

# Geo-replication

- Explain how geo-replication works
- Describe some use cases for geo-replication
- Configure geo-replication by using the UI or CLI

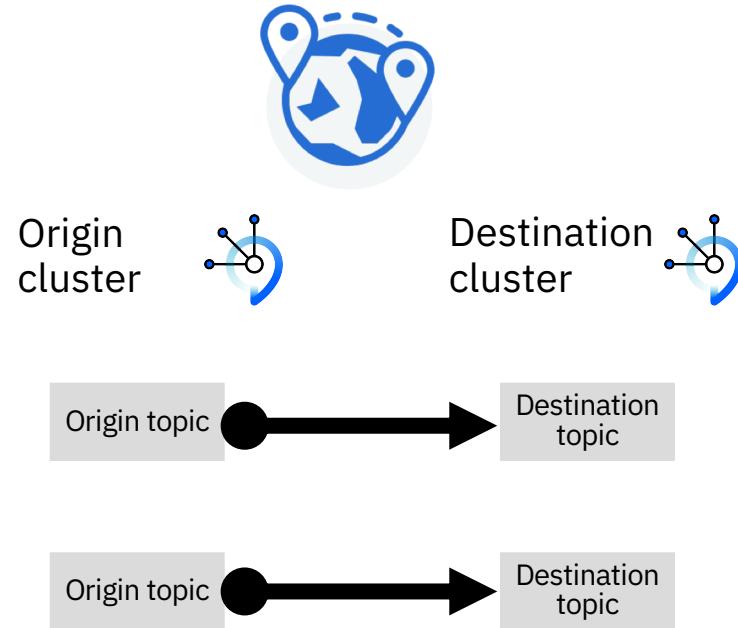
# What does it do?

Synchronizes data between clusters that are located in different geographical locations

Creates copies of your selected topics to help with disaster recovery

Copies the messages of a topic, the topic configuration, the topic's metadata, its partitions, and preserves the timestamps from the origin topic

After geo-replication starts, topics are kept in sync



# Use cases

What topics you choose to replicate and how depend on the topic data, whether it is critical to your operations, and how you want to use it

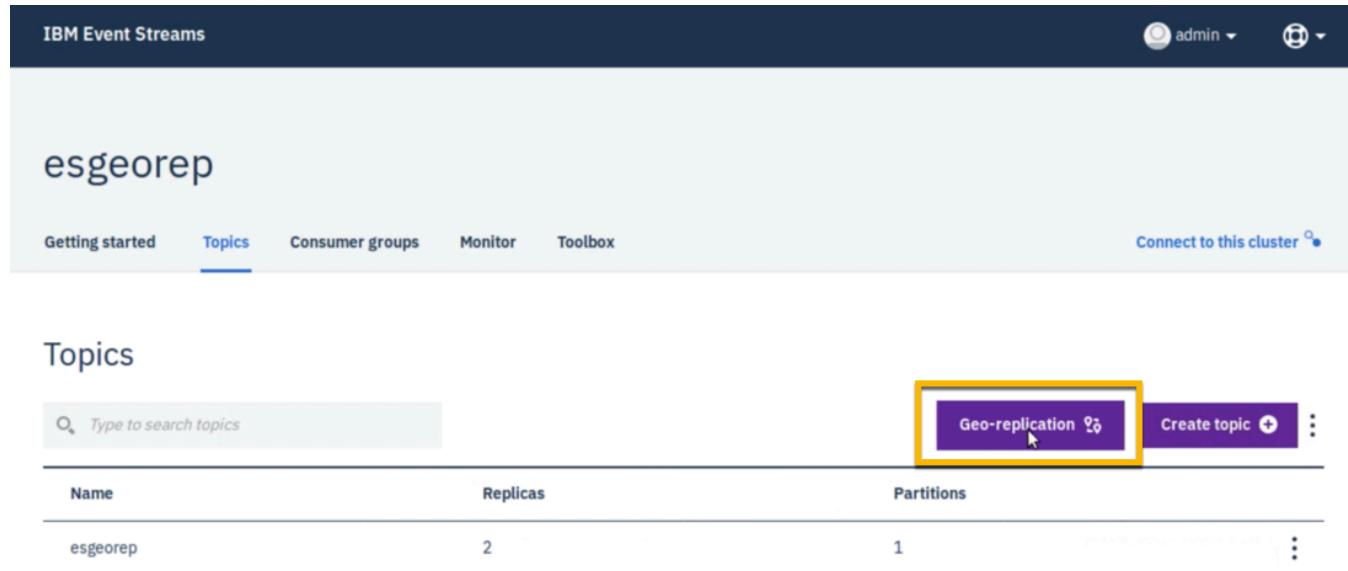
For example, you might have transaction data for your customers in topics. Such information is critical to your operations to run reliably, so you want to ensure they have back-up copies to switch to when needed.

Another example would be storing of website analytics information, such as where users clicked and how many times they did so.



# Configuring geo-replication

You can set up geo-replication by using the IBM Event Streams console or CLI (cloudctl)



The screenshot shows the IBM Event Streams console interface. At the top, there's a dark header bar with the text "IBM Event Streams" on the left, and "admin" and a user icon on the right. Below the header, the title "esgeorep" is displayed. The main navigation menu includes "Getting started", "Topics" (which is underlined, indicating it's the active tab), "Consumer groups", "Monitor", and "Toolbox". To the right of the menu, there's a "Connect to this cluster" button. The main content area is titled "Topics" and features a search bar with the placeholder "Type to search topics". Below the search bar is a table with three columns: "Name", "Replicas", and "Partitions". A single row is shown for the topic "esgeorep", which has 2 replicas and 1 partition. In the top right corner of the "esgeorep" row, there's a purple button labeled "Geo-replication" with a gear icon, which is highlighted with a yellow box. To the right of the button are "Create topic" and a three-dot menu icon.

Name	Replicas	Partitions
esgeorep	2	1

# Planning considerations

1. Prepare the destination cluster; set the number of geo-replication workers
2. Identify the topics to be copied
3. Decide whether you want to include message history in the geo-replication, or only copy messages from the time of setting up geo-replication
4. Decide whether the replicated topics on the destination cluster should have the same name as their corresponding topics on the origin cluster, or if a prefix should be added to the topic name



# Preparing destination clusters

- Configure the number of geo-replication workers on the destination cluster
- The number of workers depend on the number of topics you want to replicate, and the throughput of the produced messages
- For high availability reasons, ensure that you have at least two workers on your destination cluster
- You can configure the number of workers at the time of installing IBM Event Streams, or you can modify an existing installation

For a new installation, using the Event Streams console:

Go to the **Geo-replication** section and specify the number of workers in the **Geo-replicator workers** field

Using the CLI:

During installation, add the `--set replicator.replicas=<number-of-workers>` parameter to your `helm install` command

# Configuring the number of geo-replication workers

Go to the **Geo-replication** settings section and specify the number of workers in the **Geo-replicator workers** field

Click **Install**

The screenshot shows the IBM Cloud Private configuration interface. At the top, there is a navigation bar with links for 'Create resource', 'Catalog', 'Docs', and 'Support'. The main content area has two sections:

- Message indexing settings**: Configuration for message indexing used to enhance browsing the messages on topics. It includes a checked checkbox for 'Enable message indexing \*'.
- Geo-replication settings**: Configuration for geo-replicating topics between clusters. It includes a text input field for 'Geo-replicator workers \*' containing the value '2'.

At the bottom right of the configuration panel, there are 'Cancel' and 'Install' buttons, with the 'Install' button being highlighted by a yellow box.

# Modifying an existing installation

## Using the UI:

1. In the IBM Cloud Private console, from the navigation menu, click **Workloads > Helm Releases** and locate your Event Streams release
2. Select **More options > Upgrade**
3. Update the **geo-replication settings** and click **Upgrade**

## Using the CLI:

1. Log in to IBM Cloud Private (`cloudctl login`)
2. Use the `helm upgrade` command to modify the number of workers

For example, to set the number of geo-replication workers to 4, use the following command:

```
helm upgrade --reuse-values --set
replicator.replicas=4 destination ibm-
eventstreams-prod-1.2.0.tgz --tls
```

# Defining destination clusters

## Using the UI:

1. Log in to the Event Streams console for the destination cluster
2. Go to **Topics > Geo-replication > Origin locations** and click **Generate connection information for this cluster**
3. Copy the information to the clipboard
4. Log in to the Event Streams console for the origin cluster
5. In **Topics > Geo-replication**, click **Add destination cluster**, and paste the information

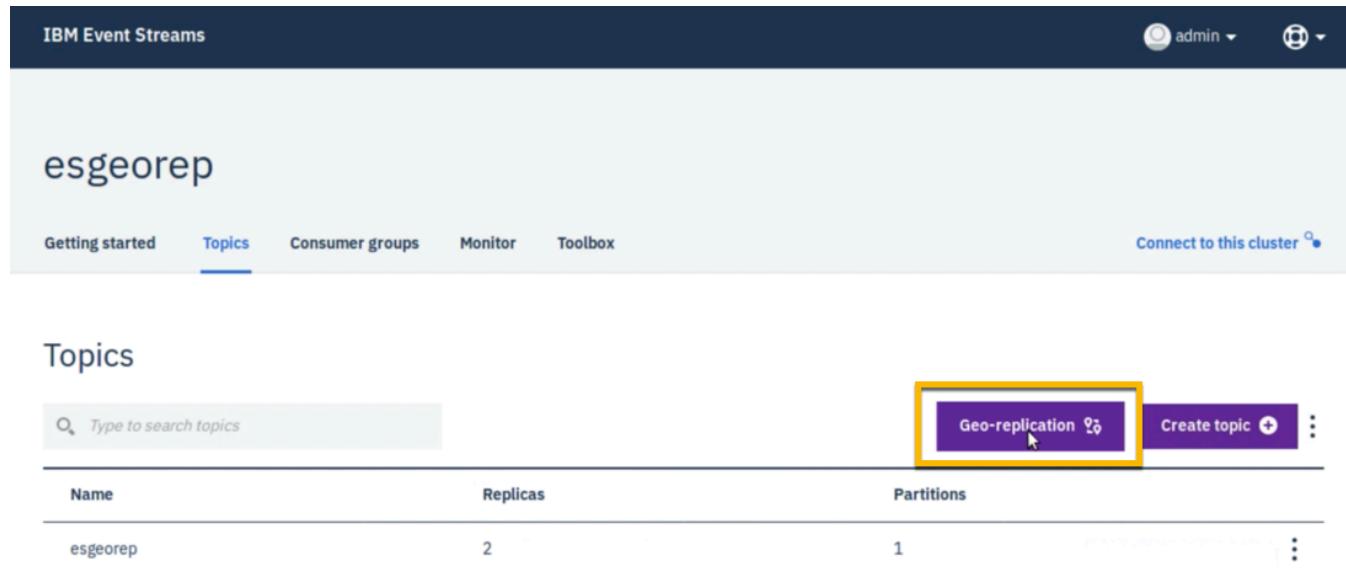
## Using the CLI:

1. On the destination cluster, log in with `cloudctl login`
2. Initialize the Event Streams CLI with `cloudctl es init`
3. Create an API key with `cloudctl es geo-cluster-apikey`
4. Log in to, and initialize the origin cluster (`cloudctl login`, `cloudctl es init`)
5. On the origin cluster, run the following command

```
cloudctl es geo-cluster-add --api-address <api-url-from-step-3> --api-key <api-key-from-step-3>
```

# Accessing Geo-replication settings

In the Event Streams console, go to **Topics** > **Geo-replication**



The screenshot shows the IBM Event Streams interface. At the top, there's a dark header bar with the text "IBM Event Streams" on the left, and "admin" and a gear icon on the right. Below the header, the title "esgeorep" is displayed. A navigation bar contains links for "Getting started", "Topics" (which is underlined), "Consumer groups", "Monitor", and "Toolbox". To the right of the navigation bar is a blue button labeled "Connect to this cluster". The main content area is titled "Topics" and features a search bar with the placeholder "Type to search topics". Below the search bar is a table with three columns: "Name", "Replicas", and "Partitions". A single row is shown for the topic "esgeorep", which has 2 replicas and 1 partition. To the right of the table is a purple button labeled "Geo-replication" with a gear icon, which is highlighted with a yellow rectangular box. There are also other buttons for "Create topic" and a vertical ellipsis menu.

Name	Replicas	Partitions
esgeorep	2	1

# Copying connection information

On the destination cluster...

Geo-replication

Geo-replication is a way to duplicate your topics to different instances of Event Streams. These are usually located in different locations to minimize the risk of data loss in the event of a cluster failure.

Origin locations ⓘ

Want to replicate topics to this cluster?

To add this cluster as a destination location on an origin cluster, you will need to generate connection information.

Generate connection information for this cluster

Destination locations ⓘ Add destination cluster

1 cluster is geo-replicating topics to this cluster  
1 topic is using this cluster for geo-replication

2

3

Destination locations ⓘ Add destination cluster +

## Geo-replication

Geo-replication is a way to duplicate your topics to different instances of Event Streams. These are usually located in different locations to minimize the risk of data loss in the event of a cluster failure.

## Geo-replication

Geo-replication is a way to duplicate your topics to different instances of Event Streams. These are usually located in different locations to minimize the risk of data loss in the event of a cluster failure.

Origin locations ⓘ

1 cluster is geo-replicating topics to this cluster  
1 topic is using this cluster for geo-replication

westgeo 1 topic

Copy connection information ⓘ

Destination locations ⓘ Add destination cluster +

# Pasting connection information

On the origin cluster...

The screenshot shows the 'Geo-replication' section of the IBM Event Streams interface. It includes a 'Want to replicate topics to this cluster?' message and a 'Generate connection information for this cluster' button. A yellow box highlights the 'Add destination cluster' button at the bottom.

1

The screenshot shows the 'Cluster connection' interface. It has tabs for 'Connect a client', 'Sample code', and 'Geo-replication' (which is selected). Below the tabs, there's a 'Back' button and a main text area: 'To replicate topics from this cluster to another, you need to get the connection information from the cluster you want to replicate to.' A text input field labeled 'Where do I find this?' contains a JSON configuration snippet. A blue circle labeled '2' points to this field. To the right, connection details are listed: API address (https://10.0.0.1:30955) and API key (wXQ...), both with green checkmarks. A 'Validate' button is also present.

2

# Selecting the topics to replicate

On the origin cluster:

Choose the topics you want to replicate by selecting the checkbox next to each, and click **Geo-replicate to destination**

Select if you want to add a prefix to the destination topic names, and then click **Create**

The screenshot shows the 'Topics' section of the IBM Event Streams interface. A search bar at the top says 'Type to search topics'. Below it is a table with columns: Name, Replicas, and Partitions. A single row for the topic 'esgeorep' is selected, indicated by a yellow box around the checkbox in the first column. At the bottom of the table, a blue circle labeled '1' is positioned above a button labeled 'Geo-replicate to destination' which is also highlighted with a yellow box.

<input type="checkbox"/>	Name	Replicas	Partitions
<input checked="" type="checkbox"/>	esgeorep	2	1

1 topic selected

Geo-replicate to destination

The screenshot shows the 'Geo-replicated topics' section of the IBM Event Streams interface. It lists one topic: 'esgeorep\_esgeorep' with the status 'Awaiting creation'. A blue circle labeled '2' is positioned above a modal dialog. The dialog header says 'You are about to create a replicator for 1 topic.' It contains two checkboxes: 'Add prefix to destination topic names' (which is checked) and 'Include message history'. At the bottom right of the dialog is a 'Create' button, which is highlighted with a yellow box. There are also 'Cancel' and 'Close' buttons.

Back to geo-replication

TOPICS	WORKERS
0	3

Geo-replicated topics

Name	Replication health
esgeorep_esgeorep	Awaiting creation

You are about to create a replicator for 1 topic.

Add prefix to destination topic names

Include message history

Create

# Topic replication status

The screenshot shows two panels of the IBM Event Streams web interface. The left panel displays a list of topics under the heading 'Topics'. It includes a search bar and a table with columns: Name, Replicas, and Partitions. One row shows 'esgeorep' with 2 replicas and 1 partition. The right panel shows the 'Geo-replicated topics' section for the 'eastgeo' geo-replication group. It lists 'esgeorep\_esgeorep' with a status of 'ASSIGNING'. A yellow callout bubble points to this status with the text: 'Status changes from Pending, to Assigning, to Running'.

IBM Event Streams

esgeorep

Topics

Type to search topics

Name	Replicas	Partitions
esgeorep	2	1 99 (1)

Back to geo-replication

eastgeo

TOPICS 1 WORKERS 3

Geo-replicated topics

Name	Replication health
esgeorep_esgeorep	ASSIGNING

Status changes from Pending, to Assigning, to Running

# Topic replication status (cont.)

For each topic that has geo-replication set up, a visual indicator is shown in the topic's row

---

→ topic	3	1	From origin	⋮
---------	---	---	-------------	---

---

# Selecting the topics to replicate by using the CLI

1. On the destination cluster, log in with  
`cloudctl login`
2. Initialize the Event Streams CLI with  
`cloudctl es init`
3. List available destination clusters, and copy the cluster ID that you want to replicate  
`cloudctl es geo-clusters`
4. List topics, and copy the names of the ones that you want to replicate  
`cloudctl es topics`

Run the following command with the destination cluster ID and topic names that you retrieved in the previous steps:

```
cloudctl es geo-replicator-create --  
destination <cluster-ID> --topics  
<comma-separated-list-of-topic-names>
```

# Switching clusters

When one of your origin IBM Event Streams clusters experiences problems and goes down, you are notified on the destination cluster UI that the origin cluster is offline

You can switch your applications over to use the geo-replicated topics on the destination cluster

After the connection is configured, your client application can continue to operate using the geo-replicated topics on the destination cluster

You can decide whether to continue processing messages on the destination cluster from the point they reached on the topic on the origin cluster, or if you want to start processing messages from the beginning of the topic

1. Log in to your destination IBM Event Streams cluster
2. Click **Connect to this cluster**
3. Go to the **Connect a client** tab, and use the information on that page to change your client application settings to use the geo-replicated topic on the destination cluster
4. Use `cloudctl es group-reset` command to specify offset, or process messages from the beginning

