



Amazon Kinesis Firehose

Amazon Kinesis Data Firehose is a service within the Amazon Kinesis suite that enables you to capture, transform, and load streaming data into AWS data stores and analytics services. Here are some key features and use cases for Kinesis Data Firehose:

1. **Streaming Data Ingestion:** Kinesis Data Firehose can ingest large volumes of streaming data from various sources, such as IoT devices, logs, clickstreams, and social media feeds.
2. **Data Transformation:** It allows you to transform streaming data in real-time before loading it into the destination. You can use AWS Lambda functions or AWS Glue ETL jobs to transform data formats, enrich data, or filter out irrelevant information.
3. **Integration with AWS Services:** Kinesis Data Firehose seamlessly integrates with several AWS services, including Amazon S3, Amazon Redshift, Amazon Elasticsearch Service, and Amazon OpenSearch Service (successor to Amazon Elasticsearch Service). This enables you to store, analyze, and visualize streaming data using familiar AWS tools.
4. **Data Delivery Guarantees:** Kinesis Data Firehose ensures reliable delivery of data to the destination with built-in fault tolerance and automatic scaling. It handles data buffering, retries, and error handling to ensure data delivery reliability.
5. **Cost Optimization:** The service helps optimize costs by automatically scaling resources based on the incoming data volume and offering options for data compression and data format conversion. You only pay for the data ingested and the data delivery throughput.
6. **Real-time Analytics:** By streaming data into analytics services like Amazon Redshift or Amazon Elasticsearch Service, you can perform real-time analytics, monitor trends, and gain insights into your data without the need for complex data loading processes.
7. **Log and Event Data Processing:** Kinesis Data Firehose is commonly used for ingesting and processing log and event data from applications, servers, and infrastructure. It simplifies the process of collecting and analyzing log data for monitoring, troubleshooting, and compliance purposes.
8. **Data Lake and Data Warehousing:** You can use Kinesis Data Firehose to load streaming data into data lakes on Amazon S3 or data warehouses like Amazon Redshift for long-term storage, batch processing, and historical analysis.

To begin with the Lab:

1. Now from the Kinesis dashboard navigate to Firehose.

Amazon Kinesis > Dashboard

Amazon Kinesis Info

Amazon Kinesis makes it easy to collect, process, and analyze data streams in real time, so you can get timely insights and react quickly to new information.

Data Streams Info
Serverless streaming data service that simplifies the capture, processing, and storage of data streams at any scale.

Total data streams
1
[Create data stream](#)

Amazon Data Firehose Info
The easiest way to reliably ingest, transform, and deliver streaming data into destinations in real time.

Amazon Kinesis Data Firehose is now called Amazon Data Firehose.
[Go to Amazon Data Firehose](#)

Managed Apache Flink Info
Get actionable insights from streaming data in real time.

Kinesis Data Analytics for Apache Flink is now called Amazon Managed Service for Apache Flink.
[Go to Managed Apache Flink](#)

- Now click on create firehose stream.

Analytics

Amazon Data Firehose

Real-time streaming delivery for any data, at any scale, and at low-cost.

Amazon Data Firehose provides the easiest way to reliably ingest, transform, and deliver streaming data into data lakes, data warehouses, and analytics services.

Getting started
Create a Firehose stream that processes and delivers streaming data to destinations.

[Create Firehose stream](#)

- Here you have to select source as Kinesis data streams and in the source, settings choose your Kinesis Stream.
- You will see that it has allotted a firehose stream name itself. Scroll down.

► Amazon Data Firehose: How it works

Choose source and destination

Specify the source and the destination for your Firehose stream. You cannot change the source and destination of your Firehose stream once it has been created.

Source | [Info](#)

Amazon Kinesis Data Streams ▾

Destination | [Info](#)

Amazon S3 ▾

Source settings

Kinesis data stream

arn:aws:kinesis:ap-south-1:878893308172:stream/demostream

[Browse](#)

[Create](#) 

Format: arn:aws:kinesis:[Region]:[AccountId]:stream/[StreamName]

- After that choose your destination bucket in S3. Then just scroll down create your stream.

Destination settings [Info](#)

Specify the destination settings for your Firehose stream.

S3 bucket

s3://demouser1221

[Browse](#)

[Create](#) 

Format: s3://bucket

New line delimiter

You can configure your Firehose stream to add a new line delimiter between records in objects that are delivered to Amazon S3.

- Not enabled
 Enabled

Amazon Data Firehose > [Firehose streams](#) > KDS-S3-eGX0s

KDS-S3-eGX0s [Info](#)

[Delete Firehose stream](#)

Firehose stream details

Status
 Active

Destination
Amazon S3

Data transformation
Not enabled

Creation time
February 18, 2024 at 22:47 GMT+5:30

Source
Amazon Kinesis Data Streams

ARN
 arn:aws:firehose:ap-south-1:878893308172:deliverystream/KDS-S3-eGX0s

Dynamic partitioning
Not enabled

Error logs status
 0 Destination error logs

- Now you have to use the same program which you have used in the previous lab for Kinesis. Use the program to send data and run that same program again.
- After some time if you will go to your S3 bucket, you will see a folder created. Now open that folder.

The screenshot shows the AWS S3 console interface. At the top, there are tabs for Objects, Properties, Permissions, Metrics, Management, and Access Points. The Objects tab is currently selected. Below the tabs, there is a header bar with various actions: Copy S3 URI, Copy URL, Download, Open, Delete, Actions, Create folder, and Upload. A search bar labeled 'Find objects by prefix' is present. The main area displays a table of objects. The table has columns for Name, Type, Last modified, Size, and Storage class. One object is listed: '2024/' which is a Folder.

- You will see this file inside your folder.
- Now you have to download this file and open it in notepad.

The screenshot shows the contents of the '2024/' folder. The 'Objects' tab is selected. The table shows one object: 'KDS-53-eGX0s-1-2024-02-18-17-19-38-6a645f0f-14cf-4d01-8c88-3c319a87467f'. The file was uploaded on February 18, 2024, at 22:54:39 (UTC+05:30) and has a size of 190.0 B. The storage class is Standard.

- And in that file you can see the log details.

The screenshot shows a text editor window with the following JSON content:

```
{"logId":1,"logDetails":"LogDetails1"}, {"logId":2,"logDetails":"LogDetails2"}, {"logId":3,"logDetails":"LogDetails3"}, {"logId":4,"logDetails":"LogDetails4"}, {"logId":5,"logDetails":"LogDetails5"}
```

As for cleanup go to your firehose and delete your delivery stream than go to Kinesis and delete that stream too.