Security Issues and Solutions in Cloud & Grid Computing

Kalyani Alisetty
Asst.Professor, MCA Department, NBN Sinhgad School of Computer Studies, Pune

Dr. K E Balachandrudu
Professor & HOD, Dept. Of CSE, PRRMEC, JNTUH, Hyderabad.

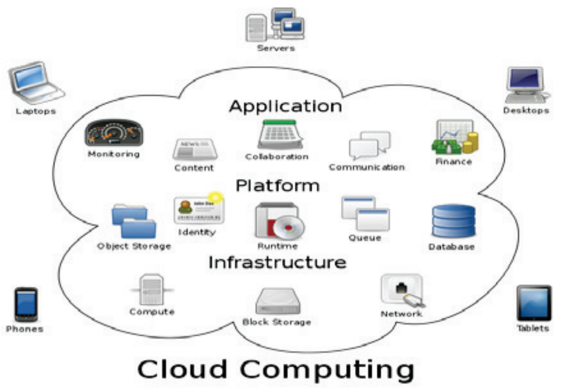
Abstract - Cloud computing is clearly one of today’s most enticing technology areas due, at least in part, to its cost-efficiency and flexibility. Cloud computing is a growing area of concern in the IT security community because cloud architectures are literally popping up all over. However, despite the surge in activity and interest, there are significant, persistent concerns about cloud computing that are impeding momentum and will eventually compromise the vision of cloud computing as a new IT procurement model. In this paper, we characterize the problems and their impact on adoption. In addition, and equally importantly, we describe how the combination of existing research thrusts has the potential to alleviate many of the concerns impeding adoption. In particular, we argue that with continued research advances in trusted computing and computation-supporting encryption, life in the cloud can be advantageous from a business intelligence standpoint over the isolated alternative that is more common today. Public clouds are available from Google.com, Amazon.com, Microsoft, Oracle/Sun, Canonical/Eucalyptus and many other vendors. Private cloud technologies, where the cloud software is loaded on local or in-house server hardware, are available from VMware, Eucalyptus, Citrix, Microsoft, and there are thousands of vendors offering â€œcloud solutionsâ€ of all sorts. A search for â€œprivate cloud hostingâ€ on Google.com produced 581,000 page results. With all of the hyperbole has come a large swell of early-adopters and developers. This paper is concerned with discovery of the vulnerabilities in the landscape of clouds, discovery of security solutions, and finding evidence that early-adopters or developers have grown more concerned with security.

General Terms - Security, Standardization, Legal Aspects.

Keywords - Cloud & Grid Computing, Security, Privacy

I. INTRODUCTION

Today, the 14th largest software company by market capitalization (Salesforce.com) operates almost entirely in the cloud, the top five software companies by sales revenue all have major cloud offerings, and the market as a whole is predicted to grow to $160B by 2011 (source: Merrill Lynch). Yet, despite the trumpeted business and technical advantages of cloud computing, many potential cloud users have yet to join the cloud, and those major corporations that are cloud users are for the most part putting only their less sensitive data in the cloud. Lack of control in the cloud is the major worry. One aspect of control is transparency in the cloud implementation - somewhat contrary to the original promise of cloud computing in which the cloud implementation is not relevant. Transparency is needed for regulatory reasons and to ease concern over the potential for data breaches. Because of today’s perceived lack of control, larger companies are testing the waters with smaller projects and less sensitive data. In short, the potential of the cloud is not being realized.



II. FEAR OF THE CLOUD

What are the “security” concerns that are preventing companies from taking advantage of the cloud? Numerous studies, for example IDC’s 2008 Cloud Services User Survey of IT executives, cite security as the number one challenge for cloud users. In this section we present taxonomy of the “security” concerns. The Cloud Security Alliance’s initial report contains a different sort of taxonomy based on 15 different security domains and the processes that need to be followed in an overall cloud deployment. We categorize the security concerns as:

Traditional security

Availability

Third-party data control

Traditional Security

These concerns involve computer and network intrusions or attacks that will be made possible or at least easier by moving to the cloud. Cloud providers respond to these concerns by arguing that their security measures and processes are more mature and tested than those of the average company. Another argument, made by the Jericho Forum, is: "It could be easier to lock down information if it's administered by a third party rather than in-house, if companies are worried about insider threats… In addition, it may be easier to enforce security via contracts with online services providers than via internal controls."

Concerns in this category include:

TS1. VM-level attacks. Potential vulnerabilities in the hypervisor or VM technology used by cloud vendors are a potential problem in multi-tenant architectures. Vulnerabilities have appeared in VMware, Xen, and Microsoft’s Virtual PC and Virtual Server. Vendors such as Third Brigade mitigate potential VM-level vulnerabilities through monitoring and firewalls.

TS2. Cloud provider vulnerabilities. These could be platform-level, such as an SQL-injection or cross-site scripting vulnerability in salesforce.com. For instance, there have been a couple of recent Google Docs vulnerabilities [26] and [40]. The Google response to one of them is here: [27]. There is nothing new in the nature of these vulnerabilities; only their setting is novel. In fact, IBM has repositioned its Rational AppScan tool, which scans for vulnerabilities in web services as a cloud security service (see Blue Cloud Initiative).

TS3. Phishing cloud provider. Phishes and other social engineers have a new attack vector, as the Sales force phishing incident shows.

TS4. Expanded network attack surface. The cloud user must protect the infrastructure used to connect and interact with the cloud, a task complicated by the cloud being outside the firewall in many cases. For instance, [38] shows an example of how the cloud might attack the machine connecting to it.

TS5. Authentication and Authorization. The enterprise authentication and authorization framework does not naturally extend into the cloud. How does a company meld its existing framework to include cloud resources?

Furthermore, how does an enterprise merge cloud security data (if even available) with its own security metrics and policies?

TS6. Forensics in the cloud. This blog posting on the CLOIDIFIN [12] project summarizes the difficulty of cloud forensic investigations: *“Traditional digital forensic methodologies permit investigators to seize equipment and perform detailed analysis on the media and data recovered. The likelihood therefore, of the data being removed, overwritten, deleted or destroyed by the perpetrator in this case is low. More closely linked to a CC environment would be businesses that own and maintain their own multi-server type infrastructure, though this would be on a far smaller scale in comparison. However, the scale of the cloud and the rate at which data is overwritten is of concern.”*

Availability
These concerns center on critical applications and data being available. Well-publicized incidents of cloud outages include Gmail (one-day outage in mid-October 2008), Amazon S3 (over seven-hour downtime on July 20, 2008), and Flexi Scale.

A1. Uptime. As with the Traditional Security concerns, cloud providers argue that their server uptime compares well with the availability of the cloud user’s own data centers. Besides just services and applications being down, this includes the concern that a third-party cloud would not scale well enough to handle certain applications. SAP’s CEO, Leo Apothecary said: *“There are certain things that you cannot run in the cloud because the cloud would collapse…Don't believe that any utility company is going to run its billing for 50 million consumers in the cloud.”*

A2. Single point of failure. Cloud services are thought of as providing more availability, but perhaps not – there are more single points of failure and attack.

A3. Assurance of computational integrity. Can an enterprise be assured that a cloud provider is faithfully running a hosted application and giving valid results? For example, Stanford's Folding Home project gives the same task to multiple clients to reach a consensus on the correct result.

Third-party data control
The legal implications of data and applications being held by a third party are complex and not well understood. There is also a potential lack of control and transparency when a third party holds the data. Part of the hype of cloud computing is that the cloud can be implementation independent, but in reality regulatory compliance requires transparency into the cloud.

All this is prompting some companies to build private clouds to avoid these issues and yet retain some of the advantages of cloud computing. For example, Benjamin Linder, Scalent System’s CEO, says : *“What I find as CEO of a software company in this space, Scalent Systems, is that most enterprises have a hard time trusting external clouds for their proprietary and high-availability systems. They are instead building internal "clouds", or "utilities" to serve their internal customers in a more controlled way.”*

BL1. Due diligence. If served a subpoena or other legal action, can a cloud user compel the cloud provider to respond in the required time-frame? A related question is the provability of deletion, relevant to an enterprise’s retention policy: How can a cloud user be guaranteed that data has been deleted by the cloud provider?

BL2. Audit ability. Audit difficulty is another side effect of the lack of control in the cloud. Is there sufficient transparency in the operations of the cloud provider for auditing purposes? Currently, this transparency is provided by documentation and manual audits. Information Security Magazine asks: *“How do you perform an on-site audit when you have a distributed and dynamic multi-tenant computing environment spread all over the globe? It may be very difficult to satisfy auditors that your data is properly isolated and cannot be viewed by other customers.”*

A related concern is proper governance of cloud-related activity. It’s easy, perhaps too easy, to start using a cloud service. One popular auditing guideline is the SAS 70, which defines guidelines for auditors to assess internal controls, for instance controls over the processing of sensitive information. SOX and HIPAA are other well-known regulations. US government agencies generally need to follow guidelines from FISMA, NIST, and FIPS. Certain regulations require data and operations to remain in certain geographic locations. Cloud providers are beginning to respond with geo-targeted offerings.

BL3. Contractual obligations. One problem with using another company's infrastructure besides the uncertain alignment of interests is that there might be surprising legal implications. For instance, here is a passage from Amazon's terms of use:

Non-Assertion. During and after the term of the Agreement, with respect to any of the Services that you elect to use, you will not assert, nor will you authorize, assist, or encourage any third party to assert, against us or any of our customers, end users, vendors, business partners (including third party sellers on websites operated by or on behalf of us), licensors, sublicensees or transferees, any patent infringement or other intellectual property infringement claim with respect to such Services.

This could be interpreted as implying that after you use EC2, you cannot file infringement claims against Amazon or its customers suggesting that EC2 itself violates any of your patents. It's not clear whether this non-assert would be upheld by the courts, but any uncertainty is bad for business

BL4. Cloud Provider Espionage. This is the worry of theft of company proprietary information by the cloud provider. For example, Google Gmail and Google Apps are examples of services supported by a private cloud infrastructure. Corporate users of these services are concerned about confidentiality and availability of their data. According to a CNN article :

For Shoukry Tiab, the vice president of IT at Jenny Craig, which uses Postinig and Google Maps, the primary concern is security and confidentiality. "Am I nervous to host corporate information on someone else's server? Yes, even if it's Google."

Note that for consumers, there were initially widespread confidentiality concerns about Gmail, but now those concerns seem to have faded. We believe this is an example of the Privacy Hump [18]:

Early on in the life cycle of a technology, there are many concerns about how these technologies will be used. These concerns are lumped together forming a "privacy hump" that represents a barrier to the acceptance of a potentially intrusive technology…. Over time, however, the concerns fade, especially if the value proposition is strong enough.

Consumers at least seem to have decided that, in this case, the dangers of placing their data in the cloud were outweighed by the value they received.

BL5. Data Lock-in. How does a cloud user avoid lock-in to a particular cloud-computing vendor? The data might itself be locked in a proprietary format, and there are also issues with training and processes. There is also the problem of the cloud user having no control over frequent changes in cloud-based services. Cog head is one example of a cloud platform whose shutdown left customers scrambling to re-write their applications to run on a different platform. Of course, one answer to lock-in is standardization, for instance Go Grid API.

BL6. Transitive nature. Another possible concern is that the contracted cloud provider might itself use subcontractors, over whom the cloud user has even less control, and who also must be trusted. One example is the online storage service called The Linkup, which in turn used an online storage company called Nirvanix. The Linkup shutdown after losing sizeable amounts of customer data, which some say was the fault of Nirvanix. Another example is Carbonate, who is suing its hardware providers for faulty equipment causing loss of customer data.

III. NEW PROBLEMS

In this section we outline new problem areas in security that arise from cloud computing. These problems may only become apparent after the maturation and more widespread adoption of cloud computing as a technology.

Cheap data and data analysis. The rise of cloud computing has created enormous data sets that can be monetized by applications such as advertising. Google, for instance, leverages its cloud infrastructure to collect and analyze consumer data for its advertising network. Collection and analysis of data is now possible cheaply, even for companies lacking Google's resources. What is the impact on privacy of abundant data and cheap data-mining? Because of the cloud, attackers potentially have massive, centralized databases available for analysis and also the raw computing power to mine these databases. For example, Google is essentially doing cheap data mining when it returns search results. How much more privacy did one have before one could be goggled?

Cost-effective defense of availability. Availability also needs to be considered in the context of an adversary whose goals are simply to sabotage activities. Increasingly, such adversaries are becoming realistic as political conflict is taken onto the web, and as the recent cyber attacks on Lithuania confirm. The damages are not only related to the losses of productivity, but extend to losses due to the degraded trust in the infrastructure, and potentially costly backup measures. The cloud computing model encourages single points of failure. It is therefore important to develop methods for sustained availability (in the context of attack), and for recovery from attack. The latter could operate on the basis of minimization of losses, required service levels, or similar measures.

Increased authentication demands. The development of cloud computing may, in the extreme, allow the use of thin clients on the client side. Rather than a license purchased and software installation on the client side, users will authenticate in order to be able to use a cloud application. There are some advantages in such a model, such as making software piracy more difficult and giving the ability to centralize monitoring. It also may help prevent the spread of sensitive data on untrustworthy clients.

Mash-up authorization. As adoption of cloud computing grows, we are likely to see more and more services performing mash-ups of data. This development has potential security implications, both in terms of data leaks, and in terms of the number of sources of data a user may have to pull data from – this, in turn, places requirements on how access is authorized for reasons of usability. While centralized access control may solve many of these problems that may not be possible – or even desirable.

IV. NEW DIRECTIONS

We now describe some elements of our vision. The core issue is that with the advent of the cloud, the cloud provider also has some control of the cloud users' data. We aim to provide tools supporting the current capabilities of the cloud while limiting cloud provider control of data *and* enabling all cloud users to benefit from cloud data through enhanced business intelligence.

Information-centric security issues

In order for enterprises to extend control to data in the cloud, we propose shifting from protecting data from the outside (system and applications which use the data) to protecting data from within. We call this approach of data and information protecting itself *information-centric*, use this terminology differently).

V. SECURITY ISSUES AND SOLUTIONS IN CLOUD COMPUTING

This paper concerns security issues and solutions in cloud computing. Cloud computing is a catch-all phrase that covers virtualized operating systems running on virtual hardware on untold numbers of physical servers. The cloud term has consumed High-Performance Computing (HPC), Grid computing and Utility Computing. The Cloud Security Alliance has adopted the definition developed by NIST; a computing in the cloud is a model exhibiting the following characteristics, on-demand self-service, Broad Network Access, Resource pooling, and Rapid elasticity and Measured service (*Cloud Security Alliance Guidance Version 2.1*, 2009). This is an area that appears to be growing larger and more pervasive as the benefits of cloud architectures become better understood. More organizations start their own cloud projects and more application developers sign on for cloud development as the hyperbole is shaken out and the real parameters of the key technologies are discovered and perfected. The basic areas of cloud vulnerability are similar to the standard issues that surround networking and networked applications. The issues specific to cloud architectures include network control being in in the hands of third parties and and a potential for sensitive data to be available to a much larger selection of third-parties, both on the staff of the cloud providers, and among the other clients of the cloud

VI. SECURITY SOLUTIONS

There are several groups interested in developing standards and security for clouds and cloud security. The Cloud Security Alliance (CSA) is gathering solution providers, non-profits and individuals to enter into discussion about the current and future best practices for information assurance in the cloud (Cloud Security Alliance (CSA) – security best practices for cloud computing, The Cloud Standards web site is collecting and coordinating information about cloud-related standards under development by other groups (CloudsStandards,). The Open Web Application Security Project (OWASP) maintains a top 10 list of vulnerabilities to cloud-based or Software as a Service deployment models which is updated as the threat landscape changes (OWASP). The Open Grid Forum publishes documents to containing security and infrastructural specifications and information for grid computing developers and researchers (Open Grid Forum).

A. Web Application Solutions

The best security solution for web applications is to develop a development framework that shows and teaches a respect for security. Tsai, W., Jin, Z., & Bai, X. (2009) put forth a four-tier framework for web-based development that though interesting, only implies a security facet in the process (Tsai, Jin, & Bai, 2009,). Towards best practices in designing for the cloud by Berre, Roman, Landre, Heuvel, Lennon, & Zeid is a road map toward cloud-centric development (Berre et al., 2009), and the X10 language is one way to achieve better use of the cloud capabilities of massive parallel processing and concurrency .

B. Accessibility Solutions

point out the value of filtering a packet-sniffer output to specific services as an effective way to address security issues shown by anomalous packets directed to specific ports or services.An often-ignored solution to accessibility vulnerabilities is to shut down unused services, keep patches updated, and reduce permissions and access rights of applications and users.

C. Authentication Solutions

Halton and Basta (2007) suggest one way to avoid IP spoofing by using encrypted protocols wherever possible. They also suggest avoiding ARP poisoning by requiring root access to change ARP tables; using static, rather than dynamic ARP tables; or at least make sure changes to the ARP tables are logged.

D. Data Verification, Tampering, Loss and Theft Solutions

Raj, Nathuji, Singh and England (2009) suggest resource isolation to ensure security of data during processing, by isolating the processor caches in virtual machines, and isolating those virtual caches from the Hypervisor cache. Hayes points out that there is no way to know if the cloud providers properly deleted a client's purged data, or whether they saved it for some unknown reason. Would cloud-providers and clients have custody battles over client data?

E. Privacy and Control Solutions

Hayes (2008) points out an interesting wrinkle here, Allowing a third-party service to take custody of personal documents raises awkward questions about control and ownership: If you move to a competing service provider, can you take a data with you? Could you lose access to a documents if you fail to pay a bill?. The issues of privacy and control cannot be solved, but merely assured with tight service-level agreements (SLAs) or by keeping the cloud itself private.

F. Physical access solutions

One simple solution, which Milne (2010) states to be a widely used solution for UK businesses is to simply use in-house private clouds. Nurmi, Wolski, Grzegorczyk, Obertelli, Soman, Youseff, & Zagorodnov show a preview of one of the available home-grown clouds in their (2009) presentation. The Eucalyptus Open-Source Cloud-Computing System (Nurmi et al., 2009).

VII. CONCLUSION

Cloud computing is the most popular notion in IT today; even an academic report [6] from UC Berkeley says “Cloud Computing is likely to have the same impact on software that foundries have had on the hardware industry.” They go on to recommend that “developers would be wise to design their next generation of systems to be deployed into Cloud Computing”. While many of the predictions may be cloud hype, we believe the new IT procurement model offered by cloud computing is here to stay. Whether adoption becomes as prevalent and deep as some forecast will depend largely on overcoming fears of the cloud.

Cloud fears largely stem from the perceived loss of control of sensitive data. Current control measures do not adequately address cloud computing third-party data storage and processing needs. In our vision, we propose to extend control measures from the enterprise into the cloud through the use of Trusted Computing and applied cryptographic techniques. These measures should alleviate much of today’s fear of cloud computing, and, we believe, have the potential to provide demonstrable business intelligence advantages to cloud participation.

II. Our vision also relates to likely problems and abuses arising from a greater reliance on cloud computing, and how to maintain security in the face of such attacks. Namely, the new threats require new constructions to maintain and improve security. Among these are tools to control and understand privacy leaks, perform authentication, and guarantee availability in the face of cloud denial-of-service attacks.

REFERENCES

[1] Security valuation cloud & Grid Environments. https://hpcrd.lbl.gov/HEPCybersecurity/HEP-Sec-Miller-Mar2005.ppt.

[2] Security issues with Google Docs. http://peekay.org/2009/03/26/security-issues-with-google-docs/.

[3] Amazon EC2 Crosses the Atlantic. http://aws.amazon.com/about-aws/whats-new/2008/12/10/amazon-ec2-crosses-the-atlantic/.

[4] Amazon S3 Availability Event: July 20, 2008. http://status.aws.amazon.com/s3- 20080720.html.

[5] Amazon's terms of use. http://aws.amazon.com/agreement

[6] An Information-Centric Approach to Information Security. ttp://virtualization.sys-con.com/node/171199.

[7] Blue Cloud. http://www-03.ibm.com/press/us/en/press release/26642.wss.

[8] EMC, Information-Centric Security. http://www.idc.pt/resources/PPTs/2007/IT&Internet_Security/12.EMC.pdf

[9] Privacy in the Clouds: Risks to Privacy and Confidentiality from Cloud Computing. http://www.worldprivacyforum.org/pdf/WPF_Cloud_Privacy_Report.pdf.

[10] Security Guidance for Critical Areas of Focus in Cloud Computing. http://www.cloudsecurityalliance.org/guidance/csaguide.pdf.

[11] Security issues with Google Docs. http://peekay.org/2009/03/26/security-issues-with-google-docs/

[12] Storm clouds ahead. http://www.networkworld.com/news/2009/030209-soa-cloud.html?page=1.