



# Kinesis Data Firehose

Amazon Kinesis Data Firehose is a fully managed service for real-time data delivery that allows you to reliably load streaming data into data lakes, data stores, and analytics services. It can ingest data from various sources, transform or process it using AWS Lambda, and then automatically deliver it to destinations like Amazon S3, Amazon Redshift, Amazon Elasticsearch Service (now OpenSearch), or even third-party services like Splunk.

Here's an overview of how Kinesis Data Firehose works:

## Key Features:

1. **Real-Time Streaming Data Capture:** Kinesis Data Firehose can continuously capture and transform streaming data in real-time from services like Kinesis Data Streams, IoT devices, or custom applications.
2. **Data Transformation:** You can use AWS Lambda to transform or enrich the data before it is delivered. This is useful for cleaning or reformatting data on the fly.
3. **Automatic Scaling:** The service automatically scales to accommodate the incoming data volume without requiring user intervention.
4. **Buffering and Batching:** You can configure Kinesis Data Firehose to buffer data for a specified period (or until a certain data size is reached) before delivering it. This allows optimization of both costs and throughput.
5. **Delivery Destinations:** Firehose supports multiple destinations, including:
  - **Amazon S3:** for durable storage.
  - **Amazon Redshift:** for analytics and querying.
  - **Amazon OpenSearch Service:** for search and visualization.
  - **Splunk:** for data analysis and monitoring.
6. **Monitoring and Security:**
  - Integration with Amazon CloudWatch allows you to monitor data flow, delivery success, and failure rates.
  - Supports encryption (both at rest and in transit), and data compression (like GZIP, ZIP), which can reduce storage and transfer costs.

## Common Use Cases:

- **Real-Time Analytics:** Capture and analyze log data, sensor data, and other real-time metrics for insights.
- **Data Lakes:** Stream large volumes of data to Amazon S3 to build a data lake.
- **Event-Driven Architectures:** Integrate with Lambda for real-time data transformations and event processing.

## Benefits:

- **Fully Managed:** No infrastructure management is required. AWS handles provisioning, scaling, and maintenance.
- **Ease of Integration:** Works seamlessly with other AWS services like S3, Redshift, and OpenSearch.
- **Cost-Effective:** Pay only for the volume of data processed and delivered, with no upfront costs.

**In this exercise, the goal is to set up a data stream using Amazon Kinesis Data Firehose and deliver the data to an S3 bucket. You'll first create an S3 bucket and a Kinesis Data Stream using simple AWS CLI commands. After confirming the resources on the AWS Console, you configure a Kinesis Data Firehose to connect the stream (as the source) with the S3 bucket (as the destination).**

**Once the Firehose is created, you test it by sending demo data through the stream, which automatically gets stored in the S3 bucket. After a few minutes, you can download and view the file in your S3 bucket, which contains the streamed data.**

**The end goal is to demonstrate how data from a Kinesis stream can be automatically delivered to an S3 bucket, making it accessible for storage, analysis, or processing.**

## To begin with the Lab:

1. In this lab, we will stream data to Amazon S3 using Kinesis Data Firehose.
2. First, we need to create an S3 bucket and Kinesis Data Stream using the commands given below. Also, you will have a separate file for the commands.

```
`aws s3api create-bucket --bucket demos3bucket-$RANDOM_ID`
```

```
`aws kinesis create-stream --stream-name demodatastream --shard-count 1`
```

```
`aws kinesis describe-stream-summary --stream-name demodatastream`
```

3. After running these commands, you can go to the AWS Console and check your resources. Below you can see that our S3 bucket and Kinesis Data Stream have been created.

Amazon S3 > Buckets > demos3bucket-481q3p

## demos3bucket-481q3p [Info](#)

Objects Properties Permissions Metrics Management Access Points

### Objects (0) [Info](#)

[C](#) [Copy S3 URI](#) [Copy URL](#) [Download](#) [Open](#) [Delete](#) [Actions](#) [Create folder](#) [Upload](#)

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

Find objects by prefix

Name	Type	Last modified	Size	Storage class
No objects You don't have any objects in this bucket.				

[Upload](#)

Amazon Kinesis > Data streams

### Data streams (1) [Info](#)

[Process data in real time](#) [Create a Firehose stream](#) [Actions](#) [Create data stream](#)

Find data streams

Name	Status	Capacity mode	Provisioned shards	Sharing policy	Data retention period	Encryption	Consumers with enhanced fan-out
<a href="#">demodatastream</a>	<a href="#">Active</a>	Provisioned	1	No	1 day	Disabled	0

- Now search for Kinesis Data Firehose and from its dashboard click on Create.

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](#)

- First, you need to choose the Source as Kinesis Data Stream and the destination as Amazon S3.

## Create Firehose stream [Info](#)

### ► 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

6. Then in the source settings choose your data Stream and leave the default name as it is.

#### Source settings

Kinesis data stream

arn:aws:kinesis:us-east-1:878893308172:stream/demodatastream

[Browse](#)

[Create](#) 

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

#### Firehose stream name

Firehose stream name

KDS-S3-ePtHl

Acceptable characters are uppercase and lowercase letters, numbers, underscores, hyphens, and periods.

7. After that for the destination choose your S3 bucket which you created earlier and then create your data Firehose.

## Destination settings [Info](#)

Specify the destination settings for your Firehose stream.

### S3 bucket

[Browse](#)[Create !\[\]\(74d4806277d7e73349d8e8c0897931e9\_img.jpg\)](#)

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

### Dynamic partitioning [Info](#)

Dynamic partitioning enables you to create targeted data sets by partitioning streaming S3 data based on partitioning keys. You can partition your source data with inline parsing and/or the specified AWS Lambda function. You can enable dynamic partitioning only when you create a new Firehose stream. You cannot enable dynamic partitioning for an existing Firehose stream. Enabling dynamic partitioning incurs additional costs per GiB of partitioned data. For more information, see [Amazon Data Firehose pricing](#).

- Not enabled  
 Enabled

8. Once your Firehose is created then you can expand the test with demo data feature.

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

KDS-S3-ePtHl [Info](#)

[Delete Firehose stream](#)

### Firehose stream details

Status  Active	Destination Amazon S3	Data transformation Not enabled	Creation time October 05, 2024 at 11:04 GMT+5:30
Source Amazon Kinesis Data Streams	ARN  arn:aws:firehose:us-east-1:878893308172:deliverystream/KD S-S3-ePtHl	Dynamic partitioning Not enabled	Error logs status  0 Destination error logs

### ▶ Test with demo data [Info](#)

Ingest simulated data to test the configuration of your Firehose stream. Standard Amazon Data Firehose charges apply.

9. Here you must click on Start sending demo data and this will start sending some data to your S3 bucket and after a few minutes you must stop sending the data.

## ▼ Test with demo data Info

Ingest simulated data to test the configuration of your Firehose stream. Standard Amazon Data Firehose charges apply.

This test runs a script in your browser to put demo data in your Amazon Data Firehose stream, which sends to your Amazon S3 destination.

```
1 {
2   "TICKER_SYMBOL": "QXZ",
3   "SECTOR": "HEALTHCARE",
4   "CHANGE": -0.05,
5   "PRICE": 84.51
6 }
```

### Step 1

Start sending demo data to your Firehose stream. If you already have data streaming to this destination, demo data is sent along with your source records.

[Start sending demo data](#)

### Step 2

Stop sending demo data to your Firehose stream after you've concluded your test to stop incurring usage charges.

[Stop sending demo data](#)

10. Then go to S3 and wait for a few minutes, so that your data can pop up. After few minutes you can see that we have the data in our S3 bucket.
11. Now go inside the folders and download the file then open it in notepad.

Amazon S3 > Buckets > demos3bucket-481q3p

**demos3bucket-481q3p Info**

[Objects](#) [Properties](#) [Permissions](#) [Metrics](#) [Management](#) [Access Points](#)

**Objects (1) Info**

[Delete](#) [Actions ▾](#) [Create folder](#)

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

[Find objects by prefix](#)  **1**

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	<a href="#">2024/</a>	Folder	-	-	-

12. This type of data you will be able to see in the file.

```

("CHANGE": -1.79, "PRICE": 99.85, "TICKER_SYMBOL": "AAPL", "SECTOR": "TECHNOLOGY"} {"CHANGE": -1.88, "PRICE": 64.15, "TICKER_SYMBOL": "TGH", "SECTOR": "FINANCIAL"} {"CHANGE": -12.32, "PRICE": 723.19, "TICKER_SYMBOL": "AMZN", "SECTOR": "TECHNOLOGY"} {"CHANGE": 0.35, "PRICE": 87.09, "TICKER_SYMBOL": "QXZ", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.15, "PRICE": 5.66, "TICKER_SYMBOL": "DEG", "SECTOR": "ENERGY"} {"CHANGE": -3.45, "PRICE": 73.96, "TICKER_SYMBOL": "WMT", "SECTOR": "RETAIL"} {"CHANGE": -0.81, "PRICE": 58.38, "TICKER_SYMBOL": "SAC", "SECTOR": "ENERGY"} {"CHANGE": -0.19, "PRICE": 99.19, "TICKER_SYMBOL": "NFLX", "SECTOR": "TECHNOLOGY"} {"CHANGE": 1.07, "PRICE": 108.92, "TICKER_SYMBOL": "AAPL", "SECTOR": "TECHNOLOGY"} {"CHANGE": -0.98, "PRICE": 65.78, "TICKER_SYMBOL": "ASD", "SECTOR": "FINANCIAL"} {"CHANGE": -0.38, "PRICE": 5.5, "TICKER_SYMBOL": "DEG", "SECTOR": "ENERGY"} {"CHANGE": 0.19, "PRICE": 52.52, "TICKER_SYMBOL": "KIN", "SECTOR": "ENERGY"} {"CHANGE": -0.61, "PRICE": 46.48, "TICKER_SYMBOL": "WFS", "SECTOR": "FINANCIAL"} {"CHANGE": -0.12, "PRICE": 1.22, "PRICE": 30.84, "TICKER_SYMBOL": "CNC", "SECTOR": "MANUFACTURING"} {"CHANGE": -0.74, "PRICE": 24.64, "TICKER_SYMBOL": "ASD", "SECTOR": "RETAIL"} {"CHANGE": -0.66, "PRICE": 5.73, "TICKER_SYMBOL": "DEG", "SECTOR": "ENERGY"} {"CHANGE": -0.47, "PRICE": 4.7, "TICKER_SYMBOL": "WSP", "SECTOR": "FINANCIAL"} {"CHANGE": -0.99, "PRICE": 187.54, "TICKER_SYMBOL": "WSP", "SECTOR": "FINANCIAL"} {"CHANGE": -0.19, "PRICE": 14.4, "TICKER_SYMBOL": "BAC", "SECTOR": "FINANCIAL"} {"CHANGE": -0.41, "PRICE": 96.08, "TICKER_SYMBOL": "DFT", "SECTOR": "RETAIL"} {"CHANGE": -0.21, "PRICE": 32.93, "TICKER_SYMBOL": "CMB", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.25, "PRICE": 12.26, "TICKER_SYMBOL": "WAS", "SECTOR": "RETAIL"} {"CHANGE": -0.21, "PRICE": 65.44, "TICKER_SYMBOL": "TGT", "SECTOR": "RETAIL"} {"CHANGE": -0.49, "PRICE": 11.77, "TICKER_SYMBOL": "HIV", "SECTOR": "ENERGY"} {"CHANGE": -1.14, "PRICE": 14.13, "TICKER_SYMBOL": "VVS", "SECTOR": "ENERGY"} {"CHANGE": -6.12, "PRICE": 81.92, "TICKER_SYMBOL": "SLW", "SECTOR": "ENERGY"} {"CHANGE": -1.21, "PRICE": 63.91, "TICKER_SYMBOL": "ASD", "SECTOR": "FINANCIAL"} {"CHANGE": -0.25, "PRICE": 14.98, "TICKER_SYMBOL": "VWS", "SECTOR": "ENERGY"} {"CHANGE": -3.67, "PRICE": 85.43, "TICKER_SYMBOL": "ALY", "SECTOR": "ENERGY"} {"CHANGE": -0.34, "PRICE": 10.34, "TICKER_SYMBOL": "NGC", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.08, "PRICE": 1.34, "TICKER_SYMBOL": "NGC", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.19, "PRICE": 5.62, "TICKER_SYMBOL": "NGC", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.01, "PRICE": 75.39, "TICKER_SYMBOL": "SLW", "SECTOR": "ENERGY"} {"CHANGE": -1.63, "PRICE": 43.59, "TICKER_SYMBOL": "VWS", "SECTOR": "ENERGY"} {"CHANGE": -0.03, "PRICE": 152.52, "TICKER_SYMBOL": "Q", "SECTOR": "RETAIL"} {"CHANGE": -6.01, "PRICE": 113.61, "TICKER_SYMBOL": "XTC", "SECTOR": "HEALTHCARE"} {"CHANGE": -9.14, "PRICE": 43.21, "TICKER_SYMBOL": "TBV", "SECTOR": "RETAIL"} {"CHANGE": -2.21, "PRICE": 78.12, "TICKER_SYMBOL": "TBV", "SECTOR": "HEALTHCARE"} {"CHANGE": -1.33, "PRICE": 117.43, "TICKER_SYMBOL": "TOP", "SECTOR": "TECHNOLOGY"} {"CHANGE": -2.78, "PRICE": 62.66, "TICKER_SYMBOL": "TGT", "SECTOR": "RETAIL"} {"CHANGE": -0.21, "PRICE": 136.42, "TICKER_SYMBOL": "DFG", "SECTOR": "TECHNOLOGY"} {"CHANGE": -0.18, "PRICE": 10.93, "TICKER_SYMBOL": "TGT", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.81, "PRICE": 22.16, "TICKER_SYMBOL": "VWS", "SECTOR": "ENERGY"} {"CHANGE": -1.02, "PRICE": 98.17, "TICKER_SYMBOL": "VWS", "SECTOR": "ENERGY"} {"CHANGE": -0.35, "PRICE": 4.93, "TICKER_SYMBOL": "DEG", "SECTOR": "ENERGY"} {"CHANGE": -3.19, "PRICE": 69.97, "TICKER_SYMBOL": "WMT", "SECTOR": "RETAIL"} {"CHANGE": -1.2, "PRICE": 31.6, "TICKER_SYMBOL": "PPL", "SECTOR": "RETAIL"} {"CHANGE": -0.38, "PRICE": 136.8, "TICKER_SYMBOL": "DFG", "SECTOR": "TECHNOLOGY"} {"CHANGE": -0.57, "PRICE": 23.47, "TICKER_SYMBOL": "ALY", "SECTOR": "RETAIL"} {"CHANGE": -0.1, "PRICE": 19.32, "TICKER_SYMBOL": "VWS", "SECTOR": "ENERGY"} {"CHANGE": -0.18, "PRICE": 193.25, "TICKER_SYMBOL": "HIV", "SECTOR": "ENERGY"} {"CHANGE": -0.71, "PRICE": 16.96, "TICKER_SYMBOL": "AZL", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.47, "PRICE": 16.58, "TICKER_SYMBOL": "BFS", "SECTOR": "RETAIL"} {"CHANGE": -0.19, "PRICE": 5.62, "TICKER_SYMBOL": "NGC", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.01, "PRICE": 188.25, "TICKER_SYMBOL": "SED", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.19, "PRICE": 174.04, "TICKER_SYMBOL": "BBM", "SECTOR": "TECHNOLOGY"} {"CHANGE": -3.3, "PRICE": 93.18, "TICKER_SYMBOL": "DFT", "SECTOR": "RETAIL"} {"CHANGE": -0.26, "PRICE": 5.48, "TICKER_SYMBOL": "KIN", "SECTOR": "ENERGY"} {"CHANGE": -1.6, "PRICE": 31.33, "TICKER_SYMBOL": "CMB", "SECTOR": "HEALTHCARE"} {"CHANGE": -1.3, "PRICE": 13.11, "TICKER_SYMBOL": "ENRG", "SECTOR": "ENERGY"} {"CHANGE": -0.16, "PRICE": 35.35, "TICKER_SYMBOL": "CMB", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.01, "PRICE": 10.0, "TICKER_SYMBOL": "SLW", "SECTOR": "ENERGY"} {"CHANGE": -0.17, "PRICE": 5.4, "TICKER_SYMBOL": "DEG", "SECTOR": "ENERGY"} {"CHANGE": -0.57, "PRICE": 152.99, "TICKER_SYMBOL": "WMT", "SECTOR": "RETAIL"} {"CHANGE": -0.17, "PRICE": 1.57, "TICKER_SYMBOL": "TOP", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.42, "PRICE": 19.6, "TICKER_SYMBOL": "PPL", "SECTOR": "FINANCIAL"} {"CHANGE": -0.84, "PRICE": -2.16, "TICKER_SYMBOL": "AAPL", "SECTOR": "TECHNOLOGY"} {"CHANGE": -2.06, "PRICE": 541.95, "TICKER_SYMBOL": "VWZ", "SECTOR": "FINANCIAL"} {"CHANGE": -0.79, "PRICE": 43.34, "TICKER_SYMBOL": "TBV", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.85, "PRICE": 82.55, "TICKER_SYMBOL": "SLW", "SECTOR": "ENERGY"} {"CHANGE": -0.08, "PRICE": 5.13, "TICKER_SYMBOL": "DEG", "SECTOR": "ENERGY"} {"CHANGE": -0.11, "PRICE": 8.21, "TICKER_SYMBOL": "WSP", "SECTOR": "RETAIL"} {"CHANGE": -0.62, "PRICE": 177.53, "TICKER_SYMBOL": "BBM", "SECTOR": "TECHNOLOGY"} {"CHANGE": -4.87, "PRICE": 108.74, "TICKER_SYMBOL": "XTC", "SECTOR": "HEALTHCARE"} {"CHANGE": -1.93, "PRICE": 179.46, "TICKER_SYMBOL": "BMM", "SECTOR": "TECHNOLOGY"} {"CHANGE": -1.38, "PRICE": 540.57, "TICKER_SYMBOL": "VWZ", "SECTOR": "FINANCIAL"} {"CHANGE": -0.39, "PRICE": 224.3, "TICKER_SYMBOL": "QXZ", "SECTOR": "FINANCIAL"} {"CHANGE": -0.32, "PRICE": 198.71, "TICKER_SYMBOL": "QAZ", "SECTOR": "TECHNOLOGY"} {"CHANGE": -1.2, "PRICE": 1.2, "PRICE": 99.75, "TICKER_SYMBOL": "WFS", "SECTOR": "ENERGY"} {"CHANGE": -0.16, "PRICE": 22.17, "TICKER_SYMBOL": "TOP", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.01, "PRICE": 112.96, "TICKER_SYMBOL": "CNC", "SECTOR": "MANUFACTURING"} {"CHANGE": -0.01, "PRICE": 10.32, "TICKER_SYMBOL": "TOP", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.01, "PRICE": 4.77, "TICKER_SYMBOL": "WSP", "SECTOR": "FINANCIAL"} {"CHANGE": -0.01, "PRICE": 1.2, "TICKER_SYMBOL": "TOP", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.01, "PRICE": 10.52, "TICKER_SYMBOL": "TOP", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.08, "PRICE": 16.66, "TICKER_SYMBOL": "BFS", "SECTOR": "RETAIL"} {"CHANGE": -0.27, "PRICE": 178.53, "TICKER_SYMBOL": "CMB", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.39, "PRICE": 53.21, "TICKER_SYMBOL": "CMB", "SECTOR": "HEALTHCARE"} {"CHANGE": -1.42, "PRICE": 33.27, "TICKER_SYMBOL": "CMB", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.37, "PRICE": 5.25, "TICKER_SYMBOL": "KIN", "SECTOR": "ENERGY"} {"CHANGE": -0.98, "PRICE": 153.93, "TICKER_SYMBOL": "MIN", "SECTOR": "RETAIL"} {"CHANGE": -0.41, "PRICE": 29.39, "TICKER_SYMBOL": "PPL", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.14, "PRICE": 116.97, "TICKER_SYMBOL": "XTC", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.14, "PRICE": 15.48, "TICKER_SYMBOL": "DEG", "SECTOR": "ENERGY"} {"CHANGE": -2.57, "PRICE": 162, "TICKER_SYMBOL": "TBV", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.25, "PRICE": 19.85, "TICKER_SYMBOL": "PLM", "SECTOR": "FINANCIAL"} {"CHANGE": -7.85, "PRICE": 91.9, "TICKER_SYMBOL": "WFS", "SECTOR": "ENERGY"} {"CHANGE": -0.13, "PRICE": 2.03, "TICKER_SYMBOL": "SED", "SECTOR": "HEALTHCARE"} {"CHANGE": -0.36, "PRICE": 107.0, "TICKER_SYMBOL": "WSP", "SECTOR": "FINANCIAL"} {"CHANGE": -6.59, "PRICE": 123.56, "TICKER_SYMBOL": "XTC", "SECTOR": "HEALTHCARE"} {"CHANGE": 0, "PRICE": 98.17, "TICKER_SYMBOL": "NFLX", "SECTOR": "TECHNOLOGY"} {"CHANGE": 3.01, "PRICE": 98.42, "TICKER_SYMBOL": "QXZ", "SECTOR": "HEALTHCARE"} {"CHANGE": 0.15, "PRICE": 11.92, "TICKER_SYMBOL": "WAS", "SECTOR": "RETAIL"}

```

13. Once you are done then delete all the resources start by deleting your S3 bucket then your Kinesis data Firehose and then the Data Stream.