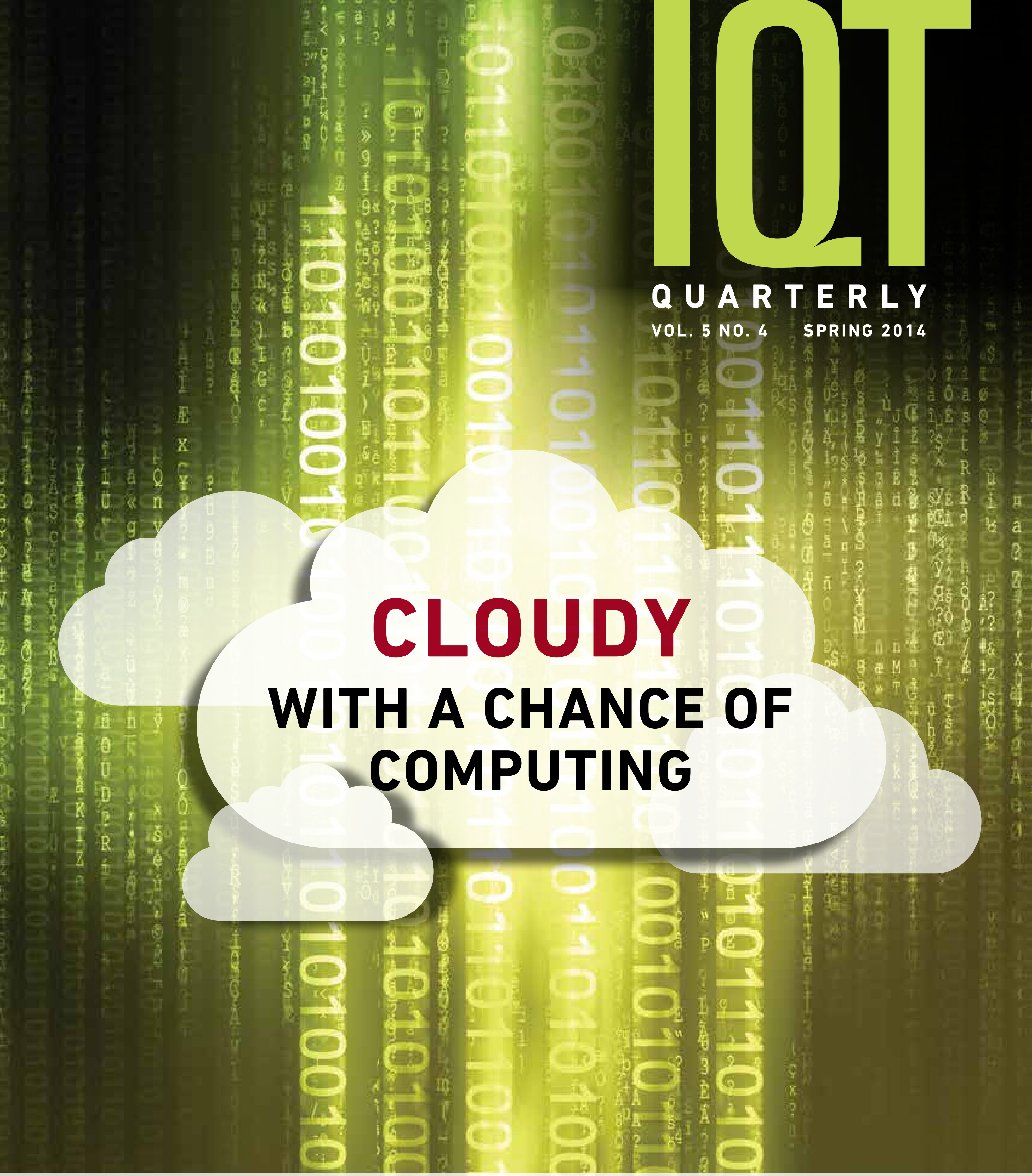
IQT QUARTERLY

VOL. 5 NO. 4 SPRING 2014



CLOUDY

WITH A CHANCE OF COMPUTING

IQT®
IN·Q·TEL

IQT Quarterly is a publication of In-Q-Tel, Inc., the strategic investment firm that serves as a bridge between the U.S. Intelligence Community and venture-backed startup firms on the leading edge of technological innovation. *IQT Quarterly* advances the situational awareness component of the IQT mission, serving as a platform to debut, discuss, and debate issues of innovation in the areas of overlap between commercial potential and U.S. Intelligence Community needs. For comments or questions regarding IQT or this document, please visit www.iqt.org, write to iqtquarterly@iqt.org, or call 703-248-3000. The views expressed are those of the authors in their personal capacities and do not necessarily reflect the opinion of IQT, their employers, or the Government.

EDITORIAL
IQT Quarterly, published by In-Q-Tel, Inc.
Editor-in-Chief: Adam Dove
Theme Editor: Greg Shipley
Contributing Editors: Brittany Carambio and Emma Shepard
Managing Editor: Lisa L. Bader

Design by Lomangino Studio LLC
Printed in the United States of America

IQT
QUARTERLY

Identify. Adapt. Deliver.™

TABLE OF CONTENTS

ON OUR RADAR

CLOUD COMPUTING IS HERE; NOW WHAT?

By Greg Shipley

In the fall of 2011, the *IQT Quarterly* tackled some of the issues surrounding the promise and hype of cloud computing. While the discussion is far from over, three years later we find ourselves with a bit more insight into some of the key questions. How secure is cloud-based infrastructure? Answer: about as secure (and insecure) as the IT infrastructure that preceded it. Does using cloud-based services save money? Answer: it depends — in some cases, yes, but in many others, no. Does cloud computing have an impact on the IT labor force? Answer: definitely, but that impact is often more about change, and less about reduction.

Most debates regarding the long-term viability of cloud computing are over; the movement has become "the new normal" in corporate America. We believe that the changes related to the rise of cloud computing will continue to have a profound impact on the future of enterprise computing. However, both questions and misconceptions remain, and much work lies ahead.

The Evolution of Enterprise IT

The story is all too familiar: a business unit or mission group has specific technology needs, becomes frustrated with its enterprise IT organization, and decides to use some of its coveted budget dollars to stand up pieces of its own IT infrastructure — independent of enterprise IT. While not often discussed, these secondary IT teams, or "shadow IT," exist in both government and corporate contexts, and in some cases have been around as long as the IT departments themselves. Their presence is not new; however, there is a rising trend in corporate America of using cloud-based service providers as a third option — another path around enterprise IT organizations. The trend raises an interesting question: have cloud-based service providers simply become the latest incarnation of shadow IT?

There's little doubt that a more "frictionless" IT environment is the end state that developers, IT operations personnel, and end users alike are all seeking. Who doesn't dream about resources on-demand, provisioning within minutes, and lower barriers to use? This is the lure of a cloud-enabled world and the reason Amazon Web Services' revenue is now measured in billions. These are achievable goals, but understanding the broader story is essential to executing against this vision.

For the technology portion of the tale, much of what drives the largest cloud providers remains a blend of traditional approaches and technology combined with some modern and significant shifts. For example, technology vendors like Cisco and Juniper continue to supply IT teams with significant quantities of network infrastructure. However, some of the largest providers are now embracing software-defined networking (SDN) concepts running on top of more generic "white box" switches; both cost and functionality are driving this change. Dell, HP, and IBM may still be selling thousands of servers into data centers, but companies like Facebook — now one of the largest purchasers of server hardware on the planet — claim to be using 100 percent Open Compute-based hardware. Cost reduction was a driver here, too: Facebook credits its Open Compute initiative with saving the company over a billion dollars in the last three years.

VMware remains the dominant virtualization player in the traditional enterprise infrastructure space, but the growing popularity and momentum behind the open source OpenStack project is undeniable. Configuration management and orchestration technologies from projects like Ansible, Chef, Puppet Labs, and SaltStack are far more prevalent in the cloud space than equivalents from the larger, legacy software vendors. These technologies also influence how new applications are developed, deployed, and scaled as the lines between developers and system administrators continue to become less defined. Finally, traditional relational database technologies are still powering thousands of cloud applications, but NoSQL-based counterparts offering graph and document-based alternatives continue to gain popularity. So does the use of object-based storage systems (e.g., Amazon S3, OpenStack Swift, Cleversafe) by a growing group within the development community. These significant shifts in technology usage will have lasting effects.

Cloud-enabled IT teams are facing new considerations and skill set requirements. For example, understanding resource constraints, and specifically bandwidth usage requirements, is even more essential. If the dynamic capabilities of cloud-enabled applications are realized, applications and workloads have the potential to be resized or moved. In a traditional model where static resources (e.g., servers) reside in a single physical data center with relatively static network connections, there are a number of variables that can affect performance. In a cloud-enabled world, that number of variables increases substantially. In effect, moving a 500 GB image file between two systems within the same data center is one thing, moving it between two data centers is quite another, and having it moved automatically is even more complicated. In some cases, having cloud-enabled applications will actually *increase* complexity; a counter-intuitive notion to some, but a harsh reality for those who are already living in the world of dynamic resource allocation.

The need for greater collaboration between facilities teams and IT personnel is another area in which organizations are becoming more mindful. For example, both Facebook and Google realized early on that their power, space, and cooling footprints would become increasingly relevant to their total delivery costs. This realization resulted in a re-engineered approach to data centers, and has saved them billions of dollars. Yet even today, most large organizations lack teams populated by facilities managers, building engineers, and IT personnel; the groups certainly communicate, but the disciplines remain far from integrated.

But perhaps the most relevant human component to the story will be the acute and growing need for cloud technology advisors, translators, and educators. Many IT personnel will need to transform into cloud ambassadors: advisors who help consult about when, where, and how services should be migrated or delivered. It's not just greater understanding of the technical "big picture" that will be important — it is the ability to help teams make the best choices and effectively serve as service brokers and enablers.

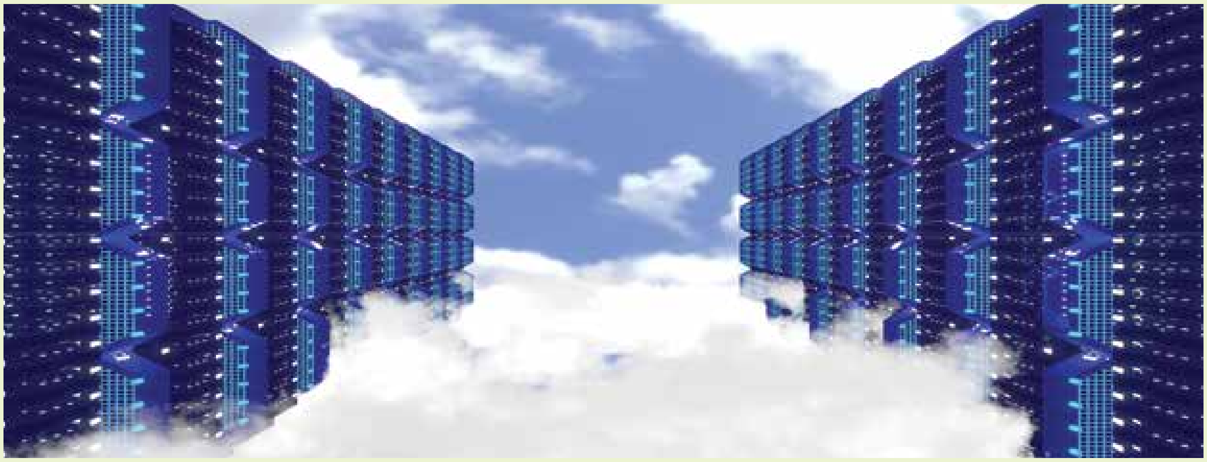
The Next Chapter

There has historically been much discussion around when, where, and why to deploy cloud related technologies, and debate about whether cloud computing models will eventually consume all of IT. Similar discussions occurred regarding the mainframe, as well as the migration to "client-server" computing models. Decades later, these technologies remain active in our computing environments, and perhaps regrettably, the domain of "legacy IT" does not appear to be leaving any time soon. As time goes on, it appears less likely that the question will be, "do we move to the Cloud?" and more likely to be, "which applications do we move, and to which cloud do we move them?" Savvy IT personnel will build out the criteria to answer these and related questions.

From lowering the cost and resources required to launch companies to inspiring entrepreneurs to build cloud-enabling technologies, the impact of public and private cloud technologies has been profound in the startup community. Questions relating to the security, cost, and functionality of the shift to Cloud will soon be replaced by more specific inquires relating to the security/cost/functionality delta between Cloud Provider X and existing infrastructure. Transparency has never been more important.

The answers to these questions will determine how much cloud, and how much computing, lies ahead. Q

Greg Shipley *(gshipley@iqt.org) is Vice President of Technical Staff within IQT's Advanced Analytics and Infrastructure Practice, where he is responsible for cloud and next generation infrastructure investments. Shipley also helps guide IQT's investments in information security areas. Prior to joining IQT, he was the founder and Chief Technology Officer for Neohapsis, an industry leader in information security and IT risk management. Shipley also ran the Chicago test lab for Network Computing magazine, was a contributing editor for Information Week magazine, and spent over a decade testing and reviewing technology on behalf of Fortune 500 companies.*



A Look Inside: Cloudy With a Chance of Computing

Cloud technologies have been widely adopted and acclaimed for their ability to optimize resources, reduce costs, decrease deployment times, and facilitate data sharing and analysis, but this relatively new technology also brings complex challenges for organizations learning how to implement and secure their cloud infrastructure. This edition of the *IQT Quarterly* examines recent advances in cloud computing.

Hemma Prafullchandra of IQT portfolio company HyTrust opens the issue with a look at the advantages, risks, and costs of cloud computing. While requirements like agility, flexibility, and elasticity are attractive benefits of the Cloud, they come with risks that must be considered. Prafullchandra offers mitigation strategies to help effectively leverage cloud infrastructure as the technology continues to mature.

Next, Justin Nemmers of CloudBolt Software discusses an often overlooked but important aspect of cloud migration: its implications for IT staffing. A shift in complexity means that organizations can no longer rely on traditional modes of operation, and they must understand that the Cloud's impact reaches far beyond its technology capabilities.

James Greene and Raghu Yeluri provide an overview of Intel's Trusted Execution Technology (TXT), a tool for safer cloud computing. TXT defines platform-level enhancements that serve as building blocks for trusted computing, data protection, and environment measurement and verification.

Dave Cole of CrowdStrike takes us behind the scenes of cloud security products. He explains the fundamental tradeoffs of traditional cloud protection solutions versus modern alternatives, and how CrowdStrike's security technology helps identify advanced threats and targeted attacks.

Next, Daniel Gwak and Greg Shipley's article examines the Open Compute Project (OCP), a community working toward efficient server, storage, and data center hardware designs for highly scalable computing. OCP offers robust, innovative alternatives to traditional computing methods, but commercial adoption and maturity still lags behind the pace of technical innovation.

Finally, we close the issue with a technology overview from IQT portfolio company Tenable Network Security, which identifies risk from vulnerabilities, compliance violations, and malware infections. Tenable has developed a service for Red Hat that allows remote, high assurance Red Hat auditing with OpenSCAP.

There is continuing discussion around cloud computing and how the Intelligence Community can securely and efficiently unleash its potential. We hope that this issue of the *IQT Quarterly* expands these discussions and encourages readers to think critically about what's next in the cloud computing landscape. Q

UNDERSTANDING CLOUD'S HIDDEN COSTS AND RISKS

By Hemma Prafullchandra

Virtualization and cloud computing are no longer emerging technologies, but in fact are disruptive technologies already in widespread use. Mandates such as "virtualize" and "cloud first" are common practices, and the return on investment is clear. However, it is important to understand the hidden costs and risks of virtualization; and it's time for organizations to implement greater maturity, transparency, and automation as they move to the Cloud.

Typically, the cloud stack consists of physical compute, storage, and network systems, with virtual infrastructure technology layered above. Depending on the chosen cloud stack (see Figure 1) a number of management, security, privacy, and compliance technologies must be used to safely operationalize the management, control, and data planes. In hybrid and public clouds, these technologies and planes may share operational responsibilities across many parties (at a minimum between the enterprise and the cloud provider), and as such, great care must be given to separate duties for people, processes, and controls. These technologies must be implemented by the enterprise and the cloud provider, regardless of the cloud service model (infrastructure as a service, platform as a service, or software as a service) being utilized. However, enterprise responsibility and control is reduced as higher service models are used, and the cloud provider's responsibilities increase as it owns and controls more of the cloud stack.

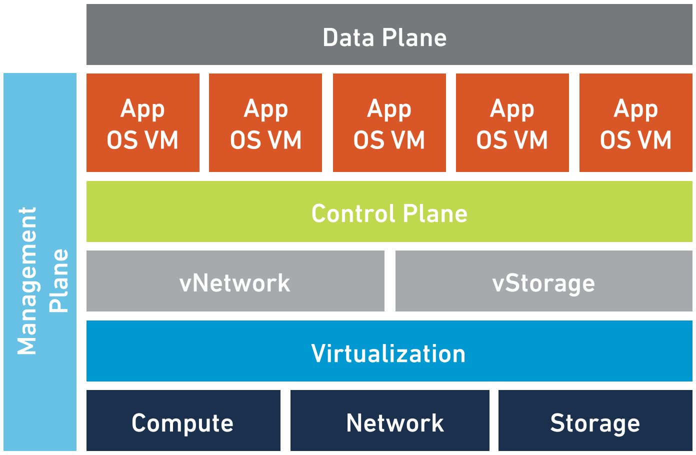


Figure 1 | Typical cloud stack

Traditional security, privacy, and compliance technologies can still be utilized in private, hybrid, or public cloud environments. These technologies include firewall, VPN, antivirus, malware detection, identity and access management, data loss prevention, and intrusion detection/prevention solutions. All of these traditional technologies provide optimized versions for specific cloud stacks with some degree of integration to those stacks for easy provisioning, planning, metering, and chargeback/billing. Most lack integrated control, visibility, and reporting that could feed into enterprise compliance and governance reporting systems.

Many cloud providers are also accredited and/or have third-party security and compliance attestation. This is excellent, but such assessments may be limited to specific data centers or regions, and remain point-in-time artifacts. Many cloud providers perform

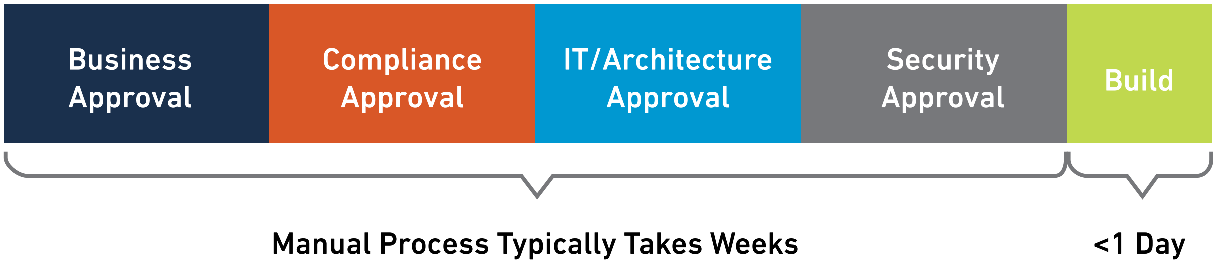


Figure 2 | Typical IT process

assessments annually and very few perform biannually, causing them to lose visibility of their current security posture very quickly.

There is far too much mistrust and unwillingness to share necessary information between the enterprise and the cloud provider for various reasons. Both groups often believe that security comes through obscurity. But in this era, building trust and being transparent can make a tremendous difference in how quickly a breach is detected and mitigated. With the growing number of sophisticated attacks, where the perpetrators have far greater resources than a single enterprise or cloud provider, organizations with high-risk systems must implement controls to ensure continuous monitoring and guarantee that this information is shared and reviewed regularly. Becoming mutually transparent is of paramount importance. Knowing the implemented architectures and controls while monitoring and sharing the ongoing findings can not only serve to lessen the consequences of a breach — this may also prevent the breach in the first place.

One of the significant hidden costs/risks associated with using cloud stacks is loss of governance. A typical IT process requires involvement and approvals from many functions in the organization (see Figure 2) — a sign of maturity and rigor in the fundamental business operations that can have significant material impact if not executed correctly. Often the process lacks automation, and forms and approvals have to be manually completed. The actual build takes less than a day with use of approved system architectures and technologies.

When a public cloud is used, the agility gained comes at a cost of bypassing internal governance controls, as shown in Figure 3. The actual build takes less than an hour if capacity and image templates are already available.

Clouds by definition are optimized for scale and multi-tenancy. Each provider cannot meet the specific requirements of each enterprise and therefore provides baseline, or "commodity" controls. For specific enterprise requirements, many providers continue to offer dedicated managed environments. Their focus is to limit the harm that a single tenant can do and to minimize internal accidental or malicious threats. Layers of defense are implemented, which makes it even harder for providers to share what is significant to a specific tenant.

The table that concludes this article summarizes some of the key benefits of leveraging cloud environments, identifies hidden costs and risks, and provides mitigation options to consider.

Many enterprises are leveraging some type of cloud infrastructure — private, hybrid, or public — but are at very different levels of maturity. There are a number of hidden costs to identify and mitigation options to consider. Transparency and greater security, privacy, and compliance automation are signs of maturity that enterprises must achieve to continue using cloud technologies without sacrificing their benefits of agility, flexibility, and elasticity.

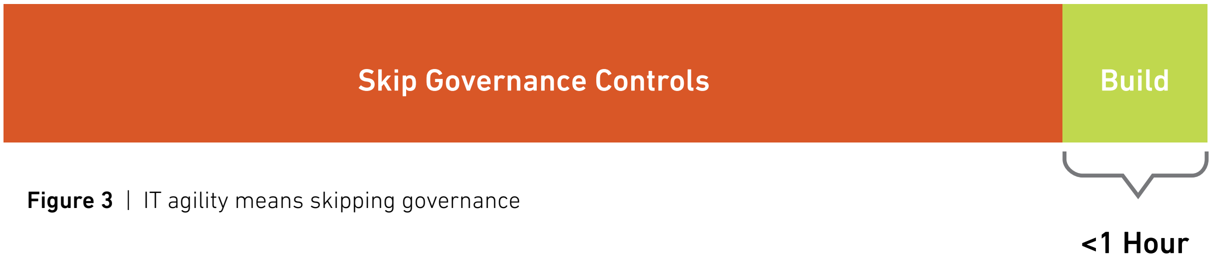


Figure 3 | IT agility means skipping governance

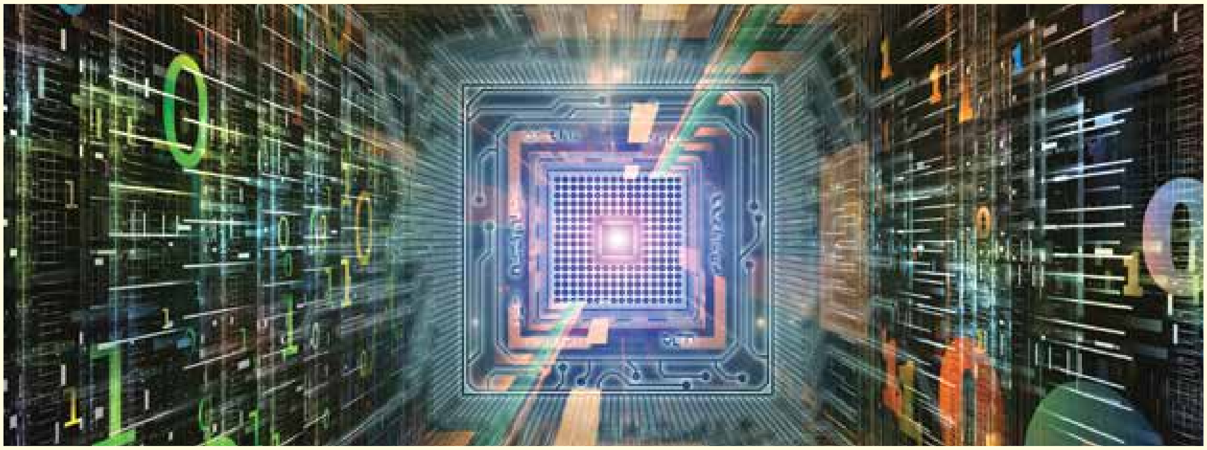
REQUIREMENT	BENEFIT	HIDDEN COST/RISKS	MITIGATION OPTIONS
IT Agility	Business lines can quickly utilize available capacity or acquire new capacity.	Bypasses internal business, IT/architecture, security, and compliance governance controls and processes.	Develop internal processes and leverage security, privacy, and compliance automation tools and technologies to automatically classify and provision necessary controls in any cloud environment.
IT Flexibility	Faster path to quickly experiment, fail, and iterate. More variation and emerging technologies available.	If used for production, huge eventual costs for internal approvals and hardening.	Establish approved cloud environments, including regions or host clusters, and workload image templates. Regularly review and verify provider environments and image templates. Update enterprise system architectures to leverage emerging technologies more frequently.
IT Elasticity	Better return on investment as additional capacity is only provisioned and utilized when needed.	Unknown where this additional capacity is being provided and whether the proper controls and safeguards are in place.	Regularly review provider process and practices. Classify workloads and ensure workloads are only deployed in allowed, integrity verified, specific cloud environments.
IT Redundancy	Business continuity and disaster recovery are easier and more cost effective to implement.	Actual ability to recover. Remaining compliant even during failovers.	Regularly verify that workloads can safely failover, and inspect that they remain under adequate protection and deployed on approved and integrity checked cloud stacks. Verify that access controls remain intact during failover and during recovery phase.
IT Maturity	Stronger physical data center capabilities and controls. Around-the-clock coverage with experienced IT staff, typically greater redundancy.	Security, privacy, and critical infrastructure directives and compliance mandates still apply. However, these were defined with traditional data centers and technologies in mind. Today, this is a shared responsibility and requires greater transparency on both sides.	Use technology that can assure workloads are: • Deployed on a compliant stack of hardware, hypervisor, and virtual images; • Continuously monitored; and • Automatically safeguarded for and on any compliance deviation. Point-in-time third-party attestation and accreditation are insufficient; regular reporting on administrative activities that affect enterprise workloads is critical.

REQUIREMENT	BENEFIT	HIDDEN COST/RISKS	MITIGATION OPTIONS
Shared IT Responsibility	Lower internal burden.	Optimized use of private, hybrid, and public clouds imply multi-tenancy (both different risk levels and owners) and frequent movement of workloads to different physical hosts.	Establish which types of business workloads can be deployed in specific private, hybrid, public clouds, and automatically associate risk level, and required security, privacy and compliance controls. Seek transparency. Regularly meet with provider and discuss separation of duties between their staff and the enterprise IT staff, and responsibility of controls implemented.
Shared IT Liability	Shift or lower insurance costs.	Enterprises must be able to maintain the same visibility and assurance as their traditional data centers, and are required to provide 'proof' no longer in their control.	Establish a regular schedule with the cloud provider to receive required visibility and assurance reporting that integrate into enterprise reporting.
IT Supply Chain	Guaranteed availability even if primary provider is suffering from technical issues.	No contractual relationship with the supply chain means no visibility and often complicated remediation processes.	Understand provider supply chain and dependencies on other third parties, including people, energy, facilities, and technology providers.

__Hemma Prafullchandra__ is the Chief Technology Officer and SVP of Products at IQT portfolio company HyTrust, where she is responsible for helping drive the company's security, privacy, and compliance product innovations and strategy. As an evangelist for what is possible, she pushes HyTrust and its ecosystem (partners, industry bodies, and customers) to enable cost-effective, secure deployment of virtualization with greater security automation.

REFERENCES

- www.hytrust.com
- NIST Special Publication 800-144
- NIST Special Publication 800-145
- PCI SSC Virtualization and Cloud Guidelines. https://www.pcisecuritystandards.org/security_standards/documents.php



THE PEOPLE SIDE OF CLOUD COMPUTING

By Justin Nemmers

The cloud-enabled enterprise fundamentally changes the way personnel interact with IT. Users are more efficient when they are granted on-demand access to resources, but these changes also alter the technical skill sets that IT organizations require to effectively support, maintain, and advance their offerings to end users. Often, these changes are not immediately obvious. Automation may be the linchpin of cloud computing, but the IT staff's ability to effectively implement and manage a cloud-enabled enterprise is critical to the organization's success and relevance.

Compounding the difficulties, legacy IT systems rarely go away, and many workloads, such as large databases, either don't cleanly map to cloud-provided infrastructure, or would be cost-prohibitive when deployed there. The co-existence of legacy infrastructure, traditional IT operations, and cloud-enabled ecosystems create a complicated dance that seasoned IT leadership and technical implementers alike must learn to effectively navigate.

As enterprise IT organizations have considered adopting cloud technologies, many fall into the trap of believing that increased automation will enable them to reduce staff. In my experience, however, staff reductions rarely happen. IT organizations that approach cloud-enabled IT as a mechanism to reduce staffing are often surprised to find that these changes do not reduce complexity in the environment, but instead merely shift complexity from the operations team to the applications team. For instance, deploying an existing application to Amazon Web Services (AWS) will not make it highly available. Instead of IT administrators using on-premises software tools with reliable access — and high-speed, low-latency network and storage interconnects — these administrators must now master concepts such as regions, availability zones, and elastic load balancers. Also, applications often need to be modified or completely redesigned to increase fault tolerance levels. The resulting deployments are still relatively complex, but they often require somewhat different skill sets than traditional IT administrators. Retraining is important for existing IT organizations because of this dramatic shift in complexity.

Governance is another common focus area that experiences significant capability gains as a result of cloud-enabled infrastructure. Automation ensures that every provisioned resource successfully completes each and every lifecycle management step 100 percent of the time. This revelation will be new to both IT operators and end users. Parts of the governance mechanism often completely break down due to end user revolt — largely because particularly onerous processes could be skipped by the administrators as they manually provisioned resources.

Cloud-based compute resources have great potential to dramatically change the computing landscape in nearly any organization. For example, one IT director worked to automate his entire provisioning and lifecycle management process, which freed up nearly three FTEs (full-time equivalents) worth of team time. Automating processes and offering end users on-demand access to resources helped their internal customers, but it also generated substantial savings for that team. That IT director recognized what many miss: the cloud offerings may shift complexity in the stack, but ultimately all of those fancy cloud instances are really just Windows and Linux systems that still require traditional care and feeding from IT. These tasks, such as Active Directory administration, patch management, vulnerability assessment, and configuration management, don't go away.

Another lesson learned here is that with shifting complexity comes dependence on new skills in the availability and monitoring realms. Lacking access to physical hardware, storage, and network infrastructure does not remove them as potential problem areas. As a result, organizations are too slowly realizing that applications need to be more tolerant of failures than they were under previous operating models. Making applications more resilient requires different skill sets that traditional IT teams need to learn in order to effectively grow into a cloud-enabled world. Additionally, when developers and quality assurance teams are getting near real-time access to necessary resources, they also tend to accelerate releases, placing an increased demand on the workforce components responsible for items such as release engineering, release planning, and possibly even marketing.

I've encountered few customers who have environments well suited for a complete migration to the public cloud. While modern IT organizations need to prepare for the inevitability of running workloads in the public or community clouds, they must also prepare for the continued offering of private cloud services and legacy infrastructures. Analyst firms such as Gartner suggest that the appropriate path forward for IT organizations is to become service brokers or providers. The subtext of that statement is that IT teams must remain in full control over who can deploy what, and where. IT organizations must control which apps can be deployed to a cloud, and which clouds are acceptable based on security, cost, capability, etc. Future IT teams should be presenting users with a choice of applications or services based on each user's role, and the IT team should worry about the most appropriate deployment environment. When this future materializes, these are skills new IT departments will need to master. Today, analyzing cloud deployment choices and recommending the approaches that should be made available are areas that typically fall outside the skill sets of many IT administrators. Unfortunately, these are precisely the skills that are needed, but many IT organizations overlook them.

The Way Ahead

While IT staff can save significant time when the entirety of provisioning and lifecycle management is automated, there are still many needs elsewhere in the IT organization. The successful approaches I've seen involved refocusing staff to value-added tasks. When IT administrators are able to spend time on interesting problems rather than performing near-constant and routine provisioning and maintenance, they are often more involved and fulfilled, and frequently produce innovative solutions that save organizations money. Changing skill sets and requirements will also likely affect existing contracts for organizations with heavily outsourced staffing.

Governance is another area where changes in the status quo can lead to additional benefits. For example, in manually provisioned and managed environments, central governance-related processes and procedures are rarely followed as closely as necessary. No matter how good the management systems, without automation and assignment, problems like virtual machine "sprawl" quickly become rampant. I've seen scenarios where end users revolt because they were finally subjected to policies that had previously been in place, but were routinely skipped by administrators

While modern IT organizations need to prepare for the inevitability of running workloads in the public or community clouds, they must also prepare for the continued offering of private cloud services and legacy infrastructures.

manually provisioning systems. Implementing automation means being prepared to retool some of the more onerous policies as needed. But even with retooled processes, automated provisioning and management nearly always provide for a higher assurance level than is possible with manual processes.

Automation in IT environments is nothing new. However, today's IT organizations can no longer solely rely on the traditional way of operating. Effective leadership of IT staff is critical to any organization's ability to successfully transition from a traditional provider of in-house services to an agile broker or provider of resources. Understanding that the Cloud has an impact on much more than just technology is a great place to start. This doesn't mean that organizations that are currently implementing cloud-enabling solutions need to hit the brakes; they just need to realize that the Cloud is not a magic solution for staffing issues. Organizations need to evaluate the potential impact of shifting complexity to other teams, and generally plan for disruption. Just as with any large-scale enterprise technology implementation, ensuring that IT staff have the appropriate skills necessary to successfully implement and maintain the desired end state will go a long way to ensuring success. Q

Justin Nemmers *is the Executive Vice President of Marketing at CloudBolt Software, Inc. CloudBolt's flagship product, CloudBolt C2, is a unified IT management platform that provides self-service IT and automated management/provisioning of on-premises and cloud-based IT resources. Prior to joining CloudBolt, Nemmers held both technical and sales-focused leadership roles at Salsa Labs and Red Hat, where he ran government services.*

Trusted Clouds: Visibility, Controls, and Compliance Capabilities to Enhance Cloud Security

By Raghu Yeluri and James Greene

By now, you are probably well aware of cloud computing. Your organization is likely using it in some form today. As a refresher, let's consider a definition of cloud computing that applies to the pooling of an on-demand virtual infrastructure, consumed as a service. This approach abstracts applications from the complexity of the underlying infrastructure, allowing IT to focus on enabling business value and innovation instead of getting bogged down by technology deployment details.

Organizations welcome the presumed cost savings and business flexibility associated with cloud deployments. However, IT practitioners unanimously cite security and compliance as primary issues that slow the adoption of cloud computing — at least for their more sensitive workloads and data sets. Fueling these concerns is a growing awareness of the new potential threats enabled by cloud implementations, such as low-level, hard to detect attacks on system BIOS (basic input/output system) and firmware, or compromises of the hypervisor or cloud operating environment. In each case, one would expect it to be very challenging to detect or mitigate these threats with existing tools, and that it would be even more challenging to assure data security if such threats manifest as a breach.

Many customers have specific security requirements mandating control over data location, isolation, and integrity. Under the current state of cloud computing, the means to verify a service's compliance are labor-intensive, inconsistent, non-scalable, or just impractical to implement. The necessary data, APIs, and tools are typically not available from the provider. For these reasons, many corporations only deploy less critical applications in the public cloud and restrict sensitive applications to traditional IT architecture running in corporate owned infrastructure.

Given the benefits of cloud computing models, this strategy of avoidance cannot continue. Businesses will want to put more sensitive and mission critical workloads into the Cloud as well. But for sensitive or regulated data workloads, more controls and security capabilities are needed to provide assurances of data protection and to support audit and reporting needs. This requirement drives the next frontier of cloud security and compliance: implementing a level of assurance at the bottom-most layers of the Cloud through the development of mechanisms to monitor and prove that the IaaS clouds' physical and virtual servers are actually performing as they should and are meeting defined security criteria.

Trust in the Cloud

In response to this need, and enabled by relatively new capabilities now included in many leading system platforms and host operating environments, organizations using or planning to use cloud services are starting to require service providers to improve security at the system hardware layer and provide greater transparency into system activities within and below the hypervisor. This means that cloud providers should be able to:

- Give organizations greater visibility into the security states of the hardware platforms running the clouds
- Provide policy-based control over where the virtual machines are and control the migration of these virtual machines based on policy specifications (such as some FISMA and DPA requirements dictate)
- Provide measured, auditable evidence that their services infrastructure complies with security policies with regulated data standards

To meet this need, vendors now offer building blocks for the development of "trustworthy clouds." These building block capabilities can be summarized as:

- Hardening the virtualization environment using known best practices
- Creating a chain of trust rooted in hardware that extends to include the hypervisor
- Using trust as part of the policy management for cloud activity (provisioning, migrations, etc.)
- Providing visibility for compliance and audit
- Using automation to bring it all together and achieve scale and management efficiency — often with hybrid cloud deployment models in mind

These building blocks become the foundations for the concept of trusted clouds. Trusted clouds address many of the previously discussed challenges and provide the ability for organizations to migrate regular and mission critical applications to leverage the benefits of cloud computing. Intel's Trusted Execution Technology (Intel® TXT) enables host boot integrity and provides an attestable infrastructure to enable visibility into the Cloud, exposing platform integrity as a new control point for virtual workloads and for audit and reporting needs.

Hardware-Assisted Security Enables Trust in the Platform

Intel TXT is a set of enhanced hardware components designed to protect sensitive information from software-based attacks. Intel TXT features include capabilities in the microprocessor, chipset, I/O subsystems, and other platform components. When coupled with an enabled operating system, hypervisor, and enabled applications, these capabilities provide confidentiality and data integrity in the face of increasingly hostile environments.

Intel TXT incorporates a number of secure processing innovations, including:

- **Trusted extensions integrated into silicon (processor and chipset):** These instructions allow for the orderly quiescence of all activities on the platform such that a tamper-resistant environment is enabled for the measurement and verification process, and allows for protection of platform secrets in the case of "reset" and other disruptive attacks.

- **Sealed storage:** Provides the ability to encrypt and store keys, data, and other sensitive information within the hardware. This can only be decrypted by the same environment that encrypted it.

- **Attestation:** Enables a system to provide assurance that the protected environment has been correctly invoked and to take a measurement of the software running in the protected space. The information is used to establish mutual trust between parties.

Intel TXT works through the creation of a measured launch environment (MLE) enabling an accurate comparison of all the critical elements of the launch environment against a hardware-protected known good source. Intel TXT measures and creates a cryptographically unique identifier for each launch-enabled component. This identifier is stored in a sealed Trusted Platform Module (TPM) on the host. Intel TXT measures BIOS and hypervisor components at launch and compares these values against the "known good" values stored in the TPM, and can indicate when an expected trusted launch has not happened. This hardware-based solution provides the foundation on which IT administrators can build trusted platform solutions to protect against aggressive software-based attacks. As shown in Figure 1, when combined with remote attestation services, this mechanism provides a powerful new control and audit capability for virtualized or cloud environments.

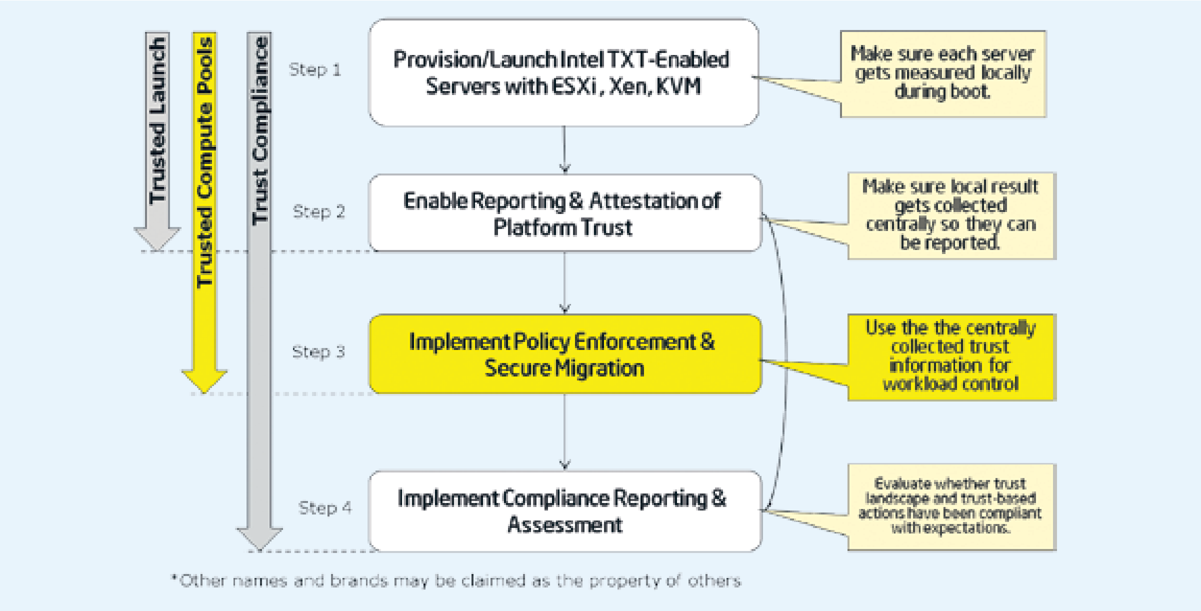

Trusted Launch
Trusted Compute Pools
Trust Compliance
Step 1
Provision/Launch Intel TXT-Enabled Servers with ESXi, Xen, KVM
Make sure each server gets measured locally during boot.
Step 2
Enable Reporting & Attestation of Platform Trust
Make sure local result gets collected centrally so they can be reported.
Step 3
Implement Policy Enforcement & Secure Migration
Use the the centrally collected trust information for workload control
Step 4
Implement Compliance Reporting & Assessment
Evaluate whether trust landscape and trust-based actions have been compliant with expectations.
*Other names and brands may be claimed as the property of others

Figure 1 | Attestation exposes platform trust status for higher-value use, such as workload control and reporting

Trusted Compute Pools

One of the most compelling examples of this is trusted compute pools (TCP). Trusted compute pools rely on establishing and propagating a new data center management attribute — "platform trust." The premise is simple: use remote attestation capabilities to challenge and verify the integrity of the various hosts in a cloud infrastructure. With this data, one can aggregate or pool the trusted systems and segregate them from untrusted resources. This, in turn, allows for the separation of higher-value, more sensitive workloads from commodity application workloads. These trusted pools allow IT to gain the benefits of the dynamic cloud environment while still enforcing higher levels of protection for their more critical and security sensitive workloads. When a trusted pool is created, systems and workloads can be tagged with specific security policies, enabling monitoring, control, and auditing for the placement and migration of workloads into, across, and outside the pool. Policies can be defined such that security-sensitive cloud services can only be launched on these resources, or migrated to other trusted platforms within these pools. This level of segregation allows for a reasonable compensating control for a cloud alternative to more traditional "air-gapped" (i.e., isolated from the rest of the data center) clusters of servers.

Once a trusted pool of platforms has been created, workloads can be selected to be placed on that pool based on their security requirements. A typical flow for workload placement would involve the following:

- A cloud subscriber requests workload to be placed in a trusted pool.
- Security management tools identify and tag workloads for classification according to security properties.
- Security management tools allow matching platform trust to workload classification according to existing policies.
- Orchestrator software determines the best server to place the workload within the trusted pool, pursuant to existing server selection and security policies; the scheduler requests an attestation of the server's integrity before the workload is placed on the server to reaffirm its boot integrity.
- A compliance record is created to register the launch of the workload in the trusted pool. This record is tied to the hardware root of trust of the server, and can be associated to a set of security controls to meet compliance requirements.

Cloud multi-tenant environments typically use virtualization capabilities to migrate virtual machines across physical hosts. Just as one takes care to assure that a workload is initially provisioned to a trusted host, it is logical to want to assure that a virtual machine is only allowed to migrate to other hosts with similar trust attributes. This goal might occur as follows:

- A migration of workload is triggered either manually or based on resource orchestrator/scheduler policies.
- The resource scheduler requests attestations of the integrity of potential target hosts to determine the set

of servers that meets the security policy associated with the workload.
- The orchestration software migrates the workload to a selected qualified server.
- A compliance record can be created to register the migration of the workload to this new location, including the attestation of the integrity at the time of selection.

Being able to prove to an audit entity that the security requirements of a given workload have been fulfilled is just as important as actually fulfilling those requirements. In the examples outlined, we see that events tracking workload placement and movement against trust boundaries can be reported and monitored on an ad hoc or programmatic basis.

Putting Data in its Place

Early cloud service providers, implementers of trusted compute pools, and their customers are requiring additional boundaries beyond platform trust to better control their workloads. A high priority boundary condition to enforce is one based on the specific physical location of a host such that workload placement can be:

- Monitored and enforced based on customer policies for boundary controls
- Verified and provided in audit and compliance reports to tenants to meet their internal and regulatory data security reporting needs

There are a few ways of attaching geolocation attributes to a platform. Conveniently, geolocation can be established within a TPM to provide hardware protections. This approach aligns naturally with trusted compute pools as the foundation of the use case for controlling and reporting compliance for workloads based on trust. Trusted compute pools with geotagging enable organizations to ensure their workloads are only executed on trusted servers located in authorized geographical areas. Such controls are specified or supported by a growing body of customer requests and regulatory mandates, such as the ability to separate customers, and workload types to address region-specific data protection requirements as defined in FISMA SP 800-53 and NIST IR 7409. Cloud service providers and ISVs (independent software vendors) are expected to extend current trusted compute pool solutions with trusted location controls to provide additional granularity of control above platform trust.

A Foundation

Trusted compute pools provide an excellent framework for trusted compute infrastructure, but they are not fully sufficient for the Cloud and for the increasing set of data and workload types that require more and more verifiable protections. The trusted, tagged pools outlined here and which are now commercially available are just the first foundational layers of providing more visibility control and compliance. A chain of trust from the hardware to the bare hypervisor, while a major advance, covers only the proverbial tip of the iceberg. It needs to be extended to explicitly support multi-tenancy and virtualized networks. Concepts such as trusted, measured VMs and enhanced encryption solutions are clear next steps.

These are areas of innovation that the ecosystem is actively engaged in today. There are a number of proof-of-concept and research projects underway to extend the basis of platform trust to deliver increasingly impactful levels of security and control for cloud workloads. These will continue to evolve to meet growing needs for protection against emerging threats and provide visibility and auditability across public and private cloud topologies. Q

__Raghu Yeluri__ is a Principal Engineer and director of security solutions architecture and development for Intel's Data Center and Cloud Products Group. In this role, he drives security solution pathfinding and development to deliver hardware-assisted security products that enable deep visibility, orchestration, and control in multi-tenant clouds. He has multiple patents filed in security, attestation, and control in virtualization and cloud computing. Yeluri holds a graduate degree in Computer Science and Engineering, and a B.S. in Electrical Engineering. Prior to Intel, he was involved in multiple artificial intelligence and knowledge-engineering startup ventures.

__James Greene__ is a senior technology lead for security technologies in the Data Center Group at Intel. In this role, he is responsible for the definition of products and usage models for data center and cloud security solutions. Greene is a frequent industry speaker and is co-author of the book, "Intel® Trusted Execution Technology for Servers: A Guide to More Secure Datacenters."

HARNESSING THE CLOUD: INFORMATION SECURITY EVOLVES

By Dave Cole

I thought I had a good handle on big data and what it meant for security. I had been part of a team that released a reputation system based on three years of massive amounts of file analysis. Further, I had been working with a lab team to do something similar for mobile apps, and in particular, their privacy attributes. We were using a high-end analytics platform and delivering content to millions of users through the Cloud in real time. Through the lens of a well-established security company, this seemed like the obvious intersection of big data and security.

Beyond the big players, security startups touting unprecedented protection due to progressive, public, cloud-hosted big data architectures are in a group that's likely larger than they would prefer to admit today. Further, established big data players like Splunk muddle the picture with security services that at least indirectly compete with some of their OEM partners and customers. There's enough market hype and confusion to fuel a few good years of pundit head-scratching and curmudgeon eye-rolling.

Given a closer look, the genuine protection yielded from the public cloud, big data, and perhaps any other new technology depends on how fully you are able to harness it. The company that designs its fundamental architecture to leverage a graph data model and the inherent scalability of public cloud understandably has a much better chance of making a compelling, new product than the organization that leverages the same technology to deliver a supplemental feature to an existing product. The latter is a tactical move that's circumvented by the adversary with a bit of time and effort. The former, more holistic approach is riskier, but this is where new, sustainable advantages are often created.

A good example of an incremental change leveraging big data analytics is the response of large antivirus vendors to the explosion of unique malware growth in the 2000s. As signature databases grew out of control, companies like Symantec and McAfee turned to a combination of big data tools and behavioral detection systems to keep pace. They added supporting file and URL reputation databases to make better detection decisions. They sped up delivery of antivirus signatures. They referenced cloud-based content in real time to help with file convictions.

While these are logical responses that yielded respectable near-term results, they failed to produce strong results against attacks that were not commodity malware. The parade of data breach headlines tells the story. Even though companies expanded their existing protection models to leverage cloud-based content (signatures, reputation, etc.), the logic is still driven by host-resident engines that assume decisions have to be made within milliseconds. This is the heritage of legacy scan-based thinking, where testing agencies grade vendors on the ability to complete an accurate scan as quickly as possible — even though real-time protection is the more common means of protecting users. Given such a short time frame, the amount of intelligence that can drive any single decision without bloating the size of the agent and updating it frequently is minimal — a design choice no reasonable endpoint security vendor has made in recent years. In the face of advanced persistent attacks that are developed over the course of days or weeks by an adversary, a system relegated to making decisions within milliseconds is designed to fail regardless of whether it has a multitude of supplemental cloud services it can call for help. Conventional solutions are destined to underperform against a determined adversary who may let an attack slowly unfold over hours, days, or weeks.

These products are also heavily, if not exclusively, focused on file execution as opposed to user or network activity. For example, the better endpoint security products will identify suspicious network activity (e.g., command and control traffic) from an executable that slipped past other defenses and convict the file as malware. They may also double-check the executable versus a cloud reputation service that has catalogued a massive number of programs to determine if it is "clean" or not. Nonetheless, even the best traditional products will not correlate executable and network behavior with expected user behavior and use this information to progressively constrain the attack based on the increasing level of certainty yielded by the series of observed system events. Stated differently, if a user normally does not send files to foreign IP addresses over RDP, why should he or she be allowed to do so, especially after a low prevalence executable was discovered on the system? Antivirus may not have enough information to stop the executable, and a data leakage prevention system likely would not have the smoking gun it needs to stymie file exfiltration — provided that it's even installed properly. They both lack sufficient context, outside of their narrowly defined scope, to be effective. A security information and event manager (SIEM) may bring this all together, but disappointingly after the fact as a narrative of events leading up to the headline-producing breach.

The gap of cloud- and big data-supplemented protection solutions is perhaps most strongly felt when attacks span multiple devices in a single environment. Even if they have enough context and time to make a good protection decision on one host, many attack campaigns stretch across multiple users and their devices. Adversaries are not betting that most attacks will succeed — they instead wager that one will make it through and allow them to spread across trusted devices after the initial breach. Since conventional solutions track state primarily on the device they reside upon, the seemingly normal administrative activity that happens following the initial attack never triggers defenses, as they cannot map the current activity to the original breach events. Specifically, what's happening on each machine is likely curious, but not worthy of conviction given the concerns around false positives and coarse-grained response options (i.e., remove and repair). But what's happening across multiple devices, if chained together, would tell a compelling story worthy of immediate response.

Even if you understand your limitations and recognize the need to re-architect to take full advantage of new technologies, it is incredibly hard to make the case for the investment. When the most visible, trusted testing organizations rate your product at 99.87 percent effectiveness across millions of malware samples, why would you invest millions of dollars to fix your product and migrate your users? The potential investment is not weighed against the sizeable revenue streams it would defend, but against the other investment opportunities presented inside the organization. Re-architecture initiatives have a track record when pitched against new product investments that is comparable to the Washington General's history versus the Harlem Globetrotters.

Being able to see the opportunity afforded by new technology, having the right people to execute, and being able to properly invest in them is a tall order for large companies, and seemingly a siren call to a legion of new startups fueled by investors who correctly surmise the vulnerable state of the incumbent players. Among the new genre of security players who have built their offerings from the ground up on a combination of public cloud, big data, and other emerging technologies, CrowdStrike stands out for its compelling design and commitment to fully harness emerging technologies. The core components of the solution are not novel; CrowdStrike uses a kernel-mode sensor for Windows and Mac endpoints as well as a cloud-based management console with a largely Amazon back end. Rather, it's how the technology is employed that creates the advantage.

In the face of advanced persistent attacks that are developed over the course of days or weeks by an adversary, a system relegated to making decisions within milliseconds is designed to fail regardless of whether it has a multitude of supplemental cloud services it can call for help.

Specifically, the host-based sensor is intentionally designed to capture the full context of activity on the device; it does not suffer from an overemphasis on file, user, or network activity. Instead of storing all this data on the device, it is shared with CrowdStrike's cloud-based state machine that has access to a full range of proprietary and third party security intelligence. The intelligence itself is slightly different, as it is focused on the adversary so that attacks can be attributed in order to answer questions of motive and identity that have previously been left to investigators. Thus, the model captures the full context of an attack and leverages as much intelligence as can be subsumed into the big data back end. The context of attacks, detected using a variety of techniques from behavioral patterns to signatures and heuristics, are stored in a graph database that carries an inherent design assumption: we cannot understand the relationship between all of the data now, but we may very well need to in the future.

The time frame for decision making in this model is necessarily different than in the rapid-fire scan and clean approach. It assumes that attacks may take place over days or weeks and that it may have to match events across multiple machines and intelligence sources to detect an attack. The primary customer benefit is purpose-built protection versus persistent, human-based attacks. A secondary benefit is that rather than looking at discrete, point-in-time detection events, customers can view an attack narrative that is told in real time as it unfolds, with a complete view of what is happening on all devices included in an attack. Lastly, response events can interject at any point in the sequence of events, or at several different points, with an intensity that matches the certainty of what's unfolding. The model is simultaneously more sensitive and more patient.

The forces of big data, cloud, mobile, virtualization, and other potent tech trends have already transformed the security industry. Not all vendor solutions have been transformed equally. The extent to which an offering has truly harnessed the advantages afforded by the recent shift in the tech landscape lives well beyond the press releases that serve to mask material differences and deep within the product design. The forces of time have been more kind to great design than they have to great marketing — we're counting on it. Q

***Dave Cole** is a seasoned product leader and a security industry veteran who has led an enterprise security startup from product concept through successful exit and headed a $2B consumer product portfolio through a period of sustained growth. He is currently VP of Products for CrowdStrike, a global provider of security technology and services focused on identifying advanced threats and targeted attacks. Prior to CrowdStrike, Cole held numerous senior positions within market-leading organizations such as Symantec and Deloitte & Touche, as well as former security startups Internet Security Systems and Foundstone.*

THE RISE OF THE OPEN COMPUTE PROJECT

By Daniel Gwak and Greg Shipley

Figure 1 | Open Compute Project hardware in use at Facebook's Prineville, Oregon data center

How do you save one billion dollars on data center costs? Mark Zuckerberg's answer: open source data center hardware. It's surprising to consider saving a billion dollars in what is quickly resembling a commodity market. After all, servers are differentiated by little more than the commodity components they cobble together: processors, hard drives, and memory — none of which are produced by server vendors themselves. Common sense business principles imply that commoditization is like sunlight to vampires when it comes to justifying margin — yet a typical rack of servers still costs more than a Ferrari.

Server vendors have continued to fight commoditization through "gratuitous differentiation" — a term used by Facebook in describing its frustration with the myriad LEDs, plastic bezels, and unnecessary circuitry that server vendors use to differentiate their box from the next. In 2011, Facebook attacked that gratuitous differentiation by designing its own minimalist servers and, according to CEO Mark Zuckerberg, saved over one billion dollars in the last three years. Those minimalist designs have been contributed to an open source community called the Open Compute Project (OCP). Since then, the Open Compute Project has gone from a custom hardware experiment at Facebook to a robust community of hundreds of contributors who have refined, improved, and adopted open source data center hardware, much the way open source software communities have evolved — accelerating the pace of innovation and mitigating adoption risk for all participants looking to capture similar savings.

However, the roots of custom data center hardware reach further back than Facebook's efforts. In April of 2009, Google revealed that it had done something radically different with its data centers. Google's solution involved liquid cooling, modular shipping containers, and an alternative UPS (Uninterruptable Power Supply) strategy based on 12-volt batteries at the server level. Refusing to purchase brand-name servers from traditional vendors, Google designed its own minimalist systems and deployed them at scale. Google's approach was such a departure from a normally staid industry that the announcement was initially mistaken for an April Fools' Day joke. But the effort turned out to be a massive success, with Google achieving nearly unheard of power usage efficiency (PUE) metrics and large reported savings. The difference from the Open Compute Project was primarily that Google held any innovation as proprietary to its business, releasing no details beyond the initial announcement and a few interesting videos.

OCP: Not Just for Servers

While most of the attention has gone to server hardware, OCP is looking to reinvent all aspects of the data center. One example is in networking, where OCP-spec top-of-rack switches are the beachhead to an effort that includes reinventing spine switches and other hardware and software solutions. In the spirit of openness, of course, these switches forego traditional closed and proprietary architectures, such as vertically integrated captive switching software, in favor of fully open technology stacks. A large part of that effort is aimed at disaggregating hardware from software, which has been vertically integrated in the networking industry to date. OCP's answer to this challenge is the Open Network Install Environment (ONIE), which defines an "install environment" for bare metal network switches. OCP switches running ONIE on bare metal network switches will give users a choice among different network operating systems that can be loaded and managed much like Linux servers. These kinds of innovations give users flexibility as well as cost savings — important variables in scaling data centers.

But Google's landmark project set a precedent in Silicon Valley. Two years later, Facebook announced that it had gone a similar route and tasked its hardware engineers with a "grid to gates"[1] redesign of the data center concept as part of a new data center to be built in Prineville, Oregon. Facebook, however, was quick to release the details and schematics of these designs to the world by contributing them to a newly established nonprofit organization it created called the Open Compute Project. The Open Compute Project was tasked with the goal of democratizing hardware and allowing for standardization across large-scale data center customers.

Today, OCP benefits from an active community of open source hardware contributors that includes hardware customers, vendors, ODMs, and technological contributors from academia. OCP hardware designs cover every major data center vertical, including servers, storage, data center design, networking, hardware management, certification of solution providers, and a newly designed rack concept referred to as Open Rack.

These hardware design contributions comprise a new way of looking at hardware that stresses four primary design goals: vanity free design, commonality, simplicity, and serviceability. An OCP committee determines a new design's acceptance into the community by measuring its adherence to these goals.

By adopting OCP hardware, Facebook has claimed 45 percent savings in capital expenditure (capex) and 38 percent less energy spent per data center — impressive results for an allegedly mature industry. But these efficiencies are realized through design changes that depart from legacy standards. These differences include rack sizes and layouts that are optimized for compute density and airflow, not compatibility with legacy systems. An OCP standard rack, for example, measures 21 inches in internal width — a departure from the legacy 19-inch standard. Other differences abound: power supplies are disaggregated and run at higher voltages for power efficiency, data center-wide UPS is foregone for more efficient rack-level UPS, and unnecessary circuitry (e.g., hardware management, video ports) are missing. Many of these changes reflect an almost entirely different way of managing hardware. In a Facebook-scale data center deployment, servers are managed through orchestration tools, not by connecting a monitor, keyboard, and mouse. Furthermore, self-healing resiliency in software allows a rip-and-replace approach to maintenance that greatly reduces the need for vendor specific management circuits. These capabilities, combined with OCP's tool-less serviceability, allow large-scale customers like Facebook to forego service contracts and service their own hardware at 20 times greater efficiency. These efficiencies have contributed to excitement around OCP in the hardware community and adoption at other large data center hardware customers like Riot Games and Rackspace. In a strong sign of momentum, the 2014 Open Compute Summit showcased talks from OCP community members that included household names, such as Microsoft, Goldman Sachs, Fidelity, Merck, and Orange.

However, challenges to broad adoption of OCP remain. Commercial maturity still lags behind technical innovation in the OCP community, with only a handful of vendors willing to sell OCP gear and Facebook dominating the demand curve. Attractive pricing and product availability remain elusive to small and medium size customers, who often rely on the very service contracts and vendor specific management tools that

[1] "Grid to Gates" refers to Facebook's effort to redesign every aspect of the data center, from the point at which power is taken from the grid until it hits the logic gates in the chips on a server's motherboard.

Figure 2 | Open Compute Project servers

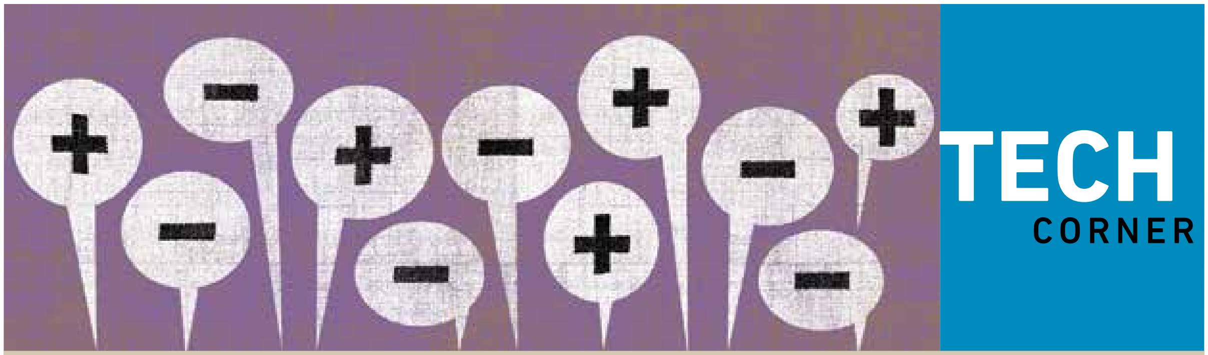
The Open Compute Project was tasked with the goal of democratizing hardware and allowing for standardization across large-scale data center customers.

OCP has endeavored to eliminate. Because of these factors, the organizations most likely to benefit from OCP adoption in the near term are large, efficiency-minded organizations that control much of their own software stack and are looking towards greenfield data center deployments.

What may change the equation for broader adoption of OCP will primarily hinge on commercial market maturity, proliferation of next generation server management tools, and the emergence of a robust ecosystem of OCP service providers and startups. In many ways, there are parallels between OCP and another open source initiative that was once at the heart of data center adoption debate: Linux. Today, Linux is a de facto choice for data center deployments. A decade from now, the same may be said for OCP. Q

Daniel Gwak *(dgwak@iqt.org) is a Senior Associate on In-Q-Tel's Investment team within the Advanced Analytics and Infrastructure Practice. Prior to IQT, Gwak worked as an investment professional at The Carlyle Group and Credit Suisse and served as an infantryman in the United States Marine Corps. He holds an M.B.A from Harvard Business School and a B.A. in Economics from Cornell University.*

Greg Shipley *(gshipley@iqt.org) is Vice President of Technical Staff within IQT's Advanced Analytics and Infrastructure Practice, where he is responsible for cloud and next generation infrastructure investments. Shipley also helps guide IQT's investments in information security areas. Prior to joining IQT, he was the founder and Chief Technology Officer for Neohapsis, an industry leader in information security and IT risk management. Shipley also ran the Chicago test lab for Network Computing magazine, was a contributing editor for Information Week magazine, and spent over a decade testing and reviewing technology on behalf of Fortune 500 companies.*



To supplement the ***IQT Quarterly***'s focus on technology trends, *Tech Corner* provides a practitioner's point of view of a current challenge in the field and insight into an effective response.

HIGH ASSURANCE OPENSCAP RED HAT AUDITING

A technology overview from IQT portfolio company Tenable Network Security

The world of network auditing is moving towards "continuous monitoring." Very soon, gone will be the days of an auditor manually testing a system to see if it meets compliance standards. Auditing technology has caught up to the complexities of modern operating systems and network devices through automation and programs like SCAP (Security Content Automation Protocol). Surprisingly, high-speed security auditing does have a security price. This comes from adding privileged agents to the operating system or giving auditing software the same access level as administrators, both of which create risk to the enterprise. Tenable Network Security has developed a technology which allows high assurance OpenSCAP audits of Red Hat Linux with a focus on privilege separation. This article describes the benefits of continuous monitoring, the security issues of performing audits with too much privilege, the NIST SCAP program, and Tenable's technology for Red Hat auditing with OpenSCAP.

A Short History of SCAP and Continuous Monitoring

After nearly twenty years of the security industry's innovation and investment in malware prevention, intrusion detection, anomaly algorithms, and firewalls, the bad guys are still getting past our defenses and into our servers to steal data. The latest craze is to open up all email and web content in a sandbox before it gets to our users and to our data.

All of these technologies are defensive tools. They prevent attacks based on knowing about an attacker's activities. However, they don't actually fix the problem of remediating vulnerabilities and preventing exploitation.

But fixing vulnerabilities is hard. Technology vendors like Microsoft, Adobe, Apple, and Red Hat do an excellent job of releasing patches that fix security issues, but it is still up to an enterprise to push these patches. This takes time and can cause a security team to waste significant effort making sure everything is updated and that nothing operational (like sending and receiving email) is interrupted.

A big issue preventing organizations from simply pushing these patches is that shortly after giving them to a user, their configurations are inconsistent. One user might need a driver for a printer which runs an application. Another might need to run older Windows 95 or Java applications. Others may have completely different browsers. These tiny differences in application usages can result in very different system configurations, which creates opportunities for software updates to fail.

Related to this are configuration issues that also overlap with vulnerabilities. A classic example of this is password length policies. An operating system can be configured to ensure that all of the user accounts have passwords of a certain length. For example, an organization could have a policy requiring all users to have a password of at least ten characters. Writing a policy like this is easy; manually testing this across hundreds or thousands of desktop computers is unthinkable. Adding a wide variety of devices like databases, routers, firewalls, Voice over IP, and other technologies to this list makes it even harder.

Almost a decade ago, the U.S. government began developing a set of protocols called SCAP (Security Content Application Protocol). These protocols allow for a variety of authorities (like government agencies, Microsoft, or you and me) to write policies which can be consumed by tools like Tenable's Nessus to further audit systems like routers and Windows devices.

The initial SCAP content was focused mostly on the Windows XP and Windows Vista operating systems, but today has expanded to Red Hat Enterprise Linux and applications like Internet Explorer. SCAP writing has also been heavily adopted by DISA for DOD policies, as well as by the Center for Internet Security (CIS), which produces best practice application hardening guides through cooperation with vendors, end users, and academia.

Red Hat has also adopted an open source project known as OpenSCAP. This is included in recent Red Hat operating systems and allows an administrator to work with SCAP content written for Red Hat to assist in auditing and managing the configuration of the Linux images.

All of this technology investment over the past decade allowed the U.S. government to declare an age of "continuous monitoring." Technically, this requires every civilian federal agency to submit a list of all their systems, vulnerabilities, and configuration issues to DHS every 30 days. The reader may note that 30 days is far from "continuous," but government leaders are pushing to move this to every three days.

Minimizing the Risk from Too Much Privilege

Organizations that wish to perform continuous monitoring need a method to gather the vulnerabilities and configuration issues across their vast number of systems. There are two primary methods to get this data: deploying agents on each system or deploying scanners with credentials to log in to each system.

In both cases, agents and scanners tend to run with administrator privileges, which makes them a target for attack and something that organizations spend a great deal of time securing.

For high assurance networks, performing security tests is very difficult because of the unknown nature of auditing. Security researchers who look at vulnerabilities all day typically think of performing these audits with full administrator access. If there is a problem with the configuration file of a certain application, then the audit typically needs to run with the same level of privilege as that application. If that application is run by the administrator, then the remote scanner or running agent needs the ability to execute any command or program on the system.

Administrators on high assurance networks deploy multiple levels of authentication and security on their systems. These often include limiting the number of programs that a user or application can execute. This limits the methods that an attacker can try, and also creates an audit trail of failed commands that are useful for detecting attackers.

This creates a problem with performing continuous monitoring in high assurance environments. Auditors need to be able to perform the audits required to demonstrate compliance, and the administrators don't want to give the auditors any type of access that creates a security risk.

What is needed is true separation of duties. The auditors should be responsible for creating and vetting the content and administrators should be able to review these audits and run them. This type of problem is exactly what Tenable aimed to address when developing its solution for Red Hat Enterprise Linux operating systems.

RSCAP: High Assurance Auditing of Red Hat with OpenSCAP

Tenable set out to develop a service for Red Hat which would allow full remote SCAP auditing without the need for an untrusted third party agent or the need to audit with complete access to the operating system. Tenable developed the "Remote SCAP" service known as RSCAP. This program supports separation of duties, authentication with public keys (not passwords), and caching of previous results for faster audits.

There are three components to RSCAP: an assessment daemon, a remediation daemon, and a communications daemon. Communication between these processes is performed through a secure file-based mechanism, and only the communications daemon is exposed to the network.

Tenable Network Security supports and enables secure network auditing as fast as necessary. The sooner that you can detect a security problem, the quicker you can fix it.

RSCAP also supports an experimental remediation function that uses the same approach of separation of duties to implement changes to the operating system. This allows administrators to receive trusted SCAP content and use it to fix security issues on the system. The auditors producing these security fixes also receive a level of trust that their fixes have been deployed securely.

The RSCAP technology will eventually be open sourced and shared with the SCAP community, and Tenable's hope is that it will be adopted by Red Hat as part of its core operating system. It is also part of Tenable's overall strategy for auditing Red Hat systems with Nessus. Nessus currently supports Linux SCAP audits through an upload of a dissolvable agent, and has been extended to support audits with systems running RSCAP. SCAP content leveraged against a system running RSCAP or a system being audited with credentials will report the same results.

Conclusion

Tenable Network Security supports and enables secure network auditing as fast as necessary. The sooner that you can detect a security problem, the quicker you can fix it. This saves time, effort, and resources. It also keeps the security issues away from potential hackers.

Tenable is also very supportive of government customers who wrestle with the security issues of centralized credentialed scanning and attack vectors introduced by third party agents. The RSCAP project helps organizations minimize the risk from doing their audits at an ever increasing pace.

***Tenable Network Security** is an IQT portfolio company offering continuous monitoring solutions to identify risk from vulnerabilities, compliance violations, and malware infections though system scanning, network traffic monitoring, and log analysis. To learn more, visit www.tenable.com.*

IN THE NEWS

The *IQT Quarterly* examines trends and advances in technology. IQT has made a number of investments in innovative cloud technologies, and several companies in the IQT portfolio are garnering attention for their unique solutions.

Cleversafe

Cleversafe's technology splits data into slices, encrypts it, and spreads it across multiple locations, making data storage and retrieval cheaper and more secure than traditional methods. The company's new Slicestor storage appliance has recently been discussed in publications including *The Chicago Tribune.* Slicestor, combined with a new software release, accelerates data storage with a mix of faster processors, solid state storage, and memory caching. Cleversafe has been an IQT portfolio company since September 2010 and is located in Chicago, IL. **www.cleversafe.com**

Pure Storage

Pure Storage is an all-flash enterprise storage company. The company recently introduced an innovative new maintenance model, Forever Flash, that has been touted for simplifying ownership and reducing total cost over the lifespan of equipment. Pure Storage has garnered attention from tech publications including *ITWire, Channelnomics*, and *CRN*, which named CEO Scott Dietzen to its 2013 Top 100 Executives List. The company is based in Mountain View, CA and has been a part of the IQT portfolio since May 2013. **www.purestorage.com**

Socrata

Socrata is a cloud software company focusing on open data and government performance products. The company recently announced a partnership with New York City to make it easier and more cost effective to transparently disseminate spending data. The company had a record-breaking 2013 in revenue, customer, and employee growth, and won multiple industry awards for its open data and performance measurement programs. Socrata is based in Seattle, WA, and joined the IQT portfolio in September 2013. **www.socrata.com**

Teradici

Teradici is a developer of industry-leading PC-over-IP (PCoIP) technology that is deployed in virtual and cloud environments, zero clients, hardware accelerators, standalone workstations, and mobile devices. The company recently partnered with Amazon, which is using Teradici's PCoIP technology for its fully managed desktop computing service, Amazon WorkSpaces. Teradici is based in Burnaby, British Columbia, and was recognized in 2013 as one of Canada's fastest growing tech companies on the Deloitte Technology Fast 50. Teradici became an IQT portfolio company in January 2011. **www.teradici.com**

IQT®

IN·Q·TEL