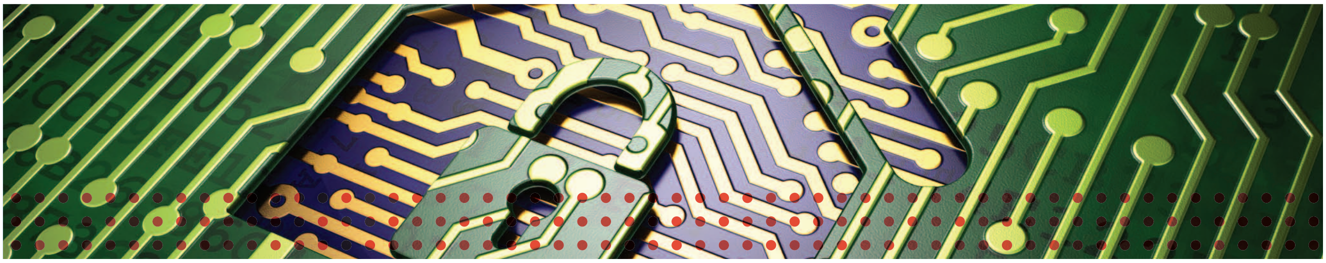

Security and Privacy in Cloud Computing

Zahir Tari, RMIT University

Significant research and development efforts in both industry and academia aim to improve the cloud's security and privacy. The author discusses related challenges, opportunities, and solutions.

The cloud has fundamentally changed the landscape of computing, storage, and communication infrastructures and services. With strong interest and investment from industry and government, the cloud is being increasingly patronized by both organizations and individuals. From the cloud provider's perspective, cloud computing's main benefits include resource consolidation, uniform management, and cost-effective operation; for the cloud user, benefits include on-demand capacity, low cost of ownership, and flexible pricing. However, the features that bring such benefits, such as sharing and consolidation, also introduce potential security and privacy problems. Security and privacy issues resulting from the illegal and unethical use of information, and causing disclosure of confidential information, can significantly hinder user acceptance of cloud-based services. Recent surveys support this observation, indicating that security and privacy concerns prevent many customers from adopting cloud computing services and platforms.

In response to such concerns, significant research and development efforts in both industry and academia have sought to improve the cloud's security and privacy. Here I give a quick (and incomplete) overview of new challenges, opportunities, and solutions in this area, with the purpose of stimulating more in-depth and extensive discussion on related problems in upcoming issues of this magazine.

Identifying New Threats and Vulnerabilities

An essential task in cloud security and privacy research is to identify new threats and vulnerabilities that are specific to cloud platforms and services. Several recent reports have explored such vulnerabilities. For example, in 2009, researchers from the University of California, San Diego, and the Massachusetts Institute of Technology demonstrated leakage attacks against Amazon's Elastic Compute Cloud (EC2) virtual machines (VMs).[1] More specifically, the researchers showed that it's possible to probe and infer the overall placement of VMs in the EC2 infrastructure. Furthermore, an attacker can launch a malicious EC2 instance and then determine whether that instance is physically colocated with a targeted (victim) instance. When the attacker's instance is successfully colocated with the

2325-6095/14/$31.00 © 2014 IEEE

victim, it can launch a side-channel attack by monitoring the status of shared physical resources such as level-1 and level-2 caches, and thus infer the victim's computation and I/O activities.

A follow-up study showed that it's possible to extract private keys via the cross-VM side channel in a lab environment.[2] In another study, researchers from the College of William and Mary reported that side-channel attacks aren't just a potential risk, but a realistic threat.[3] They created a covert channel via another shared resource (the memory bus) that had a level of reliability and throughput of more than 100 bps in both lab and EC2 environments.

These risks represent a small subset of known cloud-specific vulnerabilities and threats. However, they motivate us to think further about new adversary models, trust relations, and risk factors relative to cloud computing stakeholders. In the examples, the cloud provider isn't trusted because of its resource sharing and VM consolidation practices. Hence, the cloud provider doesn't provide a desirable level of isolation and protection between tenants in the cloud, allowing them to attack each other.

Protecting Virtual Infrastructures

Virtual infrastructures are infrastructure-level (virtual) entities, such as VMs and virtual networks, created in the cloud on behalf of users. Side-channel attacks target these virtual infrastructures. Researchers have proposed several solutions to defend against cross-VM side-channel attacks. Düppel, for example, aims to disrupt cache-based side channels. In this self-defensive approach, the target VM's guest operating system injects cache access noise (that is, flushes) so the collocated attack VM can't infer cache access patterns.[4] This solution doesn't require modifying the underlying hypervisor or cloud platform. To defend against memory bus-based side channels, a simple and practical approach is to prevent a VM from locking the memory bus and let the hypervisor emulate the execution of atomic instructions that would otherwise require memory bus locking.[5]

Other attacks against virtual infrastructures include malware attacks against tenant VMs. The cloud presents a new opportunity to defend against these attacks. More specifically, the cloud provides a uniform and tamper-resistant platform to deploy system monitoring and antimalware functions. The uniformity is reflected by the cloud provider's consistent installation, configuration, and update of antimalware services for all hosted tenants. It's tamper resistant because monitoring and detection of malware attacks can be performed from outside the hosted VMs, either by the underlying hypervisor or by the more privileged management domain (for example, Domain 0 of Xen). In CloudAV, a production-quality system that reflects the antivirus-as-a-service idea, a group of in-cloud antivirus engines analyzes suspicious files submitted by agents running in client machines (including VMs) and collectively detects malware in them.[6] VMwatcher, a virtualization-based malware-monitoring and detection system, moves commodity, off-the-shelf antimalware software from the inside to the outside of each tenant VM.[7] This way, the antimalware software is out of the malware's reach, preventing the malware from detecting, disabling, or tampering with it. Malware targeting a tenant VM—at either the user or kernel level—can be detected and prevented using such an "out-of-the-box" antimalware service.

A networked virtual infrastructure can consist of multiple VMs connected by a virtual network. With the rapid advances in software-defined networking (SDN), the cloud increasingly supports such networked virtual infrastructures. SDN decouples the control and data-forwarding functions of a physical networked infrastructure, such as a datacenter network. The SDN control plane performs control functions such as routing, naming, and firewall policy enforcement, and the SDN data plane follows the control plane's decisions to forward packets belonging to different flows. Such decoupling makes it easy to optimize the control and data planes without them affecting each other. However, the SDN paradigm raises security issues. Researchers have reported that it's possible to launch attacks against the SDN architecture, incurring excessive workload and resource consumption to both the control and the data plane.[8] Although researchers are developing defenses against such attacks, we need more generic, scalable solutions that make the SDN architecture secure, robust, and scalable, which would support virtual infrastructure hosting in the cloud.

Protecting Outsourced Computation and Services

Many organizations have been increasingly outsourcing services and computation jobs to the cloud. A client that outsources a computation job must verify the correctness of the result returned from the cloud, without incurring significant overhead at its local infrastructure—the extreme being to execute the job locally, which would nullify the benefit of outsourced job execution. Such verifiability is important to achieving cloud service trustworthiness and hence has become a topic of active research. Encouragingly, researchers have in recent years developed techniques and real systems to bring the vision of a "verifiable cloud service" closer to reality. For example, the Pantry system composes and outsources proof-based verifiable computation with untrusted storage.[9] It achieves theoretically sound verifiability of computation for realistic cloud applications, such as MapReduce jobs and simple MySQL queries.

In addition to computation outsourcing, the cloud can support network service/function outsourcing. Example network functions include traffic filtering, transcoding, firewall policy enforcement, and network-level intrusion detection. Seyed Kaveh Fayazbakhsh and his colleagues noted that, similar to computation outsourcing, a major challenge is to verify (at end points of network connections) that the "middle boxes" in the cloud correctly execute outsourced network functions with satisfactory performance.[10] They also proposed a framework for verifiable network function outsourcing (vNFO) that aims to achieve verifiability, efficiency, and accountability of outsourced network functions. Such a framework will pave the way for deploying trusted network middle boxes, in addition to end points (that is, VMs), in the cloud, enriching the cloud ecosystem.

Protecting User Data

User data is another important cloud "citizen." To protect user data in the cloud, a key challenge is to guarantee the confidentiality of privacy-sensitive data while it's stored and processed in the cloud. This problem assumes a somewhat different trust model, in which the cloud is not fully trusted because of operator errors or software vulnerabilities. As a result, the cloud provider shouldn't be able to see unencrypted or decrypted sensitive data during the data's residence in the cloud. (In other words, sensitive data should remain encrypted while in the cloud.) However, such a requirement can limit the usability of (encrypted) data when a cloud application processes it. Fortunately, researchers at the University of California, Santa Barbara, observed that many cloud applications can process encrypted data without affecting the correctness of the data execution. These researchers proposed Silverline, which identifies data that the application can properly process in encrypted form.[11] Such data will remain encrypted and hence maintain its confidentiality to the cloud provider. The cloud user will perform data decryption locally once the encrypted data is returned from the cloud as application output.

In-cloud data confidentiality poses even greater challenges. For example, even if the application data is encrypted, the access patterns exhibited by the corresponding applications can reveal sensitive information about the nature of the original data, weakening the data's confidentiality. Hence a challenge is to achieve confidentiality of data access patterns in the cloud—a problem called *oblivious RAM* (ORAM). Recently, researchers reported a breakthrough in achieving both practical and theoretically sound ORAM.[12] The solution, called Path ORAM, is elegant by design and efficient in practice.[12] In fact, Path ORAM has been implemented as part of a processor prototype called Phantom,[13] which achieves realistic performance for real-world applications. This is a significant step toward ultimate deployment of ORAM-enabled machines for sensitive data processing in the cloud.

Securing Big Data Storage and Access Control

In the recent past, more research has focused on cloud-based big data applications. Many consider the cloud to be the most promising platform for hosting, collaborating on, and sharing big data. The challenge is to secure the storage and access to this data to preserve its integrity, confidentiality, authenticity, and nonrepudiation while facilitating availability.

Interesting solutions to increase the accountability of data sharing have been proposed for cloud-based distributed systems. Smitha Sundareswaran and his colleagues, for example, proposed a decentralized accountability framework with logging capabilities using the programmable capabilities of Java Archive files.[14] The advent of many types of big data, such as electronic health records and sensor data, have spurred research on secure access and sharing with greater accountability. Recently, researchers have proposed solutions for increasing accountability and secure access to cloud-based health data,[15] as well as robust cryptographic access control methods to increase the storage security of privacy-sensitive big data. Guojun Wang and his colleagues proposed hierarchical attribute-based cryptography to facilitate secure access to users in large-scale cloud storage systems.[16] More recently, researchers have designed more advanced solutions (for example, homomorphic cryptography[17]) for secure cloud-based storage systems to facilitate secure distributed access.

Given emerging trends in big data, we need more research on efficient, scalable, and accountable privacy-preserving mechanisms that can address application-specific requirements.

Call for Contributions

The magazine welcomes articles that discuss new challenges, opportunities, and solutions in the area of cloud security and privacy—in particular, articles that relate to data, storage, computation, and communication. Enabling techniques include cryptography, virtualization, data management and analytics, software-defined networking, fault tolerance and recovery, and forensics. I'd like to hear from practitioners about their lessons and experience in developing, deploying, and using cloud security and privacy solutions and services. I also welcome reports from academia on cutting-edge research and development, new vulnerabilities and challenges, and new or even controversial ideas and visions.

References

1. T. Ristenpart et al., "Hey, You, Get Off of My Cloud: Exploring Information Leakage in Third-Party Compute Clouds," *Proc. ACM Conf. Computer and Comm. Security* (CCS 09), 2009, pp. 199–212.
2. Y. Zhang et al., "Cross-VM Side Channels and Their Use to Extract Private Keys," *Proc. 19th ACM Conf. Computer and Comm. Security* (CCS 12), 2012, pp. 305–316.
3. Z. Wu, Z. Xu, and H. Wang, "Whispers in the Hyper-space: High-speed Covert Channel Attacks in the Cloud," *Proc. Usenix Security Symp.*, 2012.
4. Y. Zhang and M.K. Reiter, "Düppel: Retrofitting Commodity Operating Systems to Mitigate Cache Side Channels in the Cloud," *Proc. 20th ACM Conf. Computer and Comm. Security* (CCS 13), 2013.
5. B. Saltaformaggio, D. Xu, and X. Zhang, "Bus-Monitor: A Hypervisor-Based Solution for Memory Bus Covert Channels," *Proc. 6th European Workshop on Systems Security* (EuroSec 13), 2013.
6. J. Oberheide, E. Cooke, and F. Jahanian, "CloudAV: N-Version Antivirus in the Network Cloud," *Proc. 17th Usenix Security Symp.*, 2008, pp. 91–106.
7. X. Jiang, X. Wang, and D. Xu, "Stealthy Malware Detection Through VMM-Based 'Out-of-the-Box' Semantic View Reconstruction," *Proc. ACM Conf. Computer and Comm. Security* (CCS 07), 2007, pp. 128–138.
8. S. Shin and G. Gu, "Attacking Software-Defined Networks: A First Feasibility Study," *Proc. ACM SIGCOMM Workshop Hot Topics in Software Defined Networking* (HotSDN 13), 2013, pp. 165–166.
9. B. Braun et al., "Verifying Computations with State," *Proc. 24th ACM Symp. Operating Systems Principles* (SOSP 13), 2013, pp. 341–357.
10. S.K. Fayazbakhsh, M.K. Reiter, and V. Sekar, "Verifiable Network Function Outsourcing: Requirements, Challenges, and Roadmap," *Proc. ACM Workshop Hot Topics in Middleboxes and Network Function Virtualization* (HotMiddlebox 13), 2013, pp. 25–30.
11. K.P.N. Puttaswamy, C. Kruegel, and B.Y. Zhao, "Silverline: Toward Data Confidentiality in Storage-Intensive Cloud Applications," *Proc. 2nd ACM Symp. Cloud Computing* (SoCC 11), 2011, article 10.
12. E. Stefanov et al., "Path ORAM: An Extremely Simple Oblivious RAM Protocol," *Proc. ACM Conf. Computer and Comm. Security* (CCS 2013), 2013, pp. 299–310.
13. M. Maas et al., "PHANTOM: Practical Oblivious Computation in a Secure Processor," *Proc. ACM Conf. Computer and Comm. Security* (CCS 13), 2013, pp. 311–324.
14. S. Sundareswaran, A.C. Squicciarini, and D. Lin, "Ensuring Distributed Accountability for Data Sharing in the Cloud," *IEEE Trans. Dependable and Secure Computing*, vol. 9, no. 4, 2012, pp. 556–568.
15. Y. Tong et al., "Cloud-Assisted Mobile-Access of Health Data with Privacy and Auditability," IEEE J. *Biomedical and Health Informatics*, vol. 18, no. 2, 2014, pp. 419–429.
16. G. Wang, Q. Liu, and J. Wu, "Hierarchical Attribute-Based Encryption for Fine-Grained Access Control in Cloud Storage Services," *Proc. 17th ACM Conf. Computer and Comm. Security*, 2010, pp. 735–737.
17. W. Lu, A.L. Varna, and M. Wu, "Confidentiality-Preserving Image Search: A Comparative Study between Homomorphic Encryption and Distance-Preserving Randomization," *IEEE Access*, vol. 2, 2014, pp. 125–141.

ZAHIR TARI *is a full professor of distributed systems at RMIT University, Australia. His research interests include system performance (for example, Web servers, P2P, and cloud computing) and system security (for example, SCADA and cloud). Tari received a PhD in computer science from the University of Grenoble, France. In addition to serving on the* IEEE Cloud Computing *editorial board, he's an associate editor of* IEEE Transactions on Computers *and* IEEE Transactions on Parallel and Distributed Systems. *Contact him at zahir.tari@rmit.edu.au.*