John Harauz, j.harauz@computer.org
Lori M. Kaufman, lori.kaufman@ieee.org
Bruce Potter, bpotter@pontetec.com

Data Security in the World of Cloud Computing

Lori M. Kaufman
BAE Systems

In the 1990s, the world was introduced to the Internet, and we began to see distributed computing's power realized on a large scale. Today, we have the ability to utilize scalable, distributed computing environments within the confines of the Internet, a practice known as *cloud computing.*

This environment strives to be dynamic, reliable, and customizable with a guaranteed quality of service.[1] Within this system, users have a myriad of virtual resources for their computing needs, and they don't need a complete understanding of the infrastructure. Cloud computing's advent has made the declaration by Scott McNealy, Sun Microsystems' founder, that "The network is the computer" a reality and given the old Sun marketing motto a new life.

In this new world of computing, users are universally required to accept the underlying premise of trust. In fact, some have conjectured that trust is the biggest concern facing cloud computing.[2] Nowhere is the element of trust more apparent than in security, and many believe trust and security to be synonymous. Here, I examine some security issues and the associated regulatory and legal concerns that have arisen as cloud computing emerges as a primary distributed computing platform.

Background of the Cloud

The concept of cloud computing has been evolving for more than 40 years. In the 1960s, J.C.R. Licklider introduced the term "intergalactic computer network" at the Advanced Research Projects Agency. This concept served to introduce the concept that the world came to know as the Internet. The underlying premise was a global interconnection of computer programs and data.

The term "cloud" originates from the telecommunications world of the 1990s, when providers began using virtual private network (VPN) services for data communication. VPNs maintained the same bandwidth as fixed networks with considerably less cost: these networks supported dynamic routing, which allowed for a balanced utilization across the network and an increase in bandwidth efficiency, and led to the coining of the term "telecom cloud." Cloud computing's premise is very similar in that it provides a virtual computing environment that's dynamically allocated to meet user needs.

From a technical perspective, cloud computing includes service-oriented architecture (SOA) and virtual applications of both hardware and software. Within this environment, it provides a scalable services delivery platform. Cloud computing shares its resources among a cloud of service consumers, partners, and vendors. By sharing resources at various levels, this platform offers various services, such as an *infrastructure cloud* (for example, hardware or IT infrastructure management), a *software cloud* (such as software, middleware, or traditional customer relationship management as a service), an *application cloud* (application, UML modeling tools, or social networks as a service), and a *business cloud* (for instance, business processes as a service) (see www.thecloudcomputing.org/2009/2/). Cloud computing itself is a field within *service computing*, a cross-discipline that bridges the gap between business and IT services. This discipline aims to enable IT services and computing technology to perform business services more efficiently and effectively (see http://tab.computer.org/tcsc/).

The UC Berkeley Space Sciences Laboratory's SETI@home (Search for Extra-Terrestrial Intelligence) project began in 1999 as an attempt to implement distributed computing through computers connected via the Internet to search for intelligent life beyond Earth. This implementation's success demonstrated the viability of using the Internet as a host for grid computing applications. Concurrent with this project, others were also developing their own variants of cloud computing.

Salesforce.com introduced one of the first practical cloud computing implementations in 1999 and established the concept of delivering enterprise services through a Web site. In 2002, Amazon Web Services launched a suite of cloud-based services, including storage, computation, and even human intelligence through the Amazon Mechanical Turk. It followed up this accomplishment in 2006 with its Elastic Compute Cloud (E2C) service, which provides a commercial service through which users can rent computers and run their own applications. AT&T also entered the cloud computing realm when it acquired USinternetworking (USi) in 2006. USi was an application service provider for more than 30 countries. In 2008, AT&T introduced Synaptic, which combined USi's five Internet data centers in the US, Europe, and Asia to serve as regional gateways within its cloud.

The US government projects that between 2010 and 2015, its spending on cloud computing will be at approximately a 40-percent compound annual growth rate and will pass $7 billion by 2015.

Today, the latest example of cloud computing is Web 2.0; Google, Yahoo, Microsoft, and other service providers now offer browser-based enterprise service applications (such as webmail and remote data backup). Now that cloud computing has emerged as a viable and readily available platform, many users from disparate backgrounds (for example, financial institutions, educators, or cybercriminals) are sharing virtual machines to perform their daily activities. This environment requires an implicit level of trust as well as an explicit level of vigilance to ensure success.

Security and Responsibility

Within the cloud computing world, the virtual environment lets users access computing power that exceeds that contained within their own physical worlds. To enter this virtual environment requires them to transfer data throughout the cloud. Consequently, several data storage concerns can arise. Typically, users will know neither the exact location of their data nor the other sources of the data collectively stored with theirs. To ensure data *confidentiality*, *integrity*, and *availability* (CIA), the storage provider must offer capabilities that, at a minimum, include

- a tested encryption schema to ensure that the shared storage environment safeguards all data;
- stringent access controls to prevent unauthorized access to the data; and
- scheduled data backup and safe storage of the backup media.

Security is implicit within these capabilities, but further fundamental concerns exist that need attention. For example, is security solely the storage provider's responsibility, or is it also incumbent on the entity that leases the storage for its applications and data? Furthermore, legal issues arise, such as e-discovery, regulatory compliance (including privacy), and auditing. The range of these legal concerns reflects the range of interests that are currently using or could use cloud computing. These issues and their yet-to-be-determined answers provide significant insight into how security plays a vital role in cloud computing's continued growth and development.

To overcome these and other concerns, we must develop a security model that promotes CIA. This model could enable each cloud to offer a measure of its to-date and projected CIA, but the obvious difficulty is that obtaining security data is difficult, if not impossible. This problem has existed since computing's advent due to financial, business, and national security concerns. It might be exacerbated in cloud computing because the need to provide data confidentiality can also impact incident reporting.

Who Will Use Clouds and Proffer Security?

Cloud computing users range from individuals and small businesses to Fortune 500 firms and governments. According to a September 2008 survey from the Pew Research Institute, nearly 69 percent of Americans use cloud computing services (such as webmail and online data backup sites).[3] In India, companies such as Ashok Leyland, Tata Elxi, Bharti, Infosys, Asian Paints, and Maruti are either piloting or using cloud computing. Additionally, nearly 1,500 companies in India already use blended (voice-chat-data) cloud-based communication services from vendors such as Cisco WebEx and Microsoft.[4] The US government projects that between 2010 and 2015, its spending on cloud computing will be at approximately a 40-percent compound annual growth rate (CAGR) and will pass $7 billion by 2015.[5]

A major selling point for cloud computing is that it offers significant computing capability that otherwise might not be affordable. For example, a startup might not have the resources to purchase in-house computers or ensure the necessary security, but the cloud offers a cost-effective alternative. Similarly, well-established entities might see the cloud as an effective way to reduce costs and improve IT capabilities. Although these two examples might be at the extremes, they describe the range of entities that will be partnering in a cloud computing world. These two cited examples imply that cloud providers will deliver the needed security. However, this paradigm might not be appropriate for all industries.

The data you can find in a cloud ranges from public source, which has minimal security concerns, to private data containing highly sensitive information (such as social security numbers, medical records, or shipping manifests for hazardous material). If a business's primary function is to provide services that don't require sensitive data, then the security that they expect or need is less than that required for a business that processes sensitive information. Consider a scenario in which banks, other financial institutions, or businesses that process highly sensitive data decide to use a cloud. Such data's nature explicitly dictates that its storage must employ a high degree of security. Additionally, the public trust that surrounds the handling of this data reflects an assumed level of high security. Does using a cloud environment alleviate the business entities of their responsibility to ensure that proper security measures are in place for both their data and applications, or do they share joint responsibility with service providers? The answer to this question will most likely involve legal interpretation. These examples demonstrate that the governance required in managing stored data and its associated applications is a function of the data itself. For example, the storage of personal information on the cloud creates its own set of regulatory concerns that indirectly impact security. Among these concerns are the following questions:

- Who has jurisdiction over data as it flows across borders?
- Can governments access that information as it changes jurisdiction?
- Is there more risk in storing personal information in data centers that belong to a single entity rather than in multiple data centers?

The answers to these and other questions lie within the realm of yet-to-be-written law. If the past decade of legal wrangling involving the Internet is any guide, then resolving these security- and regulatory-related concerns will take years, and will without question significantly influence cloud computing's evolution.

A Tempting Target for Cybercrime

Cybercrime's effects are felt throughout the Internet, and cloud computing offers a tempting target for many reasons. As previously discussed, startups typically have limited resources; to alleviate the expense of developing a secure computing environment, they might turn to cloud computing to deflect cybersecurity concerns. To support their clouds' integrity, large providers typically require that users place 100 percent of their data within the provider's cloud. Providers such as Google and Amazon have the existing infrastructure to deflect and survive a cyberattack, but not every cloud has such capability. Clouds can comprise multiple entities, and in such a configuration, no cloud can be more secure than its weakest link. If a cybercriminal can identify the provider whose vulnerabilities are the easiest to exploit, then this entity becomes a highly visibile target. The lack of security associated with this single entity threatens the entire cloud in which it resides. If not all cloud providers supply adequate security measures, then these clouds will become high-priority targets for cybercriminals. By their architecture's inherent nature, clouds offer the opportunity for simultaneous attacks to numerous sites, and without proper security, hundreds of sites could be comprised through a single malicious activity.

If a cybercriminal can identify the provider whose vulnerabilities are the easiest to exploit, then this entity becomes a highly visibile target.

Security and Technology

To advance cloud computing, the community must take proactive measures to ensure security. A movement exists to adopt universal standards (for example, open source) to ensure interoperability among service providers. Included in this effort are attempts to develop security standards to ensure data's CIA. Even though the community at large is aware of the need for security and is attempting to initiate robust measures, a realm of security concerns transcends these efforts.

As with most technological advances, regulators are typically in a "catch-up" mode to identify policy, governance, and law. Cloud computing presents an extension of problems heretofore experienced with the Internet. As mentioned, legal decisions will ultimately determine who "owns" the responsibility for securing information shared within clouds. To ensure that such decisions are informed and appropriate for the cloud computing environment, the industry itself should establish coherent and effective policy and governance to identify and implement proper security methods. To facilitate such policy and governance's emergence, the US National Institute of Standards and Technology (NIST), an agency of the Commerce Department's Technology Administration, has

Computing in SCIENCE & ENGINEERING
p. 54 Hardware Multiagents | p. 70 GPU Computing
High-Performance Computing Education
September/October
IEEE
computer society

created a cloud computing security group. This group envisions its role as promoting "the effective and secure use of the technology within government and industry by providing technical guidance and promoting standards" (see http://csrc.nist.gov/groups/SNS/cloud-computing/index.html). NIST has recently released its draft "Guide to Adopting and Using the Security Content Automation Protocol" (SCAP; see http://csrc.nist.gov/groups/SNS/cloud-computing/index.html), which identifies a "suite of specifications for organizing and expressing security-related information in standardized ways, as well as related reference data, such as identifiers for software flaws and security configuration issues."[4] Its application includes maintaining enterprise systems' security. Interestingly, a major concern included in SCAP is the lack of interoperability among system-level tools. It states that

> many tools for system security, such as patch management and vulnerability management software, use proprietary formats, nomenclatures, measurements, terminology, and content. For example, when vulnerability scanners do not use standardized names for vulnerabilities, it might not be clear to security staff whether multiple scanners are referencing the same vulnerabilities in their reports. This lack of interoperability can cause delays and inconsistencies in security assessment, decision-making, and remediation.

This concern is but one of many SCAP has noted that needs action.

In addition to NIST's efforts, the industry itself can affect an enterprise approach to cloud security. If it applies due diligence and develops a policy of self-regulation to ensure that security is effectively implemented throughout all clouds, then this policy can serve to facilitate law-making as well. By combining industry best practices with the oversight NIST and other entities are developing, we can effectively address cloud computing's future security needs. To achieve a recognized and actionable security policy, SCAP recommends that organizations demonstrate compliance with security requirements in mandates such as the US Federal Information Security Management Act (FISMA). By adhering to this approach, the policy needed to ensure cloud security can provide effective governance to both industry and lawmakers. □

References

1. L. Wang et al., "Scientific Cloud Computing: Early Definition and Experience," *Proc. 10th Int'l Conf. High-Performance Computing and Communications* (HPCC 08), IEEE CS Press, 2008, pp. 825–830.
2. J. Urquhart, "The Biggest Cloud-Computing Issue of 2009 is Trust," *C-Net News*, 7 Jan. 2009; http://news.cnet.com/8301-19413_3-10133487-240.html.
3. J.B. Horrigan, "Cloud Computing Gains in Currency," 12 Sept. 2008, http://pewresearch.org/pubs/948/cloud-computing-gains-in-currency.
4. S. Singh, "Different Cloud Computing Standards a Huge Challenge," *The Economic Times*, 4 June 2009; http://economictimes.indiatimes.com/Infotech/Different-cloud-computing-standards/articleshow/4614446.cms.
5. "US Federal Cloud Computing Market Forecast 2010–2015," tabular analysis, publication: 05/2009.

Lori M. Kaufman *is a deputy chief technology officer at BAE Systems. Her research interests include cybersecurity, software assurance, and biometrics. Kaufman has a PhD in electrical engineering from the University of Virginia. Contact her at lori.kaufman@ieee.org.*