

Security on distributed systems: Cloud security versus traditional IT

M. Azua Himmel
F. Grossman

Cloud computing is a popular subject across the IT (information technology) industry, but many risks associated with this relatively new delivery model are not yet fully understood. In this paper, we use a qualitative approach to gain insight into the vectors that contribute to cloud computing risks in the areas of security, business, and compliance. The focus is on the identification of risk vectors affecting cloud computing services and the creation of a framework that can help IT managers in their cloud adoption process and risk mitigation strategy. Economic pressures on businesses are creating a demand for an alternative delivery model that can provide flexible payments, dramatic cuts in capital investment, and reductions in operational cost. Cloud computing is positioned to take advantage of these economic pressures with low-cost IT services and a flexible payment model, but with certain security and privacy risks. The frameworks offered by this paper may assist IT professionals obtain a clearer understanding of the risk tradeoffs associated with cloud computing environments.

Introduction

Cloud computing technologies are transforming our world, and many individuals consume these services without even being aware they are doing so. For example, every time someone performs a Google** search, posts a comment on a favorite social networking site, or uses a cell phone, he or she is using a cloud. For the purpose of this paper, we define cloud computing (or simply, “cloud”) as involving an IT (information technology) delivery model that provides compute, storage, and network services as an on-demand service, usually virtualized, with shared resources offered as a utility over a network. The ubiquity of cloud computing for retail applications of IT is evident with the rapid growth of cloud-enabled smartphones, which, for example, experienced 75% growth worldwide in 2010 [1] and continuous growth acceleration overseas in 2011 (e.g., see [2]). Other factors, such as the commoditization of hypervisor technology and the rise of new operating systems (OSs) and middleware for clouds, have created new security surfaces for attacks. Massive storage clouds, built for the insatiable need to

collect and store an ever-increasing amount of data, are generating new challenges in the area of regulatory compliance. In addition, financial pressures on IT organizations to produce more services with reduced budgets are creating significant demands for the adoption of alternative models that could potentially lower the cost of IT. However, cloud standardization forces the use of a limited number of IT configurations that sometimes do not fit the needs of large corporations [3, 4].

In this paper, we perform a qualitative analysis to assist in the creation of a model that can help explain the current security, compliance, and business risks associated with cloud computing. Our qualitative process makes use of one-on-one interviews to collect the insights of 68 cloud and security experts. The interviews followed a flexible format to allow the free exchange of information and constructive interaction. Within the context of this paper, we use the word “risk” to refer to the aggregate risk of security, compliance, and business risks that corporations might experience using cloud services. The phrase “traditional IT” refers to IT services that are hosted on premise and are not virtualized or at least not virtualized using high-density configurations.

Digital Object Identifier: 10.1147/JRD.2013.2287591

© Copyright 2014 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the Journal reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied by any means or distributed royalty free without further permission by computer-based and other information-service systems. Permission to republish any other portion of this paper must be obtained from the Editor.

0018-8646/14 © 2014 IBM

Traditional IT is often single tenant, not completely automated, and it has significant differences across VM (virtual machine) installations. In addition, this paper refers to “compliance” as the law or government regulation that imposes a process required to do business. Examples of common compliance regulations include FISMA (Federal Information Security Management Act), HIPAA (Health Insurance Portability and Accountability Act), SAS 70 (Statement on Auditing Standards No. 70), and PCI-DSS (Payment Card Industry Data Security Standard) specifications.

In this paper, we gathered sufficient evidence, based on the interviews with subject-matter experts, to help identify a useful interpretation of the current state of cloud computing risks compared to traditional IT risks as well as to help IT professionals make appropriate decisions regarding cloud adoption. For those readers who may wish to review past research published in the area of cloud computing security, they may consult [5–11] for further examples and background on the subject.

Methodology

We selected a Delphi methodology [12], a structured communication technique, because it has been successful with research problems where there is incomplete and scarce information available. Also, the Delphi methodology has been shown to be effective where precise analytics is not applicable, and the analysis of the subjective judgment of individuals as a group or collection of experts is the best available source of information. One of the advantages of the Delphi method is that opinions tend to converge after successive rounds of feedback [13]. However, for this paper, we used a modified Delphi process that consisted of one-on-one interviews, Delphi iterations to build consensus, and a quantitative analysis that was built using a calibration spreadsheet. The experts that participated in this study are well-known IT professionals with many years of experience and expertise with cloud computing and security. The consensus meetings were conducted as short consecutive conversations about specific topics to come to an agreement on cloud risk vectors, the evaluation of cloud risks compared to traditional IT risks, and cloud framework taxonomies. As part of the Delphi method, we quantified the cloud risk vectors using a process we designated as the “calibration” cycle. This process took place over a six-week period, during which the results from the Delphi efforts and insight from the one-on-one interviews were merged to obtain agreement among the 68 experts. The calibration spreadsheet instrument included a row for each of the cloud risk vectors identified by the Delphi process and a radio button to quantify each vector as “increases risk,” “no change in risk,” or “decreases risk,” compared with traditional IT. The summary of the modified Delphi process and example

of the compliance calibration spreadsheet instrument are illustrated on **Figure 1**. In addition, the consensus process used a pyramid paradigm and led to agreement about the ranking of the top eight cloud risks and relative cloud risks compared with traditional IT. After the modified Delphi process was completed, a survey of 204 IT professionals was conducted to compare the risk perceptions and opinions of the 68 experts with a larger population of IT professionals. The target audience for the survey was IT professionals with several years of experience, but not necessarily experts on cloud security.

Rationalization of key cloud risks

After the consensus and calibration process, we divided the data between two main categories: those aspects that increased and decreased cloud risks compared to traditional IT. We created this division to qualitatively validate the hypothesis that cloud computing has new risks not present in traditional IT, and as such, cloud risks are not equal to traditional IT risks. However, as might be expected, we discovered that clouds and traditional IT have intersecting risks. For example, vectors that exist in both clouds and traditional IT risk sets include: human factors, authentication, authorization, SLAs (service-level agreements), DoS (denial-of-service) attacks, and cryptography key management, just to mention a few. Risk vectors that only exist in the set of cloud risk vectors are as follows: cloud image management, multi-tenancy, cloud automation scripts, and cloud management software. On the other hand, we encounter risks that tend to be unique to the set of traditional IT risk vectors such as management of physical servers and laptops, management of physical firewalls, rigid hardware taxonomy, and upfront capital investment. Almost every cloud and security expert we interviewed mentioned hypervisor technology and the new surface attack associated with cloud configurations as potential security risks. They also mentioned multi-tenancy, automation and standardization, authentication and authorization, device endpoints, concentration of value, and human factors as important security, compliance, and business risks. These eight key risk factors drove the creation of the “pyramid of cloud” risks vectors (**Figure 2**).

Hypervisor risks

The risk factor associated with hypervisors was one of the most common topics of concern shared by experts interviewed for this study. The only experts who did not agree with this concern were those with in-depth expertise on mainframes, and they asserted that mainframe LPARs (logical partitions)—partitioning a physical machine into multiple logical partitions using the IBM POWER* chip architecture—as the “correct” way to accomplish virtualization. Unfortunately, none of the

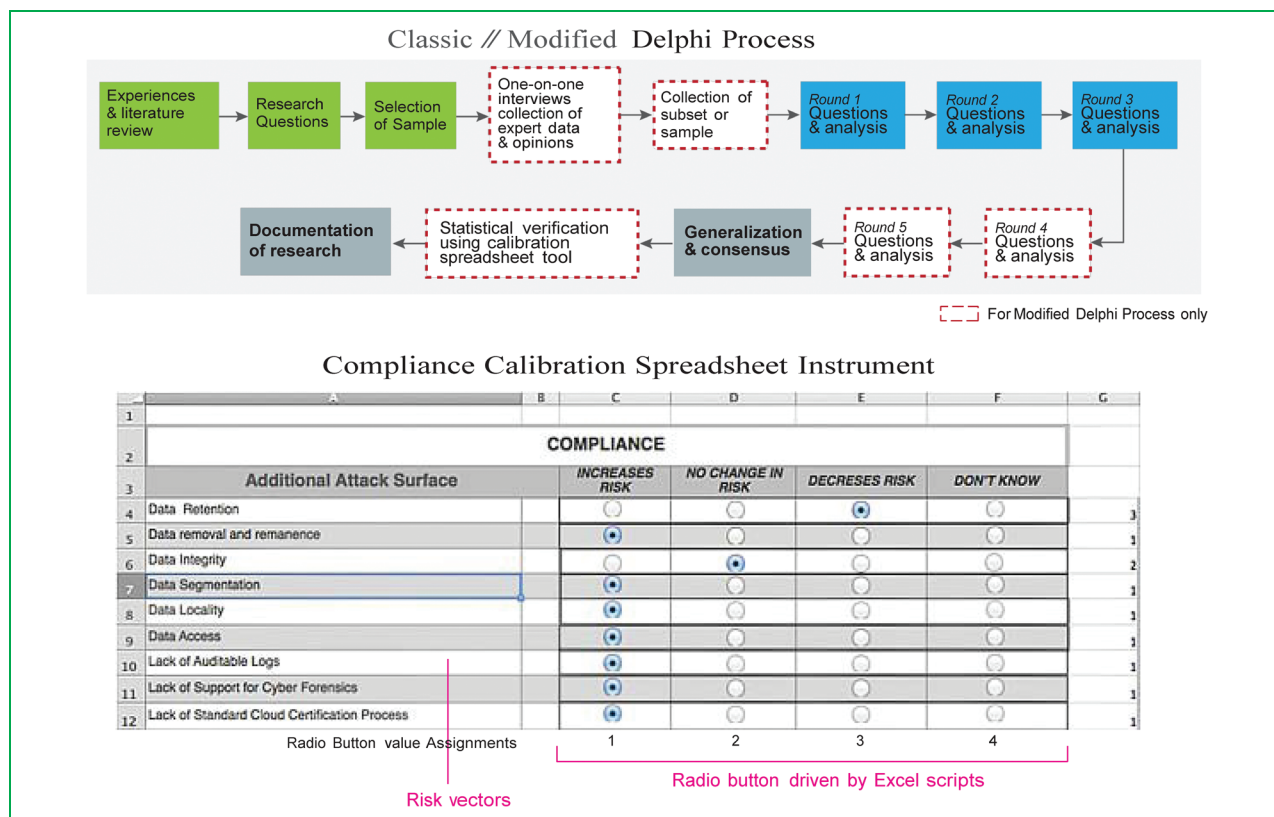


Figure 1

Methodology. The Modified versus Classic Delphi Research process is shown schematically at the top of the figure, and the Compliance Calibration Spreadsheet Instrument is shown at the bottom of the figure.

cloud computing offerings available today uses mainframe technology due to cost constraints; however, this could provide better isolation. For example, the IBM z10* mainframe is based on the IBM POWER chip design, which is very different from the commoditized Intel X86** microprocessor. One of the advantages of the mainframe LPAR virtualization is that it has access to memory addresses that do not overlap. This eliminates the possibility of cross-VMs side-channel leakage risks typical of X86 commodity hardware. In addition, LPARs have separate registers for the hypervisor and the OS, which eliminates the risk of escaping the hypervisor during a buffer overflow [14]. The risk associated with hypervisors can be easily mitigated through more inexpensive alternatives such as using extra-large VMs that consume the entire physical host, which eliminates most of the cross-VM side-channel leakages. As one would expect, this hypervisor mitigation strategy involves security and cost tradeoffs. Users could implement a cloud system with very good hypervisor isolation using LPARs, but this configuration could significantly exceed the cost of using extra-large VMs on commoditized hardware.

The cloud system is still a relatively new IT model that has a long maturity road ahead and many of these experts felt that one of the biggest risks were unknown factors, since the software components and model have not existed sufficiently long to measure their risks in a quantitative or statistical way. In roughly the last three years, clouds have made the transition from a relatively unknown paradigm to extreme IT popularity, with 60% of enterprise customers attempting or planning to move to cloud technology to achieve greater efficiencies [15, 16]. Cloud technology is usually a combination of commoditized hardware with new software incarnations of old concepts such as hypervisors. However, the business model of (1) paying just for what one uses; (2) web services in conjunction with SOA (service-oriented architecture), and (3) new automation technologies, created a very beneficial combination that enabled cloud computing. This phenomenon fueled the creation of many new applications as a service, and millions of users are now consuming clouds with software that is not yet mature. It is understandable that we constantly read and hear about security risks and compliance issues related to cloud solutions.

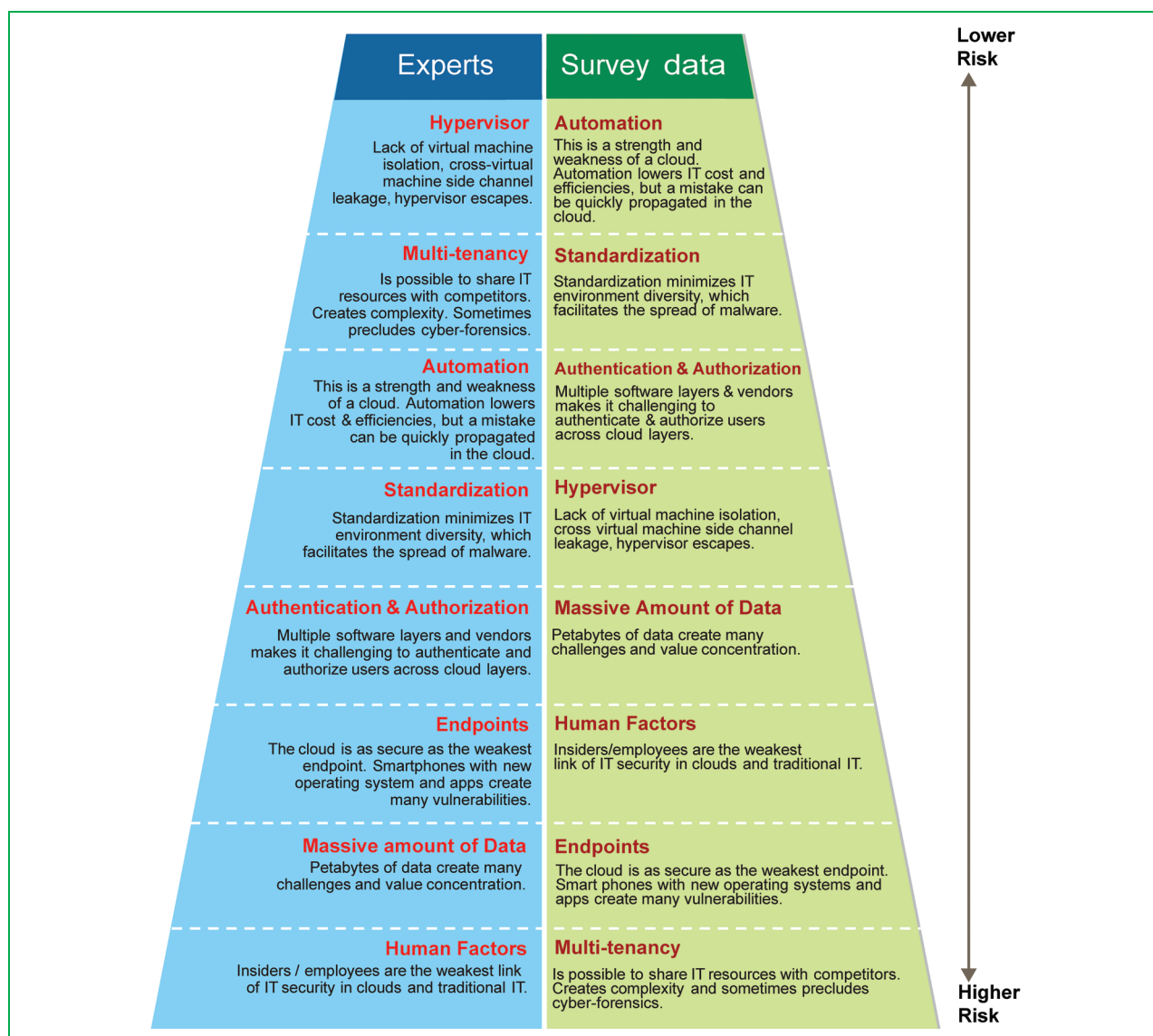


Figure 2

Pyramid of cloud risks comparison between experts and survey participants. Differences on the pyramids highlight several of the possible misconceptions relating to cloud security risks.

Multi-tenancy

The risk factor of multi-tenancy is related to issues and challenges associated with sharing the same physical IT resources such as CPU, memory, storage, network switches, firewalls, and other hardware and software components. Unfortunately, multi-tenancy could potentially lead to sharing resources with competitors who might engage in activities that exploit cross-VM isolation vulnerabilities to obtain confidential information. In addition, if a customer is subpoenaed by a court to provide data records and logs in a multi-tenant environment, there is a risk that a court

might obtain some data from “innocent” bystander VMs hosted in the same physical server. In addition, multi-tenancy combined with virtualization makes it difficult to implement tracking tools that can enable cyber-forensics on cloud environments and, as such, may lead to a failure to comply with some regulations.

Compliance has become more important in the last 10 years due to a significant increase in regulatory compliance laws. This has posed significant challenges to traditional IT infrastructure, as well as to clouds. However, clouds are significantly more challenged by

compliance processes because of their potential “infrastructure obscurity.” In addition, the transient nature of cloud consumption creates challenges with respect to privacy and auditability. New regulatory compliance rules that require maintaining documents for a long period of time are also creating great pressure for the adoption of storage clouds that make it more affordable to comply with these regulations; but other risks increase if there is little or no control over storage and cyber-forensics.

Automation and standardization

Cloud computing obtains its scaling features by automating most of the IT processes such as the provisioning and deployment of new VMs, patch management, backups, load balancing, security monitoring, and other processes. Cloud system administrators manage thousands of servers through the cloud dashboard and deploy processes with the “push of a button” (i.e., very easily). But if a mistake is made in the automation scripts, such an error can be replicated very quickly. Standardization involves a risk associated with minimizing the number of types of VMs available in the cloud. Standardization helps create a more homogeneous environment, but at the same time, this can lower the barrier for the spread of viruses or malicious software in clouds.

An enormous investment in automation leads to one of the advantages that clouds have, and cloud providers use automation to gain economies of scale, managing a huge number of machines. Automation brings lower administration cost since automation enables high density of servers per administrator—some experts claim as high as 10,000 servers per administrator in IT environments such as at Google and Amazon. In private clouds with less capacity, the numbers range between 60 and 100 servers per administrator. As mentioned, automation enhances flexibility because processes can be achieved with minimal effort. In addition, division of labor is more effective when the process experts create and maintain their own automation processes, because the knowledge does not need to be shared with administrators.

These benefits sound very compelling, but this strength is also a weakness. As mentioned, automation can deliver great efficiencies, but a mistake in a script can quickly spread with disastrous consequences. Unintentional accidents, such as the Amazon re-mirroring “storm” [17], are intensified by automation, and it is reasonable to expect more cloud storms in the future. Because a large part of cloud automation is beyond the control of the cloud user, there is no specific mitigation strategy for this vector. In general, automation is considered a riskier vector than the hypervisor and multi-tenancy.

Standardization is another aspect that can be both a disadvantage and an advantage of clouds. For example,

if a server is virtualized without standardization, the result is just more of the complexity that already exists. It is the process of standardization that reduces redundant configurations and creates a limited number of reusable IT configurations that can be automated to limit the deployment process to just a few mouse clicks [18–20]. This synergy between standardization and automation was stated by Bittman [20] when discussing his design for better virtualization, and others such as Oberle and Fisher [18] have referred to the lack of standardization as one of the key barriers to cloud adoption. However, it was the work of Rings et al. [19] on their testing framework for cloud environments that most clearly illustrated the power of standardization to create simpler configurations that are easy to automate. On the other hand, standardization minimizes the diversity of the IT environment, thus enabling the spread of viruses and malware across a cloud environment. Similar to the biological world, diversity can offer protection against viruses because it poses a challenge that requires more complexity and code to assail many IT configurations rather than just a few standard VMs. This is one of the reasons why vigilance and rigorous maintenance of software using the latest software patches is especially important in cloud environments.

Authentication and authorization

The risk factor of authentication and authorization is associated with the challenges of handling a large number of users and data objects. Traditional identity and access management (IAM) technology that manages IDs and ACLs (access control lists) for 10,000 users, as well as the protection of a few million objects, simply does not work well on the cloud scale of 200 million users and 100 billion files. IAM is predicated on a centralized model that does not scale in a distributed environment such as the cloud. The speed at which access change requests occur in the cloud is much faster, and this creates more new challenges and risks. Several federated identity management systems and remote control authentication services have been created, but there is still no standard enterprise-grade tool across cloud services available today. Authorization becomes a problem in the cloud, since identification of the user needs to pass securely across many cloud layers. On the other hand, it is fortunate that new technologies such as OAuth (open standard for authorization) are able to mitigate some of these risks. New open standards such as OAuth allow users to give tokens instead of credentials to the multiple layers of cloud services, minimizing the risk of identity theft or impersonation. However, these technologies are still in the early stages, and not all cloud providers and cloud services support them at this time. For example, Google App Engine** supports OAuth [21], but the authentication is only performed with respect to the Google Accounts service, creating the risk of duplicating identities

between the enterprise LDAP (*Lightweight Directory Access Protocol*) and the Google service. In addition, to configure one's Google App web application one needs to create several XML (Extensible Markup Language) web.xml files or YAML (app.yaml markup language) files to define the way one's application should run and its access privileges. This type of methodology is very coarse-grained and does not support other business requirements such as delegate authorization. Another example provided by the experts is Microsoft Azure**, which supports SAML (Secure Assertions Markup Language) and an STS (Security Token Service) to support this distributed identity management solution. The SAML protocol supports SSO (single sign-on) and multiple authentication methods and minimizes duplication of identity between the enterprise customer and the cloud provider, but its adoption in the industry has been slow [22]. Lack of cloud standardization with respect to authentication and authorization methods have allowed the proliferation of non-interoperable proprietary solutions that generate significant security risks.

Endpoints

The emergence of mobile devices has placed significant pressure on businesses to support interfaces with new device endpoints. New mobile devices such as cell phones and tablets are part of mobile clouds and provide enormous agility to employees. At the same time, these technologies open a large number of vulnerabilities to cloud solutions. For example, healthcare providers are under great pressure to support iPad** applications to review medical records “anywhere and anytime.” This is a great challenge because many of the mobile devices are not complying with minimum-security standards, and the software they use is still early in the maturity cycle [23].

Concentration of value

With the advent of “mega IT data centers” created by IBM, Google, Microsoft, Amazon, Yahoo!, and many others, these colossal warehouses contain tens of thousands of servers creating many new risks and challenges. The large amount and dense concentration of hardware can statistically guarantee a storage or server failure everyday somewhere in the cloud [24, 25]. Also, big data centers are becoming targets for elaborate new threats because of the large potential payout. This IT consolidation is creating value concentrations that are very attractive to malicious individuals and organized crime. This value concentration trend is not expected to change anytime soon, since the shift to consolidate into huge data centers is funded by cost reductions and energy efficiencies. The more value that is concentrated in the cloud, the bigger the target cloud services will become. In addition, the expectation of and demand for data availability “anytime and everywhere” in the world have created an enormous challenge that

sometimes is contrary to data consistency. New distributed databases, not based on SQL (Structured Query Language), have become popular in the cloud with astonishing data availability at low cost, but at the price of lower consistency. Traditional relational SQL databases are tuned for accuracy and comply with ACID (atomicity, consistency, isolation, durability) rules, but they have lower data availability due to the locking mechanism of the two-phase commit. With massive amounts of data, specifically above the 10 petabytes range, the expense associated with handling and backing up the data grows exponentially. New technologies such as Hadoop**, NoSQL, Cassandra**, and IBM General Parallel File Systems (GPFS*) have been created to help handle this kind of data in the clouds. However, these kinds of tools sometimes optimize for availability instead of accuracy, and this can create integrity and security issues. In addition, massive data concentration in clouds creates a value concentration that attracts the attention of malicious individuals. Massive amounts of data create the challenge of meeting the insatiable demand for availability without degrading the data accuracy. Availability and data accuracy become competing requirements. As one partitions the data to increase availability, the consistency decreases, and the reverse happens when one reduces the data partitions to increase the consistency, but the availability decreases [26].

Human factors

Unfortunately, all experts agree that the most significant security risk affecting clouds is the human factor, since most fraudulent behaviors are due to “insider threat.” In some cases the insider risk is higher in the clouds than in traditional IT because of the value concentration characteristic of clouds. Our experts explained how cloud providers make extensive efforts to install sophisticated monitoring tools on the administrator's console and dashboard terminals. These tools monitor changes made by system administrators, in the hopes of discouraging malicious or disgruntled employees who might steal information or sabotage the system. However, this particular risk factor is expected to continue to be a problem for the foreseeable future because, as world famous hacker Kevin Mitnik [27] said, breaking the “human firewall” is easy. In addition, more than 70% of the experts interviewed for this paper cautioned about the danger of social-engineered attacks that exploit people's vulnerabilities.

Cloud Security and Compliance Risks Framework

In this section, we focus on the risk vectors that affect security and compliance in clouds. Compliance demands can include government regulations, industry procedures, and specific business requirements. Government regulatory compliance includes regulations such as the Sarbanes-Oxley Act of 2002 (SOX), which established new audit standards

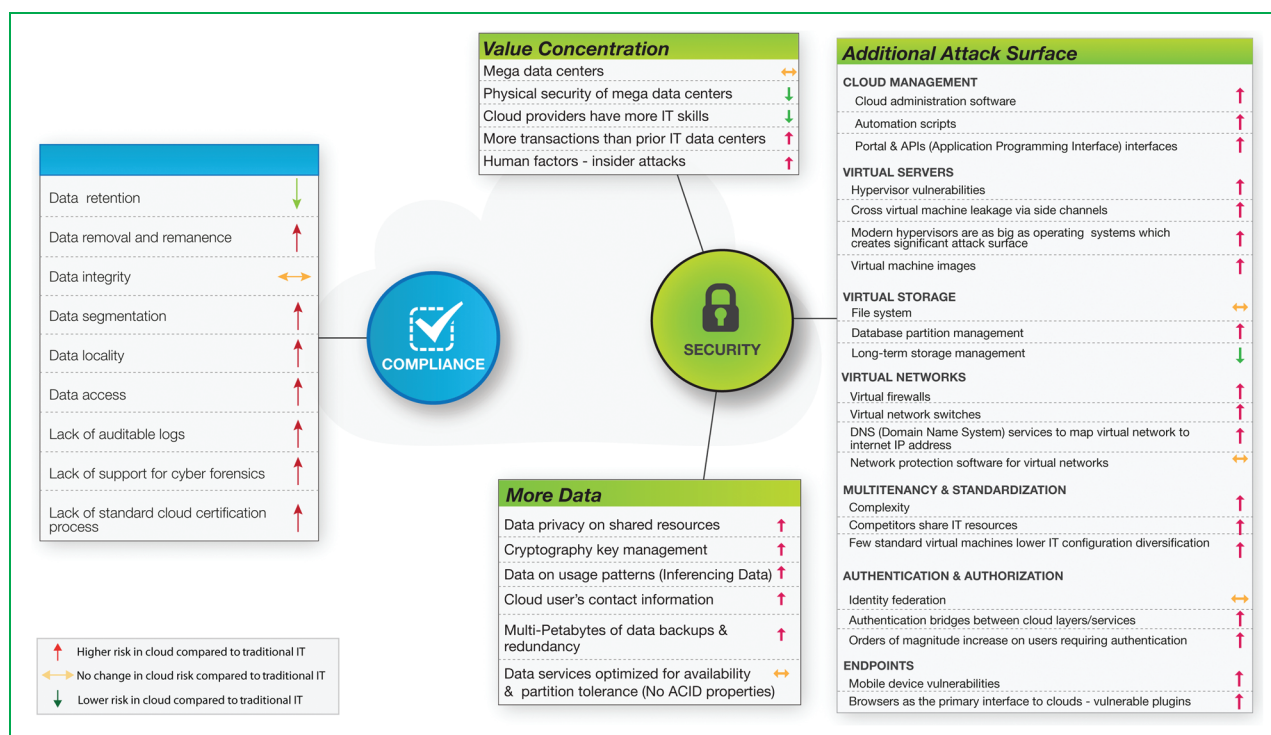


Figure 3

Cloud Security and Compliance Risks Framework. (IP: Internet Protocol.)

to increase business transparency and ethical behavior. Another example of a fairly new regulatory compliance law is the Dodd-Frank Wall Street Reform and Consumer Protection Act (Dodd-Frank), enacted in 2010. This legislation is not restricted to financial institutions and affects the governance, disclosures, and executive compensation of American-based companies. In addition, there are many industry-specific compliance regulations such as those associated with FFIEC (Federal Financial Institutions Examination Council), ISO (International Organization for Standardization) 2700X, NIST (National Institute of Standards and Technology), HIPAA, PCI-DSS, and others. The challenge is to create an IT service that can ensure the process of adhering to regulatory, commercial, and business procedures to achieve the desired compliance level.

The Cloud Security and Compliance Risks Framework (Figure 3) illustrates the key vectors affecting cloud environments, to facilitate the understanding of cloud and compliance risks. Each risk vector was assigned an arrow that depicts the risk associated with cloud compared with traditional IT. The frameworks described in this paper use a red up arrow to express higher aggregated risk associated with clouds, a horizontal yellow arrow for vectors that do not change the risk between clouds and

traditional IT configurations, and a green down arrow to express lower aggregated risk associated with clouds.

The security vectors were organized into three categories: those risks driven by attack surface, value concentration, and more data. The new attack surface is the result of additional code required for cloud administration and services. The cloud administration includes the cloud automation process, which consists of thousands of script lines that automate the instantiation, hibernation, image capturing, and end-to-end lifecycle for the virtual solutions resident in the cloud. Unfortunately, the level of maturity of cloud management software is relative low, and this inevitably tends to cause vulnerabilities.

The virtualization risk factor is divided into three types: virtual servers, virtual networks, and virtual storage. All of these refer to the ability to virtualize resources and share them across multiple tasks. This virtualization capability adds an additional attack surface, but selecting a hypervisor with a small footprint (i.e., small in terms of lines of code), such as type 1 hypervisors, can significantly reduce this type of risk. Another exposure very common in virtualized environments is VM sprawling. Since hypervisors simplify the process of creating new VM instances, most people tend to create new VMs instead of reconfiguring old ones, and do not bother to remove resources that are no longer

needed. If the cloud does not have a well-functioning process to verify patches and force updates on inactive VMs, these can quickly get out of date and become not only vulnerable viruses, but also a drag on performance. Virtualization technology also facilitates the process of capturing VM images, which significantly increases security risks on two fronts. First, the convenience of VM image creation facilitates the sharing of VMs. At the same time, malware and perhaps confidential information can also be unintentionally shared if appropriate filters or protection are not in place. The second risk that image capturing creates is fostered by the representation of the virtual system in a file. On traditional IT systems, the OS uses the BIOS (basic I/O services) interfaces to write the data stack to ROM (read only memory), maintain status information, and process the executable. In a virtualized environment, the hypervisor simulates the BIOS interfaces. Instead of using ROM, it writes to a file on disk, which is the VM image. If the image file is not encrypted and a malicious individual gains access to the disk where the VM image is resident, all the information in the VM will be accessible to the intruder, even if the VM is at rest (inactive).

Storage virtualization has the advantage of providing a lower-cost alternative to traditional storage. It is used frequently for massive amounts of data—the kind of large data that eventually is partitioned. With large data installations (> 10 TB) and the high density that virtualization creates on disks, it is virtually certain that partitions will occur. This is why new distributed databases and files systems such as the GFS** (Google File System) have been created to help handle partitions on large data storage [28]. However, all these new technologies add more attack surfaces. In addition, the technology is still immature, creating potential for vulnerabilities. However, its benefits of enhanced performance and availability are undeniable.

Clouds create overlaid networks with logical structures able to assign private addresses within the virtualized host. A virtualized network creates many benefits such as complete flexibility of the network taxonomy. A virtualized network also removes any restrictions based on physical IP (Internet Protocol) addresses provided by the network provider, and the entire network configuration can be easily serialized into a file that can then be accessed or combined with the VM image to replicate not just a particular VM but the entire IT solution. However, the strengths of the virtual networks are also their weaknesses. As a virtual network is serialized into a file containing the network specifications, the risk of unauthorized individuals being able to access the network configuration is much higher than on traditional IT configurations that use physical network devices. Like other sensitive data, network specification files need to be protected and encrypted because these files provide information that can be very useful to hackers and malicious individuals. In addition, as physical network

components such as firewalls and network switches are replaced with software components, there is a need to ensure that these components are patched regularly to avoid exposure to published vulnerabilities.

Cloud security risks can be mitigated by avoiding “rogue VMs” and using only VM images created by trusted partners. (Here, we use the phrase *rogue VM* to refer to virtual machines created using unmanaged VM images from nonqualified sources.) If possible, companies should create their own company VM images, distribute those to employees, and keep the images in a secure storage. However, if it is necessary to use VMs provided by vendors, it is important to ensure that the VM images have a signature that can be tracked back to the original vendor. If a VM image does not have a signature, it is not advisable to use it. In addition, disable or limit the use of dangerous hypervisor tools frequently used by hackers to break into a virtualized environment. Some of these tools use private communication channels between the hypervisor and the host OS that allows cross-VM communication. For example, the tools VMchat, VMCat, and VMftp use the ComChannel in the VMware** hypervisors to gain access across VMs. One should follow the configuration recommendations provided by the hypervisor vendor since a well-configured virtualized environment will be more difficult to break. In 2010, VMware provided a report indicating how to configure their hypervisors. This guide should be followed by anyone considering the use of VMware hypervisors [29].

Other additional recommendations focused on the importance of performing security audits to ensure that developers perform regular security audits on web applications to avoid the most common attacks, such as SQL-injection and cross-site scripting. In addition, regular third-party audits and penetration testing should be established to obtain an impartial opinion on the level of security a cloud provides.

The development of the compliance aspects of the framework uncovered some clear risk exposures. For example, several regulatory compliance mandates have a common requirement to securely and safely dispose of data and storage that is no longer necessary. A process must be in place to ensure there are no remnants of data that could help reconstruct the information deleted. For example, in the case of the HIPAA regulation, the removal of PHI (personal health information) is required in such a way that physical storage is completely erased. Any trace of information must truly be deleted, from both the current instance, and from any other possible instances of the data source. This kind of requirement represents a significant challenge to cloud services, since storage location and management are difficult to determine. Cloud customers with HIPAA requirements can request specific removal procedures in their SLAs to properly manage storage

containing PHI. However, since cloud providers are not currently offering audit application programming interfaces (APIs) to enable the verification of the proper removal of information, this remains a trust issue between cloud customers and cloud providers. Without audit APIs or interfaces to verify procedures followed by a cloud storage provider and a managed service provider (MSP), the cloud customer could incur audit risks. Compliance responsibilities remain with the institution rather than with vendors. For example, a cloud vendor may fail to follow the procedures documented in the SLAs, and the cloud customer may fail with respect to compliance responsibilities as a result of the vendor's poor execution of requested procedures. Since cloud providers usually do not indemnify their customers for compliance penalties, any financial liability is sustained by the business and not by the cloud vendor.

Healthcare data, government data, employee information, and financial transactions are some of the heavily regulated information that has location-specific compliance requirements. Locality requirements on data involve the need to know the actual physical site of the server and storage services containing or using the data. Many requirements specify region, county, or country where the data must reside. If location requirements are not satisfied, many regulations stipulate costly penalties. In some instances, the business can lose its license to operate. Due to the obscurity that most cloud providers maintain about their storage services, cloud customers do not have the ability to monitor or automatically audit the location of the cloud storage they are consuming. The risk of failing compliance regarding location of data is higher in cloud services than in traditional IT, where the CIO has full control over the physical location of servers and storage.

Another compliance concern is the ability to perform cyber forensics. For example, HIPAA, PHI, and FFIEC, require fraud detection, and this demands IT services that can support cyber-forensic capabilities. When working with sensitive data and transactions that must support regular audits, businesses need cyber forensics. When a legal dispute arises, the business needs to produce sufficient evidence to satisfy court requests for information. Cyber-forensics requires auditable logs, data access history details, hardware configuration, VM images, network topology, and many other data points that can help reproduce the entire environment. Many months of archival data are usually required to establish a sufficient pattern as evidence of negligence or fraud. However, most clouds do not keep the necessary information to reproduce prior environments. The dynamic aspect of clouds in which VMs are constantly instantiated and destroyed, combined with the lack of auditable logs and incomplete data points, makes it very difficult to reproduce an environment. The retention of data for a long period of time comes at significant cost, and IT managers need to make tradeoffs

between cost and compliance risks. However, since current cloud forensics tools are not mature, and reproducing the cloud environment remains a challenge, retaining additional data will not help with the technical immaturity of forensics tools. However, it is expected that cloud forensic tools will get better, and retaining data could become very helpful.

Cloud Business Risks Framework

The Business Risks Framework, illustrated in **Figure 4**, is divided into Cost, Efficiency, Control, Availability, and Legal Complexity. Cost (e.g., reduced capital investment and reduced operational costs) is one of the very few factors that experts unanimously agreed upon as having a positive effect for lowering "business risks." The experts seemed to have an unwavering belief about the significant benefits that the cloud model brings to businesses.

The cloud business model eliminates the need for capital investment, significantly reduces operational cost, and provides payment flexibility since one pays only for what one consumes. In addition, the skills and training required to maintain cloud VMs is significantly lower than the expertise and insight necessary to support traditional IT. This particular point on skills has been supported not only by the experts consulted for this paper but in many other sources [22, 30]. The flexibility of payment according to what is used is the model employed by such cloud providers as Amazon, GoGrid, and Rackspace. We were not able to find any divergent opinion regarding the benefits of having a flexible payment structure for clouds.

Understanding how to calculate the cost of VMs based on a given scenario is of key importance because the price of VMs varies significantly based on the configuration and provider. After calculating the cost of the VMs, the next step is to evaluate the overall cost associated with the cloud when compared with the current in-house IT data center. The cost comparison of cloud versus traditional IT must consider what workloads would run more cost effectively in cloud environments than traditional IT and, also, what characteristics make a workload expensive and not suitable for clouds. The findings of this paper were not surprising, but instead very consistent with expectations that workloads with variable or "bursty" demand pattern are good candidates for cloud services. The "bursty" pattern can also be created by predictable peaks caused by differences in demands due to time-of-day, day-of-the-week, or cyclical patterns such as tax season. For example, in the United State, most home-banking websites experience time-of-day peaks from 4:00 p.m. to 7:00 p.m. The first wave of demand starts on the East Coast and moves west as time progresses.

With predictable workloads, traditional IT tends to overprovision the IT infrastructure to accommodate the worst-case scenario, which typically includes the highest

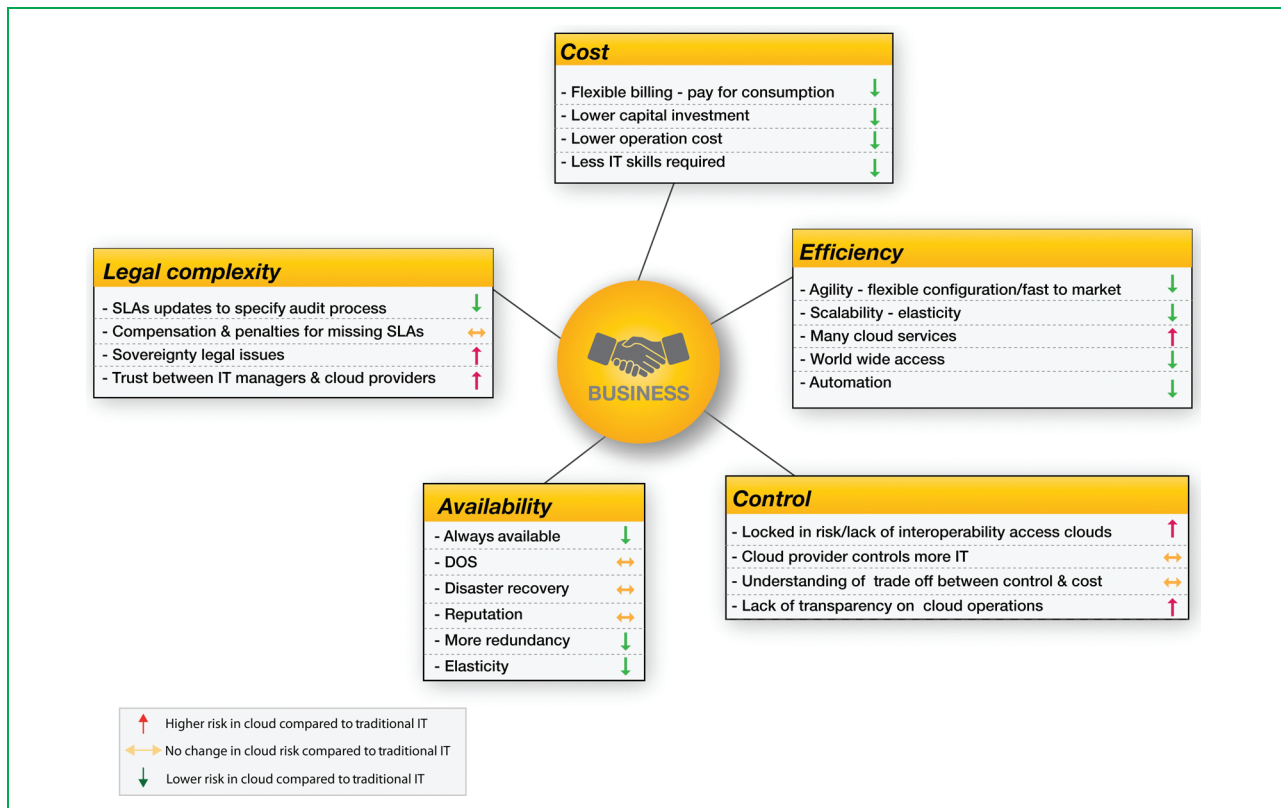


Figure 4

Cloud Business Risks Framework.

possible demand, plus a buffer for safety. The problem with this kind of approach is that the capital expense (capex) is substantial, but most of the equipment is idle for long periods. The difference between the average needs and the peaks can be significant with peak-to-average ratio (PAR) of about two or three, which means 100% to 200% increase at peak from the average volume. Predictable patterns associated with business cycles, sales promotions, or special annual events such as Christmas and tax season can generate PAR values of four to ten. Solutions with unpredictable “bursty” patterns are also good candidates for clouds because the elasticity of the cloud mitigates the risk of running out of capacity.

Another business advantage of clouds is to be able to choose from many workloads to maintain systems at high capacity. This contrasts with traditional IT, where lack of virtualization, no multi-tenancy, and a small number of instances works against optimizing workloads to maintain high utilization. In traditional IT, most workload configurations remain static because of the difficulties associated with moving services. The situation is very different in clouds. Automation and virtualization technology facilitates the move of VMs to aggregate complementary

workloads, thus maximizing the usage of IT resources. For example, solutions from different time zones that have staggered demand peaks can be combined to maintain higher hardware utilization. Clouds tend to operate at close to 95% utilization compared with traditional IT at 15% to as low as 5% utilization [31]. With this gap in hardware utilization, it is not surprising that cloud providers are able to offer resources at such unprecedented low cost.

In traditional IT, downtime due to hardware failure is very different than downtime in the clouds. The issue is not about when the server or VM in the cloud will go down, but how quickly the cloud can recover from a failure event. In traditional IT, the failure of a physical server can be a dramatic event, and sometimes it takes a significant amount of time to recover. Even in the case of high-availability systems with mirrored databases, the recovery process to replace the failed system and recover the high-availability configuration takes some significant effort. Recovery may require only a few hours, or it may take days or even weeks. On the other hand, recovery from failures in the cloud occurs within several minutes due to the automatic scripts and usage of standard VM images. Events that are catastrophic in a traditional data center are a routine

occurrence in the cloud environment, but with no tangible significance to the user. This is a paradox, because failures are more common in clouds than in traditional IT; however, due to significant redundancy, the risk of failure is greatly mitigated. One may ask why clouds have more failures than traditional IT? These failures are the result of massive amounts of systems and storage running at 95% utilization in mega data centers. Reviewing the equation of the probability of failure over time period $P(t)$, we can see that as the mean time to failure (MTTF) decreases, the probability of a failure occurring increases: $P(t)_{\text{failure}} = 1 - e^{-kt}$, where $k = 1/\text{MTTF}$. The exponential failure distribution equation helps us model the rate of failure over time based on MTTF [32].

Because clouds tend to have lower MTTF than other data centers, the failure rate in clouds is higher than in traditional IT [33, 34]. One may also ask why does the MTTF in clouds tend to be lower than for traditional IT? This is an issue that many computer systems and storage manufacturers are studying. Several data references available today from Intel [32] and others [24, 33] suggest that a byproduct of the high utilization of resources in cloud environments is the shortening of the MTTF. Based on their experiences dealing with large data centers, experts participating in this study have the opinion that lower MTTF is caused by two factors: clouds use resources more intensively, and the hardware in most clouds is commodity hardware with intrinsic lower MTTF. Both of these factors can contribute to the lower overall MTTF in clouds [32]. At the Intel Developer Forum conference in 2011, Sam Siewert and Greg Scott showed similar statistics [32]. This is in contrast to traditional IT, where hardware runs idle 85% of the time, and the storage and servers are usually of premium quality, with higher MTTF claimed by manufacturers. The main factor that is driving the lower MTTF in clouds is not clear at this point, but it is a common observation measured by cloud providers.

As suggested, clouds provide the illusion of “always available,” despite the constant failure of systems because of the large redundancy and automation they possess. Automation enables very fast recovery, which helps minimize system downtime. The shorter the recovery time, the better the availability. Similarly, the more redundancy a system has, the better its availability [30].

From a legal perspective there are many issues associated with cloud contracts due to unclear and restrictive laws, frequent trans-border operations (e.g., borders between countries), and lack of precedents to guide litigation. Most standard cloud contracts are based on “as is” warranties, which means the service is provided with no promises of any kind. There is no guarantee that the cloud service will be appropriate or that it will meet the customer’s expectations. Not all cloud providers offer SLAs, and those that specify

assurances are usually based on limited obligation and availability. Also, lack of standards is a problem for cloud contracts because there is no unified way to offer cloud services and there are no standardized benchmarks to help quantify the quality of service.

Trans-border data flows are common when data is resident in multiple countries and the cloud service and customer are located in different countries. Data flow that crosses the borders of a country is subject to the jurisdiction of multiple countries, and can create costly litigation fees because of unclear and contradictory laws. Trans-border data flow has many potential legal issues that can arise because of inappropriate handling of data, disparities between IT regulations depending on country, and ambiguity about obligations in case of a dispute. To avoid some of these problems, IT managers should fully negotiate cloud contracts to ensure the agreements satisfy the needs of the business and avoid ambiguity about roles, responsibilities, and processes.

Conclusion

The frameworks described in this paper are expected to be used by IT managers as a way to rationalize the many risks associated with cloud computing and to help increase the understanding of cloud risk vectors. In addition, this paper supports the hypothesis that the set of cloud risks includes some new risks, as well as already existing risks in traditional IT. New cloud-specific security risks, such as multi-tenancy, extreme levels of automation, and high value concentration in mega IT data center are some of the new risks tilting the overall security risks to be higher on clouds than traditional IT. Further, cloud compliance risks related to the inability to perform proper cyber forensics, data sovereignty, and data remanence (e.g., residual data that remains after attempts to remove the data) are a few of the many new risks associated with clouds that are increasing the compliance security risks in clouds as compared to traditional IT. However, some risks remain exclusively with traditional IT, such as high upfront capital investment for new IT services, and the ever-increasing operational cost of proprietary configurations. Based on the substantial data collected, we can conclude that many cloud computing risks are distinctly different from traditional IT risks, that traditional IT still has unique risks that differ from cloud computing, and that there are several risks that are shared across both environments, creating a risk intersection between cloud and traditional IT sets of risks. In **Figure 5** the reader can observe the intersection between the two sets of risk vectors, the blue circle for cloud and the green circle for traditional IT, but there are still substantial unique risks associated with each set. At the bottom left corner of Figure 5, we can see that from the business perspective, the traditional IT set of risk vectors (green circle) is substantially larger than

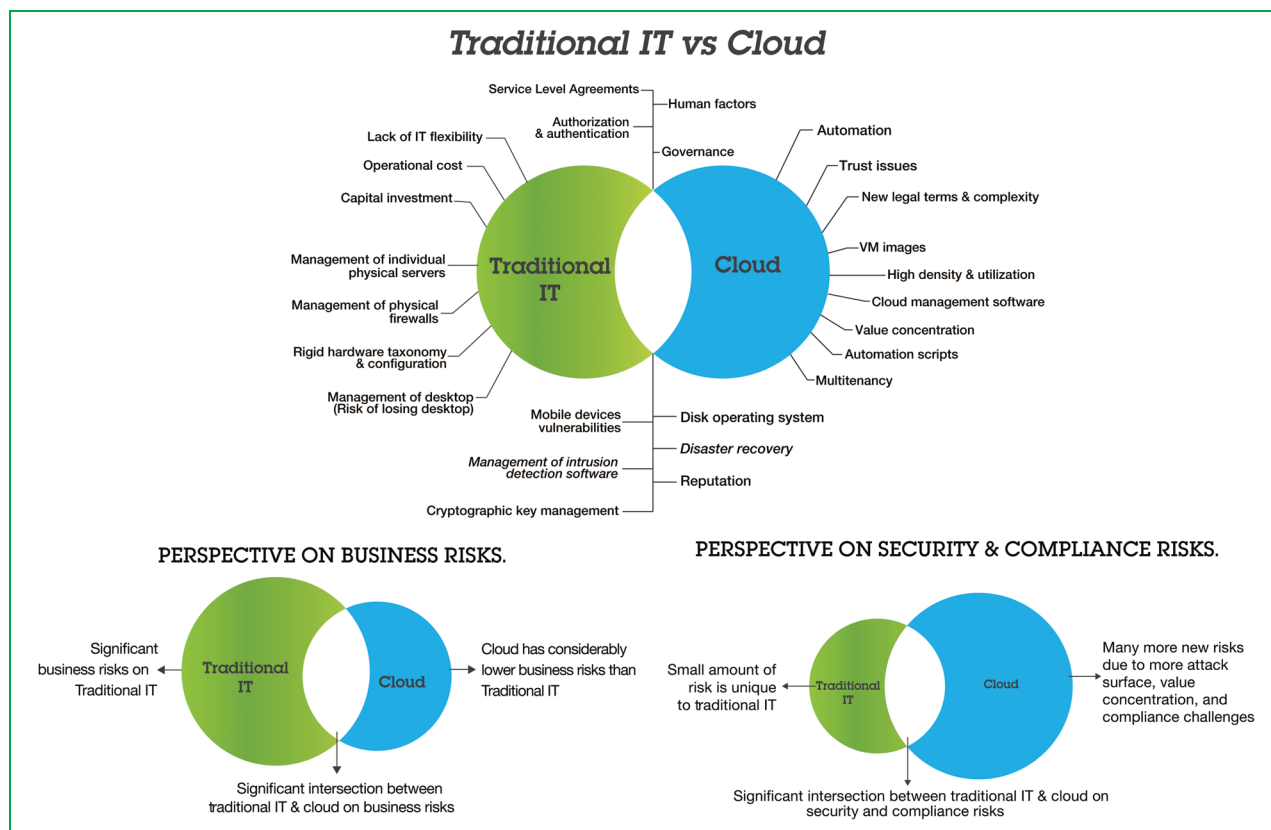


Figure 5

Traditional IT and cloud risks sets intersect, as shown on the top diagram. From the perspective of business risks, traditional IT has considerable more risks than clouds. However, from the perspective of security and compliance, clouds have more risks than traditional IT.

the cloud set of risk vectors (blue circle), illustrating the conclusion of this paper that from the business perspective clouds are a lower risk option than traditional IT. As shown in the bottom right corner of Figure 5, dealing with the security and compliance perspective, the cloud set of risk vectors (blue circle) is substantially larger than the traditional IT set of risk vectors (green circle), illustrating the second conclusion of this paper, namely, from the security and compliance perspective, clouds tend to have substantial higher risks than traditional IT.

We also concluded that based on the many mitigation strategies available for cloud, the fast pace with which cloud technology is advancing, and the significant economic advantages of the cloud business, we can conservatively predict that many of the cloud risks described in this paper will diminish over time, and clouds will become a much safer place and probably the preferred IT delivery model for outsourcing services.

*Trademark, service mark, or registered trademark of International Business Machines Corporation in the United States, other countries, or both.

**Trademark, service mark, or registered trademark of Google, Inc., Intel Corporation, Microsoft Corporation, Apple, Inc., Apache Software Foundation, VMware, or Scrum Alliance in the United States, other countries, or both.

References

1. S. K. Crook, C. J. Kolodgy, S. Hudson, and S. D. Drake, "Worldwide mobile security 2011–2015 forecast and analysis," IDC, Framingham, MA, USA, Rep., Mar. 2011.
2. C. Arthur, "Half of UK population owns a smartphone," in *Guardian*, London, U.K., 2011. [Online]. Available: <http://www.guardian.co.uk/technology/2011/oct/31/half-uk-population-owns-smartphone>
3. S. Ramgovind, M. M. Eloff, and E. Smith, "The management of security in cloud computing," in *Proc. ISSA*, Aug. 2010, pp. 1–7.
4. C. Liebert and M. Posey, "Security highlights from the 2010 U.S. WAN manager survey," IDC, Framingham, MA, USA, Rep., Jan. 2011. [Online]. Available: <http://www.idc.com/research/viewdocsynopsis.jsp?containerId=226221§ionId=null&elementId=null&pageType=SYNOPSIS>
5. T. Mather, S. Kumaraswamy, and S. Latif, *Cloud Security and Privacy—An Enterprise Perspective on Risks and Compliance*. Sebastopol, CA, USA: O'Reilly Media, 2009.
6. M. Christodorescu, R. Sailer, D. L. Schales, D. Sgandurra, and D. Zamboni, "Cloud security is not (just) virtualization security: A short paper," in *Proc. ACM Workshop Cloud Comput. Sec.*, New York, NY, USA, 2009, pp. 97–102.

7. B. R. Kandukuri, R. Paturi V, and A. Rakshit, "Cloud security issues," in *Proc. IEEE Int. Conf. Services Comput.*, Bangalore, India, 2009, pp. 517–520.
8. L. Edward, K. Amokrane, D. Lourdeaux, and J. Barthes, "An ontology for managing a virtual environment for risk prevention," in *Proc. 1st Int. Conf. Integr. Intell. Comput.*, 2010, pp. 62–67.
9. C. Cachin and M. Schunter, "A cloud you can trust: How to ensure that cloud computing's problems—Data breaches, leaks, service outages—Don't obscure its virtues," in *Proc. ICITST*, Dec. 2001, pp. 214–219, IEEE.
10. P. Zech, "Risk-based security testing in cloud computing environments," in *Proc. 4th IEEE Int. Conf. Softw. Testing, Verification Validation*, Berlin, Germany, Mar. 21–25, 2011, pp. 411–414.
11. T. Ristenpart, E. Tromer, H. Shacham, and S. Savage, "Hey, you, get off of my cloud: Exploring information leakage in third-party compute clouds," in *Proc. ACM Conf. CCS*, Chicago, IL, USA, Nov. 9–13, 2009, pp. 199–212.
12. G. J. Skulmoski, F. T. Hartman, and J. Krahn, "The Delphi method for graduate research," *J. Inf. Technol. Educ.*, vol. 6, no. 1, pp. 1–21, 2007.
13. C. Okoli and S. D. Pawlowski, "The Delphi method as a research tool: An example design considerations and applications," *Inf. Manage.*, vol. 42, no. 1, pp. 15–29, Dec. 2004.
14. D. Morton, "IBM mainframe operating systems: Timeline and brief explanation for the IBM system/360 and beyond," IBM, Armonk, NY, USA, Rep., Sep. 2011. [Online]. Available: <http://www.demorton.com/Tech/OSTL.pdf>
15. T. J. Bittman and L. Leong, "Virtual machines will slow in the enterprise, grow in the cloud," Gartner, Stamford, CT, USA, Rep. No. G00210732, Mar. 4, 2011.
16. E. C. G. T. Sacconaghi and M. R. Shah, "IT hardware: Does a move to large public clouds and data centers hurt or help server OEMs?" Bernstein Res., London, U.K., Rep., Apr. 6, 2011.
17. Richard, "Amazon says AWS outage was a 'Re-mirroring Storm,'" in *Cloud News Daily*, Accessed Dec. 14, 2011. [Online]. Available: <http://cloudnewsdaily.com/2011/04/amazon-says-aws-outage-was-a-re-mirroring-storm/>
18. K. Oberle and M. Fisher, "ETSI CLOUD—Initial standardization requirements for cloud services," in *Proc. GECON*, 2010, pp. 105–115.
19. T. Rings, J. Grabowski, and S. Schulz, "On the standardization of a testing framework for application deployment on grid and cloud infrastructures," in *Proc. Adv. Syst. Testing VALID Lifecycle Conf.*, 2010, pp. 99–107.
20. T. J. Bittman, "The road map from virtualization to cloud computing," Gartner, Stamford, CT, USA, Rep. No. G00210845, Mar. 3, 2011.
21. S. A. Almulla and C. Y. Yeun, "Cloud computing security management," in *Proc. ICESMA*, Sharjah, United Arab Emirates, 2010, pp. 1–7.
22. D. Chappell, *Introducing the Azure Services Platform: An Early Look at Windows Azure, .Net Services, SQL Services, and Live Services*, Oct. 2008, Report.
23. *IBM X-Force 2011—Mid-Year Trend and Risk Report*, IBM Corporation, Armonk, NY, USA, Sep. 2011.
24. E. Pinheiro, W. Weber, and L. A. Barroso, "Failure trends in a large disk drive population," in *Proc. 5th USENIX Conf. FAST*, 2007, p. 2.
25. F. Brown and R. Ragan, "Pulp google hacking—The next generation search engine hacking arsenal," in *Proc. Black Hat*, Las Vegas, NV, USA, 2011, pp. 1–70. [Online]. Available: https://media.blackhat.com/bh-us-11/Brown/BH_US_11_BrownRagan_Pulp_Google.pdf
26. J. Browne, *Brewer's CAP Theorem*, Accessed Jan. 11, 2009. [Online]. Available: <http://www.julianbrowne.com/article/viewer/browsers-cap-theorem>
27. K. D. Mitnick and W. L. Simon, *The Art of Deception: Controlling the Human Element of Security*. Hoboken, NJ, USA: Wiley, 2003.
28. S. Ghemawat, H. Gobioff, and S. Leung, "The google file system," in *Proc. Symp. Oper. Syst. Principles*, Lake George, NY, USA, Oct. 19–22, 2003, pp. 29–43.
29. *VMware vSphere 4.0—Security Hardening Guide*, VMware, Palo Alto, CA, USA, May 2010, pp. 1–70, Report.
30. B. S. G. Linden and J. York, "Amazon.com recommendations: Item-to-item collaborative filtering," *IEEE Internet Comput.*, vol. 7, no. 1, pp. 76–80, Jan./Feb. 2003.
31. M. Azua, *The Social Factor : Innovate, Ignite, and Win through Mass Collaboration And Social Networking*. Upper Saddle River, NJ, USA: IBM Press, 2009.
32. S. Siewert and G. Scott, "Next Generation Scalable and Efficient Data Protection," in *Proc. Intel Develop. Forum*, 2011, pp. 1–30. [Online]. Available: http://ecee.colorado.edu/~siewerts/SF11_STOS004_101F.pdf
33. S. Subashini and V. Kavitha, "A survey on security issues in service delivery models of cloud computing," *J. Netw. Comput. Appl.*, vol. 34, no. 1, pp. 1–11, Jan. 1, 2011.
34. B. Schroeder and G. A. Gibson, "Disk failures in the real world: What does an MTTF of 1,000,000 hours mean to you?" in *Proc. 5th USENIX Conf. FAST*, San Jose, CA, USA, 2007, pp. 1–16.

Received March 13, 2013; accepted for publication April 16, 2013

Maria Azua Himmel *Barclays Bank, London, United Kingdom (maria.azua@barclays.com)*. Dr. Azua is Managing Director and Global Head of Infrastructure Engineering for Barclays Bank and is responsible for engineering strategy, design, and implementation of technology in support of Barclays business units worldwide. Prior to Barclays, Dr. Azua was a veteran IBM executive of 23 years. Her last position at IBM was Vice President of Integrated IT and Technology for IBM Global Process Services. In that role, she managed the worldwide integrated IT in support of the IBM business process outsourcing service. Prior to this role, Dr. Azua was the Vice President of Advanced Cloud Solutions and Innovation for the IBM Global Technology Services division and responsible for the Latin America region for IBM cloud services and support. In 2009, she was part of the team propelling IBM into the cloud business with her role of Vice President of Cloud Enablement for the IBM Enterprise Initiatives organization, responsible for the development and deployment of the Common Cloud Platform Living Lab share service, as well as the IBM Smart Business Development and Test Cloud*. She is the author of *The Social Factor - Innovate, Ignite, and Win Through Mass Collaboration and Social Networking*, a book about social networking based on her pioneering leadership in the development of social networking tools as part of the IBM CIO (Chief Information Officer) team. Her contributions to social networking tools, B2B (business-to-business) solutions, and IT transformation are widely recognized throughout IBM and the wider IT community.

Fred Grossman *Pace University, White Plains, NY 10606 USA (grossman@pace.edu)*. Dr. Fred Grossman has been a professor of Computer Science and Information Systems for more than 40 years, and has been involved in software development for 40 years. He is a professor and Program Chair of the Doctor of Professional Studies in Computing in the Seidenberg School of Computer Science and Information Systems of Pace University. His principal research interests are in software engineering, agile methodologies and processes, automated systems development, automated programming, very-high-level language design, and the integration of information systems and organization strategies. He has trained and coached agile teams in academic and industrial settings, and is a certified ScrumMaster*. He has been active in the software industry as a founder of the New York Software Industry Association, has started several software companies, and has extensive consulting experience in computing and information technology. Dr. Grossman has a B.S. degree in mathematics from Polytechnic University, an M.S. degree in mathematics from New York University Courant Institute, and a Ph.D. degree in computer science from New York University Graduate School of Engineering.