CHAPTER 7

Creating a More Secure Datacenter and Cloud

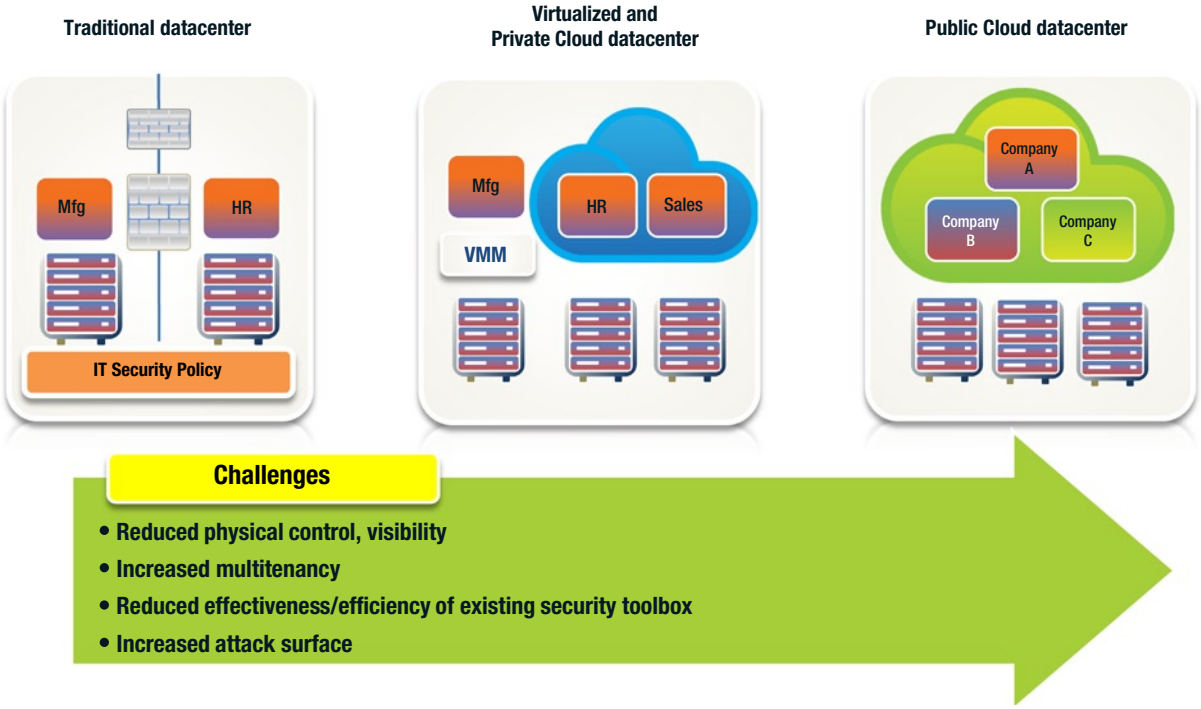
Every cloud has its silver lining but it is sometimes a little difficult to get it to the mint.

—Don Marquis

This book has discussed the utilities and capabilities enabled by Intel TXT for both datacenter and cloud computing deployment scenarios in numerous places. This chapter will provide more detailed and focused discussion on enabled use models and give examples of available enhanced security for both public and private clouds. It will explain the benefits of integrity, control, and visibility for cloud deployments and discuss various ways that the datacenter and user can take advantage of these attributes to benefit their business.

When Datacenter Meets the Cloud

It is easy and somewhat fashionable to discuss cloud computing models in excited tones as a panacea or a silver bullet that will solve all the challenges and woes of IT. Certainly, there is indeed tremendous interest in cloud computing models, and companies (including Intel) are realizing the benefits of enhanced business agility and cost reduction in their IT environments through adoption of cloud computing technologies and techniques. And analysts continue to recognize and forecast strong growth for cloud products and services. But against this background, we have to recognize that there is still a positively massive investment from businesses of all sizes in what we would think of as "traditional" datacenter IT models. These investments—like mainframe and minicomputer IT models from days of yore—are not going to go away any time soon. A rational perspective can easily see that customers will find ways to adopt new cloud computing models where it makes sense for their business, while continuing to leverage investments in their traditional IT estate, and finding ways to drive new value and efficiency from these investments. In the end, it is natural to expect that many customers will end up implementing some amalgamation of architectures and IT approaches that span traditional, virtualized, and cloud datacenters. As shown in Figure 7-1, while this allows new IT capabilities, it also introduces a number of new security challenges.



Source: Intel Corporation

Figure 7-1. *IT delivery models evolve to provide numerous options, but create new security challenges*

As you may expect, many of these challenges are driven by the lessening of physical controls—either by moving data beyond a company's four walls and into a public provider's infrastructure or by replacing physical protections such as security appliances with virtual ones. The sharing of infrastructure, in either a public way where the sharing of IT resources with anyone is likely or in a private implementation where the sharing may only be among various business units, may still be problematic. In short, as we look from left to right in the diagram, we see a reduction in control, the reduction in efficiency or effectiveness of traditional security tools, and at least the perception of increased risk of vulnerability.

The challenge of addressing these new security concerns will fall to industry and IT managers alike. In some cases, solutions that migrate from traditional deployments to new models will be the solution of choice. For example, consider how firewalls and a number of other security products that historically have been sold as discrete physical appliances have largely evolved and are now also often available as "virtual appliances" to meet new deployment and use models. But in other cases, entirely new protections and capabilities must be introduced to meet the challenges of new threat vectors, mitigate new risks, and enable appropriate security operation, audit, and compliance models. These new challenges are where technologies such as Intel TXT and its use models really shine.

The Cloud Variants

Before we get more deeply into the security solutions and use models, we should clarify our definition of the cloud. In order to avoid duplicate work and reinvention of the wheel, let's revisit a description that the US government has put in place as a definition of cloud computing attributes as a way of refreshing our perspective and establishing our baseline of understanding.

What is the "cloud"? For simplicity, we can focus on the definition published by the standards-setters at the National Institute of Standards and Technology (NIST) in their Special Publication 800-145 (SP 800-145).[1] The NIST definition establishes a *cloud* as an infrastructure that provides five essential services. These are excerpted as follows:

- *On-demand self-service.* A consumer can unilaterally provision computing capabilities, such as server time and network storage, automatically as needed, without requiring human interaction with each service provider.
- *Broad network access.* Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, tablets, laptops, and workstations).
- *Resource pooling.* The provider's computing resources are pooled to serve multiple consumers using a multitenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or datacenter). Examples of resources include storage, processing, memory, and network bandwidth.
- *Rapid elasticity.* Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time.
- *Measured service.* Cloud systems automatically control and optimize resources use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

These descriptions of essential capabilities are quite helpful to guide our understanding of the foundational attributes a cloud must provide. But more discussion is needed still, for we have seen that not all clouds are created equal, with different deployment models and different service models gaining traction in the market. These deployment/delivery models and service models will definitely impact the security capabilities that are needed to give an IT manager or a corporate security manager confidence in deploying workloads beyond a traditional IT model into new virtual or cloud datacenter models.

Cloud Delivery Models

Let's first take a look at the various cloud delivery models. Again, we can repurpose previously published work to establish our baseline for this discussion. In a paper that the authors helped create, Intel IT has published a suitable description of private, public, and hybrid cloud models that we can review here in Table 7-1, with excellent descriptions of the security challenges and considerations included.

[1] http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf.

Table 7-1. *A Brief Overview to Compare Cloud Delivery Models*

Model	Description	Advantages and Disadvantages
Private	• An internal infrastructure that leverages virtualization technology for the sole use of an enterprise behind the firewall • Can be managed by the organization or by a third party • Located on-premises (internal private cloud) or off-premises on shared or dedicated infrastructure (external private cloud)	• Provides the most control over data and platform • Potential for multitenancy of business units to cause compliance and security risk • May lack agility for bursting when additional performance or capacity is required
Public	• Resources dynamically provisioned over the Internet, via web services, or from a third-party provider • Located off-premises, typically on a shared (multitenant) infrastructure • May offer dedicated infrastructure as a response to growing security concerns	• Potential for greater cost savings if infrastructure owned and managed by public provider • Loss of control of data and platform • Potential for multitenancy with other organizations to cause security risk • Third-party security controls possibly not transparent (and may cause unknown risks)

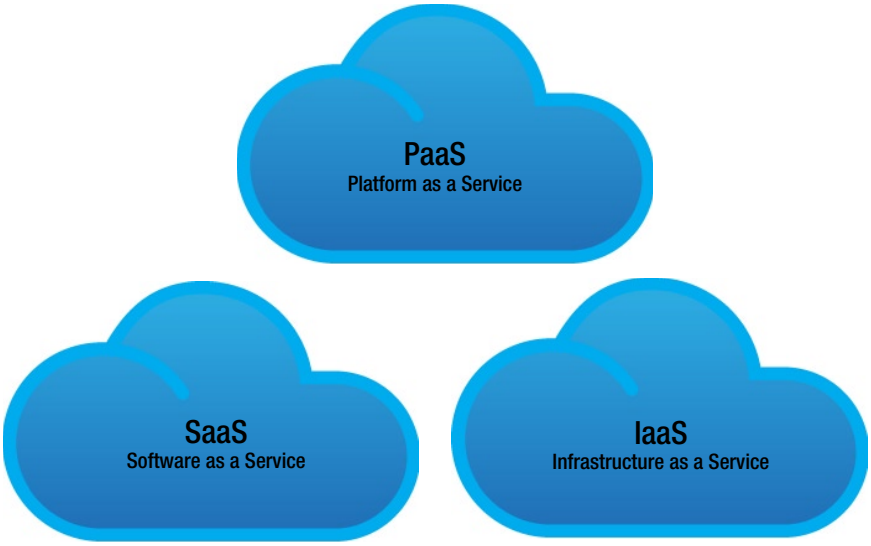
Hybrid	• A combination of private and public cloud services
	• Organizations that often maintain mission-critical services privately with the ability to cloud burst for additional capacity or add selective cloud services for specific purposes
	• Located on-premises and off-premises depending on the architecture and specific services
		• Often a compromise:
		  • Retention of physical control over the most mission-critical data, but relinquishing that control when additional capacity or scale is required during peak or seasonal periods
		  • May involve retention of physical control for mission-critical data at all times while taking advantage of public cloud provider services for less sensitive areas
		• Potential for complexity to cause unknown vulnerabilities (and unknown risks)

■ **Note** Adapted from the Intel IT Center *Planning Guide for Cloud Security*

As you can see, a sizable portion of the challenges are the result of the loss of physical control over workloads and data that occurs as one moves from a private cloud model to a public or hybrid model. This is likely not surprising given that there were similar concerns from customers as they looked at the cloud and virtualization relative to their traditional datacenters. In that scenario, they lost security capabilities and efficiencies through approaches such as physical isolation and discrete appliances as they moved to virtual shared infrastructures. Public and hybrid clouds exacerbate such concerns as they add the element of customers giving up physical control and possession of the workloads and data, as well. So they will need new protections to compensate for this, as well as new ways to view, control, and monitor how their data and workloads are being protected. And in the cases of data and workloads subject to compliance mandates and regulation, they need tools to help audit and prove these protections are in place.

It must be noted that there are few such regulations specifically calling for platform trust today, for it would be impractical to legislate controls and protections that are not widely available or implemented. Part of the motivation for this book is to help stimulate such deployment. But there are a number of regulations and controls that platform trust generally helps address today. These include sources such as the Federal Risk and Authorization Management Program (FedRAMP), the Cloud Security Alliance (CSA) Cloud Controls Matrix, NIST BIOS protections guidelines, and more. And as the maturity, awareness, and ubiquity of trusted platform solutions grows, it is natural to expect that more specific mandates for such protections will get incorporated into policies and regulations for cyber security in various deployment models.

To complete the cloud discussion, we still need to look at the various cloud service models to assess the security implications to determine what new protections and capabilities these might require. Once again, the authors would like to repurpose established definitions created by Intel IT to simplify the discussion, as these match terms often used in the industry. The primary cloud service models discussed in the market and by analysts are shown in Figure 7-2 and include Platform as a Service (often abbreviated as PaaS), Software as a Service (SaaS), and Infrastructure as a Service (IaaS).



Source: Intel Corporation, Excerpted from IT@Intel Brief "*Intel Cloud Computing Taxonomy and Ecosystem Analysis*"

Figure 7-2. *Summary of the primary models of cloud services*

The Intel IT paper defines these service models as follows:

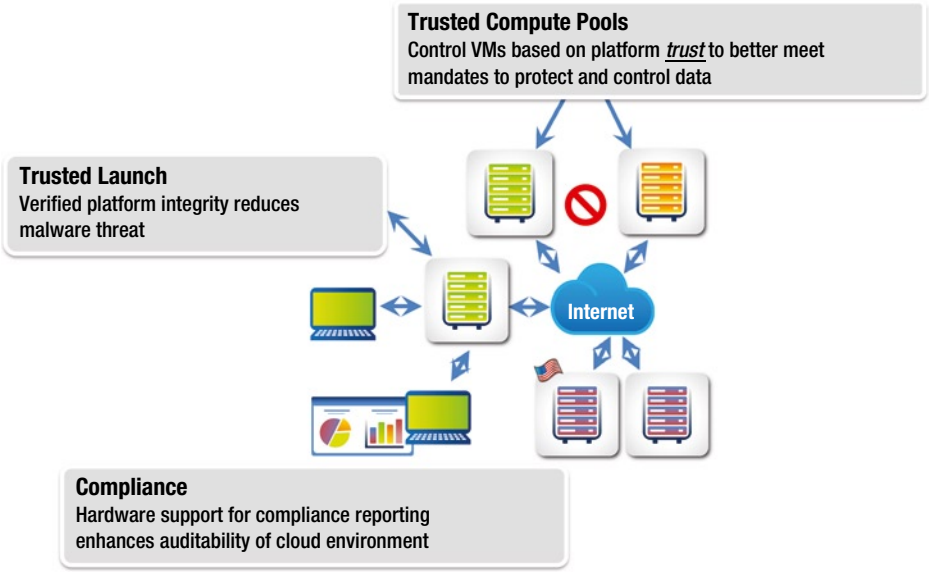
- SaaS is a model of software deployment in which an end user or enterprise subscribes to software on demand. SaaS applications are built with shared back-end services that enable multiple customers or users to access a shared data model.
- PaaS is the delivery of a cloud computing platform for developers. It facilitates development and deployment of applications without the cost and complexity of buying and managing the underlying hardware and software layers. PaaS provides all the facilities required to support the complete life cycle of building and delivering web applications and cloud services over the Internet.
- IaaS is the delivery of technology infrastructure—such as network, storage, and compute—as a service, typically through virtualization. Users subscribe to this virtual infrastructure on demand as opposed to purchasing servers, software, datacenter space, or network equipment. Billing is typically based on the resources consumed.

Once again we now have a set of service models that give customers significant options for the types of infrastructure they consume—and what services they are receiving from the cloud service provider. These models also pose different levels of control—in terms of who is responsible for what. For example, in a SaaS model, a SaaS provider such as `Salesforce.com` is responsible for much of the application infrastructure as well as the physical hosting infrastructure. The SaaS customer more or less is only responsible for the data and controlling access appropriately in this model. Alternately, an IaaS provider such as Amazon EC2 or Rackspace will typically only provide availability of basic compute, network, and storage resources on an allocated or "pay as you go" model—with no application or operating environment provided. Of course, these lines may blur over time, and providers could indeed grow from

a pure IaaS play to offer applications or more complete platform offerings, but that would basically only change their label—as long as they also evolve their security controls capability as well. For as you can see, these deployment models expose real security implications for customers as they need to consider what they are putting into the cloud infrastructure and how they can protect it and meet any policy compliance requirements. Intel TXT and its use models offer an opportunity to help provide new visibility and controls into this chasm as well as to help provide bridges and common capabilities across traditional physical datacenters and the emerging virtual and cloud IT infrastructures. And in time will likely also be useful in providing a common control capability that can be used across cloud providers, which will be useful for those companies that turn to multiple cloud providers for various services.

Intel TXT Use Models and the Cloud(s)

This book has discussed the enablement of Intel TXT in multiple chapters and in many different dimensions by this point. It is now a good opportunity to take a closer look at the impact of trusted computing use models to make sure the reader has a similarly strong understanding of how this technology can improve the security posture of their IT environments (physical, virtual, or cloud). Figure 7-3 provides a snapshot of the three leading use models for platform trust based on Intel TXT. Each use will be explained in more detail further in the chapter, but the basic premise is that the trusted launch process provides value in assuring platform integrity—lowering the threats and costs of certain classes of stealthy malware. The trusted pools use model extends that value by using the platform integrity enforced and reported by Intel TXT via attestation to be used to control workloads in the cloud. Lastly, the compliance use model extends this value yet again by providing an auditable infrastructure for verifying that the platform and workload controls are in place.

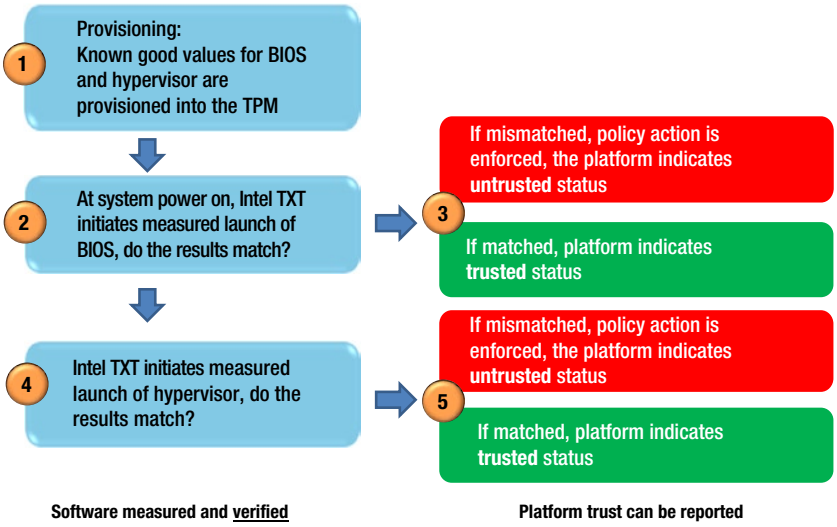


Source: Intel Corporation

Figure 7-3. *Summary of the primary use models of Intel TXT–enabled servers*

The Trusted Launch Model

The initial and foundational use model for Intel TXT is one that has been outlined in much of the first four chapters of this book; this is the ability to execute a trusted launch on a server. This fundamental capability to establish the integrity of a server has benefits for traditional datacenters as well as virtualized and cloud based servers. After all, everyone benefits from a more malware-free environment and Intel TXT, running on servers built with Intel® Xeon® processors, works with enabled software to help protect systems from BIOS and firmware attacks, malicious rootkit installations, and other malware attacks, while providing hardware-based verification that can be used for meeting compliance requirements. As detailed elsewhere in this book, the solution works by providing a processor-based, tamper-resistant environment that compares firmware, BIOS, and operating system or hypervisor code to known good configurations to establish a measured, trusted environment prior to launch. If trust is not verified, Intel TXT identifies that the code has been compromised, which lets you protect the system and remediate the problem. Figure 7-4 summarizes the high-level steps in the trusted launch process.



Source: Intel Corporation

Figure 7-4. *Steps of the trusted launch use model for Intel TXT–enabled servers*

As shown, the process starts with the establishment of the "known good" values of the approved BIOS and hypervisors that should run on the platform. These are provisioned to the TPM module, as discussed in earlier chapters of this book. At power on, the BIOS values are measured. If the BIOS measurement values are the same as the known good values stored in the TPM, then the hypervisor can be measured. But if the BIOS value results don't match, the platform will be proven untrusted. Similarly, if the measured hypervisor values match the known good values in the TPM, then the platform will be proven trusted and secrets can be unsealed, and so forth. If the hypervisor values do not match, then once again the platform will be proven untrusted. If a platform was expected to be trusted but failed these trust checks, then IT can be notified and remediation actions can begin, and other tools using the platforms can be aware and take appropriate measures.

By starting with a root of trust and measured launch environment (MLE), Intel TXT offers you better protection from malware and important visibility into the integrity of your system than would be available from a platform that cannot provide a root of trust or which is only protected by software. There is growing recognition in the industry press, among the analyst community, and from computer security specialists such as the NIST (which has published Special Publications such as SP 800-147B *BIOS Protection Guidelines for Servers*)[2] that discuss the threats from low-level platform attacks and how mitigations such as a hardware root of trust can help address these threats.

As discussed in the opening chapters of this book, and as we focused on in more detail in Chapter 6, this use model requires the most basic but most limited ecosystem enablement to activate. In summary, trusted launch enablement has impact in server hardware and BIOS, and in a suitable operating system or hypervisor that is capable of a measured launch.

Trusted Compute Pools: Driving the Market

The next use model has added benefits for customers that are deploying virtual and cloud IT architectures—as it allows the reintroduction of physical control capabilities into these increasingly shared and abstracted compute models. Trusted compute pools (TCP) are physical or logical groupings of computing platforms in a datacenter that have demonstrated integrity in key controlling components (such as BIOS and hypervisor) in the launch process. Intel TXT provides a hardware-based mechanism for verifying and reporting on platform trust as a foundation for the creation of trusted pools.

Platform trust status is attested to at launch (in the process outlined in Chapter 5) and if the launch was trusted, that platform is added to the trusted pool. Within this pool, systems and workloads can be tagged with security policies, and the access and execution of applications and workloads are monitored, controlled, and possibly audited. The most obvious premise is that highly confidential and sensitive applications and workloads would be constrained by policy to only run on systems that have proven to be trusted. Figure 7-5 outlines the basic steps to show how trusted pools could be used to enable workload controls in a virtual or cloud deployment.

[2] http://csrc.nist.gov/publications/drafts/800-147b/draft-sp800-147b_july2012.pdf.

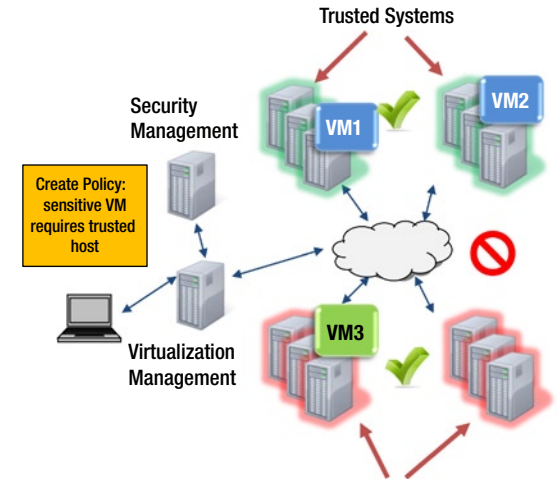


Figure 7-5. *Core components of the trusted pools use model for Intel TXT–enabled servers*

The basic steps for enabling a trusted pool start with having some mix of trusted and untrusted systems, as shown in Figure 7-5. From there, it is a matter of five general steps to create the operational trusted pools:

1. Virtualization management software can identify and report platforms that demonstrate integrity via Intel TXT using attestation mechanisms.
2. Security management software allows identification of sensitive workloads, in this case creating "labels" for the sensitive workloads, depicted in this figure as different shaded VMs.
3. Security management software can read platform trust status for the various available hosts from the virtualization management software—providing insight into the security capabilities or attributes of the hosts.
4. Security management software allows linkage of platform capability to workload classification via policy. In this example, a policy is created that specifies that sensitive VMs depicted at the top right in the figure can only run on trusted hosts.
5. Security management software policy can control VMs based on platform trust to better protect data, blocking deployments or migrations of these sensitive workloads into untrusted systems while allowing deployment or migrations among trusted hosts.

The trusted pools use model has gained perhaps even greater market interest than the basic platform trusted launch use model. This is perhaps not surprising since the need for new controls to address the security challenges of the new virtual and cloud use models is so great. Leading companies and agencies in governments, financial services, healthcare, and other industries, as well as cloud service providers that focus on these vertical market segments, have taken the lead and have done some of the initial deployments that serve as case studies for their peers to follow. Companies such as Florida Crystals, DuPont, and the Taiwan Stock Exchange have published testimonials outlining how their initial implementations have delivered security benefits and enhanced their confidence in cloud deployment models in their businesses. This list of success stories is poised to grow as the ecosystem of enabled technologies expands. It will also be fueled by Intel CloudBuilder reference architectures and OEM and ISV solution deployment guides, as well as books such as this one, to help customers understand how to implement solutions in their compute estate. Of course, as these controls get proven out by these ongoing deployments, it is easy to envision industry and regulatory codification of these mechanisms into data and platform protection guidelines and mandates.

To refresh and summarize: the trusted pools use model will indeed require a more significant set of enabled products and technologies to allow this robust, policy-driven security capability across physical, virtual, and cloud infrastructures. As outlined in Chapter 6, it will require the same platform (hardware, BIOS, and operating environment) enabling as the trusted launch use model. It has additional enabling implications for management and policy definitions and enforcement capabilities in terms of comprehending platform trust and integrity values, and implementing workload and data controls based on this information.

Extended Trusted Pools: Asset Tags and Geotags

While the market has rapidly grasped the concept of trusted pools as a new control mechanism and boundary model for enabling more security for their virtual and cloud IT architectures, some leading ISVs and end-user customers are taking an even more visionary approach and working with Intel on a natural extension for this model. The basic thinking behind this is that if Intel TXT–enabled platforms can provide visibility and new control and reporting mechanisms in the cloud, based on platform trust and attestation capabilities, could a trusted platform also store and report additional information that would provide further valuable control capabilities? As it turns out, the answer is "yes."

There are two general types of "new" control capabilities that customers desire. The first one is some type of geographic descriptor (what we often refer to as a *geotag*). After all, one of the natural concerns to address for the cloud is to be able to determine the answer to "Where is my stuff?" In a cloud without boundaries, this is impossible to answer. In a cloud that can be marked with trust and geographic description information, answers to this question can be made trivial—providing new confidence to customers. Given the large and growing number of regulations that stipulate location controls—particularly for privacy-related workloads and government data, this adds a significant breakthrough value. Now, workloads and this kind of data control that fall under the auspices of such regulation are now possibly open to cloud deployments.

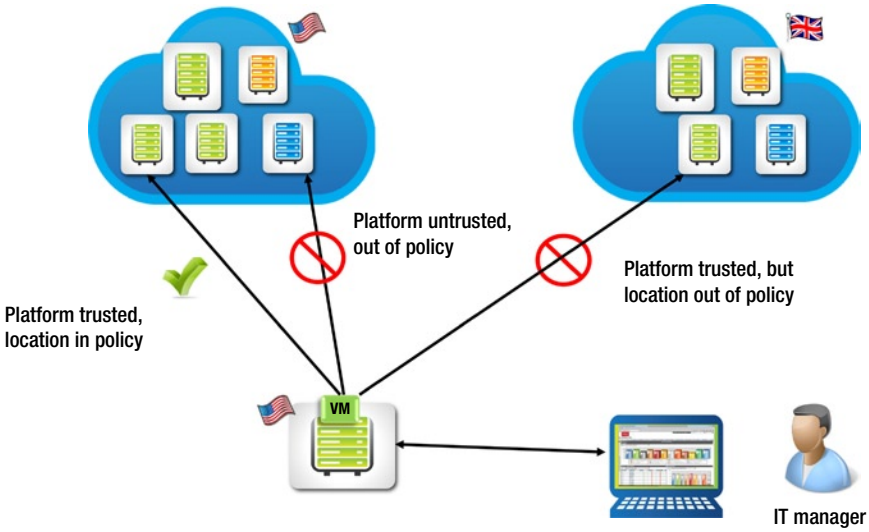
The other related type of control that customers have asked for is what we refer to as an *asset tag*. An asset tag is essentially just another set of descriptors that a customer may want to provision into their infrastructure to provide boundaries for controlling workloads and data. One could see an example where a cloud service provider may want to tag some of its compute infrastructure as belonging to different geographic availability zones or as systems with higher performance characteristics or even as dedicated to certain customers or customer types. These scenarios could provide solutions for customers paying for premium SLAs, or if they are servicing competitive customers that want assurances that their competition does not share a common infrastructure with their sensitive workloads. Similarly, by implementing asset tags, such as organization name or department, a customer could implement boundaries in a private cloud deployment. This could allow the IT or security organizations to keep data from different business units or organizational entities from commingling on common infrastructures. For example, this could be useful if a company wanted to make sure that financial, transaction processing, or human resources or other data did not become inadvertently exposed to other systems and data sources—but still wish to gain the benefits of virtualization and cloud computing models.

These tags are merely small text strings that can be populated or provisioned into the TPM. As such, the same Intel TXT–related trust infrastructure (for example, TPM and attestation services) that can store and report trust values can incrementally be provisioned with these additional geotag or asset tag descriptors. This would provide two benefits:

- Assurances that the tag descriptor values are coming from a trusted identifiable entity.
- The opportunity to leverage common attestation tools or services to gather, verify, and provide both trust and other descriptor values to the management, policy, and control infrastructure for use in deploying and securing workloads across infrastructures.

These benefits offer both security (and after all, if you can't trust the platform, you can't trust it to tell you where it was, what uses are appropriate, or anything else about it) and operational efficiency benefits that would be hard to replicate in other ways.

Figure 7-6 provides an example illustrating how platform trust and geolocation information can enable servers to provide enhanced, more granular control capabilities for critical or sensitive workloads. Of course, the greatest benefit comes when these are fully operational and the controls are driven by policy across the physical, virtual, and cloud environments.



Source: Intel Corporation

Figure 7-6. *Extending the trusted pools use model with geolocation and other asset descriptors provides additional visibility, control, and compliance capabilities*

In this example, an IT manager can access platform trust status and an additional descriptor of a geotag from a number of platforms. Behind the scenes, Intel TXT provides the trusted state and geotag descriptor via attestation capability, while untrusted systems may be from platforms that are not enabled with Intel TXT, and on these systems trust cannot be verified. In any case, the IT manager can create a policy that dictates that the workload is sensitive and must reside on both

- A trusted host (to better protect it from malware)
- A host located in the United States (perhaps due to company policy or where the data may be subject to Federal Information Security Management Act (FISMA) boundary control regulations, for example). Note that many other governments have similar regulations for assuring data stays within their governance domains.

Intel TXT and attestation capabilities provide the required insight into the platform to allow actionable data control policies in multiple dimensions—trust and location—using a common set of technologies.

So what does one really get from this extension of trusted pools use models with asset tags or geotags? Let's summarize and consider potential examples from the preceding scenario. The benefits include:

- *Increased visibility*. IT managers gain a hardware-based mechanism to verify platform integrity (trust) status and to store/report other asset descriptors, such as location, for use in their security controls portfolio.

  Example: IT infrastructure can attest to know which platforms have proven integrity, and which have not. IT can get assurances from trusted platforms regarding where cloud-based systems are located or other customer or cloud service provider–defined attributes are in order to implement data/workload controls.

- *Enhanced control*. IT managers can use platform integrity (trust) status and asset descriptor information to control virtual workloads.

  Example: Platform trust and other asset information can be used to implement policies that restrict sensitive workloads. It can be used to enforce policies to control migration or bursting to trusted systems and systems in specific geographical locations, as shown in the preceding example and illustration.

- *Automated compliance support*. IT or information security managers can attest that platform integrity (trust) status and asset descriptor information meet policy and verify that controls are in place and operational.

  Example: A governance, risk, and compliance (GRC) software suite can verify that platforms are trusted as expected and that workload controls for trust and location are established and enforced. In the preceding example, these GRC tools would gather platform trust attestations, as well as record that workloads are being placed in accordance with geographic restrictions and are issuing warnings when these policies were not adhered to.

From a governmental perspective, once again, the US Department of Commerce's NIST organization has been at the forefront of defining desirable and useful new controls to enable the cloud to be a more suitable environment for government workloads. NIST collaborated with Intel, RSA, and others to build a model that expanded on the trusted pools concept to implement location descriptor–based controls on top of trust-based controls to manage and measure compliance for workloads and data in the cloud. The resulting recommendation from this proof-of-concept model, NIST Interagency Report 7904 *Trusted Geolocation in the Cloud: Proof of Concept Implementation*,[3] was published as a draft in January 2013. From there it was presented to a broad set of governmental agencies and opened to comments from other industry participants. Interest and feedback has been very positive and will likely lead to continued enhancement and refinement of the model.

This use model is slightly less mature than the trusted pools use models. The tools for provisioning these tags or descriptors in the Intel TXT–enabled platforms are still nascent, and customers and service providers will need to define the processes and taxonomies for managing the tag values that represent the boundaries for control. But the strong market interest and lure of attracting regulated workloads to the cloud promises to drive rapid maturation of solutions in this space. In a twist that might surprise some, the platform and operating system/hypervisor tools will likely lag behind the ability of the management and policy tools to implement these extended control capabilities. This is, of course, contrary to the maturation model we have seen with the initial use models for Intel TXT.

Compliance: Changing the Landscape

We have seen that with new threats and new IT architectures, new controls for data, workload, and infrastructure are needed. And it is only natural that new mechanisms to enforce security policy and audit compliance to these security requirements are also required. As discussed previously, Intel TXT provides new enforcement mechanisms to support enhanced security for the datacenter, virtualized server environments, and the cloud. The hardware-based system integrity enforcement capabilities of an Intel TXT platform also provide a reporting mechanism to provide visibility into system status. This provides a method to implement and monitor controls and reporting tools across cloud and physical infrastructures via attestation. These Intel TXT, TPM, and attestation features assist in the verification and audit procedures, locally and remotely, to facilitate compliance automation. This is a critical capability for remedying what were previously considered insecure IT infrastructures to make them suitable for use with more critical and sensitive workloads.

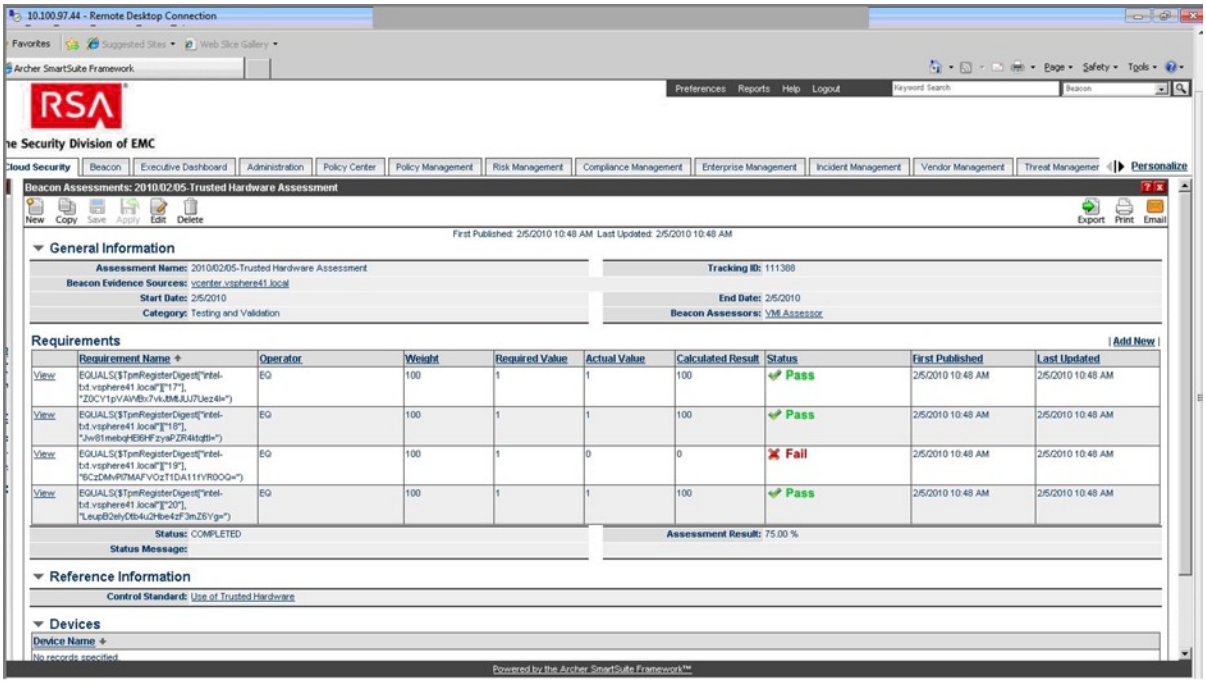
Compliance is a topic that we have touched upon briefly in previous chapters. In an increasing number of situations, it is not enough to provide protection for a type of data or component of infrastructure. It is often equally important to be able to monitor and prove that the protection is in place. Traditionally, this is often done with

[3] http://csrc.nist.gov/publications/drafts/ir7904/draft_nistir_7904.pdf.

security monitoring, logging, and compliance tools. It is also often done with a labor- and time-intensive audit process, one that gets much worse in virtual and public cloud implementations where an audit spans beyond physical, on-premise situations.

Platform trust is in an interesting position because it is only now starting to gain traction as a mandated or recommended protection (after all, it would be counterproductive to mandate a protection that was not readily available in the marketplace). But as these mandates develop and spread—such as the mapping of the trust and geolocation controls from the NIST IR 7904 recommendation, to the FedRAMP controls recommended for US government cloud purchases. For other regulation, such as the Health Insurance Portability Accountability Act (HIPAA) and International Organization for Standardization (ISO) 27000 series security control standards, trust use models will be ready to enable compliance.

Intel has long been collaborating with the ecosystem to implement controls based on Intel TXT and to show how they can be viewed, monitored, and reported remotely using common security compliance tools. In fact, in early 2010, Intel and RSA teamed with VMware to demonstrate compliance in the cloud, with a public demonstration and a published solution brief titled *Infrastructure Security: Getting to the Bottom of Compliance in the Cloud*[4] that addressed the need for controls and reporting on controls for cloud implementations. The screenshot in Figure 7-7 provides an excellent example. In the demonstration that this represents, the concept was to enact policies for restricting workloads subject to NIST SP800-53 (often referred to as FISMA) regulations to trusted hosts and related boundary controls for US locations. The RSA Archer GRC console was able to test and report on the trust status of the hosts in the demonstration configuration. The console provided a top-level compliance report, as well as a mechanism to get additional information that would be useful for audit purposes.



Source: Intel Corporation and RSA

Figure 7-7. *Screenshot of compliance dashboard reporting on platform trust*

[4] http://www.rsa.com/innovation/docs/CCOM_BRF_0310.pdf.

It has been interesting for the authors to observe how the market has responded to the use models for Intel TXT through time as we work to enable the industry and evangelize solutions for the market. Perhaps we have been too technology- or threat-focused and not as operationally aware, but the response to compliance use models has been surprisingly strong. Whereas the incremental technology is somewhat minor, the value customers perceive from compliance is quite large. Once again, this is likely due to the large void in controls created by virtualization and cloud architectures, as well as the added audit challenges and resultant cost and time burdens these create.

As trust-enabled solutions get enabled by the security ISV ecosystem, deployed in IT, and supported in the market by leading cloud service providers, compliance use models will be as important a factor in their purchase and deployment justification as the controls themselves. In some cases, it will be even more so as compliance changes the buying equation in an interesting way. Let us explain.

Typically, selling security is about managing risk—real and perceived. Many protections and controls are justified based on the premise that it will reduce or eliminate threats and risk. This is an equation that is often highly subjective. What one customer believes is a real threat or risk, another customer might find irrelevant or a corner case. Mandates that specify or recommend specific or classes of controls start to remove some of this subjective judgment. And in many cases, these governmental, industry, or corporate mandates add an incremental onus on the ability to verify that the controls are in place. Tools that facilitate this across IT architecture types are increasingly essential to making security operationally efficient.