CONFLUENT

Considerations for Getting Started with Confluent

Checklist

Preface

How to use this guide: This checklist is intended to be a 'print out and use' utility which allows you to plan for your upcoming Confluent project by asking and answering the right questions from the start. As you review each item, frame it with the question 'Have we thought about ...'

This is not meant to be an exhaustive list. We understand every customer's journey is different so we encourage you to take this framework and amend it according to your unique requirements.

As you start to plan your engagement, there are plenty of useful resources to guide you. Here are a few of our recommendations:

- The Confluent Data In Motion Blog
- Support Guides:
  - Getting Started
  - Welcome to Support
  - Confluent Cloud Onboarding Checklist

Section 1: Enablement

- [ ] Have we set up a corporate / team repository for knowledge sharing (best practices, whitepapers, contacts, roles and responsibilities etc).
- [ ] Have Confluent's top recommended resources (support, community, developer.io etc) been shared across all teams/individuals?
- [ ] Have we downloaded and reviewed the Guide to Data In Motion to plan our journey and identify knowledge gaps?
- [ ] Have any gaps in enablement been identified and do we have a plan to resolve those gaps?
- [ ] Do we know how to add the necessary individuals to the support portal and when to use it?
- [ ] Have all the necessary teams / individuals received sufficient Confluent enablement so that:
    - [ ] Each understands the Confluent Cloud concepts (e.g. Cluster Types, retention, access restriction)?
    - [ ] Each is aware of the various Confluent Components (e.g. ksqlDB, Connect, Schema Registry) and the problems they solve?
    - [ ] Each individual has enough knowledge to identify which components which are to be fully managed and those which will be self-managed?
    - [ ] Each understands how to get assistance they need through the support portal?
    - [ ] There is sufficient knowledge of security concepts? For example:
        - [ ] Do we need to set up SSO and do we know how?
        - [ ] Are we aware of the difference between Cloud UI users vs. service accounts?
        - [ ] Have we had sufficient enablement to setup authorization with ACLs or RBAC?
        - [ ] Do we understand the various networking options Confluent Cloud provides and feel confident in selecting the right options for our use case?
- [ ] Have our developers been educated on how to work with events? More specifically:
    - [ ] The implications of accepting default parameters in their producer/consumer code.
    - [ ] Concepts such as at least once, at most once, exactly once. What is needed and how to code for it.
    - [ ] How to write code to meet business requirements for throughput vs. latency and durability vs. availability.
    - [ ] How to decouple Producer vs. Consumer with schema registry and protect against bad messages.
- [ ] Are the relevant individuals enabled in the various components of Confluent Cloud and Confluent Platform (e.g. Connect, ksqlDB, Schema Registry)?
    - [ ] If Kafka Connect is to be used do we understand the important concepts (e.g. Monitoring, Tuning,Error handling, DLQs, SMTs)?
    - [ ] If ksqlDB is to be used, do we understand concepts such as Tables and Streams, Push and Pull queries, Joins, and Pipelines?
    - [ ] If schema registry is to be used, do we understand concepts such as schema compatibility settings, schema validation, supported formats (e.g. JSON, PROTOBUF)?

Section 2: Roles and Responsibilities

- [ ] Have individuals and/or teams been identified for the following tasks?
  - [ ] Operations (e.g. Monitoring, alerting)
  - [ ] Automation
  - [ ] infrastructure management (e.g. Upgrades)
  - [ ] Change Control
  - [ ] Other
- [ ] Have the right people been identified for defining business requirements and SLAs (uptime, latency, retention)?
- [ ] Have the right people been identified to provide requirements for:
  - [ ] Encryption
  - [ ] Masking / tokenization
  - [ ] Secrets
  - [ ] Authorization (ACLs vs. RBAC)
  - [ ] Authentication
  - [ ] Auditing

Section 3: Infrastructure

- [ ] Have all external systems which will need to connect to Kakfa (upstream and downstream) been identified?
    - [ ] Are these systems available to source/sink to with Kafka Connect?
- [ ] Have network boundaries / security policies been reviewed and the appropriate cloud solution identified?
- [ ] Have requirements been captured and shared for the following?
    - [ ] Encryption
    - [ ] Masking / tokenization
    - [ ] Secrets
    - [ ] Authorization (ACLs vs. RBAC)
    - [ ] Authentication
    - [ ] Auditing
- [ ] Is automation of cluster management required?
    - [ ] Are we aware of the supported tools / mechanisms within Confluent Cloud?
        - [ ] Terraform
        - [ ] Confluent Cloud CLI
        - [ ] Admin Client API
- [ ] Will downtime be permitted for maintenance (upgrades etc)?
- [ ] Do we need to enforce a strategy for topic naming?
    - [ ] Do we feel enabled to take the right approach for this?
- [ ] Do we need to enforce a standard for provisioning of Confluent Components?
    - [ ] For example, do we have customers/tenants of our Confluent Cloud environment who will want us to provision or deprovision Clusters, Topics, Connectors, ksqlDB applications, Schemas?
- [ ] Are we aware of how to review what components exist within Confluent Cloud, where they sit within our pipeline, and what the flow of data looks like?

Section 4: Operations

- [ ] SLAs:
  - [ ] What are requirements for uptime, latency and throughput?
  - [ ] What are the requirements for data retention?
  - [ ] Do we know how to translate our SLAs into monitoring alerts?
  - [ ] Are we able to let the right people know when workloads are too high or too low (for example, can we differentiate between a “slow day” and a problem upstream)?
- [ ] Monitoring:
  - [ ] Have we set the appropriate monitoring thresholds to alert on?
  - [ ] What tool will be used for alerting/monitoring?
- [ ] Support:
  - [ ] Have the right people been added to the support portal?
  - [ ] Has the appropriate support level been set?
  - [ ] Do you know the SLAs for the various support levels and how to escalate?
- [ ] Process and Procedures:
  - [ ] Are run-books needed and who will produce them?

Section 5: Define and test your use-cases

- [ ] Security:
  - [ ] Encryption
  - [ ] Authorization
  - [ ] Authentication
  - [ ] Secrets
- [ ] Functionality:
  - [ ] Connectivity to external systems
  - [ ] Stream Processing
  - [ ] Other
- [ ] Performance:
  - [ ] Throughput / Latency
  - [ ] Durability / Availability
  - [ ] Scaling
- [ ] SLAs:
  - [ ] Retention
  - [ ] Recoverability
  - [ ] RPO/RTO