

Alibaba Cloud

Realtime Compute Introduction

Document Version: 20200921

Legal disclaimer

Alibaba Cloud reminds you to carefully read and fully understand the terms and conditions of this legal disclaimer before you read or use this document. If you have read or used this document, it shall be deemed as your total acceptance of this legal disclaimer.

1. You shall download and obtain this document from the Alibaba Cloud website or other Alibaba Cloud-authorized channels, and use this document for your own legal business activities only. The content of this document is considered confidential information of Alibaba Cloud. You shall strictly abide by the confidentiality obligations. No part of this document shall be disclosed or provided to any third party for use without the prior written consent of Alibaba Cloud.
2. No part of this document shall be excerpted, translated, reproduced, transmitted, or disseminated by any organization, company or individual in any form or by any means without the prior written consent of Alibaba Cloud.
3. The content of this document may be changed because of product version upgrade, adjustment, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and an updated version of this document will be released through Alibaba Cloud-authorized channels from time to time. You should pay attention to the version changes of this document as they occur and download and obtain the most up-to-date version of this document from Alibaba Cloud-authorized channels.
4. This document serves only as a reference guide for your use of Alibaba Cloud products and services. Alibaba Cloud provides this document based on the "status quo", "being defective", and "existing functions" of its products and services. Alibaba Cloud makes every effort to provide relevant operational guidance based on existing technologies. However, Alibaba Cloud hereby makes a clear statement that it in no way guarantees the accuracy, integrity, applicability, and reliability of the content of this document, either explicitly or implicitly. Alibaba Cloud shall not take legal responsibility for any errors or lost profits incurred by any organization, company, or individual arising from download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, take responsibility for any indirect, consequential, punitive, contingent, special, or punitive damages, including lost profits arising from the use or trust in this document (even if Alibaba Cloud has been notified of the possibility of such a loss).
5. By law, all the contents in Alibaba Cloud documents, including but not limited to pictures, architecture design, page layout, and text description, are intellectual property of Alibaba Cloud and/or its affiliates. This intellectual property includes, but is not limited to, trademark rights, patent rights, copyrights, and trade secrets. No part of this document shall be used, modified, reproduced, publicly transmitted, changed, disseminated, distributed, or published without the prior written consent of Alibaba Cloud and/or its affiliates. The names owned by Alibaba Cloud shall not be used, published, or reproduced for marketing, advertising, promotion, or other purposes without the prior written consent of Alibaba Cloud. The names owned by Alibaba Cloud include, but are not limited to, "Alibaba Cloud", "Aliyun", "HiChina", and other brands of Alibaba Cloud and/or its affiliates, which appear separately or in combination, as well as the auxiliary signs and patterns of the preceding brands, or anything similar to the company names, trade names, trademarks, product or service names, domain names, patterns, logos, marks, signs, or special descriptions that third parties identify as Alibaba Cloud and/or its affiliates.
6. Please directly contact Alibaba Cloud for any errors of this document.

Document conventions









Style	Description	Example
 Danger	A danger notice indicates a situation that will cause major system changes, faults, physical injuries, and other adverse results.	 Danger: Resetting will result in the loss of user configuration data.
 Warning	A warning notice indicates a situation that may cause major system changes, faults, physical injuries, and other adverse results.	 Warning: Restarting will cause business interruption. About 10 minutes are required to restart an instance.
 Notice	A caution notice indicates warning information, supplementary instructions, and other content that the user must understand.	 Notice: If the weight is set to 0, the server no longer receives new requests.
 Note	A note indicates supplemental instructions, best practices, tips, and other content.	 Note: You can use Ctrl + A to select all files.
>	Closing angle brackets are used to indicate a multi-level menu cascade.	Click Settings> Network> Set network type .
Bold	Bold formatting is used for buttons, menus, page names, and other UI elements.	Click OK .
Courier font	Courier font is used for commands	Run the <code>cd /d C:/window</code> command to enter the Windows system folder.
<i>Italic</i>	Italic formatting is used for parameters and variables.	<code>bae log list --instanceid</code> <i>Instance_ID</i>
[] or [a b]	This format is used for an optional value, where only one item can be selected.	<code>ipconfig [-all -t]</code>
{ } or {a b}	This format is used for a required value, where only one item can be selected.	<code>switch {active stand}</code>

Table of Contents

1.What is Alibaba Cloud Realtime Compute for Apache Flink? -----	05
2.Service types -----	08
3.System architecture of Realtime Compute in exclusive mode -----	09
4.Scenarios -----	11

1. What is Alibaba Cloud Realtime Compute for Apache Flink?

Alibaba Cloud Realtime Compute for Apache Flink provides an end-to-end, high-performance platform to process big data in real time based on Apache Flink. It is widely used to process streaming data or offline data.

Features

- **Powerful real-time computing functions**

Alibaba Cloud Realtime Compute for Apache Flink integrates a wide range of functions to simplify the development process. These functions include:

- A powerful engine is used. This engine offers the following advantages:
 - Provides Flink SQL that enables automatic data recovery from failures. This ensures accurate data processing when failures occur. For more information, see [Flink SQL overview](#).
 - Supports a variety of built-in functions, such as string, date, and aggregate functions.
 - Enables accurate control over computing resources. This provides isolation between the jobs of different tenants.
- Realtime Compute for Apache Flink outperforms Apache Flink by three to four times when measured by key performance metrics. For example, in Realtime Compute for Apache Flink, the data processing delay is reduced to seconds. The throughput of a job reaches millions of data records per second, and a cluster can contain thousands of nodes.
- Realtime Compute for Apache Flink integrates various cloud-based data stores, such as DataHub, Log Service, ApsaraDB for RDS, Tablestore, and AnalyticDB for MySQL. Realtime Compute for Apache Flink can read data from and write data to these systems with minimal data integration.

- **Managed real-time computing services**

Unlike open source or user-developed streaming data services, Realtime Compute for Apache Flink is a fully managed stream processing engine. You can query streaming data without the need to deploy or manage any infrastructure. With Realtime Compute for Apache Flink, you can use streaming data services with a few clicks. Realtime Compute for Apache Flink integrates services such as data storage, data development, data administration, monitoring, and alerting. This allows you to use cost-effective streaming data services for trial and migrate your data for deployment. Realtime Compute for Apache Flink also supports complete tenant isolation. Tenant isolation and protection extend from the top application layer to the underlying infrastructure layer. This helps ensure the security and privacy of your data.

- **Low costs in labors and compute clusters**

Alibaba Cloud has made many improvements to the SQL execution engine, which allows you to create jobs more cost-effectively than open source Flink jobs. Realtime Compute for Apache Flink is more cost-effective than open source stream frameworks in both development and production costs. For example, you must consider the following costs for project budget:

- Labor costs on writing Flink jobs with complex business logic by using Java code
- Costs on job debugging, testing, optimization, and publishing

- Long-term O&M costs of open-source software such as Flink or ZooKeeper

Realtime Compute for Apache Flink allows you to fully focus on your business without the need to consider these cost issues.



Product positioning

- Realtime Compute for Apache Flink is able to:
 - Collect data about page views (PVs) and unique visitors (UVs) in real time.
 - Collect data about the average traffic flow at a traffic checkpoint within a certain period of time, such as five minutes.
 - Collect and display the pressure data of hydroelectric dams.
 - Report alerts for financial thefts in online payment services based on fixed rules.
- Realtime Compute for Apache Flink has limits in the following scenarios:
 - The stored procedure of Oracle databases cannot be replaced by Realtime Compute for Apache Flink because they are designed to handle issues in different fields.
 - Spark jobs cannot be seamlessly migrated to Realtime Compute for Apache Flink. You can rebuild and then migrate the real-time computing part of Spark jobs from Spark to Realtime Compute for Apache Flink. After the migration, the costs on the O&M and development of Spark jobs are reduced.
 - Realtime Compute for Apache Flink does not support alerting by multiple complex rules engines. If a single data record has an alert in which multiple complex rules are specified, the alert continues to change when the system is running. We recommend that you use the rules engine system to solve this issue.

Realtime Compute for Apache Flink uses Flink SQL and user-defined functions (UDFs) to provide services. It provides an end-to-end development tool for data warehousing developers and data analysts to perform streaming data analysis, statistics, and processing. You can write Flink SQL to analyze streaming data without the need to be involved in the development of the underlying code.

Terms

Term	Description
compute cluster	A compute cluster is a distributed cluster system that hosts computing tasks of Realtime Compute for Apache Flink and runs on YARN. Realtime Compute for Apache Flink has two modes: exclusive mode and shared mode. For more information, see Overview .
web console	Realtime Compute for Apache Flink provides a complete set of integrated development platform (IDE) tools to implement end-to-end data storage, data development, data administration, monitoring, and alerting functions to help you develop your business.

Term	Description
<p>project</p>	<p>In Realtime Compute for Apache Flink, a project is a basic unit used to manage clusters, jobs, resources, and users. You can join existing projects as a RAM user or create projects.</p> <div data-bbox="557 407 1383 521" style="background-color: #e0f2f7; padding: 5px;"> <p> Note Projects of Realtime Compute for Apache Flink allow concurrent operations by multiple RAM users.</p> </div>
<p>CU</p>	<p>In Realtime Compute for Apache Flink, a CU is a basic unit of jobs, with specified CPU cores, memory, and I/O capabilities. A job of Realtime Compute for Apache Flink can use one or more CUs.</p> <p>One CU represents 1 CPU core and 4 GB memory. The processing capability of one CU depends on the complexity of business operations:</p> <ul style="list-style-type: none"> • For simple operations such as single-stream filtering and string conversion, one CU can process 10,000 data records per second. • For complex operations such as operations that use a JOIN clause, GROUP BY clause, or window function, one CU can process 1,000 to 5,000 data records per second. <div data-bbox="557 949 1383 1064" style="background-color: #e0f2f7; padding: 5px;"> <p> Note The specific processing capability of CUs in Realtime Compute for Apache Flink depends on your business.</p> </div>

2. Service types

Before you purchase Realtime Compute for Apache Flink, you must understand the features of each service type to select suitable instances for your business.

Features

Realtime Compute for Apache Flink has five service types. The following table describes the features of each type.

Service type	Product line	Release status	Deployment mode	Compute engine	Development platform
Fully-managed Flink (ACK-based)	New product line	Public preview	Kubernetes-based deployment	Flink	VVP
Semi-managed Flink (ACK-based)	New product line	Released	Kubernetes-based deployment	Flink	VVP
Semi-managed Flink (EMR-based)	New product line	To be released	YARN-based deployment	Flink	VVP
Blink exclusive cluster (original product line)	Original product line	Released	YARN-based deployment	Blink	Bayes
Blink shared cluster (original product line)	Original product line	Phased-out	YARN-based deployment	Blink	Bayes

Selection suggestions

- **Fully-managed Flink (ACK-based):** If your company or team wants to focus on business development without the need to spend extra time on cluster O&M, we recommend that you use this service type.
- **Semi-managed Flink (ACK-based):** If your company or team wants to gain full control over all the cluster resources on the basis of business development and is familiar with Alibaba Cloud Container Service for Kubernetes (ACK), we recommend that you use this service type.
- **Semi-managed Flink (EMR-based):** If your company or team wants to gain full control over all the cluster resources on the basis of business development and is familiar with YARN or Alibaba Cloud E-MapReduce, we recommend that you use this service type.
- **Blink exclusive cluster (original product line):** If you have purchased a Blink exclusive cluster of Realtime Compute for Apache Flink, you can continue to use this service type.
- **Blink shared cluster (original product line):** This service type is phased out. You can only scale out or scale in the existing Blink shared clusters. We recommend that you do not use this service type.

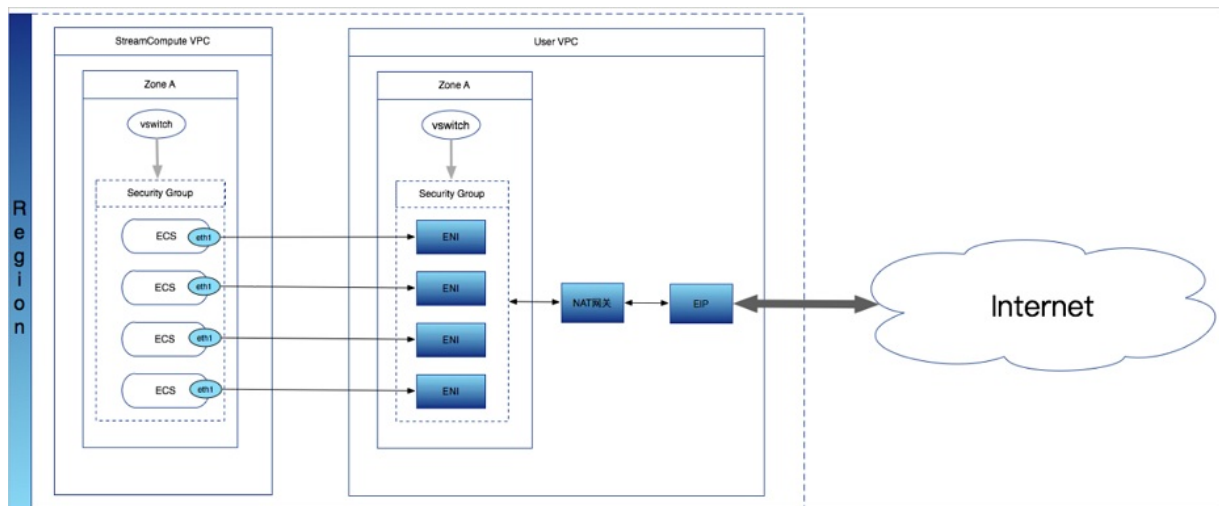
3. System architecture of Realtime Compute in exclusive mode

This topic describes the system architecture of Realtime Compute in exclusive mode.

exclusive mode system architecture benefits

System architecture

The following figure shows the system architecture of Realtime Compute in exclusive mode.



- For Realtime Compute in exclusive mode, all your purchased ECS instances are fully hosted in the VPC of your Realtime Compute cluster. In this mode, you cannot log on to ECS instances.
- When you create a Realtime Compute cluster, Realtime Compute applies for an Elastic Network Interface (ENI) under your Realtime Compute account. You can use this ENI to access all resources in your VPC.
- To enable Internal access for your Realtime Compute cluster, you can bind a network address translation (NAT) gateway and an Elastic IP address (EIP) to the ENI. For more information, see [Associate an EIP with a NAT gateway](#) and [Overview](#).
- The ENI belongs to an independent security group under your Realtime Compute account. To access services of another security group in the VPC, you must configure inbound and outbound rules for the security group.

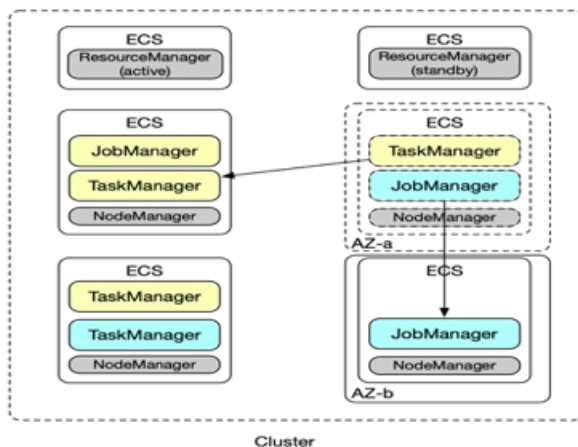
Note You are charged for the use of the ENI only when your Realtime Compute cluster accesses the Internet.

Benefits

- End-to-end real-time data computing development
 - Provides a real-time data processing capability based on Flink SQL, which implements automatic data recovery. This ensures accurate data processing if failures occur.
 - Supports multiple built-in functions, such as string, date, and aggregate functions.
 - Supports various window types, such as tumbling, sliding, and session windows.

- Implements accurate control over computing resources, which ensures resource isolation for jobs.
- Provides the following key performance metrics that are superior to those of open source Flink:
 - The data computing latency can be measured in sub-seconds.
 - The throughput of a single job can reach millions of records per second. A single cluster can consist of thousands of servers.
- Deeply integrates all types of cloud data storage systems such as DataHub, Log Service, ApsaraDB for RDS, Tablestore, and AnalyticDB for MySQL so that you can conveniently read and write data from and to these systems.
- Fully managed Realtime Compute service
 - Uses a fully managed stream computing engine.
 - Allows you to run and query streaming data without the need to provision or manage any infrastructure.
 - Allows you to enable streaming data processing services with one click.
 - Integrates features such as data storage, data development, data O&M, and monitoring and alerts. This reduces both the trial and migration costs of stream processing.
 - Effectively isolates and protects the managed and running services of different tenants.
- Low manpower and cluster costs
 - Significantly optimizes the SQL execution engine to deliver computing jobs that are more cost-effective than native Flink jobs.
 - Significantly reduces development and operation costs, which are much lower than those of open source stream frameworks.
- High availability

Assume that an ECS instance or a job of Realtime Compute fails. You can use the JobManager and TaskManager on an available ECS instance in the same zone to achieve high availability of jobs. You can also use the JobManager and TaskManager on an available ECS instance in another zone or region to achieve high availability in zones.



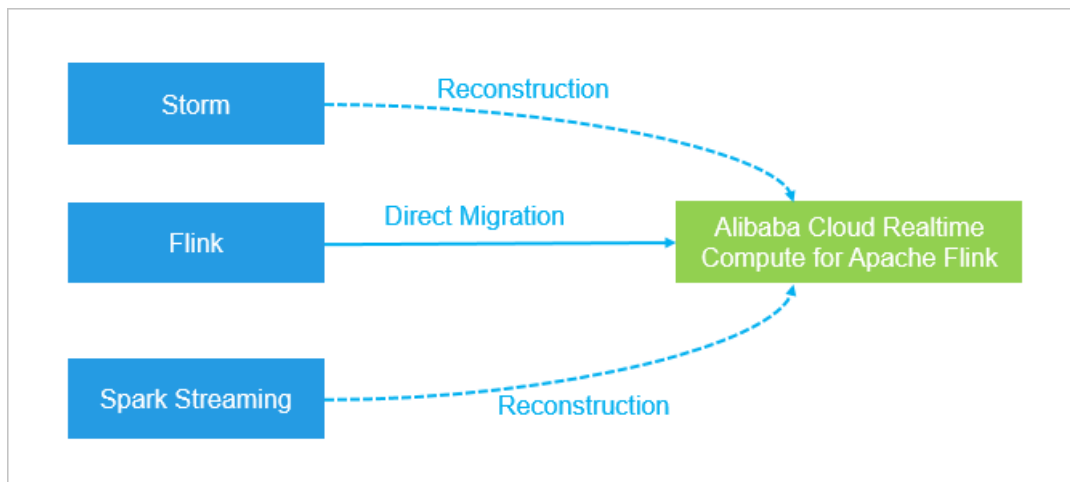
4.Scenarios

Realtime Compute for Apache Flink is widely used for real-time big data computing. This topic describes how to apply Realtime Compute for Apache Flink to different enterprise departments and technologies.

existing stream processing system enterprise department technology

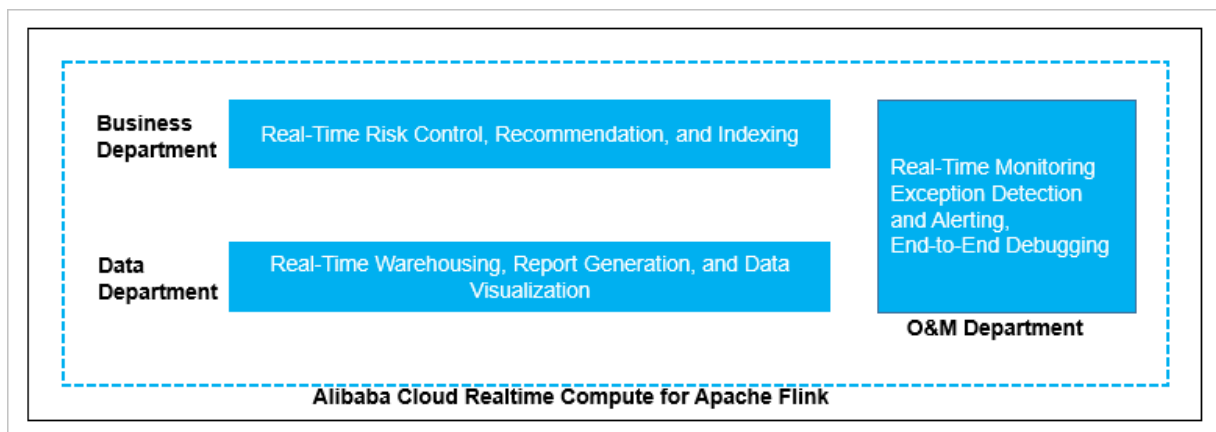
Existing stream processing system

If you have installed the Flink system on your local server, you can directly migrate it to Realtime Compute for Apache Flink. If you have installed the Storm or Spark Streaming system on your local server, you can rebuild and then migrate it to Realtime Compute for Apache Flink.



Enterprise departments

Realtime Compute for Apache Flink provides different features for the following departments:



- Business department: real-time risk control, real-time recommendation, and real-time indexing of search engines.
- Data department: real-time data warehousing, real-time reports, and real-time dashboards
- O&M department: real-time monitoring, real-time exception detection and alerting, and end-to-end debugging.

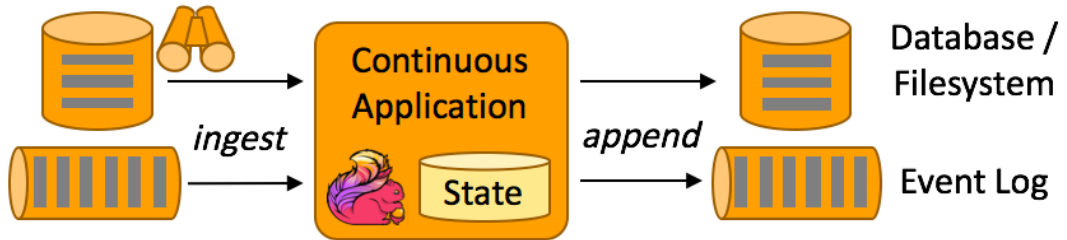
Technologies

From the technical perspective, Realtime Compute for Apache Flink is suitable for the following scenarios:

- Real-time extract, transform, load (ETL) and data streams

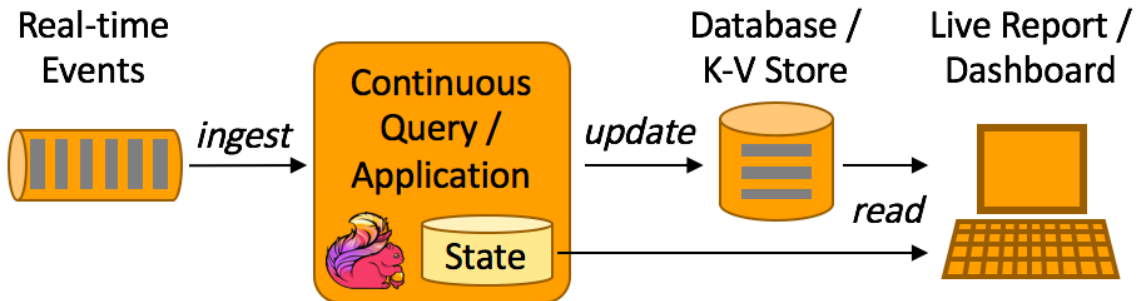
Data is delivered from point A to point B by using the real-time ETL procedure and data streams. During data delivery, data cleansing and integration may be required, such as real-time indexing in the search system and ETL procedure in real-time data warehousing.

Real-time Events



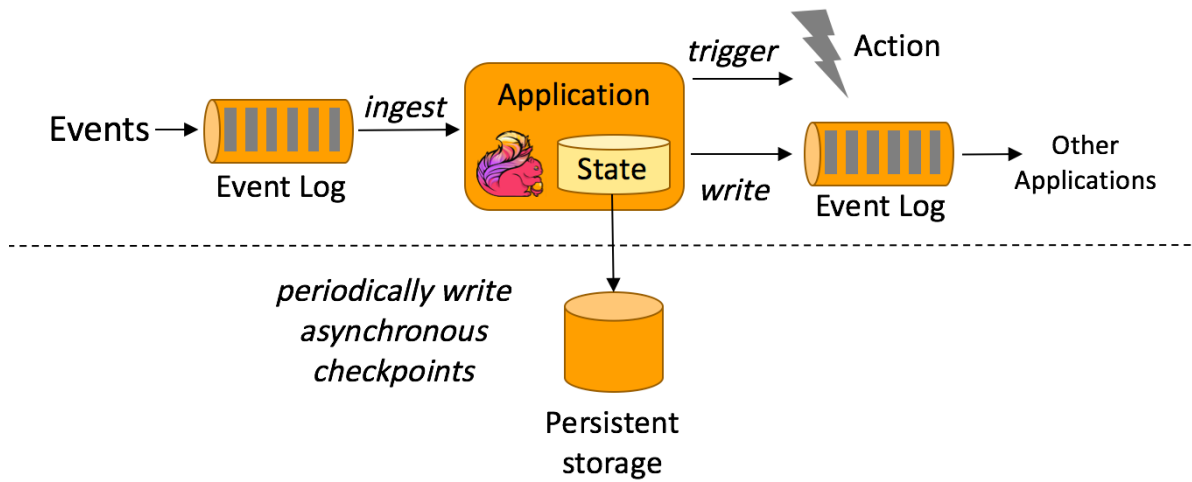
- Real-time data analysis

Data analysis is a process to extract and integrate required information from raw data to achieve your business objectives. For example, you can view the top 10 products sold per day, the average turn-around time in the warehouse, the average document click rate, and the open rate for push notifications. Real-time data analysis allows you to view real-time reports or dashboards.



- Event-driven applications

An event-driven application is a system that processes or reacts to subscription events. Event-driven applications depend on internal states and respond to suspicious events detected during fraud detection or in the risk control system or O&M exception detection system. If the behavior of a user triggers a risk control point, the system captures the event and analyzes the current and previous user behavior to determine whether to perform risk control over the user.



Note The preceding technology flowcharts are obtained from [Apache Flink official website](#).