

# Kafka Connect

1. Terraform apply (updated vm size to 2xStandard\_D4\_v3):

```
MINGW64:/c/Users/yura1/Desktop/New folder/hw/test/kafka/task2/m11_kafkaconnect_json_azure-master/terraform
azurerm_kubernetes_cluster.bdcc: Still creating... [03m00s elapsed]
azurerm_kubernetes_cluster.bdcc: Still creating... [03m10s elapsed]
azurerm_kubernetes_cluster.bdcc: Creation complete after 3m13s [id=/subscriptions/04f2456f-rope-nguk/providers/Microsoft.ContainerService/managedClusters/aks-dev-northeurope-nguk]
azurerm_role_assignment.aks_acr_pull: Creating...
kubernetes_namespace.confluent: Creating...
kubernetes_namespace.confluent: Creation complete after 0s [id=confluent]
azurerm_role_assignment.aks_acr_pull: Still creating... [00m10s elapsed]
azurerm_role_assignment.aks_acr_pull: Still creating... [00m20s elapsed]
azurerm_role_assignment.aks_acr_pull: Creation complete after 24s [id=/subscriptions/04f2456f-heurope-nguk/providers/Microsoft.ContainerRegistry/registries/acrdevnortheeuropenguk/providerId/3875-99e4-efce5c3233b6]

Warning: Deprecated Resource

  with kubernetes_namespace.confluent,
  on main.tf line 114, in resource "kubernetes_namespace" "confluent":
114: resource "kubernetes_namespace" "confluent" {

Deprecated; use kubernetes_namespace_v1.

Releasing state lock. This may take a few moments...

Apply complete! Resources: 9 added, 0 changed, 0 destroyed.

Outputs:

acr_login_server = "acrdevnortheeuropenguk.azurecr.io"
aks_api_server_url = <sensitive>
aks_kubeconfig = <sensitive>
aks_name = "aks-dev-northeurope-nguk"
client_certificate = <sensitive>
resource_group_name = "rg-dev-northeurope-nguk"
storage_account_name = "stdevnortheeuropenguk"
storage_primary_access_key = <sensitive>
```

2. Confluent Platform pods (decreased number of replicas for zookeeper to 1, also reduced amount of allocated resources for some modules):

```
yura1@pc MINGW64 ~/Desktop/New folder/hw/test/kafka/task2/m11_kafkaconnect_json_azure-master (main)
$ kubectl get pods -o wide
NAME                               READY   STATUS    RESTARTS   AGE     IP          NODE
ADINESS GATES
confluent-operator-8446f997d8-xmwpv   1/1    Running   0          31m    10.244.0.9  aks-default-40341999-vmss000001
one>
connect-0                            1/1    Running   0          4m4s   10.244.0.0  aks-default-40341999-vmss000001
one>
controlcenter-0                      1/1    Running   0          2m4s   10.244.1.114 aks-default-40341999-vmss000000
one>
elastic-0                            1/1    Running   3 (2m55s ago) 3m33s  10.244.1.70  aks-default-40341999-vmss000000
one>
kafka-0                             1/1    Running   1 (2m20s ago) 3m4s   10.244.1.243 aks-default-40341999-vmss000000
one>
kafka-1                             1/1    Running   0          3m4s   10.244.1.87  aks-default-40341999-vmss000000
one>
kafka-2                             1/1    Running   0          3m4s   10.244.1.15  aks-default-40341999-vmss000000
one>
ksqldb-0                            1/1    Running   0          2m4s   10.244.0.92  aks-default-40341999-vmss000001
one>
schemaregistry-0                     1/1    Running   0          2m4s   10.244.1.192 aks-default-40341999-vmss000000
one>
zookeeper-0                         1/1    Running   0          4m5s   10.244.0.178 aks-default-40341999-vmss000001
one>
```

### 3. Control center at localhost:

The screenshot shows a browser window with the URL `localhost:9021/clusters` in the address bar. The page has a dark blue header with the Confluent logo. Below the header, the word "Home" is displayed in large, bold, black letters. Two status boxes are present: a green box on the left showing "1 Healthy clusters" and a red box on the right showing "0 Unhealthy clusters". A search bar with a magnifying glass icon and the placeholder text "Search cluster name or id" is located below the status boxes. The main content area features a card for the cluster "controlcenter.cluster", which is labeled as "Running". The "Overview" section provides metrics: Brokers (2), Partitions (457), Topics (59), Production (48.73KB/s), and Consumption (39.33KB/s). The "Connected services" section shows 1 ksqlDB cluster and 1 Connect cluster.

# Home

1

Healthy clusters

0

Unhealthy clusters



Search cluster name or id

## controlcenter.cluster

Running

### Overview

Brokers	2
Partitions	457
Topics	59
Production	48.73KB/s
Consumption	39.33KB/s

### Connected services

ksqlDB clusters	1
Connect clusters	1

Cluster overview

## Overview

Brokers

Topics

Connect

ksqlDB

Consumers

Replicators

Cluster settings

Health+

### Brokers

2  
Total

39.43K

Production (bytes / second)

29.13K

Consumption (bytes / second)

### Topics

59  
Total

457

Partitions

0

Under replicated partitions

0

Out-of-sync replicas

### Connect

1  
Clusters

0  
Running

0  
Paused

0  
Degraded

0  
Failed

## 4. Kafka topic creation:

```
PS C:\WINDOWS\system32> Start-Process powershell -WindowStyle Hidden -ArgumentList 'kubectl port-forward connect-0 8083:8083 *-> $null'
PS C:\WINDOWS\system32> kubectl exec kafka-0 -c kafka -- bash -c "/usr/bin/kafka-topics --create --topic expedia --replication-factor 3 --partitions 3 --bootstrap-server kafka:9092"
Created topic expedia.
PS C:\WINDOWS\system32>
```

- [Cluster overview](#)
- [Brokers](#)
- [Topics](#)
- [Connect](#)
- [ksqlDB](#)
- [Consumers](#)
- [Replicators](#)
- [Cluster settings](#)
- [Health+](#)

# expedia

[Overview](#)   [Messages](#)   [Schema](#)   **Configuration**

## General settings

name	expedia
partitions	3
min.insync.replicas	2
cleanup.policy	delete
retention.ms	604800000
max.message.bytes	1048588
retention.bytes	-1

[Edit settings](#)

[Show full config](#)

## 5. Data loaded to Azure Container:

[Home](#) > [Resource Manager | Resource groups](#) > [rg-dev-northeurope-nguk](#) > [stdevnortheeuropenguk | Containers](#) >

 data
...
...

Search
Add Directory
Upload
Refresh
Delete
Copy
Paste

**Overview**

 Diagnose and solve problems

 Access Control (IAM)

> Settings

 data

Authentication method: Access key ([Switch to Microsoft Entra user account](#))

Search blobs by prefix (case-sensitive)

Showing all 1 items

	Name	Last modified
<input type="checkbox"/>	 topics	1/14/2026, 10:29:42 PM

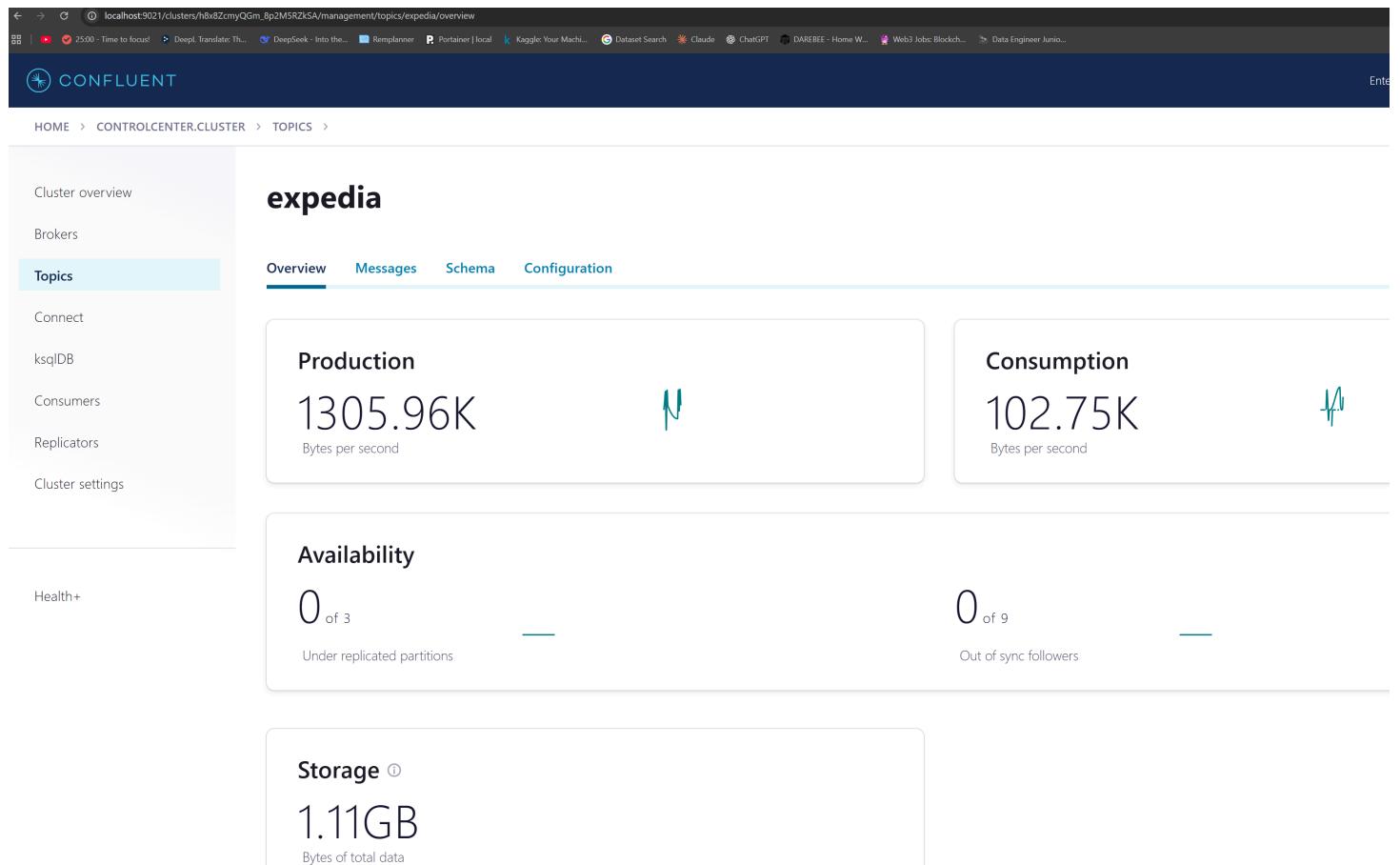
## 6. azure-source-cc.json prepared:

```
m11_kafkaconnect_json_azure-master > terraform > {} azure-source-cc.json > abc name
1  {
2    "name": "expedia2",
3    "config": {
4      "connector.class": "io.confluent.connect.azure.blob.storage.AzureBlobStorageSourceConnector",
5      "tasks.max": "2",
6
7      "azblob.account.name": "stdevnortheeuropenguk",
8      "azblob.account.key": "XXXXXXXXXXXXXXXXXXXXXXXXXXXXXX",
9      "azblob.container.name": "data",
10
11     "format.class": "io.confluent.connect.azure.blob.storage.format.avro.AvroFormat",
12
13     "bootstrap.servers": "kafka:9071",
14     "topics": "expedia",
15     "topics.dir": "topics",
16
17     "transforms": "maskDate",
18     "transforms.maskDate.type": "org.apache.kafka.connect.transforms.MaskField$Value",
19     "transforms.maskDate.fields": "date_time,srch_ci,srch_co",
20     "transforms.maskDate.replacement": "0000-00-00 00:00:00"
21   }
22 }
```

## 7. Port Forward and connector execution:

```
Windows PowerShell
PS C:\Users\yural> Start-Process powershell -WindowStyle Hidden -ArgumentList 'kubectl port-forward connect-0 8083:8083
*>> $null'
PS C:\Users\yural>
PS C:\Users\yural> Invoke-RestMethod
-Method POST -Uri http://localhost:8083/connectors -ContentType "application/json" -InFile .\azure-source-cc.json
name      config
----      -----
expedia2 @{connector.class=io.confluent.connect.azure.blob.storage.AzureBlobStorageSourceConnector; tasks.max=2; azb...
```

## 8. Result:



localhost:9021/clusters/h8x0ZcmyQGm\_8p2MSR2kSA/management/topics/expedia/message-viewer

25:00 - Time to focus DeepL Translate: Th... DeepSeek - Into the... Remplanner P. Portainer | local Kaggle: Your Machi... Dataset Search Claude ChatGPT DARBEE - Home W... Web3 Jobs Blockchain Data Engineer Junio...

Enterprise trial ends in 29 days

**CONFLUENT**

HOME > CONTROLCENTER.CLUSTER > TOPICS >

Cluster overview
**expedia**
Produce a new message

Brokers
Topics
Connect
ksqlDB
Consumers
Replicators
Cluster settings

Overview
Messages
Schema
Configuration

Production  
0 B/s

Consumption  
0 B/s

Total messages  
2,344,465

Retention time  
1 week

Filter by timestamp, offset, key or value
All partitions
Latest
Max 50 results

33 messages shown
 Auto-refresh on
[CSV](#)
[JSON](#)

Timestamp	Offset	Partition	Key	Value
1768428385003	564200	1	**	{"id":1114443,"date_time":"0000-00-00 00:00:00","site_name":2,"posa_container":null,"user_location_country":1,"user_location_r...
1768428385003	564201	1	**	{"id":1114549,"date_time":"0000-00-00 00:00:00","site_name":2,"posa_container":null,"user_location_country":66,"user_location...
1768428385003	564202	1	**	{"id":1114553,"date_time":"0000-00-00 00:00:00","site_name":11,"posa_container":null,"user_location_country":205,"user_locatio...
1768428385003	564203	1	**	{"id":1114574,"date_time":"0000-00-00 00:00:00","site_name":2,"posa_container":null,"user_location_country":66,"user_location...
1768428385003	564204	1	**	{"id":1114581,"date_time":"0000-00-00 00:00:00","site_name":2,"posa_container":null,"user_location_country":66,"user_location...

Date fields masked by MaskField transformer:



# Message details

Timestamp	Offset	Partition
1/14/2026, 11:06:25 PM (1768428385003)	564200	1

Key Value Headers

```
1  ▾ {  
2      "id": 1114443,  
3      "date_time": "0000-00-00 00:00:00",  
4      "site_name": 2,  
5      "posa_container": null,  
6      "user_location_country": 1,  
7      "user_location_region": 395,  
8      "user_location_city": 29254,  
9      "orig_destination_distance": 255.9227,  
10     "user_id": 107773,  
11     "is_mobile": 0,  
12     "is_package": 0,  
13     "channel": 10,  
14     "srch_ci": "0000-00-00 00:00:00",  
15     "srch_co": "0000-00-00 00:00:00",  
16     "srch_adults_cnt": 1,  
17     "srch_children_cnt": 0,  
18     "srch_rm_cnt": 1,  
19     "srch_destination_id": 38412
```

```
19     "srch_destination_id": 20410,  
20     "srch_destination_type_id": 6,  
21     "hotel_id": 1348619730946  
22 }
```