ELT DATA Information Security Guide

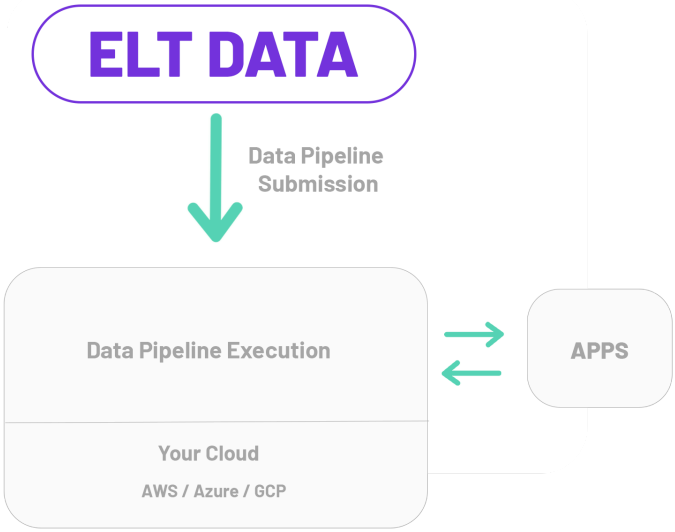
Version	Date (Y-M)	Author	Summary
1.0	2023-11	Ishan Rastogi	Initial release
2.0	2024-01	Ishan Rastogi	Added sections on security standards, PDF generation and email process
2.1	2024-05	Kiran Hosakote	Minor edits and updates

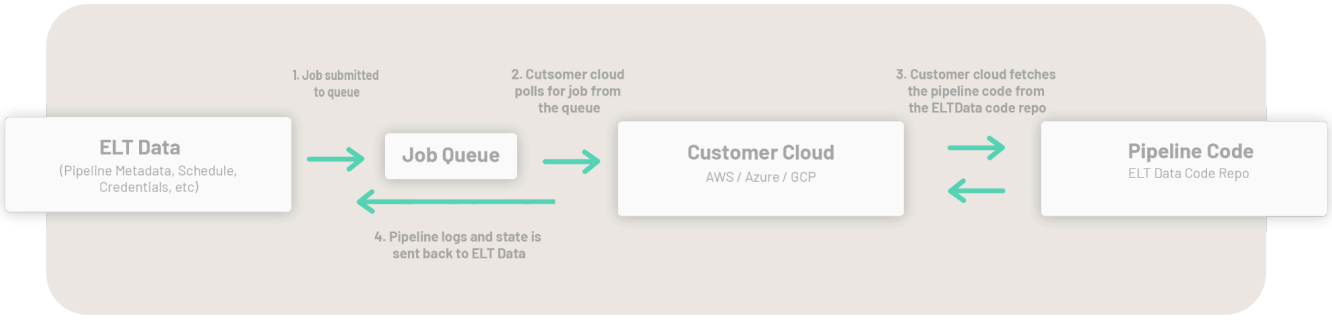
DATA PIPELINE EXECUTION MODEL

Before we dive into the security of the ELT Data system, it is important to understand the ELT Data pipeline execution model.

ELT Data orchestrates and provides the code to run the data pipelines. The execution of the pipelines takes place in the user's cloud.

The following diagrams illustrate the execution model.





When an ELT Data job is run, servers are started in the user's cloud. These servers are transient in nature. The data is stored in the cloud storage (S3, Azure Blob, GCS) in the customer's cloud. As such the data transfer takes place directly between the customer's application and cloud.

ELT Data orchestrates this entire process from outside without having access to customer's data. Customer's data remains within the cloud and is governed by their data policies.

The code transmits metadata (status, failure, duration, logs) about the pipeline back to ELT Data to enable the orchestration.

> **ELT Data orchestrates and provides the code for the data pipelines.**
>
> **The execution of the data pipeline takes place in the customer's cloud.**

SECURITY STANDARDS AND COMPLIANCE

Security Standards

ELT Data https://app.eltdata.io ("application") is hosted in the AWS US East 1 Region. It is designed, built and deployed by **Vega Solutions** LLC (Vega) to the following security standards.

1. **CIS AWS Foundations Benchmark** which is defined by the Center for Internet Security as an objective, consensus-driven security guideline for AWS Cloud Providers.

2. **AWS Foundational Security Best Practices standard** which is defined by AWS as a set of controls that detect when deployed accounts and resources deviate from security best practices. This standard provides actionable and prescriptive guidance on maintaining and improving the organization's security position.

SOC 2 and HIPAA Compliance

SOC 2: On an ongoing basis, Vega runs a SOC 2 assessment from AWS Audit Manager on the ELT Data application and infrastructure. It also runs a monitoring tool on its code repositories for code vulnerability testing. The evidence collected from these activities along with other controls for data security and privacy is used to generate periodic SOC 2 assessment reports. Please email contact@eltdata.io for the latest assessment report.

HIPAA: As defined in 45 CFR 160.103, Vega Solutions LLC is a Business Associate, not a HIPAA covered entity. When a covered entity subscribes to a ELT Data license, Vega will execute a Business Associate Contract with them to enable ELT Data to access their data in compliance with HIPAA.

APPLICATION SECURITY

Application Roles

There are four roles defined in the application, which are used as follows:

- System - Used for audit logs and automated internal tasks.
- Anonymous - Used for actions performed anonymously, e.g., password reset.
- User - End-user signups, to provide access to setup and run data pipelines.
- Admin - Access to user management, logging controls, service health checks.

Login - JWT Tokens

On presentation of credentials, users are presented with JWT tokens and logins are done in a stateless manner.

- JWT tokens are signed and cannot be altered by an end-user.
- JWT tokens are signed using a Base64 encoded string.
- The securing keys have a length of 512 bits.

CSRF Protection

ELT Data uses JWT token-based authentication. The tokens are passed in Authorization Headers and not in cookies.

Due to this the CSRF attacks are generally mitigated because the browser does not automatically include the JWT in the request headers.

However, CSRF attacks can still be carried out via XSS vulnerabilities for which we have enabled AWS Web Application Firewall (WAF). AWS WAF blocks any request that seems susceptible to XSS

SSL Enabled

The core application runs behind a proxy, with SSL access only. The SSL Certificates are renewed every 11 months.

INFRASTRUCTURE SECURITY

ELT Data **https://app.eltdata.io** runs entirely in the AWS cloud. However the data pipeline execution happens within the user's cloud. For this the users need to grant permissions to ELT Data to start and stop servers. No other permission to any other resource is needed. Please refer to the technical documentation for the exact scope of permissions required.

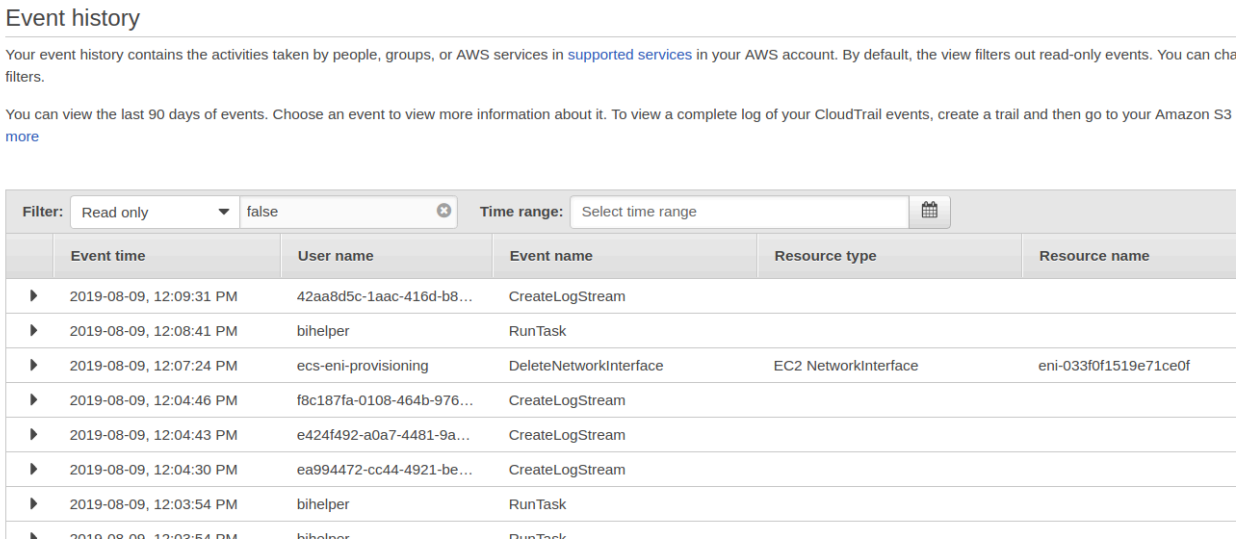
Virtual Private Cloud

ELT Data is hosted in the AWS US East 1 Region in its own Virtual Private Cloud within the AWS cloud. All infrastructure resources reside in this AWS VPC, providing complete isolation of the network and resources from all other AWS users. In addition to the VPC, the databases and internal services reside in a private subnet.

The only site open to the public internet is ELT Data, where users log into the application. No other microservices are accessible to the public internet.

Logging and Alerting

Infrastructure Events Snapshot: The logs for access to the infrastructure are stored, and can be used to establish any trail within the infrastructure.



Application Events Snapshot: The logs for access to https://app.eltdata.io are stored in the database and are available for complete audit of user access.

08/08/19 09:07:55	admin	AUTHENTICATION_SUCCESS
08/08/19 06:40:03	admin	AUTHENTICATION_SUCCESS
07/08/19 17:04:12	admin	AUTHENTICATION_SUCCESS
07/08/19 16:48:06	admin	AUTHENTICATION_SUCCESS
07/08/19 15:19:16	admin	AUTHENTICATION_SUCCESS

Infrastructure monitoring services:

In addition to the above we have the following AWS services running 24*7 to monitor infrastructure security:

- AWS Security Hub keeps a check on the best practices for security and notifies us of any deviations from them.
- AWS Guard Duty fires events in case of anomalies in infrastructure access or in case of unauthorized access attempts.
- AWS WAF blocks unauthorized and malicious web access.

Access to Infrastructure Resources

Applications within the ELT Data VPC are treated as users when they need to access ELT Data resources. Separate users **(IAM)** / roles are created for such services and each service is allowed access only to the required resources.

Data Retention and Encryption

ELT Data saves the pipeline execution metadata but has no access to the customer data. The customer data is retained in customer's cloud environment and is subject to customer's data policies.

Key Rotation Policies

AWS Systems Manager is used to access development and production servers. No SSH Keys are used for any development or for accessing production servers.

The token encryption keys (user authentication) are rotated once in six months.

AWS Resources Compliance Report

The AWS Resources Compliance Report is generated using the rules defined by AWS Config. Please write to contact@eltdata.io for a detailed report.

CODE REPOS AND VULNERABILITIES

ELT Data's codebase is hosted in Github. The code repositories are integrated with Snyk. Snyk is a security platform which continuously monitors the codebase for vulnerabilities and reports daily on any new ones detected. Automated Pull Requests are raised and reviewers assigned in case of any detection.

In addition to their own feed, Snyk monitors the following vulnerability databases:

- Common Vulnerabilities and Exposures (CVE): Provides "...an identification number, description, and at least one public reference for publicly known cybersecurity vulnerabilities." Launched in 1999, CVE is a standardized resource for other tools and services to track and evaluate vulnerabilities.

- National Vulnerability Database (NVD): A U.S. government repository of standardized vulnerability management data, including impact metrics such as CVSS. It uses CVE as one of its inputs.

STAFF DEVICES

All ELT Data devices are scanned every day for the following attributes:

1. Antivirus
2. Disk Encryption
3. Updated Device OS
4. Screen lock

The above attributes are mandatory for all devices and any deviation is reported to the system administrators.

- End of document -