

[Amazon OpenSearch Service](#) > [Ingestion: Pipelines](#) > Create pipeline

Step 1

Pipeline settings and configuration

Step 2

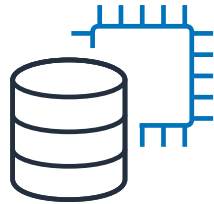
Review and create

Pipeline settings and configuration

▼ Getting started

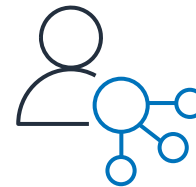
Create integration with Zero-ETL with DynamoDB

This pipeline exports a DynamoDB table to an OpenSearch cluster and processes DynamoDB streams to keep the table and cluster in sync for create, update, and delete events. [Learn more](#)



1. Select a DynamoDB table

Create or select a DynamoDB table to replicate to an index. Point-in-time recovery (PITR) and DynamoDB



2. Setup permissions

Create an IAM role with a policy that allows the pipeline to access your data. [Learn more](#)



3. Select OpenSearch domains or collections, S3 as destination

Select [OpenSearch destination](#). Optionally you can



[Alt+S]



easy viewing of
resources for
configuring pipeline.
[Learn more](#)

json, ndson or
parquet. [Learn
more](#)



Pipeline settings [Info](#)

Amazon OpenSearch Ingestion facilitates serverless pipelines to transform, enrich, and normalize the data prior to indexing and storing it in Amazon OpenSearch Service.

Pipeline name

Enter a unique pipeline name (across current AWS account in the current region).

my-pipeline-01

The name must start with a lowercase letter and must be between 3 and 28 characters. Valid characters are a-z (lowercase only), 0-9, and - (hyphen).



You can not edit the pipeline name after a pipeline is created.

Persistent Buffer - optional - new [Info](#)

A persistent buffer stores your data in a disk-based buffer across multiple AZs to add durability to your data.

☐ Enable persistent buffer

We recommend using the persistent buffer for HTTP and OTel sources.

Pipeline capacity [Info](#)

A single Ingestion OpenSearch Compute Unit (OCU) represents billable compute and memory units. You



[Alt+S]



2

ingestion-OCU

4

ingestion-OCU

Reset to default

The maximum capacity limit for a pipeline is 96. The maximum capacity limit for stateful pipelines (for example, those that use the service_map or aggregate processor) is 48.



Pipeline configuration [Info](#)

Upload file

Configuration blueprints

Browse resources

YAML

JSON

```
12         # Remove the export block if only stream is needed
13         export:
14             # REQUIRED for export: Specify the name of an existing S3 bucket
15             s3_bucket: "arn:aws:s3:::sagemaker-kb-528217753045"
16             # Specify the region of the S3 bucket
17             s3_region: "us-west-2"
18             # Optionally set the name of a prefix that DynamoDB export data is written to
19             s3_prefix: "ddb-to-opensearch-export/"
20         aws:
21             # REQUIRED: Provide the role to assume that has the necessary permissions
22             sts_role_arn: "arn:aws:iam::528217753045:role/cdk-hnb659fds-deploy-role-us-west-2"
23             # Provide the region to use for aws credentials
24             region: "us-west-2"
25         sink:
26         - opensearch:
27             # REQUIRED: Provide an AWS OpenSearch endpoint
28             hosts: [ "https://v2o7b0de7ar8g0a43jnd.us-west-2.aoss.amazonaws.com" ]
29             index: "rag-index"
30             index_type: custom
31             document_id: "${getMetadata(\"primary_key\")}"
32             action: "${getMetadata(\"opensearch_action\")}"
33             document_version: "${getMetadata(\"document_version\")}"
34             document_version_type: "external"
35         aws:
36             # REQUIRED: Provide a Role ARN with access to the domain. This role must have the necessary permissions to write to the OpenSearch domain.
```



[Alt+S]



```
43 # Specify a name here to create or update network policy for
44 # network_policy_name: "network-policy-name"
45 # Optional: Enable the S3 DLQ to capture any failed requests in
46 dlq:
47
```



Pipeline configuration validation successful.

Re-run validation

Network [Info](#)

Choose internet or VPC access. To enable VPC access, we use private IP addresses from your VPC, which provides an inherent layer of security. You control network access within your VPC using security groups. Internet endpoints are publicly accessible.

Network



VPC access -
recommended



Public access

Tags - optional [Info](#)

You can add tags to describe your pipeline. A tag consists of a case-sensitive key-value pair. For example, you can define a tag with a key-value pair of Environment Name-Development. [Learn more](#)

No tags associated with the resource.



[Alt+S]



Log publishing options - (optional) [Info](#)

☒ Publish to CloudWatch logs (recommended)

OpenSearch Ingestion logs provide access to information about errors and warnings raised during the operation, ingestion and transformation of Data Prepper pipelines and can be useful for troubleshooting.

Log group

☒ Create new group

☐ Select existing group

New log group name

/aws/vendedlogs/OpenSearchIngestion/<pipeline-name>/audit

We recommend naming your log group as a path such as /aws/vendedlogs/OpenSearchService/pipelines/mypipeline. This will make it easier to apply a CloudWatch access policy where you can grant permissions to all log groups under a specific path such as /aws/vendedlogs/OpenSearchService/pipelines.

Cancel

Next