companies all now had private clouds. With the rise of the "key chest" model of media ownership, individuals relied on remote hosting for not only their audio, video and graphic files, but for their personal documents as well. Internet and wireless telephony had now almost totally replaced traditional land-line phones. Happily, the lessons of the Japanese tsunami had been learned, and all of this data and software was redundantly hosted in at least two locations, with automatic "failover" protocols in place.

Nor did the hosting of data end there. After Congress resolved the copyright ownership of "orphan" works of authorship in 2012, the digitization of the world's books was completed. Even the largest libraries began pulping all but volumes of the greatest historical significance. Budgets previously spent on bricks, mortar, shelving and physical books were now spent on acquiring access rights and the means to deliver millions of remote, digitized works to the eReaders of library patrons everywhere.

With the costs and benefits of central hosting of information so compelling, local storage of information had become as rare as an AOL dial up account. Experts estimated that c. 85% of all of the world's important data and software was now hosted in twenty-three gigantic data farms that collectively consumed a spectacular 9% of the global output of electricity. Together, this new system of instant, global, open access was widely and justifiably acknowledged to be one of the great achievements of the modern world.

9/11/21: Within minutes of each other on the twentieth anniversary of the 9/11/01 terrorist attacks, twenty-three individuals scattered around the United States, Europe, China, Japan, South Korea and Australia put their plan into operation. This time, instead of boarding four commercial airliners, each fired up the engine of an ultralight kit aircraft that could be legally flown without obtaining a pilot's license. Some took off from dirt roads in the desert, others from pastures. Naturally, none was legally required to file a flight plan, and all flew scarcely a hundred feet above the ground. Each was accompanied by a package about the size and shape of a piece of roll aboard luggage, courtesy of an illicit arms dealer in a former Soviet republic. On each package was a digital readout, synchronously counting down to zero.

Not long after, the pilots approached their destinations. Each looked to the readout on his package, and began a gradual ascent, calculated to position his aircraft approximately 1500 feet above one of the massive server farms by the time the countdown reached zero.

At the appointed time, twenty-three Cold War era tactical, battlefield nuclear weapons exploded in blinding flashes of light and energy. Like uncommonly ferocious funeral pyres, the mushroom clouds that roared into the sky marked the annihilation of the collective knowledge of humanity.

In a wink of an eye, the earth had been cleansed of the heresies of modern civilization. The world was once again as Mohammed had known it.

In the first winter that followed, a billion people starved to death.

Technology and societal fragility: Impossible? Leaving aside the availability of the nuclear weapons, accomplishing everything else described above would be trivially simple. Unfortunately, there are unconfirmed reports that some "suitcase size" Soviet era nuclear weapons are unaccounted for, and tactical weapons of a similar size certainly still exist. Bringing such lightweight weapons (for example) across the Canadian border in any of a number of ways (e.g., using an ATV to travel through fields or woods, or using a snow mobile to cross a frozen lake in the winter) could be readily accomplished.

Nor are radical Islamic fundamentalists a necessary element to drive such a story line. Rather than religious terrorists, one could easily substitute North Korea, Muammar Quaddafi (should re survive in power), or tomorrow's megalomaniac or successor to the morality of Pakistan's A.Q. Kahn.

Finally, while tactical nukes provide a more vivid image, the same impact could be accomplished with far less drama, using conventional explosives against a server farm's connections to the electric grid and backup generators (if any). In an even less dramatic - and more likely - scenario, it would be enemy computer hackers that would launch a two part cyber attack against one or more nations. The first wave would neutralize the cyber defenses of the server farms in order to give the second wave the short time needed to destroy the data hosted on the millions of servers that were now unprotected. As with a neutron bomb, a society's ability to function would be destroyed while its hardware and other infrastructure remained eerily intact.

Still too far fetched? Perhaps not. As demonstrated by the Stuxnet worm (presumably) unleashed by the Israelis against Iranian nuclear fuel centrifuges, havoc can be wrought on even the most secure, top secret, and complex computer systems by undetected computer worms. How much more vulnerable might a server farm be, with its millions of simultaneous inputs and outputs to and from the Internet, and gigabits per nanosecond of data flow? In April of this year, U.S. authorities announced that they had disabled a botnet that had captured over two *million* PCs, all of which together could have been used to mount such an assault. The Stuxnet botnet is estimated to comprise 12 million computers, but the same result could also be achieved from the inside by a single disloyal (or even innocent) employee with an infected thumb drive.

Indeed, it should be noted that the only reason the Stuxnet worm attack came to public attention was because it attacked innocent "bystander" networks as well as its Iranian target. Perhaps future cyber worms or Trojans created for one purpose may propagate catastrophically to systems throughout the world.

In any of these scenarios, application and operating system software would be equally vulnerable to destruction. As a result, not only could data disappear, but the means to reenter and manipulate data could vanish as well. As could also the software that manages airlines, financial institutions, and just about everything else.

Whether a nation attacked in such a fashion could get back on its feet without descending into anarchy would depend upon four critical factors: the nature and extent of the data and software destroyed; whether backups of the destroyed

resources existed; whether sufficient undamaged infrastructure remained upon which backup copies of affected data and software could be restored; and whether all of this could be achieved by broadly distributed IT professionals relying on a highly compromised telecommunications system.

The network vulnerability effect. In order to grasp how vulnerable we are becoming, it is necessary to more fully understand the nature of networks, and the dependencies to which reliance upon them can lead.

One of the great discoveries of the industrial and information ages is the power of the "network effect." At the core of that discovery was the realization that the more people connect to a network, the more valuable that network is likely to become to each connected person. While postal service provided an early example, it was the railway that most dramatically demonstrated the importance of a networked economy. After standard track gauges were agreed upon, local railways could be spliced together into national, and eventually continental networks. Once this standard-enabled process began, every station and every industrial and agricultural siding that was added increased the value of a nation's railway system as an engine of growth. Within a few decades of the first deployment of locomotives on steel tracks, railways became the most pervasive means of long distance transportation for people and freight in those nations that were entering the Industrial Age.

Later, telecommunications systems followed a similar evolution from local, to national, to international connectivity, revolutionizing commerce and news dissemination along the way. Today, the spread of the Internet and the Web are transforming the world even more broadly.

But with value also comes dependence. In each of the examples noted above, if the network becomes unavailable, then the activities that rely upon the network must cease, except to the extent that they can be shifted over to alternative systems – assuming such systems exist.

In the case of transportation, multiple alternative networks do exist, in the form of air, highway, railway and water transport. Each of these networks has become more specialized vis-à-vis what it carries and where, and therefore the loss of any of these networks would cause serious disruption. But the development and wide adoption of standardized pallets and cargo containers makes trading cargoes feasible among many types of river, rail, ocean, and highway carriers. These alternatives have sufficient capacity to prevent a total stoppage of those commercial and societal functions that are most dependent on any single transportation system that might be immobilized by (for example) a strike. Moreover, no single cause or catastrophe could totally destroy any of these individual networks, much less all of them.

Highway transport would be particularly difficult to disrupt on a national basis, because (unlike railways and airports) it has few choke points that could be destroyed to break up large portions of the entire network. Even where major highways intersect, local roads connect adjacent exits, and also connect the same destinations. The highway system also lacks central control mechanisms, because each vehicle has its own driver, capable of making independent routing decisions.

The main points of vulnerability that the motor vehicle system does include are its bridges, but even the destruction of many bridges would leave regional transportation largely intact.

Because this multi-modal transportation system comprises a network of networks, it therefore exhibits several important characteristics that limit the risks of becoming dependent upon a transportation-based society and economy:

- It has *low vulnerability*, because each network has few points at which an interruption would have cascading systemic effects. And very few of the chokepoints for one system are common to another (e.g., few railways and highways cross rivers on the same bridges, airports don't land on freight yards, and railway systems don't intersect on runways).

- It has *redundancy*, because it offers alternative means of transportation between the same destinations.

- It has *resiliency*, because even with significant damage to its infrastructure, the transportation network of networks could continue to operate at reduced levels, with much of the traffic from any affected network transferring to a lesser, or differently, affected network.

As a result, national transportation systems have historically been at risk to impairment, but not immediate or total disablement, even in the event of heavy air or ground attack.

Historically, power systems have also exhibited relatively high degrees of resiliency and durability, due to the relatively wide distribution of generating facilities and control, and the ability to "load balance" across these networks. So also with telecommunications, where land lines, radio and satellite communications offer a degree of redundancy, at least for crucial communications.

The Internet, on the other hand, is evolving in ways that will decrease rather than increase the vital qualities of low vulnerability, redundancy, and resiliency. Packets of information can only travel over the telecommunications system, and can only be processed by computers connected to that system. And unlike the analog voice communications of the past, modern telecommunications involve enormous movements of digitized data.

Were the Internet to somehow be rendered unavailable, virtually everything in society today would grind to a halt, because the system can no longer effectively revert to voice-based data transmission, even if phone lines somehow remained in working order. And even voice communications are shifting to VoIP (voice over Internet Protocol) and wireless technology – meaning that the day may not be far off when the loss of the Internet backbone would render even telephones mute. Indeed, the announcement by Microsoft on May 10 that it will acquire VoIP service provider Skype for $8.5 billion signals that this process will accelerate.

Unfortunately, the Internet is also becoming more vulnerable and less resilient. Most notably, the software and servers that support it are vulnerable to attack by hackers, terrorists and national enemies, as are the systems and databases that

are connected to it. The Internet can even be shut off (or severely constrained) on a nation by nation basis, as has already occurred in countries such as Egypt and Iran.

Despite these realities, we are allowing our dependence upon the Internet to increase logarithmically as more and more essential services are deployed across it, from finance, to government, to the smart grid, to health care, to supply chains, to streaming video. If it is not yet completely accurate to say that disabling the servers that serve the Internet would bring the world to a stop, that day is certainly just around the corner.

To give a single example, let us examine how the rise of the Internet has affected the vulnerability, redundancy and resilience of the transportation network. Because each of the existing sub-networks has now become variously dependent on the Internet for command and control, the loss of that capability would paralyze, or nearly paralyze, each of these systems: the reservations and freight systems of airlines would become severely constrained, because the capacity no longer exists to take reservations by phone, even assuming that the airlines' internal computer systems could still communicate between locations. Trucking firms, bus lines and railways are similarly dependent on the Internet to manage their operations internally, and to communicate with their customers externally. And while passenger cars could continue to travel, the likelihood of obtaining gas or diesel fuel along the way would decline rapidly once the operation of refineries and fuel delivery companies became critically impaired.

In other words, a vital national network of redundant networks that once exhibited high reliability due to its low vulnerability and high redundancy and resilience is becoming increasingly dependent on a single network that can far more easily be attacked.

An overall collapse of the transportation system could therefore result whether the crisis begins with a stoppage of fuel deliveries to gas stations, coal to power stations, or of any of a number of other essential elements. Since all of these deliveries are already, or are now in the process of becoming totally dependent on the Internet, disabling the Internet would bring all to a halt. The network giveth, the network taketh away.

Still, as dangerous as the interruption of networks can be, the risk pales in comparison to the impact that would result from the loss of the underlying data that the Internet hosts. If the power grid were to crash tomorrow, it could be repaired. If even the communications system were to be brought down (but not destroyed), society could get back on its feet once communications were restored. But if the data that scientists, or doctors, or engineers, or silicon chip designers were to be destroyed, then the only way that the needs of society that are dependent on that data could be restored would be to recreate that data – from the ground up.

If the concept of a totally digitized world seems preposterous, think of the world of just twenty years ago, with its music on CDs, its pictures on film, and its records in endless ranks of file cabinets. Now consider the fact that the first Amazon Kindle was released on November 19, 2007, and that the Gartner Group now projects that

11 million eReaders – excluding Apple iPads – will be sold in the U.S. alone in 2011.[1] Google and Amazon are already offering cloud hosting of music, and Apple is expected to announce its own service soon. And on May 11 of this year, Google announced that major vendors will now start selling "Chromebooks" - inexpensive laptops that run Google's stripped-down Chrome operating system and browser, and nothing else. Everything you do will be in, and stay, in the cloud. What will the world be like, not next year, but twenty years hence?

Lastly, consider the fact that while patents (also now archived electronically) give some insight into innovation, in most cases they rarely provide useful blueprints for modern products. Most of what goes into complex semiconductor chip designs, computer algorithms, chemical processes, material science compositions and all the rest of the intellectual property that underlies our modern, technological world comprises closely-guarded trade secrets. If the limited number of records that include explanations and design information for these technologies were to be destroyed, how many of them could be fully restored, absent great and immediate effort by leading experts in each field?

What next? Given the potential consequences of a significant loss of data or a protracted interruption of the Internet, why are we not more concerned? The reasons are many: the scope of the risk we are assuming is only starting to become evident; following the Y2K debacle, crying "wolf!" about a potential IT catastrophe is not a career-enhancing pastime; every day the Internet offers more beguiling riches to be reaped; the benefits of being first to market may seem to outweigh those of ensuring robust security before bringing a new product to market; and the fact that even if experts were to conclude tomorrow that the Internet could never be made sufficiently durable, we would still be unlikely to turn back from our current course.

The question then becomes, what can we do to lower our risk?

Returning to the short list of desirable network characteristics developed above, it would be difficult to introduce physical redundancy into the system without enormous cost. And even if such redundancy could be introduced, the redundant system would be equally vulnerable to attack unless it was secured in a sufficiently different way. In the case of software and data, introducing redundancy is more economically feasible, although ensuring the survival of redundant elements (especially against cyber attack) would remain challenging.

Happily, there is still much that can – and must – be done to dramatically increase the degree to which the Internet, and the processes and data that it supports, is protected.

Resilience reversal. Ironically, one reason the Internet came into existence not in order to create systemic risk, but in an effort to lower it. More

[1] Van Camp, Jeffrey, E-Reader Sales to Jump 68 Percent in 2011, Says Gartner, DigitalTrends.com (December 9, 2011), at: http://www.digitaltrends.com/mobile/e-reader-sales-to-jump-68-percent-in-2011-says-gartner/ All Web pages cited in this article were last accessed on May 22, 2011.

specifically, one of the goals in expanding the precursor to the modern Internet – the Advanced Research Projects Agency Network (ARPANET) - was to devise and deploy a technology that would make communications more fault tolerant in the event of network disruptions. Rather than being at risk that communications between any two points could become impossible if a single transmission line were to be severed, the ARPANET technology enabled any message to take any of many available paths to its destination, direct or otherwise, utilizing the full scope of the telecommunications network. Through the development of protocols, even greater resilience was created, since each message would be broken up into myriad tiny "packets" of information, each of which could take a different route to the same destination, and then be reassembled and restored. As the project proceeded, the value of the new technology to sustain communications through the ultimate disruptive event – a nuclear attack – was realized.

At a modular level, the Internet is as simple in its elements as it is vast in its magnitude. Most persuasively, it is billions of computers (and now mobile devices) linked together by a preexisting telecommunications system that has been upgraded to carry the new demands placed upon it. That optimization includes the addition of millions of servers, called routers, that can interpret and act on the Internet addressing protocols that allow a message sent from point A to find its way to point B and there be reassembled and directed to the right recipient. Wireless routers and cellular antennae have expanded its reach, and traffic, further.

Because of the very distributed nature of its design, the Internet, like the road system, has historically exhibited very low vulnerability and very high resilience. However, as the demands upon that system have increased (most recently by the proliferation of streaming video), larger and larger point-to point fiber optic "pipes" have been installed to carry the exploding traffic. Major fiber optic lines also exist between regions and continents. Just as the creation of the highway system concentrated more traffic on fewer roadways, these major carriage lines funnel more Internet traffic on fewer physical cables.

The places where these major traffic lines begin and end, and the equipment that operates at those distributory points of retransmission, therefore represent points of vulnerability in the same way that bridges, intersections and switching yards do on highways and railway systems. To the extent that the servers that direct traffic become highly concentrated, the vulnerability to physical attack also increases. Finally, there is something called the "DNS Root Zone" (currently maintained by ICANN, an NGO that operates under authority of a Memorandum of Understanding with the U.S. Department of Commerce) that controls addressing at the national level. A cyber attack against the DNS Root Zone, or against the root domain of an individual country, could disrupt or interrupt the Internet traffic for that nation.

Vulnerability has increased at the node level of the Internet as well. While the number of connections to the Internet continues to expand, the number of significant data storage sites linked to the Internet in some sectors is beginning to decline, thereby increasing vulnerability and decreasing resilience. Most notably, this is because of the emergence of the "cloud computing" business model, made possible by the fact that massive amounts of data can now be moved back and forth between cloud service providers and enterprise users on a near real time basis.

In the cloud computing service model, a customer no longer hosts application software or its own data.[2] Instead, those resources are moved "into the cloud." That is to say, they are relocated to the remote servers provided by the cloud services provider that are in turn connected back to the customer via the Internet. For a fee, the services provider maintains the servers and software, and also becomes responsible for the physical and cyber security of its customers' data and software. When all works well, the customer retains all of the same IT capabilities it had before, at a lower cost than if it had been purchasing, maintaining, and upgrading its own on-site hardware and software. Many analysts, as well as enterprise and government CIOs, are now convinced that converting to cloud services can dramatically lower IT costs: a recent report projects that spending on cloud computing will total $222.5 billion by 2015.[3]

The U.S. Federal Government has also become a true believer in cloud computing. On February 8, 2011, U.S. CIO Vivek Kundra announced that a new Federal Cloud Computing Strategy would bind all agencies to a "Cloud First" policy that would obligate them to consider cloud solutions alternatives in all new procurement activities:

> Following the publication of this strategy, each agency will re-evaluate its technology sourcing strategy to include consideration and application of cloud computing solutions as part of the budget process. Consistent with the Cloud First policy, agencies will modify their IT portfolios to fully take advantage of the benefits of cloud computing in order to maximize capacity utilization, improve IT flexibility and responsiveness, and minimize cost.

The Strategy document states that the Office of Management and Budget has determined that up to $20 billion, or 25%, of the Federal government's annual IT expenditures, could be "migrated to cloud computing solutions."[4]

While economically attractive, the rise of cloud computing offers the potential for massive concentrations of data and software to be created. At the individual customer level, systems and data that is now located at many locations may now be hosted at fewer, or even a single location. And at the marketplace level, the software and data of hundreds, or potentially thousands, of large enterprise as well

---

[2] A wide variety of business models, not all of which involve remote hosting, have collectively been referred to as "cloud services," or "cloud computing." As used in this article, cloud services applies only to those services that involve the remote hosting of a customer's data and/or software. For an overview of the range of services that the broad definition has been applied to, see Badger, Lee et al., Cloud Computing Synopsis and Recommendations: Recommendations of the National Institute of Technology [draft], NIST Special Publication 800-146 (May 2011), *at:* http://csrc.nist.gov/publications/drafts/800-146/Draft-NIST-SP800-146.pdf The NIST draft offers the following definition:

> *Cloud computing* allows computer users to conveniently rent access to fully featured applications, to software development and deployment environments, and to computing infrastructure assets such as network-accessible data storage and processing.

[3] Cloud Computing: a Global Market Report, Global Industry Analysts, Inc. (April 2010), summarized *at:* http://cloudcomputing-vision.com/1000/report-suggests-cloud-computing-services-market-reach-us2225-billion-2015/

[4] Federal Cloud Computing Strategy, pp. 1 - 2, at: http://www.cio.gov/documents/Federal-Cloud-Computing-Strategy.pdf Security issues are addressed at pp. 26 – 28, and relevant government security resource documents are listed on pp. 37 – 38.

as government agencies and small and medium size enterprise (SME) customers may now be located at the same place.

From a cyber attack perspective, this transition is equivalent to the urbanization of a population. To use an epidemiological comparison, while a highly rural population might largely escape the effects of a plague, the same disease might ravage the inhabitants of a city. So also as regards a population's vulnerability to air attack. When it, too, becomes more highly concentrated, data and software that might once have been at low risk of a physical or cyber attack may now become an obvious and vulnerable target.

And yet the concentration of resources is considered to be a benefit of a move to the cloud services model. For example, since 1997 IBM has reduced its data centers from 432 to only 12, by moving the data into the cloud. And U.S. CIO Kundra hopes to close 800 Federal data centers by 2015.[5]

Because of these dual forces – concentration of data flows and concentrations of servers, routers, software and data - the architectural risk abatement achieved by the creators of the ARPANET has been reversed. And because the Internet and IT systems have become so crucial to the operations of all aspects of the modern world, the systemic risk has increased by multiple orders of magnitude. The need to restore the original low vulnerability and high resilience of the Internet has therefore become critical. That need will increase in direct proportion to the continuing achievement of concentration.

For this reason, the existing Internet infrastructure and the directions in which its evolution is headed must be reexamined, and IT sustainability "frameworks" of standards and best practices developed that can guide the hardening of that infrastructure, as well as its secure upgrading and maintenance as new technological advances are developed and deployed.

An(other) inconvenient truth: As usual, the recognition that cyber threats warrant the deployment of more robust defenses arrived after the danger could no longer be easily ignored. In the private sector, cyber security only began to attract widespread attention after identity theft became more widespread and significant commercial damages began to accrue, despite the fact that hobbyist hackers have been penetrating even the most sensitive systems ever since the Internet first provided points of public access.

On April 21, 2011, one of the first examples of the possible consequences of moving to the cloud was provided when customers of Amazon's cloud service program abruptly experienced unexplained difficulties. Suddenly, the Web sites and services of a variety of businesses, from social media startups to the Web site of the venerable *New York Times*, suffered interruptions that these business owners were powerless to address. Some, such as the social media site Reddit, were off line for several days. Amazon ultimately gave ten days of free hosting to its many affected customers.

[5] US CIO Unveils Government Shift to Cloud Computing, Cybersecurity News (February 14, 2011), *at:* http://cybersecuritynews.org/2011/02/14/us-cio-unveils-government-shift-to-cloud-computing/

Interestingly enough, some Amazon customers apparently avoided disruptions because they had purchased redundancy services from Amazon to protect themselves from just such an event, providing an important insight into how businesses assess and deal with risk. Economic considerations have always led some decision makers to either underestimate risk, or economize by spending less than necessary to eliminate perceived risk. Playing to this business reality, rather than making its base service offering more reliable, Amazon offered a cut rate product to attract more customers to its new line of business, and offered more protection for an additional fee. Many companies, including those whose entire business model depended on constant Internet access, settled for the base service, apparently willing to save money by taking on risk that they could have paid to avoid.

A more troubling example of the same dynamic can be found in the construction of the Japanese reactor complex referred to in the disaster scenario presented above. While that complex had been consciously designed to withstand the effects of both earthquakes and tsunamis, the greatest magnitude of each had been seriously underestimated – despite the fact that Japan lies in an extremely active earthquake zone, that modern seismometers have only been deployed for a limited period of time (and therefore the historical record of seismic data is limited), that very little about the projected scope of tsunamis was known at all at the time the facility was built, and that a meaningful percentage of the population and economic base of Japan lies within the zone of danger should a catastrophic failure of a reactor occur.

Most recently, a global survey of IT security managers responsible for critical infrastructures such as power grids, oil, gas and water, revealed that 40 percent expected a cyber attack within the year, and that 30 percent were not prepared for one – the same percentage that believed that the danger was increasing. Worse, the rate of adoption of security measures was badly lagging the threat increase, despite the fact that 70 percent of the respondents had already found malware on their systems, 80 percent had already experienced large-scale "denial of service attacks," and 25 percent had even been victims of real or threatened cyber security extortion – an increase of 25 percent from the previous year. This is particularly troubling in the electrical industry, given its critical infrastructural role, and the fact that the move towards deployment of smart grids will increase both vulnerabilities to cyber attacks, as well as widen the potential impact of a successful attack.[6]

Although awareness of cyber security risks is rising, it is important to note that the private sector has primarily been concerned with cyber fraud and the theft of information, and not with vulnerability to non-economically motivated attacks by terrorists and national enemies. This means that private sector managers are far more concerned about threats such as the surreptitious copying of financial and personal data, rather than the corruption or destruction of information and systems. In consequence, the efforts of private sector risk mangers are targeted

[6] In The Dark: Crucial Industries Confront Cyber Attacks, McAfee and the Centre for Strategic and International Studies (2011), *at:* http://www.mcafee.com/us/resources/reports/rp-critical-infrastructure-protection.pdf

first and foremost at defending systems against penetration, and detecting intrusions as quickly as possible when they occur. A cyber attack intended to destroy data might therefore not be detected until the destruction was already in process.

As the cloud computing model becomes more pervasive, however, commercial interests should expect to become just as susceptible to non-commercial attacks, because their data and systems may come to represent more valuable and vulnerable targets to an enemy of the state than government IT resources. Until this realization hits home, private sector risk managers will in particular fail to protect their IT assets to physical, as well as cyber, attacks.

What needs to be done: Whatever the actual mix of factors, the risk of a catastrophic Internet or data disaster is not on the radar screens of most risk managers. The sooner the risk is acknowledged, however, the more economically, efficiently and easily it can be addressed. Given the current surge in interest in cloud computing, the creation of IT sustainability standards should be a matter of particular interest to the new cloud services providers. Once the very substantial investments needed to build out this new architecture have been made, resistance will certainly rise to hardening these sites against physical attack.

Unlike some other current IT-based infrastructural challenges, such as a transition to a Smart Grid or the implementation of a national system of electronic health records (EHRs), creating the standards necessary to ensure the survival of a nation's data and IT infrastructure would be less challenging, due to the fact that so many of the elements would involve physical, rather than IT architectural standards.

Indeed, while the concept of developing IT sustainability standards for the protection of society at the national level is novel, the need for such standards at smaller scale is already well recognized, and has repeatedly been recognized and implemented in a variety of settings. The following are examples:

- ***Disaster Recovery***: Over the last decade, businesses have come to realize that their operations are ultimately dependent on the survival and availability of their IT systems and data: both must continue to exist, and their managers and employers must be able to access them. The 9/11 disasters provided a vivid example of the consequences that major financial, legal, and Fortune 500 businesses might suffer if systems and data access were to be interrupted for even a few days. The result has been a move towards the creation of so-called business continuity, or disaster recovery, plans that include ensuring that all data is backed up at a different location, that systems can be accessed from an alternative (and often many) locations, that management can continue to communicate effectively with workers, and so on. The goal is that a business should be able to continue to operate regardless of whether any location, including its headquarters, is destroyed, damaged, or becomes inaccessible.

- ***Seed banks:*** Curiously, perhaps the most thorough scheme for the preservation of irreplaceable data has been conceived and executed not for electronic data, but for DNA information. And the storage medium in

question is not physical media, but seeds. The name of the endeavor is the Millennium Seed Bank Project, coordinated by the Royal Botanic Gardens at Kew, and the goal is to collect, freeze, securely store, and periodically (every ten years) test the continuing viability of the seeds of every species of plant on earth. Every seed is stored in at least two locations, each of which is intensely "hardened" against any possible eventuality. For example, the Norway-maintained Svalbard Global Seed Vault is located in an abandoned coal mine on the island of Spitsbergen, only 810 miles from the North Pole.

These two examples together include the core elements of what would be required to allow a modern society to "reboot" itself following the occurrence of the type of disaster described earlier: a means of preserving all data and application software that might otherwise be destroyed; the means to periodically ensure that the stored data is still complete and viable; and the means to "fail over" to the archival data set rapidly in order to limit the adverse consequences of a cyber attack or national disaster.

A basic sustainability standard framework. In all but a few instances, formal standards and existing practices already exist that can be used to achieve maximum security and resilience for an Internet-based IT infrastructure. For example, the Department of Defense in the U.S. already has practices and standards for building facilities with ascending levels of security and survivability against physical attack. Similarly, a host of information and communications technology (ICT) security standards has already been developed by multiple standard setting organizations (SSOs), and more are being developed all of the time. Multiple SSOs are already working on security standards for the cloud as well.[7]

At a high level, the main task at hand is therefore to identify practices that can be employed at the design stage that can limit vulnerability and maximize resilience, and then, where possible, identify and assemble existing standards and practices into frameworks that can be deployed to minimize the risk that remains. The final step in such a process is to perform a "gap analysis" to determine where existing practices and standards are inadequate to complete the defensive and regenerative plan. Any voids identified must then be filled with upgrades to existing standards and practices, or by developing new standards. Where new work is required and an existing, appropriate SSO is either not available or not interested in the work, it may be necessary to launch one or more new organizations. And indeed, this is the process that has been adopted by the public-private partnership that is creating the standards frameworks that will enable the SmartGrid in the U.S.[8]

At a next level of detail, the steps to be taken might be as follows:

---

[7] For an overview of existing cyber security standards, see: Updegrove, Andrew, Security Standards and the Internet: Keeping the Cyberbarbarians from the Gate, Standards Today, Vol. VIII, No. 4 (June – July, 2009), at: http://www.consortiuminfo.org/bulletins/jun09.php#feature For an ongoing news feed regarding cloud computing standards, bookmark: http://www.consortiuminfo.org/news/archive.php?page=1&Category=2&SubCat=Cloud%20Computing

[8] For an overview of the SmartGrid effort, see Standards and the Smart Grid: the U.S. Experience, and the other articles to be found in the April – May 2009 issue (Vol. VIII No. 3) of Standards Today, *at:* http://www.consortiuminfo.org/bulletins/apr09.php

- ***Architectural analysis***: First and foremost, a systemic review would determine how to reduce vulnerability (e.g., to make the Internet once again more like the highly secure and resilient highway system). This would be achieved in part by identifying those concentrations (at the cable, hub and cloud services levels), both existing and anticipated to arise in the future, that could lead to increases in risk. Computer modeling along the lines already utilized by those that maintain the power grid could be employed to identify choke points and determine the minimum redundancy of routings and other resources needed to maintain an adequate Internet service level in the case of hypothetical types and extents of attack. Ranking the importance of points of vulnerability would allow later processes that would identify what must be protected, the available means of doing so, and the economic and performance costs of introducing those protections.[9]

- ***Selection of methodologies:*** Standards must be deployed in the real world, and therefore they must be selected in anticipation of how they will be received by those that must implement them. While ensuring that at least one viable mechanism exists to address a given point of vulnerability would obviously be essential, as a generality it is beneficial to create standards that are not mechanism-specific, so that the market can compete and innovate in meeting those requirements in the most efficient and economical fashion. A balance would therefore be needed between choosing specific requirements where viable alternatives are not deemed to exist (e.g., below ground storage of crucial archival data) and more general requirements, such as maintaining a specific level of security (e.g., by encryption, firewalls, thorough authentication, etc.), without specifying the means of achieving it. Examples of methodologies would be hardened storage locations, wide geographic distribution of small amounts of data and software, sequestration from the Internet, encryption, incident detection monitoring, assignment of responsibility for implementation, and so on). These processes would be informed by important non-security considerations, such as costs of implementation, anticipated market resistance, the need to ensure that the mechanisms adopted could be updated as technology and business practices continue to evolve, and more.

- ***Implementation plan:*** The costs of implementing such frameworks would be very significant. To achieve meaningful risk abatement at the national level, compliance in many areas would need to be high. A combination of approaches would therefore be needed in order to achieve that goal, including inclusion of appropriate requirements in government procurement
- contracts, imposition of regulatory requirements on certain crucial market participants, incentive programs to spur implementation in non-regulated industries, and more. A phase in plan would also be required, with highest priority given to hardening the most essential data and services.

[9] It should be noted that some believe that no degree of security effort can make the Internet as we know it sufficiently secure, due to its inherent design. If they are right, then the better course, however expensive, would be to replace it. See, for example, Markoff, John, Do We Need a New Internet? *The New York Times* (February 14, 2009), *at:*
http://www.nytimes.com/2009/02/15/weekinreview/15markoff.html

- ***Auditing:*** Given the importance of the goals and the breadth of the marketplace that would be required to comply, it may be necessary to include a compliance mechanism to ensure that those that are expected to implement the resulting Cybersecure Data and Internet Sustainability (CDIS) Framework in fact do so, do so effectively, and maintain their compliance over time. Rather than accomplish this result through a government agency, it may be preferable to follow the type of certification approach taken by the PCI Security Standards Council (a client of the author). The Council has developed systemic security standards for merchants, banks, hardware and software vendors, and others in order to protect the security of the credit and debit card payment chain. In order to facilitate compliance with these standards, the Council certifies inspectors and test labs that in turn test software, hardware, merchants, forensic examiners, and others. Products and service providers that have successfully passed required tests are eligible to be listed at a public registry maintained by the Council.[10]

Because concentrations of data, software and backbone services will increase risk, it would be highly advantageous to promptly develop and implement regulations against further concentrations that would be deemed to represent the highest risk, especially where the technical and other means to offset that risk appear limited. It would not be inappropriate to compare such regulations to those adopted after the Great Depression of the 1930s, and the more recent Great Recession, in order to bar certain practices deemed to be too hazardous to the financial system, or to create financial firms that have become "too big to fail."

What might a resulting CDIS Framework look like? The following is a high level overview of some of the major topics that would require attention, each giving rise to its own suite of standards.

Triage standards:

- ***Internet systems:*** So much of the essential infrastructure of the Internet (servers, operating systems and control software) as has been determined to be necessary to sustain the ongoing availability of the Internet (plus a significant safety margin) would require either the highest level of physical protection or a sufficient level of both redundancy and geographic distribution to lower its vulnerability to physical attack to an acceptable level. The same resources would need protection from cyber attack. Appropriate standards would define the elements requiring such protection, the level of protection required, the degree to which backup systems would be required, and (where appropriate) the methods of protection deemed to be adequate to ensure preservation.

- ***Essential services and data:*** Essential services and goods that are vulnerable to cyber or physical attack would also be identified, and the level of protection needed by each specified, with varying levels being applied within service providers depending upon function, as appropriate. Examples

[10] The Council's Web site is here: https://www.pcisecuritystandards.org/ Links to its many standards and guidance documents can be found here: https://www.pcisecuritystandards.org/security_standards/documents.php

of essential services would include financial markets, utilities and hospitals, while essential goods might include various high tech, defense and medical supplies. Protecting essential services would involve preserving not only the hardware and software resources needed to deliver them, but also the records upon which those services rely. Data deemed to be most essential would receive the highest priority of protection, while more ephemeral information could be consigned to less costly storage sites. Detailed methodologies would be developed to drive the prioritization of data and make appropriate designations that are linked to specified practices, standards and safeguards. Government regulations would specify which designations, and which related protections, should apply to private sector data that is deemed to be crucial to the maintenance of an operational society. Individual businesses would be free to adopt stricter, but not less strict, safeguards for such data.

Risk types: The nature and parameters of risks would need to be defined in order to create the standards needed to address them. These risk standards would be updated as threats in the field evolved. Types of risk would include (for example), bomb blast, dirty bomb, cyber attack, as well as natural disasters.

Method standards:

- ***Archiving:*** Because the risk of a successful cyber attack will presumably never be reduced to zero, it will be necessary to archive vast amounts of data to storage media that is then disconnected entirely from the Internet. Moreover, successive backups would need to be maintained, because any single backup might contain a worm that could work its way through the archive, or activate itself after the archive had been used to restore damaged systems. For the same reason, each archival copy should be partitioned in order to increase the chance that any infection could be contained. A high priority should be assigned to developing the technical details and designs for such archives, as they would represent the last line of defense against a catastrophic, and otherwise successful, cyber or physical attack.

- ***Distribution and Redundancy:*** The most critical vulnerability in the fictional scenario at the beginning of this article was not lack of defenses against air attack, but the concentration of data and software. In order to protect against physical attack, data should not be unduly concentrated, and should also be widely and redundantly distributed and archived. Note that the single most effective standard would be one that forbids the creation of chokepoints and large concentrations of data, servers and software.

- ***Physical protection:*** Normal economics would suggest that a server farm would sprawl across many acres of open land, would be housed in a structure intended solely to withstand the elements, and would be protected by light physical security (e.g., chain-link fences topped by concertina wire, plus video cameras and a security room with monitors). While such defenses might be adequate to guard against commercial espionage, they would be ineffective against even a lightly equipped attack on the ground, or provide any protection at all from the air.

Given that the risk of future wars must be assumed, in all likelihood the only appropriate and economically practical location for crucial concentrations of data, software and Internet servers would be underground. By relocating storage of data critical to society to a sufficiently deep location, attacks by terrorists could be made all but impossible with minimum ongoing expenditures, because physical access could be easily controlled. At the same time, the servers hosting the data could be immunized from the physical and electronic impact of a nuclear attack. Information with a lower triage score could permissibly be stored in multiple, redundant locations located above ground.

- ***Restoration:*** Restoring networks after a massive and successful physical or cyber attack would be an enormous task. Accordingly, not only would data need to be archived, but application software, system designs and restoration plans and protocols would require preservation as well. Periodic restoration testing of crucial IT resources would be a prudent requirement.

- ***Responsibility:*** Because some data and software will continue to be stored by more than one host, these hosts should be categorized, and each category should have stated responsibilities. For example, a "Principal Host" would have the responsibility for maintaining the ongoing, secure existence of particular data. "All Hosts" would have responsibility for protecting data, while in their custody, from being infected, and for not passing infected data to the Principal Host. In the case of data deemed to be essential to the public interest, Principal Hosts would have legal responsibilities to take required precautions. In the case of non-essential data, servers and software, the responsibility and liability of a Principal Host would be contractually defined between the parties.

- ***Cybersecurity:*** The most daunting task will not be to ensure physical security for servers, but virtual protection against cyber attack. To date, the private sector has been demonstrating a traditional risk management approach, which balances the costs of prevention against the expenses of repair and/or the costs of insuring against such expenses and any related contractual liabilities. Because the costs of eliminating the last increment of risk are usually far greater, on a percentage basis, then eliminating more garden varieties of risk, that increment is usually insured against and/or accepted as a cost of doing business, and the associated costs and liabilities are added back, on a pro rated basis, into the price of the goods and services in question.

  In the case of societally critical data and the Internet itself, however, no amount of risk should be tolerable, because even if uncompromised back up copies could be successfully re-deployed and the full operation of the Internet restored, enormous damage might already have occurred. For this reason, cybersecurity risk management concepts will need to be rethought, by placing greater importance on protection and preservation, and less on values such as ease of user authentication and access.

Each of the categories above would need to be supplemented by descending layers of detail. Due to the expense involved in deploying the resulting standards, great

care would be required not only to make the results effective, but also flexible and extensible so that future technical and other developments could be accommodated in the over-all framework.

Applying the triage concept, the requirements for each level might be as follows (with each level having sublevels that are not described):

Level 1: Blast proof underground storage; highest level cyber security protection; full backup system and data in remote but equivalent location; tested, fast fail-over to back up systems and databases

Level 2: High level of cybersecurity protection; redundant storage locations for backup systems and databases

Level 3: Protection as determined by the custodian or customer, but full cybersecurity and physical protections must be fully disclosed to customers, and public company owners of such data must disclose the same protections to their stockholders

Still at a simplistic level, a CDIS Framework might be as described in the table below.

Application Examples of CyberSecure Data and Internet Sustainability Framework

Category	Level 1 Examples	Level 2 Examples	Level 3 Examples
Finance	Trading platforms, Interactions with the Federal Reserve and Banks	Compliance and customer asset data	Marketing, billing, compliance
Airlines	Operating systems and applications software; maintenance records	Flight plans, customer account data	Personnel records
Government	Social Security and IRS records; Key Agency and military systems and data	Other public records, compliance data, statistical data	Personnel records
Large Businesses (> $1 Billion)	Shareholder records, Operating and applications software; certain industry-specific data	Tax records, compliance data	Contracts, service agreements
Mid-size Businesses ($100 MM - $1 Billion)	Shareholder records; certain industry-specific data	Tax records, compliance data	Contracts, Service agreements
SMEs (> $100 MM)	n/a	Tax records, compliance data	Contracts, Service agreements

Implementation: Designing and implementing a CDIS Framework on a national basis would represent an undertaking in scope similar to the current effort to develop and deploy the standards necessary to support a Smart Grid.

Like the Smart Grid effort, many types of stakeholders would be impacted by the need to develop a more secure IT infrastructure. In the case of the Smart Grid, the effort began with a CEO-level meeting, hosted by President Obama at the White House. Adopting the same approach would be advisable in this instance as well. Similarly, due to the breadth and depth of the practices and standards involved, the collaboration of many existing standards organizations would be required. For this reason, the equivalent of the Smart Grid Interoperability Panel would be needed to coordinate the development of the overall conceptual framework, and then the population of that framework with actual standards, both already existing and yet to be developed.

Like both the Smart Grid and the current EHR efforts, a combination of both carrots and sticks would likely be required in order to persuade the private sector to implement the resulting suite of CDIS Framework standards and practices. In the case of EHRs, the carrot is the availability (for a time) of subsidies to health industry practitioners to purchase and install EHR-compliant software, while the stick is the prospect of the penalties that will apply to non-compliant practitioners once the subsidy period has expired.

In the case of CDIS Framework standards, the corollary might be (for example) the availability of investment tax credits for cloud service providers that build server farms to CDIS Framework standards through a set date, and tax surcharges on those that fail to do so by the same date. Noncompliant vendors and service providers could also be defined as ineligible to provide IT services to the Federal Government.

The Obama Proposal: During the last session of Congress, over 50 pieces of cyber security legislation were introduced in Washington. None of these bills has achieved passage, in part because Congress has been waiting to see what the Obama administration would propose. On May 12, 2011, the President responded to an invitation from Senate Majority Leader Harry Reid and six Senate Committee Chairs by sending a broad plan to Congress intended to address cyber security in a holistic fashion by amending existing laws in some areas and proposing new laws in others. Together, the various elements of the initiative add up to an appropriately bold and comprehensive plan for securing federal networks and consumers against cyber attack.

Among other actions, the proposal would rationalize the sentencing of cyber criminals, and bring cybercrime within the coverage of the Racketeering Influenced and Corrupt Organizations Act (RICO). It would also assist businesses struggling to comply with separate security breach reporting laws in 47 states by replacing them with national compliance requirements.

Most significantly from the perspective of this article, the proposed legislation would also define and protect "critical infrastructure," and recognize that data as well as systems should fall within that definition. In a fact sheet released at the time of the announcement of the proposed legislation, the need for "Critical Infrastructure Cybersecurity Plans" is introduced as follows:

> The Nation's critical infrastructure, such as the electricity grid and financial sector, is vital to supporting the basics of life in America. Market forces are pushing infrastructure operators to put their infrastructure online, which enables them to remotely manage the infrastructure and increases their efficiency. However, when our infrastructure is online, it is also vulnerable to cyber attacks that could cripple essential services. Our proposal emphasizes transparency to help market forces ensure that critical-infrastructure operators are accountable for their cybersecurity.[11]

There are two parts of the plan that are particularly relevant to the concerns expressed in this article. The first is titled the Cybersecurity Regulatory Framework for Covered Critical Infrastructure Act ("Framework Act"),[12] which would add a new section to the Homeland Security Act of 2002.[13] The second, titled the "Department of Homeland Security Cybersecurity Authority and Information Sharing Act of 2011" ("Authority Act"), would amend Title II of the Homeland Security Act. To a considerable extent, the actions called for in these portions of the Obama plan closely mirror the recommendations made in this article. The significant exception is that the Obama plan does not recognize, or address, the risks inherent in cloud computing. In fact, the only aspect of the proposed legislation that does mention data centers is aimed at prohibiting states from adopting licensing laws that would favor in-state cloud service providers over inter-state providers.[14] While commendable from a free-market perspective, this prohibition will make it easier for dominant providers to concentrate data centers in a limited number of geographies, rather than distribute them widely, thereby decreasing vulnerability to attack.

By designating IT systems and data as "critical infrastructure," the Obama plan piggybacks conceptually on existing laws that identify and protect essential physical and other infrastructure against attacks and natural disasters. The plan does so by introducing a new definition to Title II of the Authority Act called "Critical Information Infrastructure," which it defines as:

> any physical or virtual information system that controls, processes, transmits, receives or stores electronic information in any form including data, voice or video that is—
>
> (A) vital to the functioning of critical infrastructure;

---

[11] Fact Sheet: Cybersecurity Legislative Proposal, The White House, Office of the Press Secretary (May 12, 2011), at: http://www.whitehouse.gov/the-press-office/2011/05/12/fact-sheet-cybersecurity-legislative-proposal

[12] Available *at:* http://www.whitehouse.gov/sites/default/files/omb/legislative/letters/Cybersecurity-Regulatory-Framework-for-Covered-Critical-Infrastructure-Act.pdf

[13] Homeland Security Act of 202 (6 U.S.C,. 121 35 seq). The proposed amendments, which would add new Sections 241 - 249 to Title II of the Act, may be accessed *at:* ttp://www.whitehouse.gov/sites/default/files/omb/legislative/letters/DHS-Cybersecurity-Authority.pdf

[14] The Fact Sheet referred to above includes this explanatory note:

> Data Centers. The Federal Government has embraced cloud computing, where computer services and applications are run remotely over the Internet. Cloud computing can reduce costs, increase security, and help the government take advantage of the latest private-sector innovations. This new industry should not be crippled by protectionist measures, so the proposal prevents states from requiring companies to build their data centers in that state, except where expressly authorized by federal law.

(B) so vital to the United States that the incapacity or destruction of such systems would have a debilitating impact on national security, national economic security, or national public health or safety; or

(C) owned or operated by or on behalf of a state, local, tribal, or territorial government entity.

The definition then excludes federal agency information systems.

The plan defines "Cybersecurity Threat" as:

> any action that may result in unauthorized access to, exfiltration [removal] of, *manipulation of, or impairment to the integrity,* confidentiality, or *availability of an information system or information stored on or transiting an information system*. [emphasis added]

A third definition of interest defines "protect" as:

> …those actions undertaken to secure, defend, or reduce the vulnerabilities of an information system, mitigate cybersecurity threats, or otherwise enhance information security *or the resiliency of information systems or assets*.

Together, these three definitions comprehend much of what must comprise the foundation for a truly effective cyber security regime: the recognition that it is vital to the national interest that certain IT infrastructure remain operational; that data must be protected not only against theft, but also against corruption, destruction, and unavailability; and that critical information systems must not only be made more secure, but also more resilient against attack.[15]

This goal is captured in the introductory charge (Section 243 of the Authority Act), to the Secretary of the Department of Homeland Security (DHS) to:

> … engage in cybersecurity, and other infrastructure protection activities under this title, to support the functioning of federal systems and critical information infrastructure in the interests of national security, national economic security, and national public health and safety.

The litany at the end of this charge – to support the continued functioning of critical public and private IT infrastructure "in the interests of national security, national economic security, and national public health and safety," is repeated often throughout the Authority Act, and captures the central concepts that the safety and functioning of a modern society is dependent upon the continuing viability of its IT infrastructure; that this infrastructure is susceptible to damage to cyber attack; and that this infrastructure must therefore be preemptively protected.

[15] If enacted, the definitions would appear as subsections (5), (7) and (17), respectively, of Subtitle E, Section 242, Title II.

Section 243(c) of the Authority Act mandates a number of significant activities, including developing and conducting risk assessments in concert with the private sector and agency personnel; development of new technologies with the same partners; the acquisition and deployment of these technologies and their lease and sale to the private sector – "with or without reimbursement;" and perhaps most intriguingly, the establishment of a new Federal "center" that would lead an ongoing effort to detect and defend against new cyberthreats as they emerge (one unavoidably thinks of a Center for Disease Control to protect us from the depredations of virtual viruses and other cyber diseases). This center would serve as the hub of a national effort involving the federal agencies as well as state, local, tribal and territorial authorities.

These efforts would include information sharing among all public and private parties; centralized incident reporting and the implementation of a "national cyber incident response plan;" gathering, analysis and dissemination of "timely and actionable cybersecurity threat, vulnerability, mitigation and warning information;" education and awareness building; and conducting "exercises, simulations, and other activities designed to support the national response to cybersecurity threats and incidents." The Secretary is also directed to:

> establish in cooperation with the Director of the National Institute of Standards and Technology benchmarks and guidelines for making the critical information infrastructure more secure at a fundamental level, including through automation, interoperability, and privacy-enhancing authentication;

These efforts are to be taken in cooperation with academic, foreign and international partners as well as government and private sector personnel.

In addition to its new public/private activities, DHS would now be appointed, under Section 244 of the Authority Act, as the cyber protector and supervisor of all of the federal agencies. In a provision that will be sure to attract much attention and debate, in pursuit of its duties under this directive, DHS is instructed to "acquire, intercept, retain, use, and disclose communications and other system traffic…notwithstanding any other provision of law," but subject to specific protections and constraints that follow, and others that are set forth later in Section 248(a) of the Authority Act. Sections 245 and 246 would provide immunity to non-governmental employees and private sector individuals that provide information to or otherwise cooperate with DHS in its exercise of this function. Under Section 247, these provisions would also preempt any state or local law relating to information acquisition.

Section 248, entitled "Privacy and Civil Liberties; Oversight; Penalties for Misuse," lays out extensive oversight and protections, including a requirement that appropriate policies and procedures be periodically developed in consultation with "privacy and civil liberties experts."

The Framework Act focuses more directly on critical information infrastructure, and also relies upon DHS for its implementation. The Purposes provisions set forth in Section 2 include:

> (1) enhance the cybersecurity of infrastructures determined by the Secretary to be critical to national security, national economic security, and national public health and safety;…
>
> (3) facilitate public sector and private industry consultation and development of best cybersecurity practices by encouraging a national dialogue on cybersecurity vulnerabilities affecting critical infrastructure;
>
> (4) establish workable frameworks for implementing cybersecurity minimum standards and practices designed to complement, not supplant, the scope or operation of currently available security measures;
>
> (5) to the maximum extent feasible and practicable, harmonize the designation of entities as covered critical infrastructure with existing infrastructure protection activities authorized pursuant to title II of the Homeland Security Act of 2002 (6 U.S.C. 121 et seq.);…

Other purpose clauses address inter-agency cooperation, protection of civil liberties, and promotion of efficiency and cost-effectiveness.

Sections 3 and 9 of the Framework Act would charge the Secretary of DHS with designating appropriately essential entities as "covered critical infrastructure," and therefore subject to the various compliance requirements of the Framework Act. No entity could be so designated unless its incapacity or disruption, "would have a debilitating impact on national security, national economic security, national public health or safety," or if it is "dependent upon information infrastructure to operate, or is a part of information infrastructure and critical to its operation."

Section 3 of the Framework Act provides a non-exclusive list of criteria to be used by the Secretary in making critical infrastructure designations, including interdependency with other critical infrastructure, size, and the potential impact of a failure of that entity. As recommended earlier in this article, the Secretary is also directed to establish "risk-based tiers" into which individual entities are to be placed:

> …based on the severity of, with regard to the entity, a system or asset it operates, or a service it provides:
>
> (1) the threat of a cyber attack;
>
> (2) its vulnerability to a cyber attack;
>
> (3) the extent of consequences as a result of a cyber attack; and
>
> (4) such other factors as the Secretary determines to be appropriate.

After making its determinations, the Secretary is directed to create a public list of the entities it has designated as critical infrastructure. Once the Secretary has included a given entity on that list, the designation "shall be considered a final action for purposes of judicial review in accordance with 5 U.S.C. 702," at which

point the entity may follow administrative procedures to contest that designation.

Under Section 4, DHS is directed to undertake a variety of risk mitigation actions, including identifying and prioritizing cyber risks, and reviewing and designating frameworks developed to address those risks. As with the current electronic health records (EHRs) and SmartGrid initiatives, the development of what I have referred to above as CDIS Frameworks is to occur via a public-private partnership. The effort is to commence with the Secretary requesting:

> …that representatives of organizations that coordinate or facilitate the development and use of voluntary consensus standards, representatives of appropriate voluntary consensus standards development organizations, appropriate representatives of State and local governments, agencies, and the private sector, including sector coordinating councils and information sharing and analysis centers, propose standardized frameworks for addressing cybersecurity risks.[16]

The Secretary is then directed to evaluate and designate appropriate private sector developed CDIS Frameworks for DHS use. If in the judgment of the Secretary suitable frameworks are not offered up, the Secretary is directed to invite the Director of the National Institute of Technology "to provide advice and guidance on any possible alternative framework or frameworks in consultation with appropriate public and private stakeholders." The Framework Act also specifies that "Frameworks shall not require the use of a particular measure, but shall leave the choice of particular measures to an entity to which the framework applies."

The Framework Act therefore incorporates a number of the recommendations offered above: the creation of CDIS Frameworks through public-private collaboration; the involvement of NIST; and the setting of performance standards rather than mandating the specific technical means to achieve those goals. What it fails to do is direct NIST to immediately facilitate the creation of an aggressive CDIS Framework development program based on experiences already learned from the SmartGrid Interoperability Panel.

The Framework Act also includes several other recommendations made earlier in this article, including requiring entities to develop appropriate cybersecurity plans and subjecting covered entities to inspection by appropriate, accredited evaluators (the same methodology followed by the PCI Security Standards Council, as briefly described above). Section 5 describes this process, including the requirement for annual compliance inspections of covered entities.

The Framework Act takes a soft and assistive, rather than a harsh, regulatory approach. If DHS concludes after reviewing compliance certifications and evaluator reports that a covered entity's compliance is inadequate, the Secretary is

[16] That the private sector would be asked to play the lead in cyber security standards development is a matter not just of recent convention, but also of law, dictated by the passage of the National Technology Transfer and Advancement Act of 1995 (15 U.S.C Section 3701), which formalized the U.S. "bottom up" process of standards development. For an overview of the interaction of the public and private sectors in standards development, *see:* Updegrove, Andrew, A Work in Process: Government Support for Standard Setting in the United States: 1980 – 2004, Standards Today, Vol. IV No. 1 (January 2005), *at:* http://www.consortiuminfo.org/bulletins/jan05.php#feature

authorized (Section 8) to enter into discussions with management "on ways to improve the cybersecurity plan or the evaluation, which may include the provision of technical assistance." If those discussions are not productive, the Secretary can issue a public statement of insufficiency, and take:

> (1)…such other action as may be determined appropriate by the Secretary;
>
> except that the Secretary shall not, in enforcing the provisions of this Title, issue a shutdown order, require use of a particular measure, or impose fines, civil penalties, or monetary liabilities on the owner or operator of the covered critical infrastructure as a result of such review; and
>
> (2) the Secretary shall establish an administrative review process for covered critical infrastructure to appeal a finding under this subsection that the covered critical infrastructure is not sufficiently addressing the identified cybersecurity risks.

Subsection (b) of Section 8 hints at the inclusion of another practice recommended above: mandating compliance by Federal contractors with the requirements of the Framework Act. That subsection states that:

> The Secretary shall work with the Federal Acquisition Regulatory Council established under section 1302 of title 41, United States Code, to amend the Federal Acquisition Regulation, as may be necessary and appropriate, in conjunction with the implementation of provisions under this Title.

As will be obvious, the creation of so extensive a national program will require extensive rule-making, both to guide covered entities as well as to describe in adequate detail the ways in which CDIS Frameworks will be developed and issued; the form and manner in which compliance certifications must be prepared and submitted; the mechanisms by which evaluators will be accredited; and much more. Authority for that rulemaking is provided in Section 9 of the Framework Act.

Conclusions: To date, the response of both the private and public sectors to the dangers of cyber attack has significantly lagged the emergence of cyber security threats. Similarly, while great attention has been given in the wake of 9/11 to the risks of physical attack against various types of infrastructure, little attention has thus far been given to the impact that the destruction of Internet infrastructure and server farms might have on society.

Given our long and consistent history of ignoring all types of risks until a disaster has made it impossible, or impolitic, to ignore them further, until the release of the Obama proposal it seemed highly likely that the type of systemic security initiative discussed above would not occur absent a catalytic event. The question is whether such an event would represent only an incremental disaster that fails to deliver a devastating blow, or one that would truly cripple the economy of the United States or another nation.

And there is every reason to think that the risk is real. The efficacy of asymmetric warfare – in which a weak nation exploits particular vulnerabilities in order to inflict significant damage against a far stronger opponent – has been repeatedly demonstrated in recent years. The 9/11 attacks and the use of IEDs (improvised explosive devices) in Iraq and Afghanistan both provide compelling examples. And because the defense budget of the United States is so enormous relative to that of every other nation in the world, the appeal of asymmetric warfare can only increase. Whether viewed from a cost, risk, or feasibility standpoint, an attack against the IT infrastructure of the United States will become increasingly attractive, offering the greatest return on investment.

If we allow our vulnerability to an asymmetric attack against our IT infrastructure, whether physical or cyber, to grow, we should assume that the likelihood of such an attack will grow as well. Whether that vulnerability increases dramatically or is constrained significantly will be determined by the infrastructural decisions we make today. Given that we will never be able to totally protect ourselves against cyber risks, we must not only increase our defenses, but also diminish the potential damage that a successful attack could cause. This can best be achieved by designing increased resilience into the Internet itself, and by decreasing the concentration of the IT resources that are exposed to it.

The Obama administration should therefore be highly commended for the vision and determination evident in the legislation that it has just proposed. The plans offered recognize both the gravity of the situation as well as the breadth of the infrastructure that will need to be protected in order to truly address matters of national and economic security. Unfortunately, it can be assumed that there will be stiff opposition to much of what the President has proposed: many entities will fight against passage due to the extra burdens that designation as a covered entity would imposes upon them. Proponents of less government intrusion into the private sector will object to the additional regulations. And those calling for deficit reduction will accurately note that the implementation of the Obama plan will involve significant expense and the hiring of additional DHS personnel.

Whether the Obama administration will be successful in advancing its new legislation, especially as the next presidential election approaches, remains to be seen. However, as has been demonstrated by the arguments set forth in this article, the various pieces of the Obama plan are not only appropriate but also vitally necessary. Indeed, a strong argument can be made that the plan does not go far enough. Absent a firm directive for NIST to immediately begin catalyzing the process of developing the CDIS Frameworks that will provide the essential tools for achieving true cybersecurity, many years may be wasted. Nor does the plan specifically mandate the type of systemic architectural review of our national IT infrastructure that is needed to lower systemic risk. Finally, while the plan recognizes the need to augment network resiliency, it includes no mandates to take actions that would dramatically and proactively pursue that end, such as imposing limits on the amount of data and systems that can be concentrated in one place; requiring failover backup from archived data; and mandating increased physical security.

But the Obama plan does provide a significant first step, because its scope will provoke the type of dialogue that is necessary to move the issue forward. The

public perception of cyber threats needs to move beyond the theft of cardholder data to a comprehension of the magnitude of the systemic risks generated by our increasing reliance on networks.

As with global warming, our increasing, self-inflicted vulnerability to a catastrophic physical or cyber attack represents an inconvenient truth of great concern. But unlike global warming, the means of reversing that vulnerability are affordable, and the process of addressing that risk can be completed in less than a decade. If we act now, the task will be simpler, and the costs lower. If instead we choose to increasingly concentrate our most crucial IT resources in an unprotected state, then we may well earn the unwanted distinction of becoming the first nation of the modern age to succeed in engineering its own societal collapse.