
Wayne A. Jansen, NIST

Abstract
In meteorology, the most destructive extratropical cyclones evolve with the formation of a bent-back front and cloud head separated from the main polar-front, creating a hook that completely encircles a pocket of warm air with colder air. The most damaging winds occur near the tip of the hook. The cloud hook formation provides a useful analogy for cloud computing, in which the most acute obstacles with outsourced services (i.e., the cloud hook) are security and privacy issues. This paper identifies key issues, which are believed to have long-term significance in cloud computing security and privacy, based on documented problems and exhibited weaknesses.

1. Introduction
Cloud computing has been defined as “a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction” [45]. Cloud computing can be considered a new computing paradigm with implications for greater flexibility and availability at lower cost. Because of this, cloud computing has been receiving a good deal of attention lately.

Cloud computing services benefit from economies of scale achieved through versatile use of resources, specialization, and other efficiencies. However, it is an emerging form of distributed computing still in its infancy. The term itself is often used today with a range of meanings and interpretations [16]. Three widely referenced service models have evolved [41, 62, 69]:

- Software-as-a-Service (SaaS) enables a software deployment model in which one or more applications and the computing resources to run them are provided for use on demand as a turnkey service. It can reduce the total cost of hardware and software development, maintenance, and operations.
- Platform-as-a-Service (PaaS) enables a software deployment model in which the computing platform is provided as an on-demand service that applications can be developed upon and deployed. It can reduce the cost and complexity of buying, housing, and managing hardware and software components of the platform.
- Infrastructure-as-a-Service (IaaS) enables a software deployment model in which the basic computing infrastructure of servers, software, and network equipment is provided as an on-demand service upon which a platform to develop and execute applications can be founded. It can be used to avoid buying, housing, and managing the basic hardware and software infrastructure components.

Cloud computing can be implemented entirely within an organizational computing environment as a private cloud. However, it should be clear from the service models described that a main thrust of cloud computing is to provide a means to outsource parts of that environment to an outside party. As with any outsourcing of information technology services, concerns exist about the implications for computer security and privacy, particularly with moving vital applications or data from the organization’s computing center to the computing center of another organization.

While reducing cost is a primary motivation for moving towards a cloud provider, reducing responsibility for security or privacy should not be. Ultimately, the organization is accountable for the overall state of the outsourced service. Monitoring and addressing security and privacy issues remain in the purview of the organization, just as other important issues, such as performance, availability, and recovery.

This paper looks at the main security and privacy issues pertinent to cloud computing, as they relate to outsourcing portions of the organizational computing environment. It points out areas of concern that require special attention and provides the necessary background to make informed security decisions.

2. Key Security Issues
Although the emergence of cloud computing is a recent development, insights into critical aspects of security can be gleaned from reported experiences of early adopters and also from researchers analyzing and
experimenting with available service provider platforms and associated technologies. The sections that follow highlight security-related issues that are believed to have long-term significance for cloud computing. Where possible, examples are given of problems previously exhibited to illustrate the issue.¹

The issues are organized into several general categories: trust, architecture, identity management, software isolation, data protection, and availability. Because cloud computing has grown out of an amalgamation of technologies, including service-oriented architecture, virtualization, Web 2.0, and utility computing, many of the security issues involved can be viewed as known problems cast in a new setting. Nevertheless, it represents a thought-provoking paradigm shift that goes beyond conventional norms in de-perimeterizing the organizational infrastructure—at the extreme, displacing applications from one organization’s infrastructure to the infrastructure of another organization, where the applications of potential adversaries may also operate.

3. Trust

Under the cloud computing paradigm, an organization relinquishes direct control over many aspects of security and, in doing so, confers an unprecedented level of trust on the service provider.

Insider Access. Data processed or stored outside the confines of an organization, its firewall, and other security controls bring with it an inherent level of risk. The insider security threat is a well-known issue for most organizations and, despite the name, applies as well to outsourced cloud services [4, 37]. Insider threats go beyond those posed by current or former employees to include organizational affiliates, contractors, and other parties that have received access to an organization’s networks, systems, and data to carry out or facilitate operations. Incidents may involve various types of fraud, sabotage of information resources, and theft of information. Incidents may also be caused unintentionally.

Moving data and applications to an external cloud computing environment expands the insider security risk not only to the service provider’s staff, but also potentially among other customers using the service. For example, an internal denial of service attack against the Amazon Elastic Compute Cloud (EC2) was demonstrated that involved a service user creating an initial 20 accounts and launching virtual machine instances for each, then using those accounts to create an additional 20 accounts and machine instances in an iterative fashion to grow and consume resources exponentially [60].

Composite Services. Cloud services themselves can be composed through nesting and layering with other cloud services. For instance, a SaaS provider could build its services upon those of a PaaS or IaaS cloud. Cloud service providers that subcontract some services to third-party service providers should raise concerns, including the scope of control over the third-party, the responsibilities involved, and the remedies and recourse available should problems occur. Trust is often not transitive, requiring that third-party arrangements be disclosed in advance of reaching an agreement with the service provider, and that the terms of these arrangements are maintained throughout the agreement or until sufficient notification can be given of any anticipated changes.

Liability and performance guarantees can become a serious issue with composite cloud services. The Linkup, an online storage service that closed down after losing access to a significant amount of data from its 20,000 customers, illustrates such a situation. Because another company, Nirvanix, hosted the data for The Linkup, and yet another, Savvis, hosted its application and database, direct responsibility for the cause of the failure was unclear [1].

Visibility. Migration to cloud services relinquishes control to the service provider for securing the systems on which the organization’s data and applications operate. To avoid creating gaps in security, management, procedural, and technical controls must be applied commensurately with those used for internal organizational systems. The task is formidable, since metrics for comparing the security of two computer systems are an ongoing area of research [27]. Moreover, network and system level monitoring by the user is generally outside the scope of most service arrangements, limiting visibility and the means to audit operations directly. To ensure that policy and procedures are being enforced throughout the system lifecycle, service arrangements should contain some means for gaining visibility into the security controls and processes employed the service provider, as well as their performance over time.

Risk Management. With cloud-based services, some subsystems or subsystem components are outside of the direct control of the organization that owns the information and authorizes use of system. Many people feel more comfortable with risk when they have more control over the processes and equipment involved. At a minimum, a high degree of control provides the option to weigh alternatives, set priorities, and act decisively in the best interest organization when faced with an incident. In choosing between an in-house

¹ Certain commercial products and trade names are identified in this paper to illustrate technical concepts. However, it does not imply a recommendation or an endorsement by NIST.
solution and a cloud-based implementation, the associated risks need to be assessed in detail.

Assessing and managing risk in systems that use cloud services can be a challenge. Ideally, the level of trust is based on the amount of direct control the organization is able to exert on the external service provider with regard to employment of security controls necessary for the protection of the service and the evidence brought forth as to the effectiveness of those controls [29]. However, verifying the correct functioning of a subsystem and the effectiveness of security controls as extensively as with an organizational system may not be feasible, and the level of trust must be based on other factors.

4. Architecture

The systems architecture of the software systems used to deliver cloud services comprises hardware and software residing in the cloud. The physical location of the infrastructure is determined by the service provider as is the implementation of reliability and scalability logic of the underlying support framework. Virtual machines (VMs) typically serve as the abstract unit of deployment and are loosely coupled with the cloud storage architecture. Applications are built on the programming interfaces of Internet-accessible services and typically involve multiple intercommunicating cloud components.

Attack Surface. A hypervisor or virtual machine monitor is an additional layer of software between an operating system and hardware platform, needed to operate multi-tenant VMs and applications hosted thereupon. Besides virtualized resources, the hypervisor normally supports other programming interfaces to conduct administrative operations, such as launching, migrating, and terminating VM instances. Compared with a non-virtualized implementation, the addition of a hypervisor causes an increase in the attack surface.

The complexity in VM environments can also be more challenging than their traditional counterpart, giving rise to conditions that undermine security [18]. For example, paging, checkpointing, and migration of VMs can leak sensitive data to persistent storage, subverting protection mechanisms in the hosted operating system. The hypervisor itself can also be compromised. A zero-day exploit in the HyperVM virtualization application purportedly led to the destruction of approximately 100,000 virtual server-based Websites hosted at Vaser.com [21].

Virtual Network Protection. Most virtualization platforms have the ability to create software-based switches and network configurations as part of the virtual environment to allow VMs on the same host to communicate more directly and efficiently. For example, the VMware virtual networking architecture supports same-host networking in which a private subnet is created for VMs requiring no external network access. Traffic over such networks is not visible to the security protection devices on the physical network, such as network-based intrusion detection and prevention systems [63]. To avoid a loss of visibility and protection against intra-host attacks, duplication of the physical network protections may be required on the virtual network [55].

Ancillary Data. While the focus of protection is placed mainly on application data, service providers also hold significant details about the service users’ accounts that could be compromised and used in subsequent attacks. While payment information is one example, other, more subtle information sources can be involved. For example, a database of contact information stolen from Salesforce.com, via a targeted phishing attack against an employee, was used to launch successful targeted email attacks against users of the service [36, 42]. The incident illustrates the need for service providers to promptly report security breaches occurring not only in the data it holds for its service users, but also the data it holds about them.

Another type of ancillary data is VM images. A VM image entails the software stack, including installed and configured applications, used to boot the VM into an initial state or the state of some previous checkpoint. Sharing VM images is a common practice in some cloud computing environments. Image repositories must be carefully managed and controlled to avoid problems. The provider of an image faces risks, since an image can contain proprietary code and data. An attacker may attempt to examine images to determine whether they leak information or provide an avenue for attack [66]. This is especially true of development images that are incidentally released. The reverse may also occur—an attacker may attempt to supply a VM image containing malware to users of a cloud computing system [28, 66]. For example, researchers demonstrated that by manipulating the registration process to gain a first-page listing, they could readily entice cloud users to run VM images contributed to Amazon EC2 [60].

Client-Side Protection. A successful defense against attacks requires both a secure client and a secure Website infrastructure. With emphasis typically placed on the latter, the former can be easily overlooked. Web browsers, a key element for many cloud computing services, and the various available plug-ins and extensions for them are notorious for their security problems [34, 53, 54]. Moreover, many browser add-ons do not provide automatic updates, increasing the persistence of existing vulnerabilities.
The increased availability and use of social media, personal Webmail, and other publicly available sites also has associated risks that can impact the security of the browser, its underlying platform, and cloud service accounts negatively through social engineering attacks. For example, spyware reportedly installed in a hospital via an employee’s Yahoo Webmail account sent out more than 1,000 screen captures containing financial and other confidential information to the originator before it was discovered [44]. Having a backdoor Trojan, keystroke logger, or other type of malware running on a client does not bode well for the security of the cloud or other Web-based services [15]. Organizations need to employ measures to secure the client side as part of the overall architecture. Banks are beginning to take the lead in deploying hardened browser environments that encrypt network exchanges and protect against keystroke logging [10, 11].

Server-Side Protection. Virtual servers and applications, much like their non-virtualized counterparts, need to be secured in IaaS clouds. Following organizational policies and procedures, hardening of the operating system and applications should occur to produce VM images for deployment. Care must also be taken to make adjustments for the virtualized environments in which the images run. For example, virtual firewalls can be used to isolate groups of VMs from other groups hosted, such as production systems from development systems or development systems from other cloud-resident systems. Carefully managing VM images is also important to avoid accidentally deploying images containing vulnerabilities.

5. Identity Management

Data sensitivity and privacy of information have increasingly become a concern for organizations, and unauthorized access to information resources in the cloud is a major issue. One reason is that an organization’s identification and authentication framework may not naturally extend into the cloud and may require effort to modify the existing framework to support cloud services [6]. The alternative of having two different systems for use authentication, one for internal organizational systems and another for external cloud-based systems is a complication that can become unworkable over time. Identity federation, popularized with the introduction of service oriented architectures, is one solution that can be accomplished in a number of ways, such as with the Security Assertion Markup Language (SAML) standard.

Authentication. A growing number of cloud service providers support the SAML standard and use it to administer users and authenticate them before providing access to applications and data. SAML provides a means to exchange information, such as assertions related to a subject or authentication information, between cooperating domains. SAML request and response messages are typically mapped over the Simple Object Access Protocol (SOAP), which relies on XML for its format. With Amazon Web Services, for example, once a user has established a public key certificate, it is used to sign SOAP requests to the EC2 to interact with it.

SOAP message security validation is complicated and must be carried out carefully to prevent attacks. XML wrapping attacks involving the manipulation of SOAP messages have been successfully demonstrated against Amazon’s EC2 services [17, 23]. A new element (i.e., the wrapper) is introduced into the SOAP Security header; the original message body is then moved under the wrapper and replaced by a bogus body containing an operation defined by the attacker. The original body can still be referenced and its signature verified, but the operation in the replacement body is executed instead.

Access Control. Besides authentication, the capability to adapt user privileges and maintain control over access to resources is also required, as part of identity management. Standards like the eXtensible Access Control Markup Language (XACML) can be employed to control access to cloud resources, instead of using a service provider’s proprietary interface. XACML focuses on the mechanism for arriving at authorization decisions, which complements SAML’s focus on the means for transferring authentication and authorization decisions between cooperating entities.

XACML is capable of controlling the proprietary service interfaces of most providers, and some service providers, such as salesforce.com and Google Apps, already have it in place. Messages transmitted between XACML entities are susceptible to attack by malicious third parties, making it important to have safeguards in place to protect decision requests and authorization decisions from possible attacks, including unauthorized disclosure, replay, deletion and modification [33].

6. Software Isolation

High degrees of multi-tenancy over large numbers of platforms are needed for cloud computing to achieve the envisioned flexibility of on-demand provisioning of reliable services and the cost benefits and efficiencies due to economies of scale. To reach the high scales of consumption desired, service providers have to ensure dynamic flexible delivery of service and isolation of user resources. Multi-tenancy in cloud computing is typically done by multiplexing the execution of VMs from potentially different users on the same physical server [56]. It is important to note that applications
deployed on guest VMs remain susceptible to attack and compromise, much the same as their non-virtualized counterparts. This was dramatically exemplified recently by a botnet found operating out of Amazon’s EC2 cloud computing environment [43, 67].

**Hypervisor Complexity.** The security of a computer system depends on the quality of the underlying software kernel that controls the confinement and execution of processes. A hypervisor or virtual machine monitor (VMM) is designed to run multiple guest VMs, hosting operating systems and applications, concurrently on a single host computer and to provide isolation between the guest VMs.

A VMM can, in theory, be smaller and less complex than an operating system. Small size and simplicity make it easier to analyze and improve the quality of security, giving a VMM the potential to be better suited for maintaining strong isolation between guest VMs than an operating system is for isolating processes [31]. In practice, however, modern hypervisors can be large and complex, comparable to an operating system, which negates this advantage. For example, Xen, an open source x86 VMM, incorporates a modified Linux kernel to implement a privileged partition for input/output operations, and KVM, another open source effort, transforms a Linux kernel into a VMM [31, 58, 68]. Understanding the use of virtualization by a service provider is a prerequisite to understanding the risks involved.

**Attack Vectors.** Multi-tenancy in VM-based cloud infrastructures, together with the subtleties in the way physical resources are shared between guest VMs, can give rise to new sources of threats. The most serious threat is that malicious code can escape the confines of its VMM and interfere with the hypervisor or other guest VMs. Live migration, the ability to transition a VM between hypervisors on different host computers without halting the guest operating system, and other features provided by VMM environments to facilitate systems management, also increase software size and complexity and potentially add other areas to target in an attack.

Several examples illustrate the types of attack vectors possible. The first is mapping the cloud infrastructure. While seemingly a daunting task to perform, researchers have demonstrated an approach with Amazon’s EC2 network [56]. By launching multiple VM instances from multiple user accounts and using network probes, assigned IP addresses and domain names were used to identify service location patterns. Building on that information and general technique, the plausible location of a specific target VM could be identified and new VMs instantiated to be eventually co-resident with the target.

Once a suitable target location is found, the next step for the guest VM is to bypass or overcome containment by the hypervisor or to takedown the hypervisor and system entirely. Weaknesses in the available programming interfaces and the processing of instructions are common targets for uncovering vulnerabilities to exploit [13]. For example, a vulnerability was discovered in a VMware routine handling FTP requests, allowing specially crafted requests to corrupt a heap buffer in the hypervisor, which could allow the execution of arbitrary code [57, 59]. Similarly, a serious flaw that allows an attacker to write to an arbitrary out-of-bounds memory location was discovered in the PIIX4 power management code of VMware by fuzzing emulated I/O ports [50]. A denial of service vulnerability was also uncovered in a virtual device driver, which could allow a guest VM to crash the VMware host along with other VMs active there [64].

More indirect attack avenues may also be possible. For example, researchers developed a way for an attacker to gain administrative control of VMware guest VMs during a live migration, employing a man-in-the-middle attack to modify the code used for authentication [49]. Memory modification during migration presents other possibilities such as the potential to insert a VM-base rootkit layer below the operating system [35]. Another example of an indirect attack is monitoring resource utilization on a shared server to gain information and perhaps perform a side-channel attack, similar to attacks used in other computing environments [56]. For example, an attacker could determine periods of high activity, estimate high-traffic rates, and possibly launch keystroke timing attacks to gather passwords and other data from a target server.

**7. Data Protection**

Data stored in the cloud typically resides in a shared environment collocated with data from other customers. Organizations moving sensitive and regulated data into the cloud, therefore, must account for the means by which access to the data is controlled and the data is kept secure.

**Data Isolation.** Data can take many forms. For example, for cloud-based application development, it includes the application programs, scripts, and configuration settings, along with the development tools. For deployed applications, it includes records and other content created or used by the applications, as well as account information about the users of the applications. Access controls are one means to keep data away from unauthorized users; encryption is another. Access controls are typically identity-based,
which makes authentication of the user’s identity an important issue in cloud computing.

Database environments used in cloud computing can vary significantly. For example, some environments support a multi-instance model, while others support a multi-tenant model. The former provides a unique database management system running on a VM instance for each service user, giving the user complete control over role definition, user authorization, and other administrative tasks related to security. The latter provides a predefined environment for the cloud service user that is shared with other tenants, typically through tagging data with a user identifier. Tagging gives the appearance of exclusive use of the instance, but relies on the service provider to maintain a sound secure database environment.

Various types of multi-tenant arrangements exist for databases. Each type pools resources differently, offering different degrees of isolation and resource efficiency [26, 65]. Other considerations also apply. For example, certain features like data encryption are only viable with arrangements that use separate rather than shared databases. These sorts of tradeoffs require careful evaluation of the suitability of the data management solution for the data involved. Requirements in certain fields, such as healthcare, would likely influence the choice of database and data organization used in an application. Privacy sensitive information, in general, is a serious concern [52].

Data must be secured while at rest, in transit, and in use, and access to the data controlled. Standards for communications protocols and public key certificates allow data transfers to be protected using cryptography. Procedures for protecting data at rest, however, are not as well standardized, making interoperability an issue due to the predominance of proprietary systems. The lack of interoperability affects the availability of data and complicates the portability of applications and data between cloud computing service providers.

Currently, the responsibility for cryptographic key management falls mainly on the cloud service subscriber. Key generation and storage is usually performed outside the cloud using hardware security modules, which do not scale well to the cloud paradigm. Work is ongoing to identify scalable and usable cryptographic key management and exchange strategies for use by government, which could help to alleviate the problem eventually. Protecting data in use is an emerging area of cryptography with few practical results to offer, leaving trust mechanisms as the main safeguard [22].

Data Sanitization. The data sanitization practices that a service provider implements have obvious implications for security. Sanitization is the removal of sensitive data from a storage device in various situations, such as when a storage device is removed from service or moved elsewhere to be stored. It also applies to backup copies made for recovery and restoration of service, and residual data remaining upon termination of service. In a cloud computing environment, data from one subscriber is physically commingled with the data of other subscribers, which can complicate matters. For example, with the proper skills and equipment, it is possible to recover data from failed drives that are not disposed of properly by service providers.

Data Location. One of the most common compliance issues facing an organization is data location [30, 51]. Use of an in-house computing center allows an organization to structure its computing environment and know in detail where data is stored and the safeguards used to protect the data. In contrast, a characteristic of many cloud computing services is that the detailed information of the location of an organization’s data is unavailable or not disclosed to the service subscriber. This situation makes it difficult to ascertain whether sufficient safeguards are in place and whether legal and regulatory compliance requirements are being met. External audits and security certifications can, to some extent, alleviate this issue, but they are not a panacea.

Once information crosses a national border, it is extremely difficult to guarantee protection under foreign laws and regulations. For example, the broad powers of USA Patriot Act have raised concern with some foreign governments that the provisions would allow the U.S. government to access private information, such as medical records, outsourced to American companies [5]. Constraints on the transborder flow of unclassified sensitive data, as well as the requirements on the protection afforded the data, have become the subject of national and regional privacy and security laws and regulations [12].

The main compliance concerns with transborder data flows include whether the laws in the jurisdiction where the data was collected permit the flow, whether those laws continue to apply to the data post-transfer, and whether the laws at the destination present additional risks or benefits [12]. Technical, physical and administrative safeguards, such as access controls, often apply. For example, European data protection laws may impose additional obligations on the handling and processing of European data transferred to the U.S. [9].

8. Availability

In simple terms, availability means that an organization has its full set of computing resources
accessible and usable at all times. Availability can be affected temporarily or permanently, and a loss can be partial or complete. Denial of service attacks, equipment outages, and natural disasters are all threats to availability.

**Temporary Outages.** Despite employing architectures designed for high service reliability and availability, cloud computing services can and do experience outages and performance slowdowns [41]. Several examples illustrate this point. In February 2008, Amazon’s Simple Storage Service (S3) and EC2 services suffered a three-hour outage that, in turn, affected Twitter and other startup companies using the services [38, 47]. In June 2009, a lightning storm caused a partial EC2 outage that affected some users for 4 hours [48]. Similarly, a database cluster failure at Salesforce.com caused an outage for several hours in February 2008, and in January 2009, another brief outage occurred due to a network device failure [14, 20]. In March 2009, Microsoft’s Azure cloud service experienced severe degradation for about 22 hours due to networking issues related to an upgrade [7].

At a level of 99.999% reliability, 8.76 hours of downtime is to be expected in a year. The level of reliability of a cloud service and also its capabilities for backup and recovery should be taken into account in the organization’s contingency planning to address the restoration and recovery of disrupted cloud services and operations, using alternate services, equipment, and locations. Cloud storage services may represent a single point of failure for the applications hosted there. In such situations, a second cloud service provider could be used to back up data processed by the primary provider to ensure that during a prolonged disruption or serious disaster at the primary, the data remains available for immediate resumption of critical operations.

**Prolonged and Permanent Outages.** It is possible for a service provider to experience serious problems, like bankruptcy or facility loss, which affect service for extended periods or cause a complete shutdown. For example, in April 2009, the FBI raided computing centers in Texas and seized hundreds of servers, when investigating fraud allegations against a handful of companies that operated out of the centers [70]. The seizure disrupted service to hundreds of other businesses unrelated to the investigation, but who had the misfortune of having their computer operations collocated at the targeted centers. Other examples are the major data loss experienced by ma.gnolia, a bookmark repository service, and the abrupt failure of Omnidrive, an on-line storage provider, who closed without warning to its users in 2008 [3, 24].

**Denial of Service.** A denial of service attack involves saturating the target with bogus requests to prevent it from responding to legitimate requests in a timely manner. An attacker typically uses multiple computers or a botnet to launch an assault. Even an unsuccessful distributed denial of service attack can quickly consume a large amount of resources to defend against and cause charges to soar. The dynamic provisioning of a cloud in some ways simplifies the work of an attacker to cause harm. While the resources of a cloud are significant, with enough attacking computers they can become saturated [28]. For example, a denial of service attack against BitBucket, a code hosting site, caused an outage of over 19 hours of downtime during an apparent denial of service attack on the underlying Amazon cloud infrastructure it uses [2, 46].

Besides publicly available services, denial of service attacks can occur against private services, such as those used in cloud management. For example, a denial of service attack occurred against the cloud management programming interface of the Amazon Cloud Services involved machine instances replicating themselves exponentially [60]. Internally assigned non-routable addresses, used to manage resources within the service provider’s network, may also be used as an attack vector. A worst-case possibility that exists is for elements of one cloud to attack those of another or to attack some of its own elements [28].

**Value Concentration.** The bank robber Willie Hutton is often attributed with the claim that he robbed banks “because that is where the money is” [8]. In many ways, data records are the currency of the 21st century and cloud-based data stores are the bank vault, making them an increasingly preferred target. Just as an economy of scale exists in robbing banks instead of individuals, a high payoff ratio also exists for successfully compromising a cloud.

Finesse and circumvention was Willie’s trademark and that style works well in the digital world of cloud computing. For instance, a recent exploit targeted a Twitter employee’s email account by reportedly answering a set of security questions and then using that information to access company files stored on his organizational Google Apps account [25, 61]. A similar weakness was noted in Amazon Web Services (AWS) [19]. A registered email address and valid password for an account are all that is required to download authentication credentials from the AWS Web dashboard, which in turn grant access to the account’s resources. Since lost passwords can be reset by email, an attacker controlling the mail system, or passively eavesdropping on the network thru which email containing a password reset would pass, could effectively take control of the account.

Having data collocated with the data of an organization with a high threat profile could also lead
to denial of service, as an unintended casualty from an attack targeted against that organization. Similarly, indirect effects from an attack against the physical resources of a high-profile organization’s service provider are also a possibility. For example, IRS facilities are continually targeted by would-be attackers [32, 39, 40].

9. Conclusion

In emphasizing the cost and performance benefits of the cloud, some fundamental security problems have receded into the background and been left unresolved. Several critical pieces of technology, such as a solution for federated trust, are not yet fully realized, impinging on successful deployments. Determining the security of complex computer systems is also a long-standing security problem that overshadows large scale computing in general. Attaining the high assurance qualities in implementations has been an elusive goal of computer security researchers and practitioners, and is also a work in progress for cloud computing.

Security of the cloud infrastructure relies on trusted computing and cryptography. Organizational data must be protected in a manner consistent with policies, whether in the organization’s computing center or the cloud. No standard service contract exists that covers the ranges of cloud services available and the needs of different organizations. Having a list of common outsourcing provisions, such as privacy and security standards, regulatory and compliance issues, service level requirements and penalties, change management processes, continuity of service provisions, and termination rights, provides a useful starting point [51].

The migration to a cloud computing environment is in many ways an exercise in risk management. Both qualitative and quantitative factors apply in an analysis. The risks must be carefully balanced against the available safeguards and expected benefits, with the understanding that accountability for security remains with the organization. Too many controls can be inefficient and ineffective, if the benefits outweigh the costs and associated risks. An appropriate balance between the strength of controls and the relative risk associated with particular programs and operations must be ensured.

10. References


[53] N. Provos et al., The Ghost In The Browser: Analysis of Web-based Malware, Hot Topics in Understanding Botnets (HotBots), April 10, 2007, Cambridge, MA


