PROJECT DESCRIPTION

TRUSTED GEOLOCATION IN THE CLOUD

Mike Bartock
Computer Security Division
Information Technology Laboratory
National Institute of Standards and Technology

Murugiah Souppaya
Computer Security Division
Information Technology Laboratory
National Institute of Standards and Technology

DRAFT

May 11, 2017

trusted-cloud-nccoe@nist.gov

NIST National Institute of Standards and Technology U.S. Department of Commerce

NCCoE NATIONAL CYBERSECURITY CENTER OF EXCELLENCE

The National Cybersecurity Center of Excellence (NCCoE), a part of the National Institute of Standards and Technology (NIST), is a collaborative hub where industry organizations, government agencies, and academic institutions work together to address businesses' most pressing cybersecurity challenges. This public-private partnership enables the creation of practical cybersecurity solutions for specific industries or broad, cross-sector technology challenges. Working with technology partners—from Fortune 50 market leaders to smaller companies specializing in IT security— the NCCoE develops modular, easily adaptable example cybersecurity solutions demonstrating how to apply standards and best practices using commercially available technology. Information is available at: https://nccoe.nist.gov.

This document describes a problem that is relevant to many industry sectors. NCCoE cybersecurity experts will address this challenge through collaboration with a community of interest, including vendors of cybersecurity solutions. The resulting reference design will detail an approach that can be incorporated across multiple sectors.

ABSTRACT

The motivation behind this Building Block is to improve the security of cloud computing and accelerate the adoption of cloud computing technologies by establishing an automated hardware root of trust method for enforcing and monitoring geolocation restrictions for cloud servers. A hardware root of trust is an inherently trusted combination of hardware and firmware that maintains the integrity of the geolocation information and the platform. Once the cloud platform has been attested to be trustworthy and to comply with a defined geolocation policy, then other use properties can be instantiated to support additional security capabilities that are built on this foundational hardware root of trust. These capabilities can include restricting workloads to running on trusted hardware in a trusted location; restricting communications between workloads; ensure workload data is protected at rest; applying security policies to workloads; and leveraging these capabilities across a hybrid cloud. This project will result in a freely available NIST Cybersecurity Practice Guide.

KEYWORDS

hardware root of trust; cloud computing; geolocation; migration; hybrid; encryption; network segmentation

Comments on NCCoE Documents

Organizations are encouraged to review all draft publications during public comment periods and provide feedback. All publications from NIST's National Cybersecurity Center of Excellence are available at http://nccoe.nist.gov/projects.

Comments on this publication may be submitted to: trusted-cloud-nccoe@nist.gov

Table of Contents

1. Executive Summary 1

    Purpose 1

    Background 1

2. Scenarios 2

3. Security Characteristics 4

    Stage 1: Platform Attestation and Safer Hypervisor or Operating System Launch 4

    Stage 2: Trust-Based Homogeneous Secure Migration within a Single Cloud Platform 5

    Stage 3: Trust-Based and Geolocation-Based Homogeneous Secure Migration within a Single Cloud Platform 5

    Stage 4: Data Protection and Encryption Key Management Enforcement Based on Trust-Based and Geolocation-Based Homogeneous Secure Migration within a Single Cloud Platform 6

    Stage 5: Persistent Data Flow Segmentation Before and After the Trust-Based and Geolocation-Based Homogeneous Secure Migration within a Single Cloud Platform. 7

    Stage 6: Industry Sector Compliance Enforcement for Regulated Workloads Before and After the Trust-Based and Geolocation-Based Homogeneous Secure Migration. 7

    Stage 7: Trust-Based and Geolocation-Based Homogeneous and Policy Enforcement in a Secure Cloud Bursting across Two Cloud Platforms 8

4. Relevant Standards and Guidance 11

5. Component List 12

1. Executive Summary

Purpose

Organizations often need to use cloud servers physically located within their own countries or where their data is generated and processed. Determining the approximate physical location of an object, such as a cloud computing server, is generally known as geolocation. Geolocation for cloud servers can be accomplished in many ways, with varying degrees of accuracy, but traditional geolocation methods are not secured, and they are enforced through management and operational controls that cannot be automated and scaled. Traditional geolocation methods depend on people and processes that cannot be trusted to meet cloud security needs.

The motivation behind this Building Block is to improve the security of cloud computing and accelerate the adoption of cloud computing technologies by establishing an automated hardware root of trust method for enforcing and monitoring geolocation restrictions for cloud servers.

A hardware root of trust is an inherently trusted combination of hardware and firmware that maintains the integrity of the geolocation information and the platform. The hardware root of trust is seeded by the organization, with the host's unique identifier and platform metadata stored in tamper-resistant hardware. This information is accessed by management and security tools using secure protocols to assert the integrity of the platform and confirm the location of the host.

Once the cloud platform has been attested to be trustworthy and to comply with a defined geolocation policy, then other use properties can be instantiated to support additional security capabilities that are built on this foundational hardware root of trust. One of the use cases is the ability to allow migration of workloads from an on-premise data center to a provider data center hosted on the Internet to take advantage of public cloud resources. In addition, various sets of policies can be enforced in support of the use cases. First, data protection capabilities can be implemented so that the workloads are decrypted only after the hosting platforms are measured and meet the geolocation policy. Second, the network data flows between the components of the workload can be isolated at runtime in support of an organization-defined segmentation policy in a shared environment. Finally, the security compliance properties can be automatically assessed and enforced to support industry sector-specific risk management frameworks and associated security controls.

Background

Shared cloud computing technologies are designed to be highly agile and flexible, transparently using any available resources to process workloads for their customers. However, there are security and privacy concerns associated with sharing resources and allowing unrestricted workload migration. Whenever multiple workloads are present on a single cloud server, there is a need to segregate those workloads so they do not interfere with each other, gain access to each other's sensitive data, or otherwise compromise the security or privacy of other workloads. As an example, consider two rival companies with workloads on a multi-tenancy cloud platform; each company would want to ensure that the server can be trusted to protect its information from the other company. Similarly, a single organization might have multiple workloads that need to be kept separate because of differing security requirements and needs for each workload, such as isolating a regulated workload from a public-facing workload.

Another concern with shared cloud computing is that workloads could move from cloud servers located in one country to servers in another country. Each country has its own laws for data security, privacy, and other aspects of information technology (IT). Because these laws may conflict with an organization's policies or mandates (e.g., laws, regulations), an organization may decide that it needs to restrict which cloud servers it uses based on their country.

2. Scenarios

This section proposes high-level usage scenarios that support various characteristics and requirements for cloud workloads across cloud environments, whether they are physically located on premise in a private data center or in a hosted and shared cloud provider infrastructure.

Using trusted compute pools is a leading approach to aggregate trusted systems and segregate them from untrusted resources, which results in the separation of higher-value, more sensitive workloads from commodity application and data workloads. The principles of operation are to:

Configure a part of the cloud that meets the security requirements defined by the organization.

Control access to that resource pool so the identified workloads are instantiated within that environment.

Control access and enforcement of the key management system to support the organization's data protection policy.

Control the flow of network traffic between the components within the trusted compute pool.

Enable audits and enforcement of security controls on that resource pool so it adheres to industry sector-specific compliance requirements.

These trusted compute pools allow the organization to gain the benefits of dynamic cloud environments while still enforcing higher levels of protection for more critical workloads. The ultimate goal is to be able to use trusted geolocation for deploying and migrating cloud workloads between cloud servers within private and hybrid clouds. This goal is dependent on smaller prerequisite goals, which can be thought of as capabilities the solution can provide. The following phases group the prerequisites and goals to increasingly support the proposed usage scenario:

1. **Platform Attestation and Safer Hypervisor or Operating System Launch**
   This ensures that the cloud workloads are instantiated and executed on trusted server platforms.

2. **Trust-Based Homogeneous Secure Migration within a Single Cloud Platform**
   This stage allows cloud workloads to be migrated among homogeneous trusted server platforms within a cloud owned and operated by a single provider.

3. **Trust-Based and Geolocation-Based Homogeneous Secure Migration within a Single Cloud Platform**
   This stage allows cloud workloads to be migrated among homogeneous trusted server platforms within a cloud operated by a single organization, taking into consideration geolocation restrictions. Although other trusted resource pool policies can be implemented, the geolocation usage model is selected for this scenario.

4. **Data Protection and Encryption Key Management Enforcement Based on Trust-Based and Geolocation-Based Homogeneous Secure Migration within a Single Cloud Platform**
   This stage allows the encryption keys to be released to the management platform to instantiate the cloud workloads only if they are migrated among homogeneous trusted server platforms within a cloud, taking into consideration geolocation restrictions.

5. **Persistent Data Flow Segmentation Before and After the Trust-Based and Geolocation-Based Homogeneous Secure Migration within a Single Cloud**
   Once the workloads are instantiated, the network isolation and segmentation rules between the components of the workload supporting an organization-defined policy are maintained before and after the secure migration.

6. **Industry Sector Compliance Enforcement for Regulated Workloads Before and After the Trust-Based and Geolocation-Based Homogeneous Secure Migration**
   The security controls supporting an industry sector risk management framework for regulated workloads are assessed, audited, and enforced before and after the secure migration.

7. **Trust-Based and Geolocation-Based Homogeneous and Policy Enforcement in a Secure Cloud Bursting across Two Cloud Platforms**
   This stage allows cloud workloads to be migrated among homogeneous trusted

server platforms between two clouds managed by different organizations (e.g., private cloud to a cloud hosted by a cloud service provider), taking into consideration geolocation restrictions, data protection policies, network segmentation rules, and industry sector compliance requirements. The policies assigned to the workloads are maintained and enforced before and after the secure migration.

3. Security Characteristics

The prerequisite security characteristics for each stage, along with more general information, are explained below.

Stage 1: Platform Attestation and Safer Hypervisor or Operating System Launch

A fundamental component of a solution is having some assurance that the cloud platform hosting the workload can be trusted. If the platform is not trustworthy, then the workloads are susceptible to potential attacks from the compute pool. This is a fundamental property and root of trust that is leveraged by the subsequent stages. For example, the attestation of the claimed geolocation property is not trustworthy if the platform has been tampered with. Having basic assurance of trustworthiness is the initial stage in the implementation.

Stage 1 includes the following goals:

1. **Configure a cloud server platform as being trusted**
   The cloud server platform includes the hardware firmware (e.g., Basic Input/Output System (BIOS) or modern Unified Extensible Firmware Interface (UEFI) settings) and the hypervisor or operating system base image. A hardware cryptographic module is provisioned to store the measurements of the underlying hardware and software modules.

2. **Before each hypervisor or operating system launch, verify the trustworthiness of the cloud server platform**
   The items configured in goal 1 (BIOS/UEFI and hypervisor/operating system) need to have their configurations verified before launching the hypervisor and operating system to ensure that the assumed level of trust is still in place.

3. **During hypervisor or operating system execution, periodically audit the trustworthiness of the cloud server platform**
   This periodic audit is essentially the same check as that performed as goal 2, except that it is performed frequently while the hypervisor or operating system is executing. Ideally this measurement would be part of continuous monitoring and support the compliance requirements for regulated workloads.

Achieving all of these goals will not prevent attacks from succeeding, but it will detect unauthorized changes to the hypervisor/operating system or BIOS/UEFI in near-realtime. If the firmware or software modules are tampered with or otherwise subverted, the alteration will be detected. This will prevent the workloads from launching, thus limiting damage to the information being processed within the cloud server platform.

Stage 2: Trust-Based Homogeneous Secure Migration within a Single Cloud Platform

Once stage 1 has been successfully completed, the next objective is to be able to migrate workloads among homogeneous, trusted platforms. Workload migration is a key attribute of cloud computing, improving scalability and reliability. The purpose of this stage is to ensure that any server that a workload is moved to will have the same level of security assurance as the server it is coming from.

Stage 2 includes the following goals:

1. **Deploy workloads only to cloud servers with trusted platforms**
   This indicates that stage 1, goal 3 (auditing platform trustworthiness during hypervisor or operating system execution) occurs and a workload is instantiated on a cloud server if the audit demonstrates that the platform is trustworthy.

2. **Migrate workloads on trusted platforms to homogeneous cloud servers on trusted platforms; prohibit migration of workloads between trusted and untrusted servers**
   For this scenario, homogeneous cloud servers are servers that have the same hardware architecture (e.g., Central Processing Unit (CPU) type) and the same hypervisor or operating system type, and they are under the control of a single cloud management console. If a workload has been deployed to a trusted platform, the level of assurance can only be maintained if it is migrated to hosts with identical trust levels. This goal is built upon stage 1, goal 3 (auditing platform trustworthiness during hypervisor execution) and it is performed on both the source and destination cloud server so that the migration is successful if both servers meet the defined policy.

Achieving these goals ensures that the workloads are deployed to trusted platforms, thus reducing the chance of workload compromise.

Stage 3: Trust-Based and Geolocation-Based Homogeneous Secure Migration within a Single Cloud Platform

The next stage builds upon stage 2 by adding the ability to continuously monitor and enforce geolocation restrictions.

Stage 3 includes the following goals:

1. **Have trusted geolocation information for each trusted platform instance**
   This information is represented as a cryptographic hash, and it is stored in the cloud server’s cryptographic module so that it can be verified and audited.

2. **Provide configuration management and policy enforcement mechanisms for trusted platforms that include enforcement of geolocation restrictions**
   This goal builds upon stage 2, goal 2 (migrating workloads on trusted platforms to other trusted platforms), enhancing it by adding a geolocation check during the migration process.

3. **During hypervisor or operating system execution, periodically audit the geolocation of the cloud server platform against geolocation policy restrictions.**
   This goal is built upon stage 1, goal 3 (auditing platform trustworthiness during hypervisor or operating system execution) and it audits the geolocation information against a defined policy to ensure compliance after the platform has been attested.

Achieving these goals ensures that the workloads are not migrated to a server in an unsuitable geographic location. This avoids issues caused by clouds spanning different physical locations (e.g., countries or states with different data security and privacy laws).

Stage 4: Data Protection and Encryption Key Management Enforcement Based on Trust-Based and Geolocation-Based Homogeneous Secure Migration within a Single Cloud Platform

The next stage builds upon the successful implementation of stage 3 by adding the ability to protect the workloads with encryption technology based on server trust and geolocation information.

Stage 4 includes the following goals:

1. **Have the cloud performing trust-based and geolocation-based measurement**
   The trust and geolocation information must be audited periodically so this information can be utilized to make a decision about the release of the encryption keys.

2. **Provide a Key Management Solution for Data protection**
   A key management solution is instantiated in a secure environment and integrated with the cloud management platform. Workloads will be encrypted and their keys will be stored in the data protection key management server.

3. **Provide an encryption key server**
   Workloads are decrypted by the server and instantiated if the server meets the trusted resource pool characteristics. Before a server is given access to the encryption key for a workload held by the cloud management platform, it must meet the trust and geolocation policy.

Achieving these goals ensures that the workloads are decrypted on a server that meets the trust or geolocation policy. The workloads are fully encrypted at rest in the storage devices to minimize the risk of data leakage. Even if the workloads are backed up and physically moved to another server, they cannot be decrypted by a server that is not part of the authorized trusted resource pool.

Stage 5: Persistent Data Flow Segmentation Before and After the Trust-Based and Geolocation-Based Homogeneous Secure Migration within a Single Cloud Platform

This stage builds upon Stage 3 at the minimum and it can include Stage 4 for a more robust implementation by including the capability to isolate the network data flows between the components within a workload.

Stage 5 includes the following goals:

1. **Define the network data flow policy between the various components of the workload**
   A whitelist representing authorized network communications between the different components is enumerated and documented.

2. **Associate and enforce the network data flow policy with each component**
   The management platform binds the network data flow policy to the workload and its components, and enforces the rules when the workload is instantiated.

3. **Persist the network data flow policy across migration**
   The management platform associates the policy to the workload so that the rules are enforced when the workload is moved between different trusted cloud servers.

Achieving these goals allows the organization to implement micro-segmentation policies between components of the workload independent of the cloud servers. This mechanism can detect and restrict the lateral movement from a component in the event it is compromised. This capability helps the organization meet the least privilege principle for network flow within a workload.

Stage 6: Industry Sector Compliance Enforcement for Regulated Workloads Before and After the Trust-Based and Geolocation-Based Homogeneous Secure Migration

This stage builds upon Stage 3 at the minimum, and it can leverage the capabilities from Stage 4 and 5 for a more complete implementation, by including the enforcement, continuous monitoring, and reporting of the configuration of the hosted infrastructure and the workloads needed to meet the NIST SP 800-53 moderate baseline security controls which are associated to the NIST Cybersecurity Framework.

Stage 6 includes the following goals:

1. **Enumerate the NIST SP 800-53 technical security controls applicable to securing the cloud platform and workloads**
   The moderate baseline is selected and the appropriated controls are listed to represent the measurement framework in addition to the trust and geolocation-

based characteristics, data protection capabilities, and network segmentation properties. The measurement framework includes applicable security controls from families such as access control, audit and accountability, configuration management, identification and authentication, system and communication protection, and system and information integrity. The associated NIST Cybersecurity Framework functions, categories, and sub-categories are documented based on their relationship to the applicable NIST SP 800-53 security controls.

2. **Document the architecture of the solution and the security baseline**
The components of the cloud infrastructure and the workload are enumerated and the associated security capabilities and technical mechanisms with recommended values are documented to represent the security baseline. The security mechanisms can be built natively into the cloud platform or provided by additional hardware and software components integrated into the solution.

3. **Develop a mapping of the NIST SP 800-53 technical security controls to the architecture with technical security mechanisms and assigned values**
This step combines the outputs of the previous steps to establish the relationship between the NIST SP 800-53 security controls and the security mechanisms of the solution.

4. **Implement, assess, and report on the technical security mechanisms with respect to the NIST SP 800-53 security controls**
The security baseline is instrumented across the different components from the infrastructure to the workload. The technical mechanisms are continuously assessed to demonstrate their risk compliance status to the security controls. A governance, risk, and compliance (GRC) tool can be leveraged to provide a detailed report or high-level dashboard view of the compliance posture.

Achieving these goals allows the organization to implement a technical baseline to secure the cloud platform and the workload to meet an industry sector-specific compliance framework. The technical mechanisms are continuously enforced and assessed to secure the environment over the lifecycle of the cloud platforms and workloads. These mechanisms enable the organization to manage risks and meet the compliance requirements.

Stage 7: Trust-Based and Geolocation-Based Homogeneous and Policy Enforcement in a Secure Cloud Bursting across Two Cloud Platforms

This stage builds upon stage 3 and can leverage the capabilities from stages 4, 5, and 6. It introduces the notion that the workloads are migrated between two cloud platforms that are controlled by two different entities, like a cloud operated in an organization's private data center and a cloud managed by a cloud service provider.

Stage 7 includes the following goals:

1. **Have two clouds ensure trust by performing geolocation-based migration**
   Each cloud operator can provide the capabilities described in the previous phases to ensure that the workload migration can occur successfully within their environment in support of the defined policy.

2. **Establish a trust broker that two cloud operators can use to ensure trust, geolocation-based migration, data protection, workload network segmentation, and secure baseline and compliance policy enforcement**
   A trust broker is established to arbitrate and attest that the two cloud platforms meet the trust, geolocation-based migration, data protection, workload network segmentation, and secure baseline and compliance policy requirements. The trust broker bridges the management platforms of both clouds to provide a centralized portal for enforcing, assessing, and auditing the infrastructure and workloads to meet a specified policy.

Achieving these goals ensures that workloads can be migrated between different cloud platforms to take advantage of public cloud service providers in support of the organization's business objectives of cost savings, high availability, and resiliency without compromising the security of the workloads and while still meeting the compliance requirements.

The following table summarizes the required and optional capabilities for each stage. A complete and robust implementation will include capabilities defined in all the stages.

	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6	Stage 7
Platform Attestation and Safer Launch	x	x					
Trust-Based Homogeneous Secure Migration within a Single Cloud Platform	x	x					
Trust-Based and Geolocation-Based Homogeneous Secure Migration within a Single Cloud Platform	x	x	x				
Data Protection and Encryption Key	x	x	x	x	optional	optional	optional

Management Enforcement Based on Trust-Based and Geolocation-Based Homogeneous Secure Migration within a Single Cloud Platform							
Persistent Data Flow Segmentation Before and After the Trust-Based and Geolocation-Based Homogeneous Secure Migration within a Single Cloud Platform	x	x	x	optional	x	optional	optional
Industry Sector Compliance Enforcement for Regulated Workloads Before and After the Trust-Based and Geolocation-Based Homogeneous Secure Migration	x	x	x	optional	optional	x	optional
Trust-Based and Geolocation-Based Homogeneous and Policy Enforcement in a Secure Cloud Bursting across Two Cloud Platforms	x	x	x	optional	optional	optional	x

Table 1: Capabilities for Each Stage

4. Relevant Standards and Guidance

The following resources and references provide additional information to be leveraged to develop this solution:

- National Institute of Standards and Technology (NIST), NIST FIPS 197, Advanced Encryption Standard (AES)
  http://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.197.pdf

- National Institute of Standards and Technology (NIST), NIST Federal Information Processing Standards (FIPS) 140-2, Security Requirements for Cryptographic Modules
  http://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.140-2.pdf

- National Institute of Standards and Technology (NIST), NIST Special Publication (SP) 800-53 Revision 4, *Security and Privacy Controls for Federal Information Systems and Organizations*, April 2013 (including updates as of January 22, 2015).
  https://doi.org/10.6028/NIST.SP.800-53r4

- National Institute of Standards and Technology (NIST), NIST Special Publication (SP) 800-70 Revision 3, *National Checklist Program for IT Products – Guidelines for Checklist Users and Developers*, December 2015 (including updates as of December 8, 2016).
  https://doi.org/10.6028/NIST.SP.800-70r3

- National Institute of Standards and Technology (NIST), NIST Special Publication (SP) 800-125, *Guide to Security for Full Virtualization Technologies*, January 2011.
  http://csrc.nist.gov/publications/PubsSPs.html#800-125

- National Institute of Standards and Technology (NIST), NIST Special Publication (SP) 800-128, *Guide for Security-Focused Configuration Management of Information Systems*, August 2011.
  http://csrc.nist.gov/publications/PubsSPs.html#800-128

- National Institute of Standards and Technology (NIST), NIST Special Publication (SP) 800-137, *Information Security Continuous Monitoring for Federal Information Systems and Organizations*, September 2011.
  http://csrc.nist.gov/publications/PubsSPs.html#800-137

- National Institute of Standards and Technology (NIST), NIST Special Publication (SP) 800-144, *Guidelines on Security and Privacy in Public Cloud Computing*, December 2011.
  http://csrc.nist.gov/publications/PubsSPs.html#800-144

- National Institute of Standards and Technology (NIST), NIST Special Publication (SP) 800-147B, *BIOS Protection Guidelines for Servers*, August 2014. http://csrc.nist.gov/publications/PubsSPs.html#SP-800-147-B

- National Institute of Standards and Technology (NIST), Draft NIST Special Publication (SP) 800-155, *BIOS Integrity Measurement Guidelines*, December 2011. http://csrc.nist.gov/publications/PubsSPs.html#800-155

- National Institute of Standards and Technology (NIST), *Framework for Improving Critical Infrastructure Cybersecurity,* Version 1.0, February 12, 2014. https://www.nist.gov/sites/default/files/documents/cyberframework/cybersecurity-framework-021214.pdf [accessed 12/12/16]

5. Component List

- Commodity servers with hardware crypto module
- Commodity network switches
- Hypervisors
- Operating systems
- Application containers
- Attestation server
- Orchestration and management servers
- Database servers
- Directory servers
- Software defined network
- Data encryption and key management server
- Cloud service