WHITE PAPER

snapLogic

Security and Privacy Overview

Cloud application security, data security and privacy, and password management

Overview

Security is a growing concern and should not be taken lightly across an organization. With increasing cloud applications and services, organizations must establish a series of security processes to mitigate security issues.

At SnapLogic, we understand that privacy and security are of paramount importance to our customers. In this white paper, you will learn how SnapLogic addresses security in multiple areas.

SnapLogic currently complies with the following certifications:

Compliance and certifications

SnapLogic's DevOps group and legal departments work closely together to ensure protection of customer data and compliance with applicable privacy and other laws. Procedures are jointly developed and communicated to hosting and legal staff. SnapLogic's legal department, in connection with the security team, is responsible for confirming the implementation of these procedures.

SnapLogic performs various types of internal and third-party audits to validate compliance with SnapLogic, customer, and regulatory requirements. The SnapLogic platform has been certified by third parties for SOC2 Type II, HIPAA, SSAE18 Type II, and ISAE 3402 Type 2 standards. It also complies with GDPR, European Union-U.S., and Swiss-U.S. Privacy Shield frameworks. Upon completion of each audit, a written report of the findings and recommendations is created and maintained in a secure central repository. If a non-compliance deficiency or other finding is discovered during the course of an audit, SnapLogic promptly assesses, prioritizes, mitigates, or identifies appropriate compensating controls.

Risk assessment and penetration testing

SnapLogic conducts regular internal and third-party risk assessments and penetration tests on data applications, systems, and infrastructure associated with accessing, processing, storage, communication, and/or transmission of customer or sensitive data.

Privacy policy

SnapLogic is committed to safeguarding its customers' online privacy. SnapLogic has implemented numerous technical and administrative measures that ensure customer data remains protected and secure, and is transparent regarding how SnapLogic handles data. SnapLogic's privacy policy can be obtained at **snaplogic.com/privacy-policy**.

Cloud application security

SnapLogic platform

SnapLogic's platform allows users to control access rights. SnapLogic does not collect personal information on behalf of its customers, does not control such information, and does not monitor the content of pipelines run in the SnapLogic context. Instead, SnapLogic relies on its customers to ensure that their collection and use of data and personal information via the SnapLogic platform complies with their own privacy policies and all applicable laws.

SnapLogic's system processes data in a content-agnostic manner. SnapLogic does not store any customer data within the platform. However, the metadata relating to integration resides in the cloud platform.

The comprehensive security applied by Amazon Web Services (AWS) protects the SnapLogic Intelligent Integration Platform (IIP) with all the certifications that AWS engenders. However, the SnapLogic IIP layers its own

security precautions on top. This includes protecting the platform with the latest security and resilience patches, as well as engaging third-party security audits.

The SnapLogic Intelligent Integration Platform consists of a multi-tenant cloud service for creating, managing, and monitoring integrations; and the Snaplex, the elastic execution grid for data processing that can run in the cloud, behind the firewall, and/or natively in a Hadoop cluster. SnapLogic security includes a combination of policy, procedure, and technology spanning physical network, infrastructure, platform, and data, ensuring that private and sensitive data is always protected. SnapLogic adheres to security best practices; runs regular internal security audits; and maintains policies that span operations, data passwords and credentials, facilities and networks, and connectivity.

Customer data is streamed between applications, databases, files, social networks, and big data sources via the self-upgrading Snaplex. Customer integration metadata and log files are stored on the AWS infrastructure.

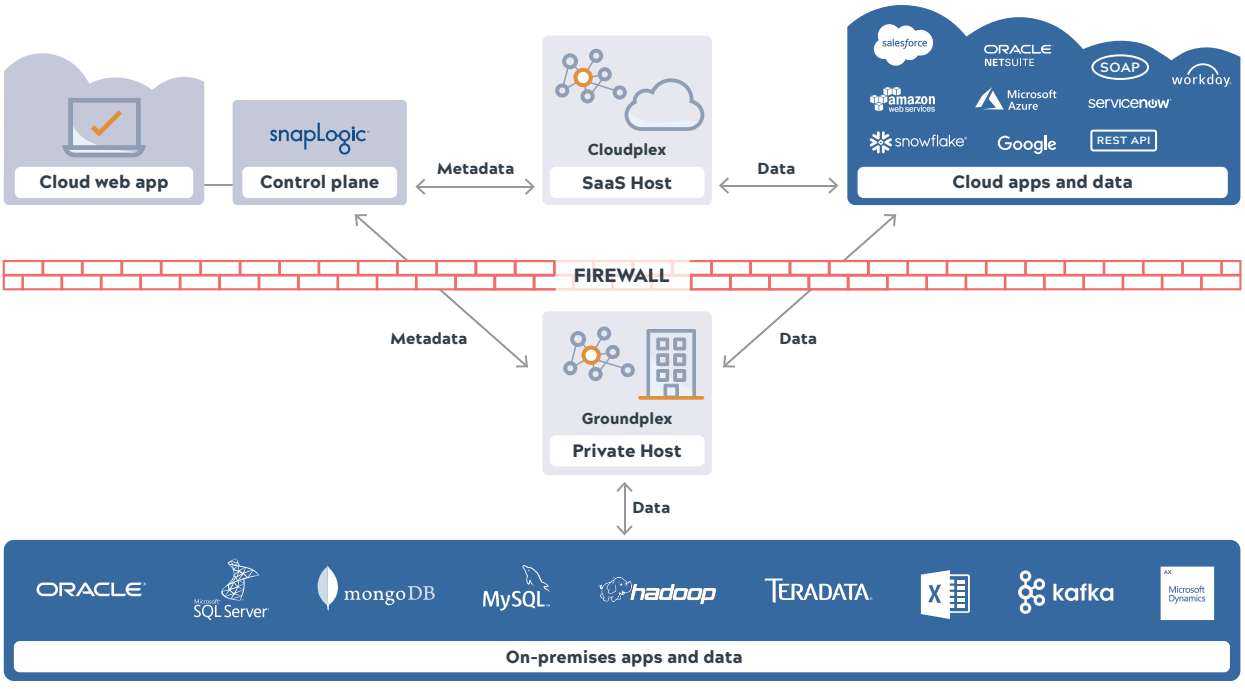
The SnapLogic Intelligent Integration Platform consists of a multi-tenant cloud service for creating, managing, and monitoring integrations.

Unified platform - designer, manager, and dashboard

The SnapLogic user makes an HTTPS connection to the control plane to interact with the Metadata stored there. The authentication can use HTTP basic authentication or be delegated to a SAML-2 based SSO, such as OpenAM, OKTA, or Ping.

When using the SnapLogic Designer, SnapLogic automatically retrieves preview data for the pipeline development process. This preview data is encrypted in the Snaplex node where the validation process is taking place with the specific session key, and passed to the browser (via the control plane, over SSL/TLS encrypted communication links) for the developer to understand the data content and shape. The set of encrypted data is held in the browser's memory cache and decrypted only when visualization of the data is requested.

For the browser session key, this arrangement is agreed upon between the browser and the control plane when the session is started. When a validation process is initiated from the browser, the key is encrypted using the organization's Public Key and passed via the control plane to the Snaplex node (session key is decrypted using the



node's Private Key) executing the validation process. As each Snap (which generates preview data) completes its output, the data is encrypted and passed to the control plane for distribution to the Designer session in the browser. The encrypted session key is never stored in the control plane, and the data is not persisted in the Snaplex node or the browser. Even in the browser, the preview data is held in the encrypted form, only being decrypted when visualization is required.

Pipeline execution and communication with endpoints

- When a pipeline execution is initiated, it may be initiated by one of the following methods:
- A manual execution by an authenticated use through the UI
- A triggered execution (which may be authenticated using basic authentication or token-based authentication)
- A scheduled execution, initiated by the SnapLogic Scheduler
- Invoked from another (already authenticated) pipeline in the runtime via the "ForEach" or "Pipeline Execute" Snaps

In the pipeline execution model, some authenticated process instructs SnapLogic to run a pipeline on a Snaplex. The pipeline and its associated metadata, accounts, and configurations must have been predefined. When the SnapLogic Control plane receives the request, it checks the target Snaplex for which node is the least-loaded, and the execution is allocated to the resulting node(s) in a randomized fashion. The instruction to execute with the packaged metadata is passed to the Snaplex node (over the available WebSocket connection) for execution. The Snaplex node will then prepare the execution, load all the necessary classes (downloading any missing or updated Snap libraries as required), and then execute.

If the pipeline has external endpoints, then the Snap will make necessary connections to the application over the protocol determined by the application. Most cloud-based applications will be using the HTTPS protocol, but local databases and applications will often use their own proprietary network protocols over TCP/IP (e.g., SQL Server, Oracle, SAP). The data retrieved from the applications is coerced into a JSON format and marshalled into the SnapLogic pipeline data flow. Data in SnapLogic is streamed between the Snaps in a pipeline, with minimal resistance during the data transfer between Snaps in memory.

Some Snaps may need to overflow data to disk. In the case of a Sort Snap, for example, the Snap will create temporary files in the designated Java tmp directory (as per the JVM configuration). These files are created to be deleted in the event of the JVM termination, in addition to being removed on completion of the Snap processing. Access to the operating system of the nodes in a Cloudplex is limited to the SnapLogic TechOps team. For customer-deployed Snaplex nodes (Groundplex), access to the operating system layer is in the customer's control.

While the pipeline is executing, each of the Snaps is logging its progress locally, and the host Snaplex node will be relaying that progress data to the SnapLogic control plane so that any query on the pipeline's progress is visible. On completion of the pipeline, its final stats are relayed back to the control plane for storage in the pipeline runtime database. This is the information visible in the Dashboard and available through the public API requests.

Data security and privacy

Encryption and data management

SnapLogic provides adequate protection of sensitive customer data through a combination of access controls and encryption.

Encryption is required if:

- Customer data is transmitted over public networks
- The use of encryption is mandated by law or regulation
- SnapLogic determines that encryption is necessary to protect customer data

SnapLogic encrypts data at the disk level with the account data stored in a server-side encrypted bucket in the Amazon S3 environment. SnapLogic does not, by default, persist customer data. At SnapLogic, we also encrypt data fields at the account level and only operate on the customer data using customer defined pipelines which may perform any operations that may be necessary on the data.

Accounts

Accounts are stored in SnapLogic in an encrypted form, doubly encrypted at rest as the encrypted file is stored in an S3 bucket, which has server-side encryption enabled.

When an account is created, the UI presents the form for the capture of the data fields, and before transmission, the data is encrypted using the browser's Web Crypto API. The crypto key used will be the public key used by the org. Once the encrypted data is ready, on submit, the encrypted data is passed via the established HTTPS connection) to the control pane. The service that is the target of the API used then passes that encrypted data to the S3 serve for persistence.

At SnapLogic, we also encrypt data fields at the account level and only operate on the customer data using customer defined pipelines which may perform any operations that may be necessary on the data.

When a pipeline is executed in a Snaplex which requires the information contained in the account data stored in S3, the pipeline metadata will have the identifier to specify which file that account detail is stored in, and that will be retrieved from the control plane for use locally within the Snaplex node. The data in the file will be decrypted using the Private key and is held locally. The data for the credentials are not persisted locally. The location of the S3 bucket used to store the encrypted account information may be geographically located; this is an option provided at org creation.

Enhanced account encryption

By default, SnapLogic orgs use keys provided by SnapLogic for the authentication of infrastructure components, such as Snaplex nodes, and encryption of the accounts used in the pipelines. With this infrastructure, the user may use a combination of cloud- and ground-based Snaplexes.

To implement enhanced security, users may use the Enhanced Account Encryption feature. With this feature, the user designates his or her own key-pair for use within their Groundplex deployment, and does not share the private key with SnapLogic. The data is encrypted with a public key before it leaves the browser, then is decrypted with a private key on the Groundplex. In this scenario, account information is only usable in the Groundplexes where the private key is available.

The key for decrypting the accounts would be unavailable for Cloudplex deployments. To use this enhanced security capability in a cloud deployment, users may choose to deploy their own cloud nodes in a provider of their own choosing, in their control.

Key rotation

Enhanced Account Encryption supports key rotation. Org admins can change the encryption key on the org, which will change account encryption within that org. Preview Data

SnapLogic uses "schema on read" to help understand the schema of the data flowing through a pipeline. In order to get a representative sample, SnapLogic tries to get up to 50 documents to establish that schema. Where "Read" type Snaps are in use, such as a database select, it will include an implicit "LIMIT 50" when trying to generate the input. For other, REST- or SOAP- based endpoints, up to the first 50 documents retrieved will be used.

Preview data encryption

When a preview request is initiated for any Snap, that preview is created on the Snaplex by executing the Snap with input data from the previous Snap(s) in the pipeline, if available. The output from the Snap is then encrypted, using the AES encryption key generated in the user's browser, specific to that session. That encrypted data is then passed to the browser, via the SnapLogic control plane, over SSL/TLS encrypted communication links, where it may only be viewed within that user session and with the session-based encryption key. As soon as the user session

is closed, or the browser is closed, the keys are gone, and the data must be regenerated for viewing. The preview data is held in the browser in an encrypted form, only decrypted for the purpose of the preview when requested in the browser. The data is not visible/clear text even when exploring the browser internals.

Preview key encryption

For the browser session key, the session start time is agreed between the browser and the control pane. When a validation process is initiated from the browser, the key is encrypted using the organization's public key and passed via the control plane to the Snaplex node executing the validation process. As each Snap completes its output, the data is encrypted and passed to the control plane for distribution to the Designer session in the browser. The encrypted session key is never stored in the control plane, and the data is not persisted in the Snaplex node or the browser. Even in the browser, the preview data is held in the encrypted form, only being decrypted when visualization is required. Although the preview data is currently cached in the SnapLogic control plane, the key is never kept there.

The metadata is secured inside the protected SnapLogic environment, and only accessed by the SnapLogic Control Plane Services - no access is permitted by any outside service.

SnapLogic metadata

The SnapLogic Metadata (definitions of pipelines, tasks, execution runtimes, etc., and not customer data) is stored in the SnapLogic Control plane (currently running on the Amazon EC2 infrastructure) using a MongoDB database. The metadata is secured inside the protected SnapLogic environment, and only accessed by the SnapLogic Control Plane Services - no access is permitted by any outside service. The metadata is not stored in an encrypted form; only sensitive data, such as account information will be encrypted. Each of the fields which make up the different types of accounts is encrypted using the keys provided, either SnapLogic's or the customer's own.

Password management

User identity and access management

The SnapLogic Intelligent Integration Platform server supports an authentication and privilege model that allows the administrator to grant, limit, or restrict access to components and pipelines. The server applies access rules to all requests, and grants or denies access depending on the type of operation attempted by the user. Users who share a particular responsibility can be assigned to groups.

Single sign on (SSO)

Single Sign On (SSO) is a convenient way for users to log into multiple software systems without needing to enter their user name and password for each system. SnapLogic supports SSO through Security Assertion Markup Language (SAML) standard. The supported authentication methods, include Open AM, OKTA, and Ping.

The SnapLogic Intelligent Integration Platform server supports an authentication and privilege model that allows the administrator to grant, limit, or restrict access to components and pipelines

The SAML standard defines how Service Providers (SPs) can communicate with Identity Providers (IdPs) to securely authenticate users. In this case, SnapLogic is the Service Provider. The communication between SnapLogic and IdP starts after the user enters their organization name and clicks the SSO Log In button on the SnapLogic home page. The SnapLogic server will use the organization name to find the associated IdP and then redirect the user's web browser to that IdP with an authentication request. The destination for the redirect is defined by the IdP and is informed of the SnapLogic service by uploading the SnapLogic metadata file that is generated when configuring the user's organization to use SSO. Finally, the authentication response is validated by the SnapLogic server using the IdP metadata, and the user is allowed to begin working in the SnapLogic Designer.

Configuring the organization to use SSO requires the exchange of metadata between the SP and the IdP.

Password security and permissions

Strict and enforced password and account creation policies reduce the chance of intruder attacks in SnapLogic's development environments and data centers. Passwords are controlled by the platform itself. Therefore, customers must enforce strict password control via the SnapLogic Administration Console.

SnapLogic includes powerful built-in end user and admin permissions matrices, as well as customizable permissions. Depending on the assigned role, users can get or be denied access to specific features, project spaces, projects and their assets. In addition, administrative permissions can be used to limit the access level of administrators. The SnapLogic "SuperAdmin," which is used by the provisioning and platform administration team, has read access only to all metadata resources, across all orgs and no visibility into user or preview data. SnapLogic carefully limits the user with that privilege; the permission is granted only with multiple senior management sign-off. All user actions are automatically audited while operating on the platform.

The SnapLogic integration platform as a service (iPaaS) has a robust enterprise-grade hierarchical model to organize and provides fine-grained access control to all system assets. Assets are defined as any resource that can be created by an organization and tracked by the system in the monitoring and reporting dashboard.