RAPID7 INSIGHT PLATFORM SECURITY

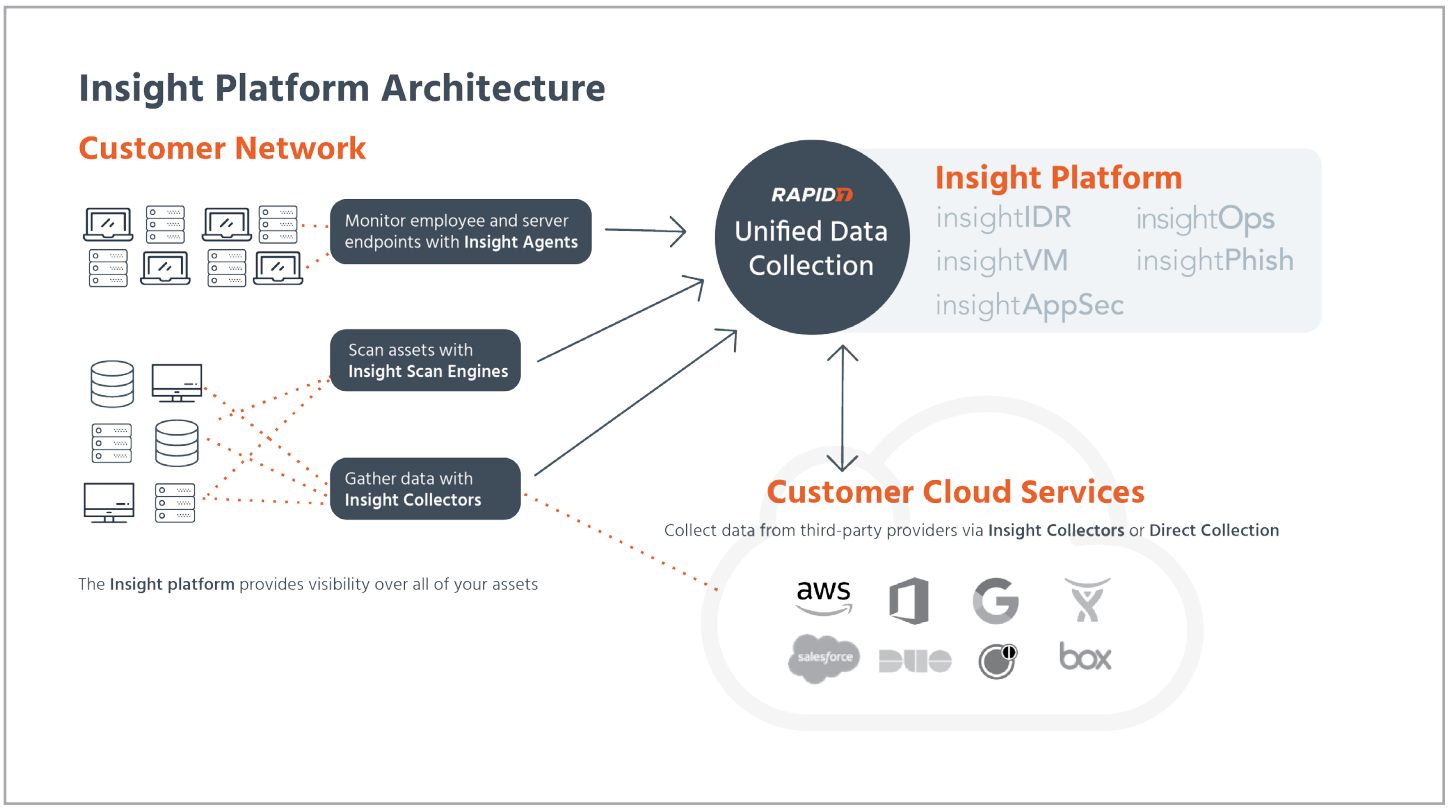
Understanding the architecture, security mechanisms, and technical foundations that make up the Rapid7 Insight platform

051618

TABLE OF CONTENTS

OVERVIEW

The Rapid7 Insight platform provides data collection, visibility, analytics, and automation to establish a shared point of view between security, IT operations, and DevOps. Our cloud platform delivers one-click access to Rapid7's vulnerability management, application testing, incident detection and response, phishing analysis and simulation, and log management solutions. This document introduces the architecture, security mechanisms, and technical foundations that make up the Rapid7 Insight platform.



DATA COLLECTION

The Insight platform offers multiple options for collecting data across your IT environment. Whether you use collectors, the Rapid7 Insight Agent, scan engines, or direct connections to our platform, our unified data collection enables your teams to collect data once and use it across multiple products on the Insight platform. Once configured, data sources continuously collect data, enabling teams to collaborate effectively as they analyze shared data, alert on risk vectors, and automate remediation and breach response.

Collection Methods

Collectors

Rapid7 uses collectors to gather information from on-premises and cloud networks and to securely transfer data to the Insight platform. Collectors sit behind the client's firewall, respond to changes in the environment, and securely transmit relevant data to the Insight platform for analysis. Collectors were designed with the following core tenets in mind:

- Collectors can be configured only by administrators.
- All data is secured during the transmission process, which uses strong encryption protocols.
- Data transferred from each separate collector is uniquely identified and stored and cannot be accessed by any third parties.

During installation, a collector generates an activation key and a fingerprint. An administrator uses the activation key to pair the collector with your Insight platform organization.* Once paired, the Insight platform will verify a collector's identity by performing a challenge-response handshake using the known shared secret (i.e., the fingerprint).

The collector relies on a TLS connection (HTTPS) to communicate with the Insight platform. Specifically, the collector is explicitly coded to trust only certificates that have a signature chain that can be vetted by one of the Java trust store Certificate Authorities (CAs). Once the challenge-response handshake is complete, the collector is ready to accept command and control instructions from the Insight platform. For security reasons, the collector always reaches out to the Insight platform; the Insight platform cannot reach through the client's firewall and initiate a conversation.

**InsightVM embeds a collector inside the console that is automatically paired using your InsightVM license key.*

Rapid7 Insight Agent

The universal Insight Agent is lightweight software you can install on any asset—in the cloud or on-

premises—to easily centralize and monitor data on the Insight platform. The Insight Agent gives you endpoint visibility and detection by collecting real-time system information—including basic asset identification information, running processes, and logs—from your assets and sending this data back to the Insight platform for analysis. The Insight Agent can be installed directly on Windows, Linux, or Mac assets. Each Insight Agent only collects data from the endpoint on which it is installed.

The Insight Agent authenticates using TLS client authentication. When you deploy the Insight Agent, the deployment includes a private SSL key representing your organization. This key is used to authenticate and authorize your agent with the Insight platform.*

The agent can communicate directly to the Insight platform, or proxy communication through Insight collectors on your network. Finding the best route to the Insight platform occurs automatically or can be configured in advanced use cases.

**For InsightOps log data, an API token is used to authenticate the Insight Agent instead of TLS client authentication. Log data is encrypted in transit via TLS.*

Scan Engines

On-premises scan engines are used by InsightVM and InsightAppSec to scan Rapid7 customers' environments by finding and remotely connecting to systems to retrieve asset information.

InsightVM scan engines perform vulnerability scans of your networks and report results back to the InsightVM console. Engines can be distributed across internal networks, public networks, and cloud providers. Scan engines provide strategic views of your network from an attacker's perspective. In deciding how and where to deploy scan engines, you choose how you would like to "see" your network. Scan engines can be configured to perform authenticated scans to check for software applications and packages and to verify patches.

InsightAppSec scan engines allow scanning internal web applications that aren't accessible to the public internet. Engines connect to the web applications you configure and report results back to the Insight platform. For security reasons, the InsightAppSec scan engine always reaches out to the Insight platform for instructions; the Insight platform cannot reach through the client's firewall and initiate a scan.

Direct Collection

Rapid7 collects some data directly via a connection with the Insight platform. The Insight platform can connect to third parties on your behalf, such as container registries for InsightVM or IMAP servers for InsightPhish. You can also send data directly to us via our APIs or with one of our software libraries or extensions, such as the InsightOps application logging libraries or the InsightPhish Outlook/O365 extension.

For details on the collection methods and specific data collected for each product, please visit help.rapid7.com.

Credential Storage

Much of the data collection for our products requires access to credentials with a high level of privilege on your networks. Credentials are stored differently depending on where and how they are used. Credentials are always encrypted before being stored in the Insight platform. Where possible—such as when collecting data with collectors—credentials can only be decrypted by on-premises components and not the Insight platform itself.

Collectors

All credentials used by the collector to obtain data from your local environment are strongly encrypted in a manner that prevents the passwords from being decrypted based on the information stored in the cloud. Every collector installation generates a unique public/private key pair that is then split across environments. The public key is uploaded to the cloud, and the private key is stored locally on the collector. When writing the private key to the local disk, the collector encrypts the private key contents. The private key can only be decrypted using information obtained via successful communication with the Rapid7 Insight platform, thus only active, live collectors with healthy communication with the cloud can access the private key.

Whenever a credential is added to the Rapid7 Insight platform, the password is encrypted with RSA PKI (4096 bit keys) for every collector using each collector's public key. It is then persisted in the database within a client-specific database schema. Once the credential is stored, the Insight platform no longer has access to the cleartext credential. When the collector needs to use a credential, the encrypted password is retrieved from the cloud, and decrypted within the collector using the collector's private key. The password is used in memory and cleared without ever being stored to disk. A collector can only request access to the credentials necessary for the event sources configured on that collector.

InsightVM Console

The InsightVM security console, the on-premises component of InsightVM, stores credentials used for authenticated scanning unless the user utilizes a credential management system that the InsightVM console can integrate with (e.g. CyberArk). Scan credentials are encrypted with a combination of RSA and 3DES. InsightVM does not store scan credentials in the Insight platform.

DATA PROCESSING AND STORAGE

Rapid7's quest to accelerate insight for security and technology practitioners requires collecting and processing an enormous amount of data. The Insight platform's analytics engine relies on various NoSQL and relational databases, as well as S3 and other AWS services to process and store your data.

Geographic Location

The Insight platform offers different regions for storage to help you comply with policies or preferences for the physical storage location of your data. Customers can select from three cloud regions. Rapid7 will not move data from the region you select, and data is not replicated across other regions.*

- North America: United States and Canada
- Europe: Germany and Ireland
- Asia Pacific: Japan (InsightVM only)
- Australia

**Log search data for InsightIDR customers provisioned before September 2017 is stored in Europe.*

Encryption at Rest

Much of the data processed and stored is encrypted at rest using various file or disk level encryption mechanisms. Data is encrypted using industry standard AES-256 encryption with keys managed through AWS's Key Management Service (KMS). Where possible, Rapid7 utilizes AWS's services to manage encryption at rest (e.g. S3, EBS, RDS, etc.). When not possible, Rapid7 utilizes block level encryption provided by LUKS.

Encryption in Transit

Data sent to and from the Insight platform, including data collected by collectors, agents, and engines; data ingested via APIs and plugins*; and interaction with the user interface is encrypted with TLS (HTTPS). Collectors, agents, engines, and plugins are configured to verify and require a valid TLS certificate issued by a trusted certificate authority.

**InsightOps libraries allow you to send data encrypted with TLS, but you can also send log data over unencrypted connections to support legacy connections.*

Data Separation

To offer you horizontally scalable solutions without any risk of one customer accessing another's data, Rapid7 designed the Insight platform around secure, multi-tenant services from its inception. Each organization is assigned its own relational database schema within database instances. Data stored in object stores or distributed file systems is tokenized using a unique UUID that logically separates each customer's data from each other.

Data Reliability

The Insight platform is composed of a collection of disparate server types that host a set of services that enable Rapid7 products. Each service is designed to scale horizontally. Each layer of the data collection and processing pipeline is designed to be fault tolerant and to continue to operate in the event of reliability issues with our cloud environment. If one component of the Insight platform is unavailable, other components will store data until the component is available again. All persisted data is stored redundantly so that the loss of a single server or an entire availability zone should not result in data loss. All infrastructure is monitored for performance, availability, and reliability. Operations staff are available 24/7 to respond to incidents.

In addition to a redundant and fault tolerant architecture, customer data is backed up in a variety of ways. Rapid7 relies on Amazon S3 for storing data backups. Backups are not replicated outside of the region customers selected when creating their Insight platform account, but data is replicated across multiple data centers within the region. S3 can withstand the concurrent loss of data in two different facilities. Artifacts from critical stages of the data processing pipeline are backed up as data is processed. Log data backup occurs in real time as it is ingested. Automatic database backups occur daily.

Physical Security

AWS Data Centers

Data is stored in AWS data centers which 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 need for such privileges. All visitors and contractors are required to present identification and are signed in and continuously escorted by staff. More information can be found here: https://aws.amazon.com/compliance/data-center/controls/.

Rapid7 Offices

There are various risk-mitigating physical and logical security controls in place, such as security guards at front desks or locked office entrances controlled by electronic badge access, automatic screen locking, and drive encryption. All visitors must check in first when they enter Rapid7 facilities and must be escorted when entering sensitive areas.

Data Destruction

If you opt to leave a Rapid7 service, you'll have the opportunity to collect and transfer any data that is possible to export.* Should you request deletion of the data, the Rapid7 team will initiate the process within 14 days. When a user is using multiple Rapid7 products and leaves or cancels a product which makes use of shared data, the shared data is not deleted if any other shared data products are still enabled for that user's account.

**Data export tools exist for InsightVM and InsightOps.*

ACCESS TO DATA

Ensuring your data is used only in a manner consistent with your expectations is a responsibility Rapid7 takes very seriously. For this reason, policies including two-factor authentication, jump hosts, service segregation, and per-service permissions ensuring least-privilege access methodologies are applied.

Who From Rapid7 Can Access Your Data?

- Sales, Marketing, and other customer support teams have access to contact information, sales data, and usage information for product support and product analytics.
- Sales and Solution Engineers only have access to your sensitive customer information, such as user, network, vulnerability, incident, or asset information, if you choose to use a production environment for a proof-of-concept.
- Support, Software Developers, and Operations Engineers have limited access to data to support application development and troubleshooting. Rapid7 does not access sensitive customer information, such as user, network, vulnerability, incident, or asset information, unless you have explicitly requested or authorized us to do so to diagnose or troubleshoot issues with our service.

Access Control

Rapid7 provisions all network and application access using the principle of least privilege. All access requests are documented and approved by system owners.

Additionally, Rapid7 requires two-factor authentication for remote management access to our jump hosts and our backend production systems and environments. In accordance with NIST recommendations, Rapid7 explicitly disallows SMS and phone call-based two-factor authentication for remote management. This two-factor authentication includes multiple factors at each step (e.g., connection to jump host AND connection to backend servers). VPN or direct corporate LAN access is required before connecting to jump hosts, and a valid jump host session is required before connecting to any production infrastructure.

APPLICATION SECURITY

Rapid7 products on the Insight platform are designed to fit securely into your environment and adhere to security best practices. Rapid7 takes several steps to protect against common attack vectors and regularly performs application security testing, vulnerability scanning, and penetration testing.

Single Sign-On

Products that authenticate via the Insight platform share the same user database and authentication mechanisms. Customers who use multiple Insight products benefit from the Insight platform's single sign-on (SSO) functionality, needing only to sign in to the Insight platform once to access all of their Insight products. Additionally, customers can customize some aspects of their user account authentication policies, such as multi-factor authentication prompt settings. Sessions expire after 30 minutes of inactivity. User account credentials for the Insight platform are hashed using Bcrypt with a high number of iterations to protect the credential.

Customers with on-premises InsightVM consoles authenticate to the Software-as-a-Service (SaaS) portion of InsightVM via their console. The InsightVM console supports local authentication, LDAP, Active Directory, SAML, and Kerberos authentication. Local two-factor authentication can be enabled as well. User account credentials used for local authentication to the security console are salted and hashed, with RSA being employed as part of the hashing process.

Role-Based Access Control (RBAC)

The Insight platform supports global roles that apply to all products* on the Insight platform, and specific roles applicable to specific products. Rapid7 is always working to add new roles and permissions to our products, including the ability to customize RBAC for your needs.

Global Roles:

- Platform Admin - Full control over all products. Platform admins have full access to user management, including adding and deleting users, viewing all data, and performing all functions. This role can also manage product trials.
- Product Admin - Full control over a single product. Product admins can view all data, perform all edit functions, and access any admin functions within their product.
- Product Read/Write User - Able to access all or most features within a product except for administration of users and some settings. Able to modify data and/or some settings.
- Product Read-Only User - Able to access some features within a product with read-only access.

**RBAC for InsightVM is controlled via the on-premises InsightVM console. InsightVM uses customizable fine grained roles and permissions that differ from the global roles above.*

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 and providing cloud infrastructure for many large enterprises and governments. Additionally, AWS's networks are multi-homed across a number of Internet service providers to achieve Internet access diversity. Insight platform services scale horizontally behind load balancers to further mitigate DDoS attacks.

Man in the Middle (MITM) Attacks

By default, all communication with Rapid7's cloud instances occur over authenticated channels*. HTTPS traffic is secured with TLS and authenticated using trusted Certificate Authorities to prevent MITM attacks.

**InsightOps libraries allows you to send data encrypted with TLS, but you can also send log data over unencrypted connections to support legacy connections.*

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.

CLOUD SECURITY ARCHITECTURE AND GOVERNANCE

In addition to designing security into each layer of our products, Rapid7 also builds security into every aspect of our cloud architecture that hosts the Insight platform. Misconfiguration of cloud infrastructure continues to be a leading attack vector against SaaS companies. This section describes how Rapid7 implements, validates, and monitors the cloud security architecture.

Least Privileged Design

The principle of least privilege and separation of duties is built into every layer of our cloud infrastructure:

- **AWS account separation and access**
  The Insight platform uses a microservice architecture consisting of several small services working together. These services are logically separated into several different AWS accounts to shrink the blast radius of security incidents. Each AWS account contains a grouping of related services that provide a single product or product feature. Developer and operations employees are granted least privilege access to each AWS account individually as needed to perform their jobs. Employees access AWS consoles and APIs via Rapid7's corporate SSO system which requires two-factor authentication. VPC networks are not peered across accounts. Cross account communication is permitted where needed using least-privilege IAM roles or authenticated REST services. No direct database access is allowed between accounts.

- **Subnet separation**
  Several different network subnets exist within a single account. Services are provisioned to the appropriate subnet for their purpose. For example, databases are placed in a subnet with no route to the Internet to mitigate the risk of data exfiltration.

- **Host-level firewalls**
  Each set of identical Insight platform services are assigned to a separate security group, which acts as an independent firewall for that service. Security groups deny network traffic by default, so all network traffic rules are whitelist-based and are defined to allow services to communicate with each other using only the specific ports and protocols necessary for them to function together. This mitigates the risk of lateral movement between instances comprising each Insight platform service.

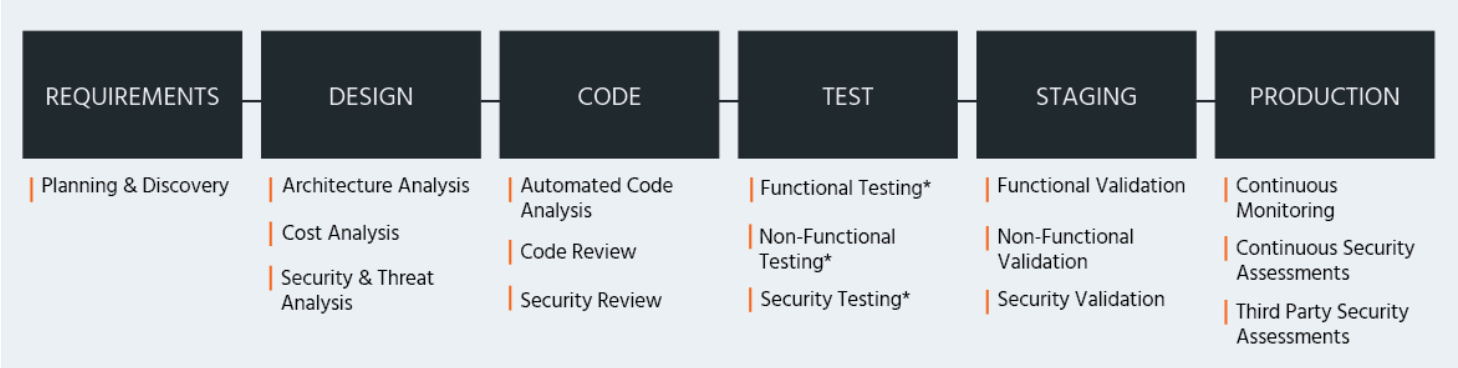
- **Service-level roles**
  When Insight platform services need to access AWS services, (e.g. S3, KMS, SNS/SQS, etc.), their access is permitted via IAM roles attached to each group of identical services. Credentials for these roles are managed by AWS and regularly rotated. These roles allow least privileged access to cloud resources.

Each service type has its own IAM role. For example, if a microservice needs access to read from S3, its IAM role would only permit reading from a specific S3 bucket. No write access would be permitted.

- **No direct access**
  Access to internal services is guarded by user authentication, IP address whitelisting, and two-factor authentication. Internal services such as jump hosts and back-office admin portals can only be accessed from the Rapid7 corporate LAN or VPN. Admin SSH access to backend services, databases, and other infrastructure must transit through a jump host.

Secure Software Development

Rapid7's Engineering teams follow a formally documented SDLC process which is based on Agile/Scrum methodologies. This process includes peer code review, automated testing, and scenario testing to ensure quality and to identify security vulnerabilities prior to shipping.



**Automated and manual*

Changes to cloud infrastructure are orchestrated in code. When making an infrastructure configuration change, these tools provide a view of what will change and a mechanism to rollback if a problem occurs. Infrastructure is regularly monitored for manual changes made outside of these tools. In the event infrastructure is changed outside of code, it can quickly be brought back in compliance.

All changes are peer reviewed. Changes require at least two engineers' approvals before being deployed to production.

Security requirements are part of the Requirements Gathering and Design phases to ensure security best practices and principles are incorporated early on in our SDLC.

Configuration Scanning

Rapid7 performs regular automated scans of our cloud infrastructure with a suite of tools to ensure our policies and best practices are in place. If a misconfiguration is found, operations engineers are alerted immediately so they can diagnose and fix the problem.

Rapid7 scans for several rules including appropriate controls from compliance benchmarks such as CIS AWS Foundations and AWS Well-Architected Framework. Some notable examples of checks scanned for are listed below:

- **Publicly accessible S3 buckets** - S3 buckets are scanned to ensure buckets aren't publicly accessible (readable or writable). Bucket Access Control Lists (ACLs) and bucket policies are evaluated. While some buckets (such as website assets, public downloads, etc.) are designed to be public, no other S3 buckets are ever permitted to be publicly readable or writable.

- **Publicly accessible resources** - Internal servers, databases, and other resources should never be accessible to the Internet. Security groups are scanned to ensure ingress from 0.0.0.0/0 is only allowed on appropriate resources and only for specific ports.

- **IAM keys** - IAM keys are scanned for age, recent access, and attached policies. Newly issued keys are reviewed. IAM policies associated with IAM keys used by third-party services are checked against approved policies. Old keys are removed when they are no longer needed.

Traceability

Rapid7 ensures cloud actions are logged and monitored. All AWS API actions are logged. Access to servers and services are logged to external systems. Logs are analyzed by Rapid7's own products (InsightIDR and InsightOps) for notable events. Rapid7's Managed Detection & Response services team works with our internal Security Operations team to monitor these events 24/7 and investigate any alerts.

COMPLIANCE

Rapid7 SOC Reports

Rapid7 can provide a SOC 2 Type II report covering InsightIDR, InsightOps, InsightVM, and InsightAppSec under NDA. This report is a representation of Rapid7's overall security posture and controls.

AWS SOC Reports

The Insight platform is hosted by AWS. You can retrieve AWS compliance reports (SOC 2, SOC 3, FedRAMP Partner Package, ISO 27001:2013 SoA etc.) here: https://aws.amazon.com/artifact/.

Third-Party Penetration Test

External penetration tests are conducted on an annual basis by a third party. Results of the reports are not provided externally, but Rapid7 can provide letters of attestation from the external firm.

Vulnerability Handling and Disclosure

Rapid7 has a defined standard operating procedure for responsible handling and disclosure of vulnerabilities that are reported in our products and web properties. In the case that a vulnerability is reported to us, Rapid7 will work with the reporter to triage and fix the vulnerability in a timely fashion. Rapid7 will also provide public acknowledgement and attribution to any reporters who request it. Additional information about this process can be found here: https://www.rapid7.com/security/disclosure/.

PCI

Rapid7 is Self Assessment Questionnaire (SAQ) compliant in alignment with our bank's PCI guidelines. Rapid7 can provide a PCI certificate upon request.

ABOUT RAPID7

Rapid7 powers the practice of SecOps by delivering shared visibility, analytics, and automation that unites security, IT, and DevOps teams. The Rapid7 Insight platform empowers these teams to jointly manage and reduce risk, detect and contain attackers, and analyze and optimize operations. Rapid7 technology, services, and research drive vulnerability management, application security, incident detection and response (SIEM), orchestration and automation, and log management for organizations around the globe. To learn more about Rapid7 or join our threat research, visit www.rapid7.com.