AWS Nitro System API & Security Claims

Amazon Web Services, Inc.
Version 1.0 – April 11, 2023

Prepared By
John Redford
Xiang Wen Kuan

Prepared For
Amazon Web Services

1 Executive Summary

Synopsis

In the last calendar quarter of 2022, Amazon Web Services (AWS) engaged NCC Group to conduct an architecture review of the AWS Nitro System design, with focus on specific claims AWS made for the security of the Nitro System APIs. The planning and execution of this work continued into the first calendar quarter of 2023.

The following is a description of the Nitro System, abridged from public documentation[1]:

> The AWS Nitro System is a combination of purpose-built server designs, data processors, system management components, and specialized firmware which provide the underlying platform for all Amazon EC2 instances launched since the beginning of 2018. Three key components of the Nitro System are:
>
> - Purpose-built Nitro Cards — Hardware devices designed by AWS that provide overall system control and input/output (I/O) virtualization independent of the main system board with its CPUs and memory.
> - Nitro Security Chip — Enables a secure boot process for the overall system based on a hardware root of trust, the ability to offer bare metal instances, as well as defense in depth that offers protection to the server from unauthorized modification of system firmware.
> - Nitro Hypervisor — A deliberately minimized and firmware-like hypervisor designed to provide strong resource isolation, and performance that is nearly indistinguishable from a bare metal server.

AWS asserts several security claims regarding how the Nitro System is designed to prevent AWS employees from accessing customer data. The claims are enumerated in the Claims portion of this report.

The Nitro System had been designed to achieve security goals covering these claims by providing a total system administration, management, and monitoring infrastructure that operates from the hardware level up through provisioning and deployment and end-of-life of systems. AWS has no secondary or alternative paths of access to Nitro EC2 host systems. This scale of thoroughness and the appropriate alignment of security goals with sustainable and achievable business practices enabled AWS to design a system that would support very strong security and customer privacy claims.

As a matter of design, NCC Group found no gaps in the Nitro System that would compromise these security claims. All designs involve trade-offs, and AWS has chosen a design where the impact of a malicious compromise would be similar to a small-scale hardware failure.

Project Logistics

The scope of analysis covered the verification of a number of security claims around the design of the administrative Nitro APIs as well as the development and administrative processes that create and manage the Nitro APIs. AWS Operators use the Nitro APIs to perform a set of well-defined tasks. This covered the development and deployment of software elements of the Nitro System, the infrastructure it uses, and the procedures for creating and deploying a Nitro System environment.

1. The Security Design of the AWS Nitro System https://docs.aws.amazon.com/whitepapers/latest/security-design-of-aws-nitro-system/security-design-of-aws-nitro-system.html.

The scope of NCC Group's evaluation included:

- The Nitro System API, its requests, functionality, development & deployment processes and the supporting infrastructure elements which it requires.
- The role of AWS Operators who have authorized access to Nitro System APIs.
- AWS non-operator employees, including developers of the Nitro System and personnel who manage and assign roles to employees.
- Sufficient review of the Nitro System and its environment to ensure no alternative systems were present or necessitated by design.

The following elements were not in scope:

- Evaluation of EC2 Control Plane services generally.
- Evaluation of Nitro hypervisor, Nitro firmware, Nitro software on Nitro cards.
- Evaluation of Nitro Cards.
- Evaluation of the physical environment and physical security controls.

Assessment Methodology

The assessment was largely conducted via interviews with the AWS Nitro development team and through access to documentation provided by AWS. NCC Group conducted interviews with multiple Distinguished Engineers on the Nitro team, including the lead engineer. These interviews covered the origin and design goals of the system as well as its operational characteristics and constraints. Additionally, information was provided in documents and by screenshare to enable NCC Group to reach a thorough understanding of the security of the Nitro System. The documentation provided covered extensive internal details of the design.

The assessment does not include in-depth review of the implementation of specific components or any hands-on testing or technical validation. The assessment of claims was based on the degree to which the Nitro System design, as assessed and observed, provided the means to support these claims and ensure they would be maintained. As this was a design-level review, any failure to meet this goal in the design itself would result in not considering a claim to be supported.

Project Limitations

This engagement represents a point-in-time evaluation of the Nitro System APIs. Security threats and attacker techniques evolve rapidly, and the results of this assessment are not intended to represent an endorsement of the adequacy of current security measures against future threats. Furthermore, statements made by NCC Group refer to the system as presented during the assessment window, and provide no assurance with regards to any future chosen or compelled technical changes or deviations in policy. While the secure boot process and interaction of Nitro Cards was a factor in analysis, the security of these systems in the presence of a physical attacker was deferred to that other work or excluded from scope.

This review is based on the attestation of AWS staff and product design documents as presented to NCC Group. While AWS supplied suitable support in that regard, NCC Group cannot attest to the accuracy of the information or associated conclusions, or whether the implementation matches the design. Any statements about what the Nitro System does is in reference to its design.

2 Nitro System Design

Introduction

NCC Group performed an architectural design review of the Nitro System that included several weeks of discussions and analysis with additional planning and review over several months. The objective of the review was to determine whether the architecture of the Nitro System would satisfy the security claims made by AWS. The consulting team considered the system from the perspective of "AWS Operators", individuals with non-public access to the environment, as well as developers of the Nitro system and other employees of AWS.

Analysis also determined that these APIs are the only means for AWS Operators to interact with instance hosts and there are no other ways to connect, login, or obtain any other type of privileged access. Analysis also determined that no other employees of AWS other than AWS Operators can make use of these APIs.

Nitro System Design Goals

Least Privilege

The design of the Nitro System adheres to the principle of least privilege. The design included disallowing:

- Access to customer data.
- Access to execute arbitrary commands or code.
- Customization of an instance with unique patches or software versions.
- Migration of data or storage to another instance or environment.
- Unrecorded activity.
- Emergency procedures to by-pass or remove security protections.

The design of the Nitro System was aligned with the essential requirement of the environments to handle small-scale loss of availability events, such as conventional hardware failures or an accident. The scope of potential abuse by a malicious user would be similar in impact to those events.

Redundancy and Zero Trust

Components of the Nitro System are designed to perform redundant and layered security controls. The design ensures that authentication and authorization are checked multiple times and multiple paths are required for software changes to be developed and deployed. Instance hosts are designed to independently verify proper access controls before performing any requested actions.

Confidentiality and Integrity

Strong encryption and signing controls are present throughout the design of the environment. The overall design of the Nitro System includes making sure that all Nitro System communication is securely encrypted and that all deployed Nitro System components are signed and validated. To reduce the likelihood of compromising the keys that protect communications and software updates, the design makes sure that the roots of trust are located on secured systems that do not rely on lower levels of access control for their security management.

Auditing and Monitoring

All Nitro System API requests are logged into a secure CloudWatch environment with layered access controls. Logs are constantly monitored for use of sensitive requests or patterns of activity to call human attention to those systems. This monitoring and human engagement is regularly tested and performs effectively in internal "red team" exercises.

Design Incentives

This aligns with the natural incentives of AWS and its employees. In some environments, a hierarchy of access naturally arises, with a handful of administrators having complete

access to many company systems, often including the systems that control access to other systems. However, AWS employees do not have any reason to have such access to customer data. It is against their business goals for their employees to ever obtain that access. The design of the Nitro System and its development and deployment all support the incentive not to have any way for any employee to obtain access to customer data. Internal AWS personnel processes that define employee identity also reflect and support these incentives.

Public Nitro System Documentation

Many details of the Nitro System design are available in several places.

- The Security Design of the AWS Nitro System.
  https://docs.aws.amazon.com/whitepapers/latest/security-design-of-aws-nitro-system/security-design-of-aws-nitro-system.html
- Model Checking Boot Code from AWS Data Centers.
  https://link.springer.com/chapter/10.1007/978-3-319-96142-2_28
- AWS re:Inforce 2019: Security Benefits of the Nitro Architecture (SEP401-R).
  https://www.youtube.com/watch?v=kN9XcFp5vUM

3 Claims

Product Security Design Claims

AWS attests that production Nitro System hosts will adhere to an explicit policy of protecting data privacy of customers with the following claims:

1. There is no mechanism for a cloud service provider employee to log in to the underlying host.
2. No administrative API can access customer content on the underlying host.
3. There is no mechanism for a cloud service provider employee to access customer content stored on instance storage and encrypted EBS volumes.
4. There is no mechanism for a cloud service provider employee to access encrypted data transmitted over the network.
5. Access to administrative APIs always requires authentication and authorization.
6. Access to administrative APIs is always logged.
7. Hosts can only run tested and signed software that is deployed by an authenticated and authorized deployment service. No cloud service provider employee can deploy code directly onto hosts.

Claim Analysis

1. There is no mechanism for a cloud service provider employee to log in to the underlying host.

By design, systems expose no mechanism that can provide access to a shell or any other such mechanism for the execution of arbitrary commands. There is no means to enable or deploy such a mechanism. No exceptional or external mechanisms exist to provide this capability.

Analysis: NCC Group finds that the architecture of the Nitro System fully supports this claim. There is no indication that a cloud service provider employee can obtain such access or any equivalent access to any host.

2. No administrative API can access customer content on the underlying host.

The administrative APIs do not perform any activity to access or reveal customer content. There are no APIs that could cause content to be moved to another location where it could be accessed. There are no APIs that would reduce or remove protections on customer content.

Analysis: NCC Group finds that the architecture of the Nitro System fully supports this claim. The administrative APIs cannot access customer content on the underlying host. That functionality does not exist.

3. There is no mechanism for a cloud service provider employee to access customer content stored on instance storage and encrypted EBS volumes.

The administrative API does not include any functionality that would provide access to customer content on instance storage. It is not possible to use the API to create conditions in which this would become possible. The instance storage elements are all encrypted at rest, as are encrypted EBS volumes.

Unencrypted EBS storage volumes do still exist as a customer choice.

Analysis: NCC Group finds that the architecture of the Nitro System fully supports this claim. There is no mechanism by which a cloud service provider employee can access customer content stored on host instances or in encrypted EBS volumes.

4. There is no mechanism for a cloud service provider employee to access encrypted data transmitted over the network.

The encryption used to protect Nitro related data and other AWS managed encryption makes use of appropriate algorithms and secure key management. The TLS 1.2 protocol is used to negotiate encrypted connections. The keys directly used by Nitro hardware are stored on locally-encrypted storage protected by a tamper-resistant TPM chip. All administrative API communications and AWS-managed communications are securely encrypted. AWS makes use of secure protocol version and algorithm variants and can rapidly migrate to newer ones as necessary.

Analysis: NCC Group finds that the architecture of the Nitro System fully supports this claim. There is no mechanism for cloud service provider employees to gain access to encryption keys or disable communication encryption.

5. Access to administrative APIs always requires authentication and authorization.

The administrative APIs require requests to contain a bearer token that provides both authentication and authorization data. These tokens are generated and provided to an authorized operator in accord with the identity of that operator and the access rights associated with them. Tokens only grant access to the resources for which they are issued and expire shortly. Access rights are described by the association of operator groups with specific APIs and sets of managed resources. Access rights are limited by quotas on resources affected in a given time period.

The rights configuration data itself is quickly auditable and not excessively complex, so users do not have excessive or inappropriate rights. The access rights configuration data is defined with the token provider service and follows the same peer review and change control procedures.

Analysis: NCC Group finds that the architecture of the Nitro System fully supports this claim. The bearer token system implemented prevents a malicious user from reusing the token outside its authorized intent. Token expiration times are long enough to avoid issues if the issuing service is disrupted, but short enough to limit misuse. Locating the configuration of access rights with the code for the system that provides bearer tokens ensures that both have equal procedural protection against malicious alteration. Quotas prevent any abuse of authorized access from impacting an unreasonable number of systems.

6. Access to administrative APIs is always logged.

All access events, including failed authentication or authorization of requests, will be logged immediately into a dedicated CloudWatch log stream created and managed by the AWS Nitro development team. These events are constantly monitored for any significantly unusual or suspicious activity or patterns of activity.

Analysis: NCC Group finds that the architecture of the Nitro System fully supports this claim. The monitoring process is configured to identify requests that would indicate abuse by a malicious agent or otherwise inappropriate use of the administrative APIs.

7. Hosts can only run tested and signed software that is deployed by an authenticated and authorized deployment service. No cloud service provider employee can deploy code directly onto hosts.

Integrity protection of the software and self-updating mechanisms of the Nitro environment is present from the point of manufacture. The process through which systems boot, acquire identity and perform higher-level software loading and management functions are secured from this original point. It is possible to functionally return a Nitro component to its initial secure state. From this condition, it can be given identity in production or non-production

environments. No movement of a component to another environment is possible, due to controls in both the Nitro components and the environments.

Analysis: NCC Group finds that the architecture of the Nitro System fully supports this claim. The procedures and practices used to develop and authorize software would not enable any malicious agent, even with AWS Operator or Nitro developer access, to inject unauthorized functionality into the system. The incentives of all authorized persons align with preventing this from occurring.

4 Documents Reviewed

The following documents were reviewed during this assessment.

Public AWS Documentation

These documents are available to the general public and recommended to obtain a more complete understanding of the Nitro System.

- The Security Design of the AWS Nitro System. https://docs.aws.amazon.com/whitepapers/latest/security-design-of-aws-nitro-system/security-design-of-aws-nitro-system.html
- Model Checking Boot Code from AWS Data Centers. https://link.springer.com/chapter/10.1007/978-3-319-96142-2_28
- AWS re:Inforce 2019: Security Benefits of the Nitro Architecture (SEP401-R). https://www.youtube.com/watch?v=kN9XcFp5vUM

Internal AWS Documentation

These AWS internal documents were kept on an AWS system for reading and no copies were made.

- **Nitro Controls Document** - Thorough design documentation for the Nitro System covering the controls to prevent operator access to customer content. The content of this document included a detailed overview of the Nitro System and its components, description of the trusted computing base (TCB), description of how Nitro APIs work, threat models, and more.
- **Nitro Security Design** - An early design document that covered the design goals and constraints of the Nitro System.
- **EC2 Public Key Infrastruture** - Documentation on the public key infrastructure that provides the foundation of trust between internal EC2 components.
- **Content of Bearer Tokens** - Specific documentation covering the structure of bearer tokens used by the Nitro System.
- **Nitro Pipeline** - Documentation on the orchestration agent used for Nitro software deployments. The Nitro Pipeline is the bridge between standard Amazon tooling, testing suite service, and other EC2 services and tooling. Parts of this document were redacted.