Controlling Data in the Cloud: Outsourcing Computation without Outsourcing Control

Richard Chow, Philippe Golle, Markus Jakobsson, Elaine Shi, Jessica Staddon
PARC
{rchow,pgolle,mjakobss,eshi,staddon}@parc.com

Ryusuke Masuoka, Jesus Molina
Fujitsu Laboratories of America
{ryusuke.masuoka, jesus.molina}@us.fujitsu.com

ABSTRACT
Cloud computing is clearly one of today's most enticing technology areas due, at least in part, to its cost-efficiency and flexibility. 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.

Categories and Subject Descriptors
K.6.5 [**Management of Computing and Information Systems**]: Security and Protection (D.4.6, K.4.2)

General Terms
Security, Standardization, Legal Aspects.

Keywords
Cloud computing, security, privacy

1. 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.

CCSW'09, November 13, 2009, Chicago, Illinois, USA.

When thinking about solutions to cloud computing's adoption problem, it is important to realize that many of the issues are essentially old problems in a new setting, although they may be more acute. For example, corporate partnerships and offshore outsourcing involve similar trust and regulatory issues. Similarly, open source software enables IT departments to quickly build and deploy applications, but at the cost of control and governance. Finally, virtual machine attacks and Web service vulnerabilities existed long before cloud computing became fashionable. Indeed, this very overlap is reason for optimism; many of these "cloud problems" have long been studied and the foundations for solutions exist.

In our vision, integrity of the cloud infrastructure is ensured through the use of Trusted Computing. In addition, we advocate the seamless extension of control from the enterprise into the cloud through the powerful combination of high-assurance remote server integrity, and cryptographic protocols supporting computation on ciphertext. With our approach, content is protected in a manner consistent with policies, whether in the enterprise or the cloud. Yet, because the protection mechanisms support computation, it is possible for all cloud participants to mutually benefit from the cloud data *in a controlled manner*. Hence, there are business intelligence advantages derived from operating in the cloud that simply don't exist otherwise. We believe that the ability to get smarter through use of the cloud is the key differentiator that will sufficiently alleviate privacy fears to ensure widespread adoption.

Organization. In Section 2, we give an overview of existing cloud computing concerns. We explore in more detail what the concerns of cloud users are, that is, what might be causing fear of the cloud.

In Section 3 we describe new problem areas in security that we see arising from the trend towards cloud computing. We present evidence that these will become real problems after the maturation and more widespread adoption of cloud computing as a technology.

Finally, in Section 4 we present our vision, some broad strategies that might be used to mitigate some of the concerns outlined in Sections 2 and 3.

2. 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 [29] of IT executives, cite security as the number one challenge for cloud users.

In this section we present a taxonomy of the "security" concerns. The Cloud Security Alliance's initial report [39] 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 [16], 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 [48], Xen [51], and Microsoft's Virtual PC and Virtual Server [47]. Vendors such as Third Brigade [46] 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 [8]).

TS3. Phishing cloud provider. Phishers and other social engineers have a new attack vector, as the Salesforce phishing incident [37] 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 [20]), Amazon S3 (over seven-hour downtime on July 20, 2008 [2]), and FlexiScale (18-hour outage on October 31, 2008 [22]).

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 Apotheker 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."* (11/24/08, searchSAP.com)

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 [15]: *"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. Auditability. 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 [28]: *"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 [44].

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 [1].

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 [3]:
10.4. 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 [50]:
For Shoukry Tiab, the vice president of IT at Jenny Craig, which uses Postini 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 (see [35]), 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 (see [14]). Coghead [13] 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 GoGrid API [25].

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 [32]. Another example is Carbonite [30], who is suing its hardware providers for faulty equipment causing loss of customer data.

3. 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 Googled?

Because of privacy concerns, enterprises running clouds collecting data have felt increasing pressure to anonymize their data. EPIC has called for Gmail, Google Docs, Google Calendar, and the company's other Web applications to be shut down until appropriate privacy guards are in place [23]. Google and Yahoo!, because of pressure from privacy advocates, now have an 18 month retention policy for their search data, after which it will be anonymized. This means that some identifying data will be removed such as IP addresses and cookie information. The anonymized data is retained though, to support the continual testing of their algorithms. Another reason to anonymize data is to share data with other parties. These may be to support research (e.g., the AOL incident [5]) or to subcontract out data mining on the data (e.g., the Netflix data set [34]).

We note that anonymizing data is a difficult problem. For example, in [33] the Netflix data set was partially de-anonymized, and in [45] the then-Governor of Massachusetts was identified as a patient of Massachusetts General Hospital from an anonymized list of discharged patients. Tools are needed for effective anonymization, which will increase in importance as clouds proliferate and more data is collected that needs to be analyzed safely or shared.

An example of indirect data-mining that might be performed by a cloud provider is to note transactional and relationship information (see World Privacy Forum Report [36]). For example, the sharing of information by two companies may signal a merger is under consideration.

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 [31]. 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.

Thin clients result in a number of opportunities related to security, including the paradigm in which typical users do not have to worry about the risks of any actions – their security is managed by the cloud, which maintains the software they run. This architecture stimulates mobility of users, but increases the need to address authentication in a secure manner. In addition, the movement towards increased hosting of data and applications in the cloud and lesser reliance on specific user machines is likely to increase the threat of phishing and other abusive technologies aimed at stealing access credentials, or otherwise derive them, e.g., by brute force methods.

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.

One example in this area is provided by Facebook. Facebook users upload both sensitive and non-sensitive data. This data is both utilized by Facebook to present the data to other users, and also utilized by third party applications that are run by the platform. These applications are typically not verified by Facebook. Hence, there is a drive to create malicious applications that run in Facebook's cloud to steal sensitive data, e.g., see [21].

4. 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

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* (note that [4], [17], [19] use this terminology differently). This self-protection requires intelligence be put in the data itself. Data needs to be self-describing and defending, regardless of its environment. Data needs to be encrypted and packaged with a usage policy. When accessed, data should consult its policy and attempt to re-create a secure environment using virtualization and reveal itself only if the environment is verified as trustworthy (using Trusted Computing). Information-centric security is a natural extension of the trend toward finer, stronger, and more usable data protection.

High-Assurance Remote Server Attestation

We have noted that lack of transparency is discouraging businesses from moving their data to the cloud. Data owners wish to audit how their data is being handled at the cloud, and in particular, ensure that their data is not being abused or leaked, or at least have an unalterable audit trail when it does happen.

Currently customers must be satisfied with cloud providers using manual auditing procedures like SAS-70.

A promising approach to address this problem is based on Trusted Computing. Imagine a trusted monitor installed at the cloud server that can monitor or audit the operations of the cloud server. The trusted monitor can provide "proofs of compliance" to the data owner, stating that certain access policies have not been violated. To ensure integrity of the monitor, Trusted Computing also allows secure bootstrapping of this monitor to run beside (and securely isolated from) the operating system and applications. The monitor can enforce access control policies and perform monitoring/auditing tasks. To produce a "proof of compliance", the code of the monitor is signed, as well as a "statement of compliance" produced by the monitor. When the data owner receives this proof of compliance, it can verify that the correct monitor code is run, and that the cloud server has complied with access control policies.

Privacy-Enhanced Business Intelligence

A different approach to retaining control of data is to require the encryption of all cloud data. The problem is that encryption limits data use. In particular searching and indexing the data becomes problematic. For example, if data is stored in clear-text, one can efficiently search for a document by specifying a keyword. This is impossible to do with traditional, randomized encryption schemes. State-of-the-art cryptography may offer new tools to solve these problems. Cryptographers have recently invented versatile encryption schemes that allow operation and computation on the ciphertext. For example, searchable encryption (also referred to as predicate encryption; see [43], [9], [42], [41], and [10]) allows the data owner to compute a capability from his secret key. A capability encodes a search query, and the cloud can use this capability to decide which documents match the search query, without learning any additional information. Other cryptographic primitives such as homomorphic encryption [24] and Private Information Retrieval (PIR) [11] perform computations on encrypted data without decrypting. As these cryptographic techniques mature, they may open up new possibilities for cloud computing security

While in many cases more research is needed to make these cryptographic tools sufficiently practical for the cloud, we believe they present the best opportunity for a clear differentiator for cloud computing since these protocols can enable cloud users to benefit from one another's data in a controlled manner. In particular, even encrypted data can enable anomaly detection that is valuable from a business intelligence standpoint. For example, a cloud payroll service might provide, with the agreement of participants, aggregate data about payroll execution time that allows users to identify inefficiencies in their own processes. Taking the vision even further, if the cloud service provider is empowered with some ability to search the encrypted data, the proliferation of cloud data can potentially enable better insider threat detection (e.g. by detecting user activities outside of the norm) and better data loss prevention (DLP) (e.g. through detecting anomalous content).

Apart from ensuring privacy, applied cryptography may also offer tools to address other security problems related to cloud computing. For example, in proofs of retrievability (e.g., [7], [49]) the storage server can show a compact proof that it is correctly storing all of the client's data.

5. 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's 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.

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.

6. REFERENCES

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

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

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

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

[5] AOL apologizes for release of user search data. http://news.cnet.com/2100-1030_3-6102793.html.

[6] Armbrust, M., Fox, A., Griffith, R. et al. Above the Clouds: A Berkeley View of Cloud Computing. UCB/EECS-2009-28, EECS Department, University of California, Berkeley, 2009.

[7] Ateniese, G., Burns, R., Curtmola, R., Herring, J., Kissner, L., Z., Peterson, and Song, D. Provable Data Possession at Untrusted Stores. In CCS. 2007.

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

[9] Boneh, B., Di Crescenzo, G., Ostrovsky, R., and Persiano, G. Public Key Encryption with Keyword Search. In EUROCRYPT. 2004.

[10] Boneh, D and Waters, B. Conjunctive, Subset, and Range Queries on Encrypted Data. In The Fourth Theory of Cryptography Conference (TCC 2007), 2007.

[11] Chor, B., Kushilevitz, E., Goldreich, O., and Sudan, M. Private Information Retrieval. J. ACM, 45, 6 (1998), 965-981.

[12] CLOIDIFIN. http://community.zdnet.co.uk/blog/0,1000000567,20006251 96b,00.htm?new_comment.

[13] Cloud Bursts as Coghead Calls It Quits. http://blogs.zdnet.com/collaboration/?p=349.

[14] Cloud computing: Don't get caught without an exit strategy. http://www.computerworld.com/action/article.do?command =viewArticleBasic&articleId=9128665&source=NLT_AM.

[15] Disaster-Proofing The Cloud. http://www.forbes.com/2008/11/24/cio-cloud-disaster-tech-cio-cx_dw_1125cloud.html.

[16] Don't cloud your vision. http://www.ft.com/cms/s/0/303680a6-bf51-11dd-ae63-0000779fd18c.html?nclick_check=1.

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

[18] End-User Privacy in Human–Computer Interaction. http://www.cs.cmu.edu/~jasonh/publications/fnt-end-user-privacy-in-human-computer-interaction-final.pdf.

[19] ESG White Paper, The Information-Centric Security Architecture. http://japan.emc.com/collateral/analyst-reports/emc-white-paper-v4-4-21-2006.pdf.

[20] Extended Gmail outage hits Apps admins. http://www.computerworld.com/action/article.do?command =viewArticleBasic&articleId=9117322.

[21] Facebook users suffer viral surge. http://news.bbc.co.uk/2/hi/technology/7918839.stm.

[22] FlexiScale Suffers 18-Hour Outage. http://www.thewhir.com/web-hosting-news/103108_FlexiScale_Suffers_18_Hour_Outage.

[23] FTC questions cloud-computing security. http://news.cnet.com/8301-13578_3-10198577-38.html?part=rss&subj=news&tag=2547-1_3-0-20.

[24] Gentry, C. Fully Homomorphic Encryption Using Ideal Lattices. In STOC. 2009.

[25] GoGrid API. http://www.gogrid.com/company/press-releases/gogrid-moves-api-specification-to-creativecommons.php.

[26] Google Docs Glitch Exposes Private Files. http://www.pcworld.com/article/160927/google_docs_glitch _exposes_private_files.html.

[27] Google's response to Google Docs concerns. http://googledocs.blogspot.com/2009/03/just-to-clarify.html.

[28] How to Secure Cloud Computing. http://searchsecurity.techtarget.com/magOnline/0,sid14_gci1 349550,00.html.

[29] IT Cloud Services User Survey, pt.2: Top Benefits & Challenges. http://blogs.idc.com/ie/?p=210.

[30] Latest cloud storage hiccups prompts data security questions. http://www.computerworld.com/action/article.do?command =viewArticleBasic&articleId=9130682&source=NLT_PM.

[31] Lithuania Weathers Cyber Attack, Braces for Round 2. http://blog.washingtonpost.com/securityfix/2008/07/lithuania_weathers_cyber_attac_1.html.

[32] Loss of customer data spurs closure of online storage service 'The Linkup'. http://www.networkworld.com/news/2008/081108-linkup-failure.html?page=1.

[33] Narayanan, A. and Shmatikov, V. Robust De-anonymization of Large Sparse Datasets. In IEEE Symposium on Security and Privacy. IEEE Computer Society, 2008.

[34] Netflix Prize. http://www.netflixprize.com/.

[35] Organizations urge Google to suspend GMail. http://www.privacyrights.org/ar/GmailLetter.htm.

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

[37] Salesforce.com Warns Customers of Phishing Scam. http://www.pcworld.com/businesscenter/article/139353/salesforcecom_warns_customers_of_phishing_scam.html.

[38] Security Evaluation of Grid Environments. https://hpcrd.lbl.gov/HEPCybersecurity/HEP-Sec-Miller-Mar2005.ppt.

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

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

[41] Shen, E., Shi, E., and Waters, B. Predicate Privacy in Encryption Systems. In TCC. 2009.

[42] Shi, E. Bethencourt, J., Chan, H., Song, D., and Perrig, A. Multi-Dimensional Range Query over Encrypted Data. In IEEE Symposium on Security and Privacy. 2007.

[43] Song, D., Wagner, D., and Perrig, A. Practical Techniques for Searches on Encrypted Data. In IEEE Symposium on Research in Security and Privacy. 2000.

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

[45] Sweeney, L. Weaving technology and policy together. J. of Law, Medicine and Ethics, 25, 2-3 (1997).

[46] Third Brigade. http://www.thirdbrigade.com.

[47] VirtualPC vulnerability. http://www.microsoft.com/technet/security/bulletin/ms07-049.mspx.

[48] VMWare vulnerability. http://securitytracker.com/alerts/2008/Feb/1019493.html.

[49] Waters, B. and Shacham, H. Compact Proofs of Retrievability. In ASIACRYPT. 2008.

[50] Why Google Apps is not being adopted. http://money.cnn.com/2008/08/19/technology/google_apps.fortune/index.htm.

[51] Xen vulnerability. http://secunia.com/advisories/26986/.