

Lesson 08 Demo 08

Creating a Kinesis Data Firehose

Objective: To create a Kinesis Data Firehose, set up an Amazon S3 bucket as the destination, and test the data delivery stream

Tools required: AWS Workspaces

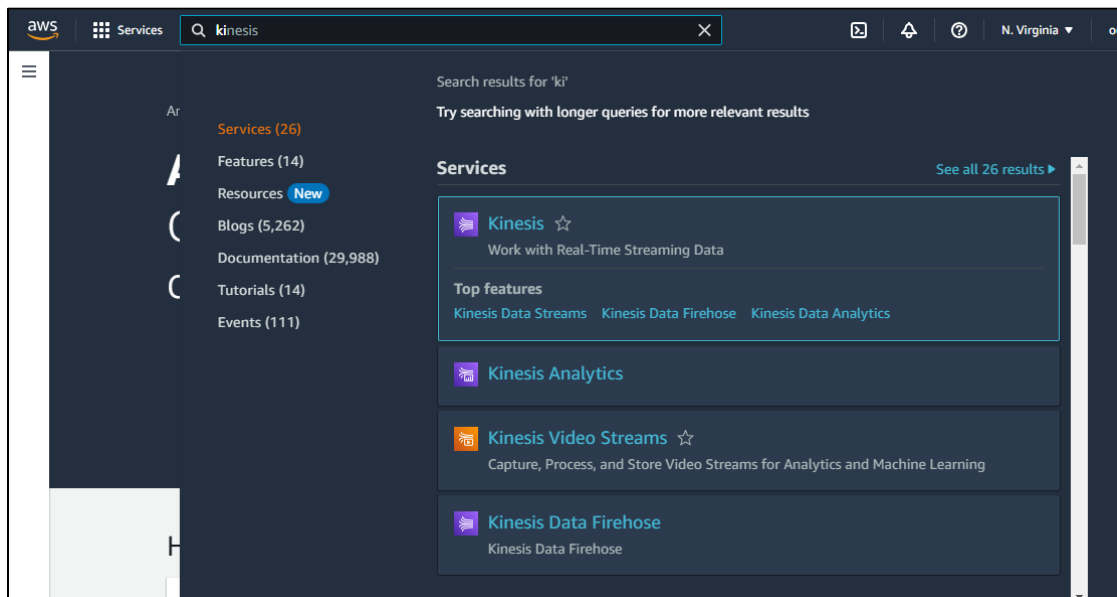
Prerequisites: Amazon account

Steps to be followed:

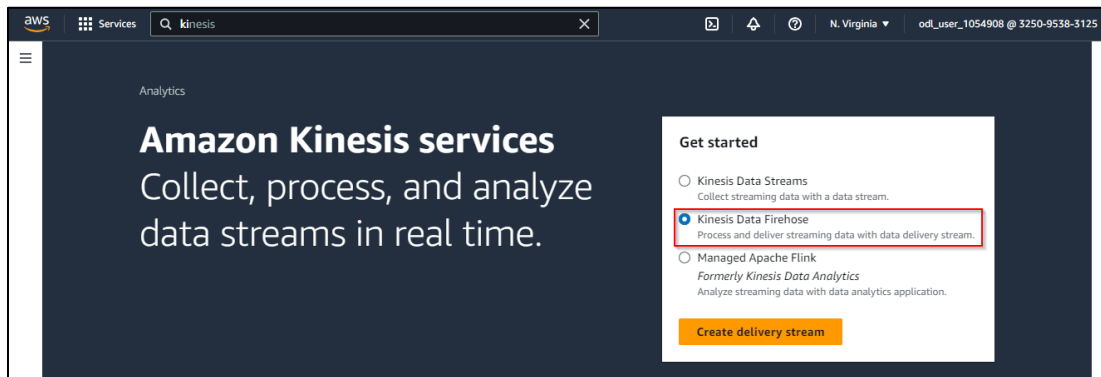
1. Create a Kinesis Data Firehose
2. Create an S3 bucket

Step 1: Create a Kinesis Data Firehose

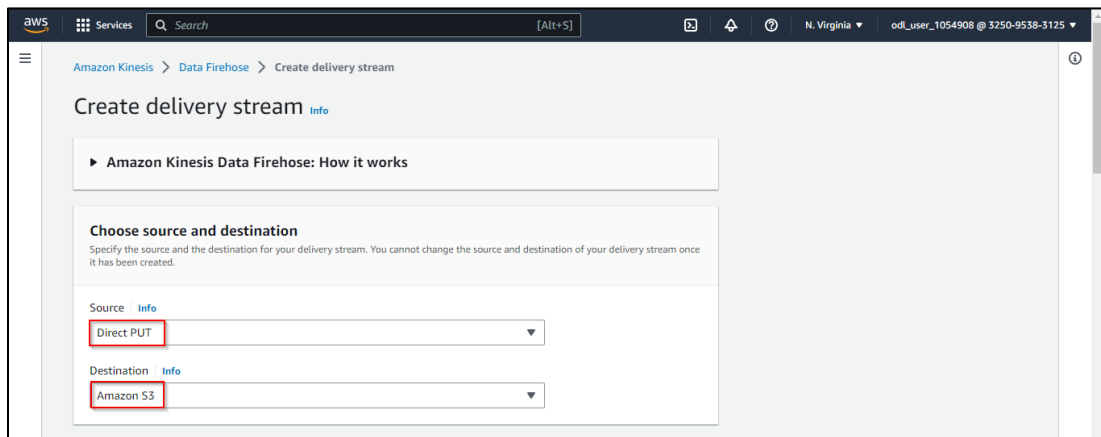
1.1 Navigate to the Amazon portal and search for and select **Kinesis**



1.2 Select the **Kinesis Data Firehose** option and click **Create delivery stream**



1.3 Choose **Direct PUT** as the source and **Amazon S3** as the destination



1.4 Click on the **Create** button to set up an S3 bucket

Destination settings [Info](#)

Specify the destination settings for your delivery stream.


S3 bucket

[Browse](#)

[Create !\[\]\(352af49c7e82e66a229b1eb98760843c_img.jpg\)](#)

Format: s3://bucket

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 delivery stream. You cannot enable dynamic partitioning for an existing delivery stream. Enabling dynamic partitioning incurs additional costs per GiB of partitioned data. For more information, see [Kinesis Data Firehose pricing](#). 

☒ Not enabled

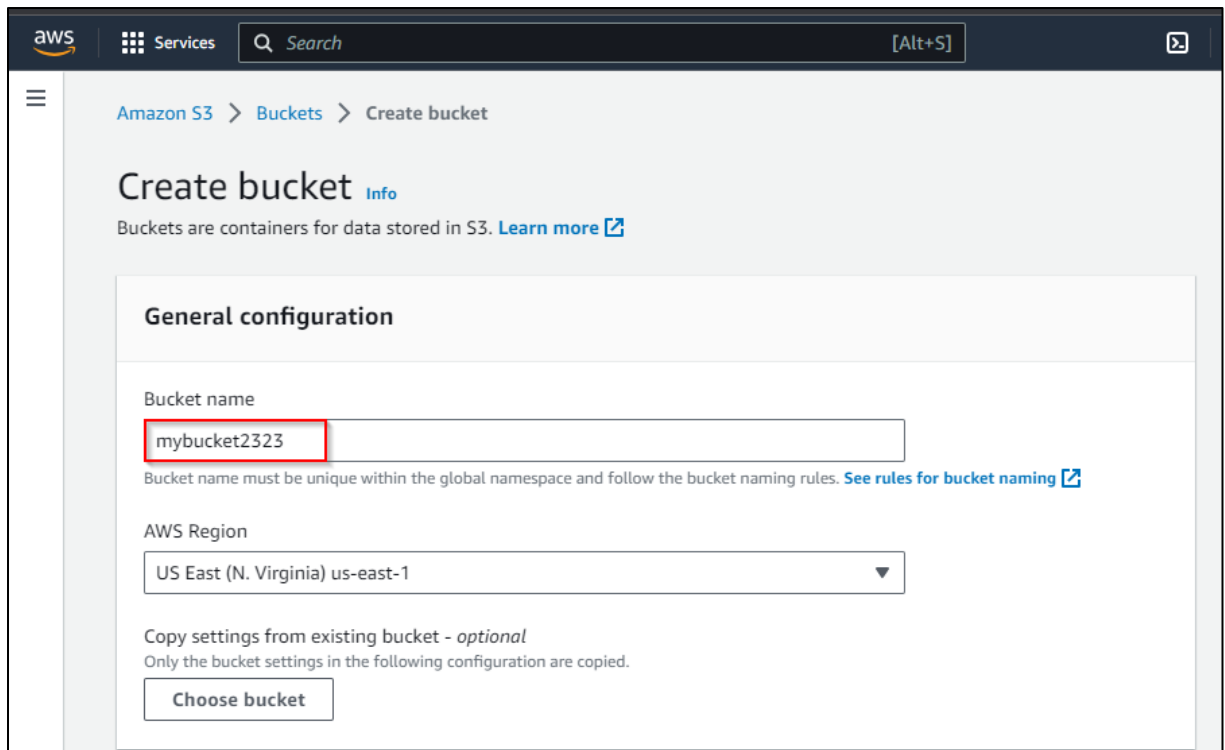
☐ Enabled

S3 bucket prefix - optional

By default, Kinesis Data Firehose appends the prefix "YYYY/MM/dd/HH" (in UTC) to the data it delivers to Amazon S3. You can override this default by specifying a custom prefix that includes expressions that are evaluated at runtime.

Step 2: Create an S3 bucket

2.1 Name the bucket **mybucket2323**



The screenshot shows the AWS Management Console interface for creating a new S3 bucket. The breadcrumb navigation at the top indicates the path: Amazon S3 > Buckets > Create bucket. The main heading is 'Create bucket' with an 'Info' link. Below this, a descriptive sentence states: 'Buckets are containers for data stored in S3. [Learn more](#)'. The 'General configuration' section contains the following fields:

- Bucket name:** A text input field containing 'mybucket2323', which is highlighted with a red rectangular border.
- AWS Region:** A dropdown menu currently showing 'US East (N. Virginia) us-east-1'.

Below the region dropdown, there is a section titled 'Copy settings from existing bucket - optional' with the subtext 'Only the bucket settings in the following configuration are copied.' and a button labeled 'Choose bucket'.

2.2 Scroll down and enable the **Unblock all public access** dialog box

Block Public Access settings for this bucket

Public access is granted to buckets and objects through access control lists (ACLs), bucket policies, access point policies, or all. In order to ensure that public access to this bucket and its objects is blocked, turn on Block all public access. These settings apply only to this bucket and its access points. AWS recommends that you turn on Block all public access, but before applying any of these settings, ensure that your applications will work correctly without public access. If you require some level of public access to this bucket or objects within, you can customize the individual settings below to suit your specific storage use cases. [Learn more](#)

☐ **Block all public access**
 Turning this setting on is the same as turning on all four settings below. Each of the following settings are independent of one another.


☐ **Block public access to buckets and objects granted through *new* access control lists (ACLs)**
 S3 will block public access permissions applied to newly added buckets or objects, and prevent the creation of new public access ACLs for existing buckets and objects. This setting doesn't change any existing permissions that allow public access to S3 resources using ACLs.

☐ **Block public access to buckets and objects granted through *any* access control lists (ACLs)**
 S3 will ignore all ACLs that grant public access to buckets and objects.

☐ **Block public access to buckets and objects granted through *new* public bucket or access point policies**
 S3 will block new bucket and access point policies that grant public access to buckets and objects. This setting doesn't change any existing policies that allow public access to S3 resources.

☐ **Block public and cross-account access to buckets and objects through *any* public bucket or access point policies**
 S3 will ignore public and cross-account access for buckets or access points with policies that grant public access to buckets and objects.

2.3 Check the **acknowledgment** box to confirm



Turning off block all public access might result in this bucket and the objects within becoming public
 AWS recommends that you turn on block all public access, unless public access is required for specific and verified use cases such as static website hosting.

☒ I acknowledge that the current settings might result in this bucket and the objects within becoming public.

2.4 Once the bucket creation is complete, click **Create bucket**

Encryption type [Info](#)

☒ Server-side encryption with Amazon S3 managed keys (SSE-S3)

☐ Server-side encryption with AWS Key Management Service keys (SSE-KMS)

☐ Dual-layer server-side encryption with AWS Key Management Service keys (DSSE-KMS)

 Secure your objects with two separate layers of encryption. For details on pricing, see [DSSE-KMS pricing](#) on the [Storage](#) tab of the [Amazon S3 pricing page](#).

Bucket Key

 Using an S3 Bucket Key for SSE-KMS reduces encryption costs by lowering calls to AWS KMS. S3 Bucket Keys aren't supported for DSSE-KMS. [Learn more](#)

☐ Disable

☒ Enable

► **Advanced settings**

ⓘ After creating the bucket, you can upload files and folders to the bucket, and configure additional bucket settings.

[Cancel](#)
[Create bucket](#)

Services

[Alt+S]

Global

odl_user_1054908 @ 3250-9538-3125

Successfully created bucket "mybucket20000"

 To upload files and folders, or to configure additional bucket settings choose [View details](#).

[View details](#)

Amazon S3 > Buckets

Account snapshot

 Storage lens provides visibility into storage usage and activity trends. [Learn more](#)

[View Storage Lens dashboard](#)

Buckets (1) [Info](#)

 Buckets are containers for data stored in S3. [Learn more](#)

Copy ARN
 Empty
 Delete
 [Create bucket](#)

Name

▲

AWS Region

▼

Access

▼

Creation date

▼

mybucket20000

US East (N. Virginia) us-east-1

Objects can be public

August 31, 2023, 15:44:21 (UTC+05:30)

The bucket has been successfully created.

2.5 Now, navigate back to the **Kinesis** dashboard and click on the **Browse** button

Destination settings [Info](#)

Specify the destination settings for your delivery stream.

S3 bucket

Browse

Create [↗](#)

Format: s3://bucket

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 delivery stream. You cannot enable dynamic partitioning for an existing delivery stream. Enabling dynamic partitioning incurs additional costs per GiB of partitioned data. For more information, see [Kinesis Data Firehose pricing](#). [↗](#)

☒ Not enabled
 ☐ Enabled

S3 bucket prefix - optional

By default, Kinesis Data Firehose appends the prefix "YYYY/MM/dd/HH" (in UTC) to the data it delivers to Amazon S3. You can override this default by specifying a custom prefix that includes expressions that are evaluated at runtime.

You can repeat the same keys in your S3 bucket prefix. Maximum S3 bucket prefix characters: 1024.

2.6 Choose the bucket name and click **Choose**

Choose a bucket in Amazon S3

Buckets (1/1)

<

1

>

Name ↗	Region
<input checked="" type="radio"/> mybucket20000	US East (N. Virginia)

Cancel

Choose

Destination settings [Info](#)

Specify the destination settings for your delivery stream.

S3 bucket

Format: s3://bucket

[Browse](#) [Create](#)

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 delivery stream. You cannot enable dynamic partitioning for an existing delivery stream. Enabling dynamic partitioning incurs additional costs per GiB of partitioned data. For more information, see [Kinesis Data Firehose pricing](#).

☒ Not enabled

☐ Enabled

The S3 bucket has now been successfully added.

2.7 Click on the **Create delivery stream** button

► **Buffer hints, compression and encryption**

The fields below are pre-populated with the recommended default values for S3. Pricing may vary depending on storage and request costs.

► **Advanced settings**

Server-side encryption not enabled; error logging enabled; IAM role KinesisFirehoseServiceRole-PUT-S3-EK3bz-us-east-1-1693475407951; no tags.

[Cancel](#) [Create delivery stream](#)

PUT-S3-EK3bz was successfully created.

Amazon Kinesis > Data Firehose > PUT-S3-EK3bz

PUT-S3-EK3bz [Info](#) [Delete delivery stream](#)

Delivery stream details

Status	Destination	Data transformation	Creation time
Active	Amazon S3	Not enabled	August 31, 2023 at 16:12 GMT+5:30
Source	ARN	Dynamic partitioning	Error logs status
Direct PUT	arn:aws:firehose:us-east-1:325095383125:deliverystream/PUT-S3-EK3bz	Not enabled	0 Destination error logs

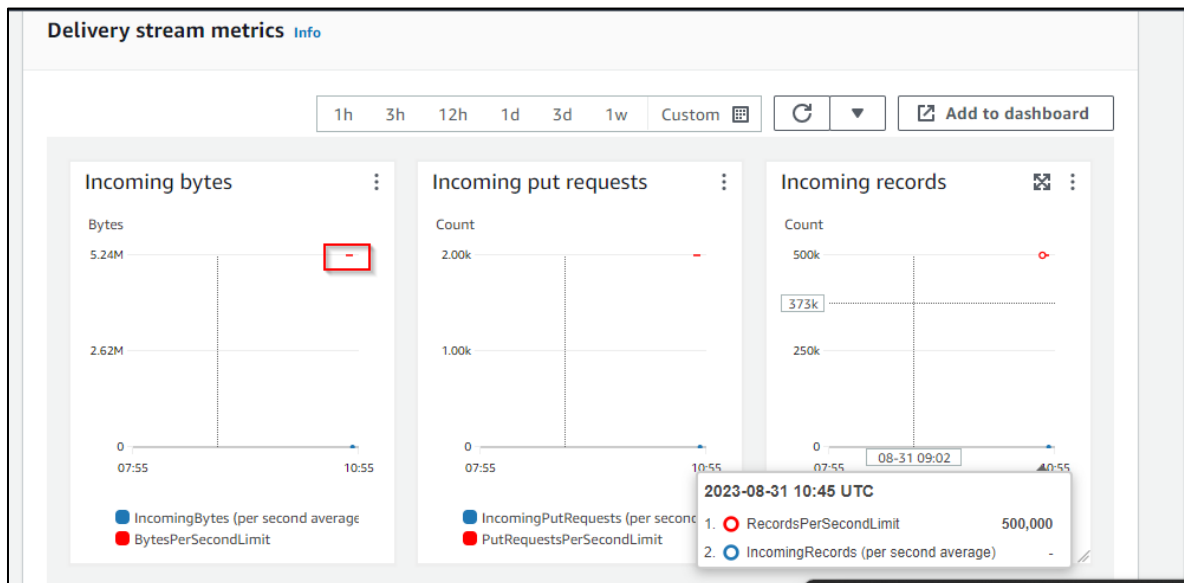
The Amazon Data Firehose has been successfully created.

2.8 Under **Test with demo data**, click **Start sending demo data**

The screenshot shows the Amazon Kinesis console interface. On the left is a navigation menu with 'Amazon Kinesis', 'Dashboard', 'Data streams', 'Data Firehose', 'Managed Apache Flink', and 'Resources'. The main content area is titled 'Test with demo data' and includes a description: 'Ingest simulated data to test the configuration of your delivery stream. Standard Amazon Kinesis Data Firehose charges apply.' Below this is a JSON script for demo data:

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

 There are two steps: 'Step 1: Start sending demo data to your delivery stream. If you already have data streaming to this destination, demo data is sent along with your source records.' and 'Step 2: Stop sending demo data to your delivery stream after you've concluded your test to stop incurring usage charges.' The 'Start sending demo data' button in Step 1 is highlighted with a red rectangular box.



The **Delivery stream metrics** information will appear here.

By following these steps, you have successfully established a data delivery stream from a source to an Amazon S3 bucket.