*Proprietary & Confidential*

CONFLUENT

# System Description of the Confluent Platform

**SOC 3**
Relevant to Security, Availability, and Confidentiality

JANUARY 1, 2022 TO DECEMBER 31, 2022

MOSSADAMS

# Table of Contents

# I. Independent Service Auditor's Report

Confluent, Inc.
899 West Evelyn
Mountain View, CA 94041

To the Management of Confluent:

## Scope

We have examined Confluent's accompanying assertion in Section II titled "Confluent's Assertion" (assertion) that the controls within Confluent's Platform (system) were effective throughout the period January 1, 2022 to December 31, 2022, to provide reasonable assurance that Confluent's service commitments and system requirements were achieved based on the trust services criteria relevant to Security, Availability, and Confidentiality (applicable trust services criteria) set forth in TSP Section 100, *2017 Trust Services Criteria for Security, Availability, Processing Integrity, Confidentiality, and Privacy* (AICPA, *Trust Services Criteria*).

Confluent uses Amazon Web Services as a subservice organization for managed hosting of infrastructure, software, and data related to the Proactive Support component. The description indicates that complementary subservice organization controls that are suitably designed and operating effectively are necessary, along with controls at Confluent, to achieve Confluent's service commitments and system requirements based on the applicable trust services criteria. The description presents the types of complementary subservice organization controls assumed in the design of Confluent's controls. The description does not disclose the actual controls at the subservice organization. Our examination did not include the services provided by the subservice organization, and we have not evaluated the suitability of the design or operating effectiveness of such complementary subservice organization controls.

## Service Organization's Responsibilities

Confluent is responsible for its service commitments and system requirements and for designing, implementing, and operating effective controls within the system to provide reasonable assurance that Confluent's service commitments and system requirements were achieved. Confluent has also provided the accompanying assertion about the effectiveness of controls within the system. When preparing its assertion, Confluent is responsible for selecting, and identifying in its assertion, the applicable trust services criteria and for having a reasonable basis for its assertion by performing an assessment of the effectiveness of the controls within the system.

## Service Auditor's Responsibilities

Our responsibility is to express an opinion, based on our examination, on whether management's assertion that controls within the system were effective throughout the period to provide reasonable assurance that the service organization's service commitments and system requirements were achieved based on the applicable trust services criteria. Our examination was conducted in accordance with attestation standards established by the American Institute of Certified Public Accountants (AICPA). Those standards require that we plan and perform our examination to obtain reasonable assurance about whether management's assertion is fairly stated, in all material respects. We believe that the evidence we obtained is sufficient and appropriate to provide a reasonable basis for our opinion.

Our examination included:

- Obtaining an understanding of the system and the service organization's service commitments and system requirements
- Assessing the risks that controls were not effective to achieve Confluent's service commitments and system requirements based on the applicable trust services criteria
- Performing procedures to obtain evidence about whether controls within the system were effective to achieve Confluent's service commitments and system requirements based the applicable trust services criteria

Our examination also included performing such other procedures as we considered necessary in the circumstances.

We are required to be independent and to meet our other ethical responsibilities in accordance with relevant ethical requirements relating to the engagement.

## Inherent Limitations

There are inherent limitations in the effectiveness of any system of internal control, including the possibility of human error and the circumvention of controls.

Because of their nature, controls may not always operate effectively to provide reasonable assurance that the service organization's service commitments and system requirements were achieved based on the applicable trust services criteria. Also, the projection to the future of any conclusions about the effectiveness of controls is subject to the risk that controls may become inadequate because of changes in conditions or that the degree of compliance with the policies or procedures may deteriorate.

## Opinion

In our opinion, management's assertion that the controls within Confluent's Platform were effective throughout the period January 1, 2022 to December 31, 2022, to provide reasonable assurance that Confluent's service commitments and system requirements were achieved based on the applicable trust services criteria is fairly stated, in all material respects.

Moss Adams LLP

San Francisco, California
March 28, 2023

# II. Confluent's Assertion

We are responsible for designing, implementing, operating, and maintaining effective controls within Confluent's Platform (system) throughout the period January 1, 2022 to December 31, 2022 to provide reasonable assurance that Confluent's service commitments and system requirements relevant to Security, Availability, and Confidentiality were achieved. Our description of the boundaries of the system is presented in Section III titled "Confluent's Description of the Boundaries of the Confluent Platform" and identifies the aspects of the system covered by our assertion.

We have performed an evaluation of the effectiveness of the controls within the system throughout the period January 1, 2022 to December 31, 2022, to provide reasonable assurance that Confluent's service commitments and system requirements were achieved based on the trust services criteria relevant to Security, Availability, and Confidentiality (applicable trust services criteria) set forth in TSP Section 100*, 2017 Trust Services Criteria for Security, Availability, Processing Integrity, Confidentiality, and Privacy (*AICPA*, Trust Services Criteria)*. Confluent's objectives for the system in applying the applicable trust services criteria are embodied in its service commitments and system requirements relevant to the applicable trust services criteria. The principal service commitments and system requirements related to the applicable trust services criteria are presented in Section III titled "Confluent's Description of the Boundaries of the Confluent Platform".

Confluent uses Amazon Web Services as a subservice organization for managed hosting of infrastructure, software, and data related to the Proactive Support component. The description indicates that complementary subservice organization controls that are suitably designed and operating effectively are necessary, along with controls at Confluent, to achieve Confluent's service commitments and system requirements based on the applicable trust services criteria. The description presents the types of complementary subservice organization controls assumed in the design of Confluent's controls. The description does not disclose the actual controls at the subservice organization.

There are inherent limitations in any system of internal control, including the possibility of human error and the circumvention of controls. Because of these inherent limitations, a service organization may achieve reasonable, but not absolute, assurance that its service commitments and system requirements are achieved.

We assert that the controls within the system were effective throughout the period January 1, 2022 to December 31, 2022, to provide reasonable assurance that Confluent's service commitments and system requirements were achieved based on the applicable trust services criteria.

# III. Confluent's Description of the Boundaries of the Confluent Platform

## A. System Overview

### 1. Services Provided

#### COMPANY OVERVIEW

Confluent was founded by the team that built Apache Kafka. Apache Kafka is a community-distributed, event-streaming platform capable of handling trillions of events a day. Initially conceived as a messaging queue, Kafka is based on an abstraction of a distributed commit log. Since being created and open sourced by LinkedIn in 2011, Kafka has quickly evolved from a messaging queue to a full-fledged event streaming platform. Confluent delivers the most complete distribution of Kafka with the Confluent Platform. The Confluent Platform improves Kafka with additional community and commercial features designed to enhance the streaming experience of both operators and developers in production, at massive scale. Confluent provides a streaming platform that enables companies to access data as real-time streams. Confluent is headquartered in Mountain View, CA with additional offices in San Francisco, CA and London, UK, and remote workers at various locations. Currently, Confluent employs approximately 2,200 employees across these locations.

Confluent believes every byte of data has a story to tell, something of significance that informs the next thing to be done. In a data-driven enterprise, how data moves is nearly as important as the data itself. With greater speed and agility, data's value increases exponentially.

#### SYSTEM DESCRIPTION

Confluent Platform provides customers with the distribution of Apache Kafka for production environments simplifying engineering operations and the administration of Kafka clusters. It complements Apache Kafka with administration, monitoring, and management tools. The Confluent Platform is primarily hosted within a customer's environment and is comprised of the following components:

- *Confluent Cloud Console*™ – A comprehensive management and monitoring system for Apache Kafka. Control Center provides:
  - The ability to monitor and manage clusters from a rich user interface
  - The ability to quickly scan through clusters for anomalies and track down messages to their sources
  - Full integration with connectors, allowing users to manage data pipelines without a line of code
  - The delivery of real-time analysis of the performance of Kafka
  - The ability to drill into topics, producers, consumers, and more to understand the activity within their data pipelines enabling organizations to govern a growing ecosystem of stream data applications

- *Confluent for Kubernetes (CFK)* – CFK is a cloud-native control plane for deploying and managing Confluent in your private cloud environment. It provides a standard and simple interface to customize, deploy, and manage Confluent Platform through declarative API.
- *Kafka Connect* – A tool for scalability and reliably streaming data between Apache Kafka® and other data systems. It makes it simple to quickly define connectors that move large data sets into and out of Kafka. Kafka Connect can ingest entire databases or collect metrics from application servers into Kafka topics, making the data available for stream processing with low latency. An export connector can deliver data from Kafka topics into secondary indexes like Elasticsearch or into batch systems such as Hadoop for offline analysis. Available connectors may be installed separately via Confluent Hub.
- *ksqlDB* – ksqlDB is the streaming SQL engine for Kafka. It provides an easy-to-use yet powerful interactive SQL interface for stream processing on Kafka, without the need to write code in a programming language such as Java or Python. ksqlDB is scalable, elastic, fault-tolerant, and real-time. It supports a wide range of streaming operations, including data filtering, transformations, aggregations, joins, windowing, and sessionization.
- *Confluent Health+* – Confluent Health+ allows users to monitor and manage the Confluent Platform environment, ensuring the health of clusters and minimizing business disruption with intelligent alerts, monitoring, and proactive support based on best practices created by the inventors Apache Kafka®.
- *Replicator* – Confluent Replicator can be deployed across clusters and in multiple data centers. Multi-data center deployments enable use cases, such as:
  - *Active-active geo-localized deployments* – allows users to access a near-by data center to optimize their architecture for low latency and high performance
  - *Active-passive disaster recovery (DR) deployments* – allows the failing over applications to use Confluent Platform in a different data center if a disaster occurs
  - *Centralized analytics* – Aggregates data from multiple Apache Kafka® clusters into one location for organization-wide analytics
  - *Cloud migration* – Uses Kafka to synchronize data between on-prem applications and cloud deployments
- *Schema Registry* – The Confluent Schema Registry provides a serving layer for metadata. Specifically, it provides a RESTful interface for storing and retrieving Avro®, JSON Schema, and Protobuf schemas. It stores a versioned history of schemas based on a specified subject name strategy, provides multiple compatibility settings and allows evolution of schemas according to the configured compatibility settings and expanded support for these schema types. It also provides serializers that plug into Apache Kafka® clients to handle schema storage and retrieval for Kafka messages sent in any of the supported formats.
- *Security Plugins* – Plugins for other services in the Confluent Platform which add extended security features. Plugins are currently provided for the Representational State Transfer (REST) Proxy and Schema Registry. Key features include principal propagation and pluggable access control. Features of the Security Plugins allow users to:
  - Propagate principals on an incoming REST Proxy request, forwarding them to Kafka
  - Automatically apply Kafka ACLs to REST Proxy requests
  - Propagate principals via SSL and SASL
  - Apply a pluggable authorizer to Schema Registry requests
  - Restrict schema evolution management to administrative users with read-only access for applications and developers.

- *Self-Balancing Clusters* – Self-Balancing automates resource workload balancing, provides failure detection and self-healing, and allows adding or decommissioning brokers as needed, with no manual tuning required. Self-balancing provides:
  - Fully automated load balancing
  - Auto-monitoring of clusters for imbalances based on a large set of parameters, configurations, and runtime variables
  - Continuous metrics aggregation and rebalancing plans, generated instantaneously in most cases, and executed automatically
  - Automatic triggering of rebalance operations based on simple configurations set on Confluent Control Center or in Kafka server properties files (Users can choose to auto-balance only when brokers are added, or anytime, which rebalances for any uneven load.)
  - At-a-glance visibility into the state of the clusters and the strategy and progress of auto-balancing through a few key metrics.

### SYSTEM BOUNDARIES

Included within the scope of this report are the production systems, infrastructure, software, people, procedures, and data supporting Confluent Platform. Customers are responsible for the implementation and management of the Confluent Platform in their hosted environment. This report is specific to the development and support of the Confluent Platform and does not include the Confluent Cloud environment.

## 2. Infrastructure

The Confluent Platform is designed, implemented, and operated to achieve specific business objectives in accordance with management-specified requirements. The purpose of the system description is to delineate the boundaries of the system, which include the services outlined above and the components described below.

Infrastructure for the Confluent Platform is unique and delimited. The primary components of the Confluent Platform are hosted on customer infrastructure, but Confluent utilizes AWS to provide Proactive Support, and to develop, build, and test the other components of the Confluent Platform software for on-premises installations. Refer to Figure 1 for the Technical Overview Diagram below for an overview of AWS services within the Confluent environment.

Proactive Support reports performance metrics back to Confluent over an outbound HTTPS connection and does not provide Confluent personnel access to customer environments. To facilitate operation of Proactive Support and development of the Confluent Platform, Confluent uses the following AWS services:

| AWS Service | Function |
|---|---|
| **Amazon CloudFront** | Amazon CloudFront is a fast content delivery network (CDN) service that delivers data, videos, applications, and APIs to customers globally with low latency and high transfer speeds. |

| AWS Service | Function |
|---|---|
| **Amazon Elastic Compute Cloud (EC2)** | Amazon EC2 provides a virtual computing environment that uses web service interfaces to perform the following functions:<br>• Launch instances of operating systems<br>• Create Amazon Machine Images (AMI) containing applications, libraries, data, and associated configuration settings<br>• Configure security and network access on the EC2 instances |
| **Amazon DynamoDB** | Amazon DynamoDB is a fully managed proprietary NoSQL database service that supports key-value and document data structures. |
| **Amazon Relational Database Service (RDS)** | Amazon RDS is a web service used to operate relational databases in the AWS cloud. |
| **Amazon Route 53** | Amazon Route 53 is a Domain Name System (DNS) web service. |
| **Amazon Simple Notification Service (SNS)** | Amazon SNS is a notification service that provides infrastructure for the mass delivery of messages. |
| **Amazon Simple Storage Service (S3)** | Amazon S3 is virtual storage used in conjunction with Amazon EC2 to store object data. Amazon S3 is also used to automatically replicate data across AWS regions. |
| **Amazon Virtual Private Cloud (VPC)** | Amazon VPC is used to provision logically isolated virtual networks in the AWS cloud. Amazon VPC is used to manage the virtual networking environment, including selection of IP address ranges, creation of subnets, and configuration of route tables and network gateways. |
| **AWS Identity and Access Management (IAM)** | AWS IAM enables users to manage access to AWS services and resources securely. Using IAM, users can create and manage AWS users and groups, and use permissions to allow and deny their access to AWS resources. |

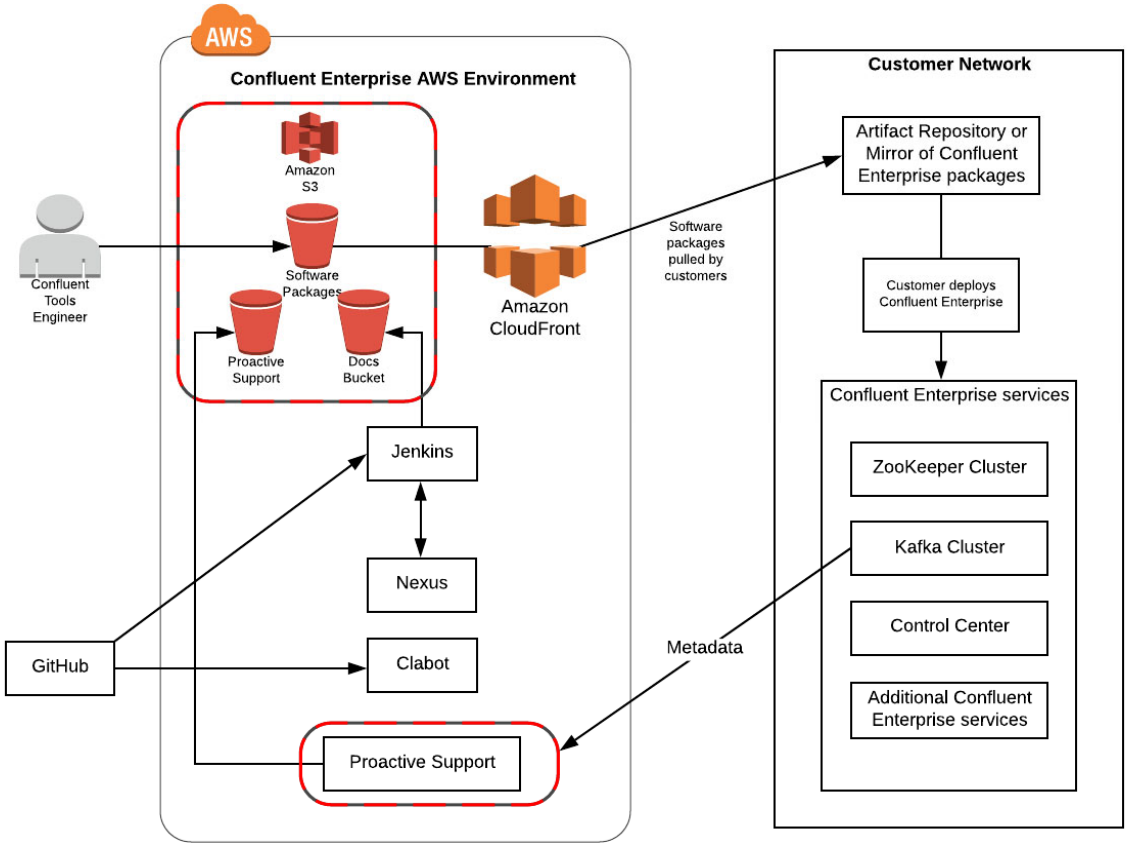


FIGURE 1: TECHNICAL OVERVIEW DIAGRAM

## 3. Software

Confluent has various software programs and tools used to support the Confluent Platform, with key programs listed below:

| Vendor Software | Function |
|---|---|
| **Auth0** | The Auth0 Customer Identity Management service is a third-party solution to add authentication and authorization services to web applications. |
| **Datadog** | Datadog is a monitoring service for cloud-scale applications, providing monitoring of servers, databases, tools, and services, through a SaaS-based data analytics platform. |
| **GitHub** | GitHub provides hosting for software development version control using Git. GitHub provides access control and several other collaboration features, such as bug tracking, feature requests, task management, and wikis for all projects. |

| Vendor Software | Function |
|---|---|
| **HashiCorp Vault** | HashiCorp Vault is an encryption service with centralized key management. |
| **Jenkins** | Jenkins is an open-source automation server that enables developers to build, test, and deploy their software. |
| **Jira** | Jira is an issue tracking and project management tool. |
| **Kubernetes** | An open-source system for automating deployment, scaling, and management of containerized applications. |
| **Okta** | Okta is an identity and access management platform providing cloud software that helps companies manage and secure user authentication into modern applications. |
| **Zendesk** | Zendesk is a cloud-based customer service platform that is designed to improve communication between the company and its customers. |

## 4. People

Confluent is organized in the following functional areas:

- *Business Systems* – Responsible for helping Confluent operate as a data-driven, strategic, and efficient company. This group also includes the internal IT group responsible for IT support at a corporate level.
- *Engineering* – Responsible for the engineering functions across Confluent products. This group is also responsible for maintaining and developing infrastructure components supporting the Confluent suite of products.
- *Finance* – Responsible for the oversight of financial processes for Confluent including fees, internal finance, and budgeting.
- *Global Technical Support* – Responsible for providing customer service, support, and training to Confluent customers.
- *Growth and Marketing* – Responsible for the promotion and demand creation for Confluent products and services.
- *Information Technology (IT)* – Responsible for managing and supporting IT systems, onboarding and offboarding processes, and asset management functions.
- *Legal* – Responsible for providing legal advice to foster sound decision-making for Confluent's operations.
- *People Operations and Recruiting* – Responsible for ensuring that Confluent attracts, hires, develops, and retains the best talent to fuel business growth.
- *Product* – Responsible for designing the systems and planning feature updates, enhancements, and bug fixes based on customer and internal requests.

- *Security Steering Committee* – Responsible for the oversight of internal control and includes members independent from control operators.
- *Trust, Security, and Reliability* – Responsible for security and compliance efforts, including production, client-facing, corporate information assets, and data.

## 5. Data

Confluent Platform is a customer on-premises software package in which the infrastructure is managed by Confluent's customers. As such, Confluent does not process, store, or transmit customer data. Customers are responsible for their own data and infrastructure hosting. The Proactive Support service records and reports metadata about the various operational metrics from the Confluent clusters in the customers' environments to help Confluent improve the overall system. Proactive Support data is transmitted from customer environments to Confluent through an outbound HTTPS connection between Confluent and the customer.

## 6. Processes and Procedures

Confluent has developed and communicated to its personnel procedures to protect service data and the company's assets. Teams are expected to adhere to Confluent policies and procedures that define data are protected through rules and requirements. These are located on the company's intranet and shared drive and can be accessed by any Confluent employee.

The policies and standards used to safeguard Confluent Platform include:

- Acceptable Use Policy
- Access Management Standard
- Asset Management Standard
- Business Continuity and Disaster Recovery Plan
- Cloud Permissions Procedure
- Cloud Security Standard
- Configuration Management Standard
- Cryptography Standard
- Data Classification and Handling Standard
- Information Security Policy
- Issue Management Standard
- Logging and Monitoring Standard
- Mobile Device Management Standard
- Risk Management Standard
- Secure System Development Standard
- Security Incident Response Standard
- Vendor Management Standard
- Vulnerability Management Standard

## B. Principal Service Commitments and System Requirements

Confluent communicates operational requirements to support the achievement of security, availability, and confidentiality through its policies and in its contracts with customers. Confluent's commitments are documented and communicated to customers through the following:

- System Requirements (https://docs.confluent.io/platform/current/installation/system-requirements.html)
- Data Processing Addendum (https://www.confluent.io/cloud-customer-dpa/)

## C. Complementary Subservice Organization Controls

Confluent's controls related to the Confluent Platform cover only a portion of overall internal control for each user entity of Confluent It is not feasible for the criteria related to the Confluent Platform to be achieved solely by Confluent Therefore, each user entity's internal controls must be evaluated in conjunction with Confluent's controls, taking into account the types of controls expected to be implemented by the subservice organization as described below.

| Complementary Subservice Organization Controls | |
|---|---|
| 1 | Access to hosted systems requires users to use a secure method to authenticate. |
| 2 | User content is segregated and made viewable only to authorized individuals. |
| 3 | Network security mechanisms restrict external access to the production environment. |
| 4 | New user accounts are approved by appropriate individuals prior to being provisioned. |
| 5 | User accounts are removed when access is no longer needed. |
| 6 | User accounts are reviewed on a regular basis by appropriate personnel. |
| 7 | Access modifications to hosted systems are approved by appropriate individuals prior to being provisioned. |
| 8 | Access to physical facilities is restricted to authorized users. |
| 9 | Production media is securely decommissioned and physically destroyed prior to being removed from the data center. |
| 10 | Encrypted communication is required for connections to the production system. |
| 11 | Access to hosted data is restricted to appropriate users. |
| 12 | Hosted data is protected during transmission through encryption and secure protocols. |
| 13 | Antivirus or antimalware solutions are installed to detect or prevent unauthorized or malicious software. |
| 14 | System configuration changes are logged and monitored. |
| 15 | Vulnerabilities are identified and tracked to resolution. |
| 16 | Security events are monitored and evaluated to determine their potential impact. |

| Complementary Subservice Organization Controls | |
|---|---|
| **17** | Operations personnel log, monitor, and evaluate incident events identified by monitoring systems. |
| **18** | Operations personnel respond to, contain, and remediate incident events, and update stakeholders, as needed. |
| **19** | System changes are documented, tested, and approved prior to migration to production. |
| **20** | Access to make system changes is restricted to appropriate personnel. |
| **21** | Operations personnel monitor processing and system capacity. |
| **22** | Environmental protections, software, data back-up processes, and recovery infrastructure are implemented. |
| **23** | System failover and backup procedures are tested. |

MOSSADAMS