Rapid7 Insight Platform:

Cloud Security Overview

Contents

01
INTRODUCTION

Whether you're a Rapid7 customer today or tomorrow, it is important to us that you understand how Rapid7 secures the data you entrust to our Insight Platform. Just as within your own security-conscious organization, the initial proposal to develop software in the cloud was met with healthy opposition and a series of debates. Since our earliest days here at Rapid7 we have a sole focus on what our customers need, and thus we did not develop our cloud solutions to purely speed our time-to-market and simply apply security later. We carefully built the Insight Platform to equip security teams with modern data processing without the significant overhead of managing the infrastructure. We worked with our customers to methodically design the security controls necessary to ensure we are reducing the risk of compromise from the first day we started to provide these benefits.

> "Trust is the basis of relationships between individuals and companies"
>
> — Corey Thomas, Rapid7 CEO

Just as no two organizations' networks are the same, each cloud service is unique. Our well-read 2013 research into publically available data in Amazon Web Services (AWS) shed some light on the topic and it serves us well to reiterate what we wrote at the time:

"It should be emphasized that a public bucket is not a risk created by Amazon but rather a misconfiguration caused by the owner of the bucket."

It will be no surprise now that simultaneous to when this research was underway, our Products organization was evaluating multiple infrastructure-as-a-service (IaaS) providers and we determined the Insight Platform would best meet our customers' needs if built on AWS. Not only had we done our research, we had published it, critiqued it, and even shared it with the vendors. Based on this extensive work, it was clear that Amazon's continual release of innovative security controls gave us advantages we couldn't have realized were we building a traditional SaaS infrastructure in our own data centers.

In keeping with our usual Rapid7 approach, we mapped out the reasons someone would attempt to compromise our cloud, what data they would seek, and what methods they would use. We know attackers well because we help you defend against them every day, so it should be only natural for us to use this same mentality for our own solutions. This exercise is regularly run at Rapid7 and it steers all of the decisions we make to exceed the standards and expectations of our customers. As this was being discussed and tested, we focused on four key aspects of importance for securing our cloud:

1. How we collect your data
2. How we process your data
3. How we scale our infrastructure
4. How we automate our delivery

These are, not surprisingly, also the four primary benefits for which we designed the Insight Platform, and each has different security controls built directly into its core. In addition to all of these measures, we do exactly as we advise our own customers: we test their effectiveness against an attack. The Rapid7 team performs regular penetration tests and web application scans, but we also require penetration tests from parties not associated with Rapid7 to ensure unbiased results. Nothing is taken more seriously by the Rapid7 cloud development teams than potential risks discovered in these security assessments.

Let us now look at each of these four areas and how these controls were thought through.

02

SIMPLIFYING DATA COLLECTION FOR OUR CUSTOMERS

Rapid7's software, no matter the solution, is built to provide value to you, and this necessitates we make it easy for you to collect the relevant data for your security use cases. Whether the data provides an understanding of your organization's exposure to an attack or suspicious user behavior, our solutions must have access to various types of data, ranging from extremely sensitive to what may seem unimportant, in isolation. In considering the risks to our customers--and associated value to attackers—we designed all data collection and transmission to lower the possibilities of interception, impersonation, data mixing, and data poisoning.

Never Hold Onto the Attacker's Ultimate Prize

Much of the data collection for our solutions requires access to credentials with a high level of privilege on your networks. In our exercises to map out the motives and goals of an attack on the Insight Platform, these credentials were the ultimate prize – with them, an attacker can impersonate a legitimate user on a customers' network and move laterally into other systems. Considering this high risk and that the credentials only have value for the data collection taking place on your networks, we designed the Insight Platform to **never** have access to them in plain text. They are encrypted on the Collectors residing on your networks before transmission, and only ever decrypted on the Collector using a combination of the Collector's private key and the separate necessary parameters, which it must obtain from the cloud.

Only Trust the Right Source of Information

In considering how an attacker could attempt to impersonate a Rapid7 Collector or the Rapid7 Insight Platform, we established a single communication model that all data transmission must follow. Before any data can be transmitted from a customer's Collector to the Rapid7 Insight Platform, the Collector must first be registered with a dedicated customer's instance in the cloud via its unique activation key and fingerprint. Upon registration, all data from this collector will only ever be accepted by the corresponding customer instance with which it was registered. Any given activation key can be used for registration only once and the collector will only trust a recipient with its data upon verifying trust certificates with a signature chain meeting very specific criteria.

Every payload must be sent using this trust relationship, and the channel is never left open. All data transmitted to each customer's dedicated cloud instance is compressed and sent to the Insight Platform over the encrypted TLS channel. Each transmission must be initiated by the customer's registered Collector after mutual authentication has occurred. The cloud cannot initiate the communication with customer Collectors; it can only wait for requests from the Collector and respond with any necessary instructions for software updates or configuration changes. Communications to the cloud must be completely ignored if they are begun without the verification of an established trust relationship. There can be no exceptions to these rules.

03

PROVIDING ADVANCED ANALYTICS THROUGH MODERN DATA PROCESSING

After more than a decade packaging our software to run on-premise, the long internal debate over whether we would be adding enough value with cloud solutions finally ended when recent technologies opened new possibilities for processing massive amounts of data. With the advancement of IaaS providers, we could take advantage of cutting edge technology without the need to manage a new server every time we wanted to experiment with a new use case. After evaluating all of the market-leading providers, we chose the one we consider the most security-minded and innovative. Amazon runs one of the world's largest networks of web sites, and since early 2006, Amazon Web Services (AWS) has provided companies of all sizes with an infrastructure platform that powers business applications of tremendous scale.

With AWS, we have the ability to develop and run the advanced analytics you need with the right processing and storage technology for each. Our solutions that take advantage of the Insight Platform rely on various NoSQL and relational databases to store and process your data. Each Rapid7 customer is assigned their own relational database schema, which houses all asset names, other human-readable descriptions, and various public keys that support broader security processes related to your infrastructure. Much of the data processed and stored is encrypted at rest using various file or disk level encryption mechanisms.

Together, Rapid7 and AWS have a comprehensive approach to ensure security and reliability of the Rapid7 service. It starts with the physical datacenter, extends through the computer, network, and storage layers of the service, and is complemented by well-defined access policies and ongoing audit and certification by 3rd parties.

Because "The Cloud" Is Still Made of Physical Servers

A secondary benefit of choosing Amazon as our IaaS provider was the mitigated risk of an attacker (or even our engineers) not having the ability to locate the physical servers running our software. AWS datacenters are housed in nondescript facilities. Physical access is strictly controlled both at the perimeter and at building ingress points by professional security staff utilizing video surveillance, state of the art intrusion detection systems, and other electronic means. Authorized staff must pass two-factor authentication a minimum of two times to access datacenter floors and all physical access by employees is logged and audited routinely. When an employee no longer has a business need for these privileges, their access is immediately revoked, even if they continue to be an employee of Amazon.

Datacenter access and information is only provided to employees and contractors who have a legitimate business

SAS 70 Type II

AWS has successfully completed a SAS70 Type II Audit, and is committed to continuously maintaining the appropriate security certifications and accreditations to demonstrate the security of their infrastructure.

AT-101 (SOC 2) Security Trust Service Principles

In addition to our own security policies, regular penetration tests, and application scans, we have a SOC II Type 2 in place for the foundation of our platform and are continuing to expand the specific compliance regimes for which we are audited.

need for such privileges. All visitors and contractors are required to present identification and are signed in and continuously escorted by staff.

Since Securing Your Own Cloud Is Enough Work

When considering other potential avenues of attack, we had to protect the Insight Platform against scenarios in which a different AWS customer gets compromised. By leveraging Amazon EC2 Security Groups in Virtual Private Clouds (VPCs), we logically isolate an extensive number of our services from both one another as well as the outside world. In addition to these controls, the AWS network provides protection against traditional network security issues including:

- *Distributed Denial of Service (DDoS) Attacks:* AWS network infrastructure leverages proprietary DDoS mitigation techniques developed as a result of running the world's largest online retailer. Additionally, AWS's networks are multi-homed across a number of providers to achieve Internet access diversity.

- *Man in the Middle (MITM) Attacks:* Amazon EC2 VMs automatically generate new SSH host certificates on first boot and log them to the instance's console. Rapid7 leverages these secure APIs to access the host certificates before logging into an instance for the first time.

- *IP Spoofing:* Amazon EC2 VMs running the Rapid7 service cannot send spoofed network traffic. The AWS controlled, host-based firewall infrastructure does not permit an instance to send traffic with a source IP or MAC address other than its own.

04

AUTO-SCALING INFRASTRUCTURE TO MEET ANY SIZE CUSTOMER'S NEEDS

To truly offer you horizontally scalable solutions without any risk of one customer accessing another's data, we designed the Rapid7 Insight Platform around secure, multi-tenant services from its inception. Each data collector has its trust established with a specific instance in the cloud and any attempt to transmit data to a different instance would be completely ignored by the application to which the data is being sent. Similarly, each end user's account is tied only to the analytics and data to which its company owns, so the users experience each application as if it were built solely for their organizations.

To prevent any data sovereignty violations, every customer's instance of an application on the Rapid7 Insight Platform can only reside in that organization's global region of choice. While data replication is instrumented within each regional cloud, it will never be pooled across them. Given the extensive measures built into each cloud for redundancy and avoiding data loss during isolated service outages, Rapid7 has opted to use each region to host effectively independent clouds running the same version of each application. The flexibility to choose where you would like your data to reside is very important to us, especially given the variation in regulations affecting each customer.

To combat the possibility of Rapid7 employees getting their accounts compromised, because we have to recognize that it's possible for anyone, The Insight Platform's supporting infrastructure is designed to be fully automated to ensure security policies are consistently applied. These policies include two-factor authentication, bastion/jump hosting, service segregation, and by-service defined permissions ensuring least-privilege and access methodologies are applied.

05
CRAFTING AUTOMATED SOFTWARE DELIVERY TOOLS

Given the fast pace of the threat landscape and new exposure discoveries, the Rapid7 Platform Delivery team instruments the necessary tools to support the continuous deployment model we designed to respond to your needs in the fastest way possible. The combination of automated Amazon Machine Image (AMI) instance baking, least-privileges required, and AWS instance roles create a non-permissive environment that mitigates the ability of an attacker to easily move throughout the Rapid7 Insight Platform if they gained access to a specific system.

To remain secure while deploying software to production throughout the day, we use AWS Instance Roles to define the specific restricted-permission sets based on the server type. Instance roles dynamically receive AWS credentials used to authenticate with other AWS resources, eliminating the need to hard code or store credentials in a configuration service. Furthermore, the AWS Instance Role credentials are automatically changed by AWS frequently. The AWS SDK encapsulates the exact key rotation logic, but it is documented by AWS that the validity of these temporary credentials never exceeds a one-hour period.

Many SaaS vendors lose hours of productivity to manage an effective patch management program across the entirety of their data centers [unless they opt to rarely patch]. Cognizant of this challenge and the unacceptable risk of running our software on vulnerable servers, the Platform Delivery team built much of the Insight Platform software deployment process on AMI baking using Chef. As soon as a patch is released, we simply need to update the impacted base AMIs and restart them. By automating this within the deployment process, it ensures that each time new software is deployed or a solution is horizontally scaled, it is running on a fully-patched, properly configured virtual server.

To ensure our solutions are available when you need them without introducing risk through direct access, every EC2 instance (regardless of server type) is granted the necessary privileges to enable centralized, real-time monitoring of our servers and automated alarm notification email delivery.

One example of these mitigating controls and how they fit into the aforementioned exercise of thinking through how an attacker would attempt to compromise your data goes as follows:

- A data normalization server instance is granted permissions to read and poll the raw data upload buckets
- Every UI (web) server instance has no permissions whatsoever to access any S3 bucket
- If an attacker were to gain access to one of our AWS EC2 servers powering an application's web interface, they would not have permission to access raw data in S3

We have open sourced many components we've built to automate and secure our platform. If you'd like to take advantage of these components for your own software development, our public github repositories make them available for both use and contribution.

06 CONCLUSION

At Rapid7, we built the Insight Platform to meet our customers' evolving needs without demanding security professionals spend their time managing hardware, architecture, or scale. Unlike the many organizations which have attempted to add security later, every design decision and process proposal from the first day was evaluated for the risk it would introduce and security measures necessary to reduce it. We constantly strive to safeguard your data while incorporating cutting-edge technologies to more effectively address your needs.

We understand the inherent trust you are placing in us from the first byte of data you collect with our solutions and take this very seriously. Each aspect of our software, third-party technologies, infrastructure, and software development lifecycle involves deliberation and is opened to criticism from other parties. We have written this paper and deployed a Trust website in an attempt to be as transparent as possible with the public without revealing enough detail to put our customers at risk. If you would like to know more than is provided here or believe you can improve upon our approach, we welcome the conversation.

07 RECOMMENDED MATERIALS

If you'd like to learn more about Rapid7's approach to security, privacy, and trust, visit https://www.rapid7.com/trust/.

If you would like to know more than is provided in this paper, our development team created a more in-depth Technical Primer we make available to our customers under NDA.

08
ABOUT RAPID7

Rapid7 is a leading provider of security data and analytics solutions that enable organizations to implement an active, analytics-driven approach to cyber security. We combine our extensive experience in security data and analytics and deep insight into attacker behaviors and techniques to make sense of the wealth of data available to organizations about their IT environments and users. Our solutions empower organizations to prevent attacks by providing visibility into vulnerabilities and to rapidly detect compromises, respond to breaches, and correct the underlying causes of attacks. Rapid7 is trusted by more than 5,300 organizations across 100 countries, including 36% of the Fortune 1000. To learn more about Rapid7 or get involved in our threat research, visit www.rapid7.com.