

Data Engineering in the Cloud

Work with data streams using Azure Stream Analytics

Xuemao Zhang
East Stroudsburg University

January 18, 2025

Outline

- With the aid of stream analytics, we utilize Event Hub to handle real-time data.
- Lab

Step 1. Create a datalake (storage account with hierarchical namespace enabled) and create a container

- 1 Create an Azure storage Gen 2 account if you do not have one. Add at least one container as well.

The screenshot shows the Microsoft Azure portal. At the top is a blue header bar with the Microsoft Azure logo, an 'Upgrade' button, a search bar with the placeholder text 'Search resources, services, and docs (G+)', and a 'Copilot' button. Below the header, the 'Azure services' section displays a row of icons for various services: 'Create a resource', 'Event Hubs', 'Stream Analytics jobs', 'Cost analysis', 'Azure Cosmos DB', 'Azure Synapse Analytics', 'Subscriptions', 'Data factories', and 'Storage accounts'. The 'Resources' section is below, with tabs for 'Recent' and 'Favorite'. The 'Recent' tab is active, showing a table of recently viewed resources.

Name	Type	Last Viewed
jobstream1	Stream Analytics job	12 minutes ago
xzhang2eventhub1	Event Hubs Namespace	21 minutes ago
xzhang2	Resource group	22 minutes ago
AzureSubscription	Subscription	3 days ago
xzhang2storage1	Storage account	4 days ago

See all

Lab

Step 2. Create an Azure Synapse workspace

- 2 Search for “Azure Synapse Analytics” in the search bar and create a Synapse workspace. Follow what we did before:
 - ▶ We need the storage account and a container
 - ▶ In Security tab, add the SQL administrator credentials and a password. You must remember these as these will be required in the future when working with Synapse Analytics.

Home > Azure Synapse Analytics >

Create Synapse workspace

Validation succeeded

Basics Security Networking Tags Review + create

Product Details

Azure Synapse Analytics workspace by Microsoft
[Terms of use](#) | [Privacy policy](#)

Serverless SQL est. cost/7B
5.00 USD

Terms

By clicking Create, I (a) agree to the legal terms and privacy statement(s) associated with the Marketplace offering(s) listed above; (b) authorize Microsoft to bill my current payment method for the fees associated with the offering(s), with the same billing frequency as my Azure subscription; and (c) agree that Microsoft may share my contact, usage and transactional information with the provider(s) of the offering(s) for support, billing and other transactional activities. Microsoft does not provide rights for third-party offerings. For additional details see [Azure Marketplace Terms](#).

Basics

Subscription	AzureSubscription
Resource group	xzhang2
Region	East US 2
Workspace name	(new) synapse729
Data Lake Storage Gen2 account	https://xzhang2storage1.dfs.core.windows.net
Data Lake Storage Gen2 file system	raw
Managed resource group	None
Role assignments	The Storage Blob Data Contributor role will be assigned on the specified Data Lake Storage Gen2 account to the workspace managed identity.

Security

- 3 Click on the “Create” button. The deployment process will take a few minutes.

The screenshot shows the Microsoft Azure portal interface. At the top, there's a blue header with the Microsoft Azure logo, an 'Upgrade' button, a search bar, and a 'Copilot' button. Below the header, the breadcrumb navigation shows 'Home > Microsoft.Azure.SynapseAnalytics-20240729111659 | Overview'. The main content area is titled 'Deployment' and shows a 'Deployment is in progress' status. On the left, there's a sidebar with 'Overview' selected, and other options like 'Inputs', 'Outputs', and 'Template'. The deployment details section shows the following information:

- Deployment name: Microsoft.Azure.SynapseAnalytics-20240729111659
- Subscription: AzureSubscription
- Resource group: kchang2
- Start time: 7/29/2024, 11:18:02 AM
- Correlation ID: 73363e11-daf9-413d-a091-32b795785fe

Below this, there's a 'Deployment details' section with a table showing the resource 'synapse729' of type 'Synapse workspace' with a status of 'Created'. The table has columns for Resource, Type, Status, and Operation details.

At the bottom, there's a 'Give feedback' section with a link to 'Tell us about your experience with deployment'.

- Once the deployment is completed, click on the “Go to resource group” button.

The screenshot shows the Microsoft Azure portal interface. At the top, there's a navigation bar with the Microsoft Azure logo and an 'Upgrade' button. Below the navigation bar, the breadcrumb 'Home >' is visible. The main heading is 'Microsoft.Azure.SynapseAnalytics-20240729111659 | Overview', with a 'Deployment' sub-heading. A search bar is present. Below the search bar, there's a left-hand navigation pane with 'Overview' selected, and other options like 'Inputs', 'Outputs', and 'Template'. The main content area displays a green checkmark icon and the text 'Your deployment is complete'. Below this, deployment details are listed: 'Deployment name : Microsoft.Azure.SynapseAnalytics-20240729111659', 'Subscription : AzureSubscription', and 'Resource group : xzhang2'. To the right, the 'Start time' is '7/29/2024' and the 'Correlation ID' is '733d0d11-'. Below the details, there are expandable sections for 'Deployment details' and 'Next steps'. At the bottom of the main content area, there is a blue button labeled 'Go to resource group'. A 'Give feedback' section at the very bottom includes a link to 'Tell us about your experience with deployment'.

5 The Resource group dashboard will be as shown in the image below.

Home > Microsoft Azure Synapse Analytics-20240729111659 | Overview >

xzhang2
Resource group

Search + Create ⚙ Manage view ✖ 🗑 Delete resource group 🔄 Refresh ⬇ Export to CSV 🔍 Open query | 🏷 Assign tags ➡ Move 🗑 Delete ⬇ Export template 📱 Open in mobile

Overview

- Activity log
- Access control (IAM)
- Tags
- Resource visualizer
- Events
- Settings
- Cost Management
- Monitoring
- Automation
- Help

Essentials

Subscription ([mouse](#)) : [AzureSubscription](#)
 Subscription ID : c375e05-a27a-41f9-ba5e-20eebfeb29c6
 Deployments : [1 Failed 19 Succeeded](#)
 Location : East US

Tags ([edit](#)) : [Add tags](#)

Resources Recommendations

Filter for any field... Type equals **all** ✕ Location equals **all** ✕ 🔍 Add filter

Showing 1 to 4 of 4 records. ☐ Show hidden types ⓘ

Name ↑	Type ↑	Location ↑
jobstream1	Stream Analytics job	East US
synapse729	Synapse workspace	East US 2
xzhang2eventhub1	Event Hubs Namespace	East US
xzhang2storage1	Storage account	East US

No grouping

- 6 Click on the name of the Synapse workspace listed under the Resources tab of your Resource Group dashboard.

Home > Microsoft.Azure.SynapseAnalytics-20240729111659 | Overview >

xzhang2
Resource group

Search

Create Manage view Delete resource group Refresh Export to CSV Open query Assign tags Move Delete Export template Open in mobile

Overview

Activity log
Access control (IAM)
Tags
Resource visualizer
Events
Settings
Cost Management
Monitoring
Automation
Help

Essentials

Subscription (moad) : [AzureSubscription](#)
Subscription ID : c375ef05-a27a-41f9-ba5e-20eebfef29c6
Deployments : [1 Failed 19 Succeeded](#)
Location : East US
Tags (edit) : [Add tags](#)

Resources Recommendations

Filter for any field... Type equals all Location equals all Add filter

Showing 1 to 4 of 4 records. Show hidden types No grouping

Name ↑↓	Type ↑↓	Location ↑↓
jobstream1	Stream Analytics job	East US
synapse729	Synapse workspace	East US 2
xzhang2eventhub1	Event Hubs Namespace	East US
xzhang2storage1	Storage account	East US

- 7 The Synapse workspace page will appear as shown in the image below.

The screenshot shows the Azure Synapse workspace 'synapse729' overview page. The breadcrumb navigation is 'Home > Microsoft.Azure.SynapseAnalytics-20240729111659 | Overview > xzhang2'. The page title is 'synapse729 Synapse workspace'. Below the title is a search bar and a row of buttons: '+ New dedicated SQL pool', '+ New Apache Spark pool', '+ New Data Explorer pool (preview)', 'Refresh', and 'Reset SQL admin passw...'. A left-hand navigation pane lists: Overview (selected), Activity log, Access control (IAM), Tags, Diagnose and solve problems, Settings (with sub-items: Microsoft Entra ID, Properties, Locks), Analytics pools, Security, Monitoring, Automation, and Help. The main content area is titled 'Essentials' and displays a table of workspace details:

Resource group (move)	: xzhang2	Network
Status	: Succeeded	Primary #
Location	: East US 2	Primary #
Subscription (move)	: AzureSubscription	SQL adm
Subscription ID	: c375ef05-a27a-41f9-ba5e-20eebfeb29c6	SQL Micr
Managed virtual network	: No	Dedicat
Managed Identity object ...	: 46444f17-6726-4ef5-b12d-ba5b9e1b97f2	Serverles
Workspace web URL	: https://web.azure.synapse.net/workspace=%2fsubscriptions%2fc375ef05-a27a-41f9-ba5e-20ee...	Develop
Tags (edit)	: Add tags	

Below the 'Essentials' section is a 'Getting started' area with two cards:

- Open Synapse Studio**: Start building your fully-integrated analytics solution and unlock new insights. [Open in](#)
- Read documentation**: Learn how to be productive quickly. Explore concepts, tutorials, and samples. [Learn more in](#)

- 8 Click on the “+ New dedicated SQL pool” option.

Home > Microsoft.Azure.SynapseAnalytics-20240729111659 | Overview > xzhang2 >

synapse729
Synapse workspace

Search

+ New dedicated SQL pool + New Apache Spark pool + New Data Explorer pool (preview) Refresh Reset SQL admin passw

Overview

- Activity log
- Access control (IAM)
- Tags
- Diagnose and solve problems
- Settings
 - Microsoft Entra ID
 - Properties
 - Locks
- Analytics pools
- Security
- Monitoring
- Automation
- Help

Essentials

Resource group (move)	xzhang2	Networki
Status	Succeeded	Primary #
Location	East US 2	Primary #
Subscription (move)	AzureSubscription	SQL adm
Subscription ID	c375ef05-a27a-41f9-ba5e-20eebf29c6	SQL Micr
Managed virtual network	No	Dedicat
Managed Identity object ...	46444f17-6726-4ef5-b12d-ba5b9e1b97f2	Serverles
Workspace web URL	https://web.azureynapse.net?workspace=%2fsubscriptions%2fc375ef05-a27a-41f9-ba5e-20ee...	Developn
Tags (edit)	Add tags	

Getting started

Open Synapse Studio
Start building your fully-integrated analytics solution and unlock new insights.
[Open c#](#)

Read documentation
Learn how to be productive quickly. Explore concepts, tutorials, and samples.
[Learn more c#](#)

Lab

- 9 Add a name for the dedicated SQL pool and set the performance level to the bare minimum as shown in the image below.
- Click on the “Review + create” button.

[Home](#) > [Microsoft.Azure.SynapseAnalytics-20240729111659](#) | [Overview](#) > [xzhang2](#) > [synapse729](#) >

New dedicated SQL pool ...

[* Basics](#) [* Additional settings](#) [Tags](#) [Review + create](#)

Create a dedicated SQL pool with your preferred configurations. Complete the Basics tab then go to Review + Create to provision with smart defaults, or visit each tab to customize. [Learn more](#) ⓘ

Dedicated SQL pool details

Name your dedicated SQL pool and choose its initial settings.

Dedicated SQL pool name *	<input type="text" value="sqlpool"/> ✓
Performance level ⓘ	<div><div></div><div>DW100c</div></div>
Estimated price ⓘ	<div><div>Est. Cost Per Hour 1.20 USD View pricing details</div></div>

- 10 If it fails, open the Synapse Studio and go to Manage tab to add a dedicated SQL pool.

Microsoft Azure | Synapse Analytics | synapse779

Synapse live | Validate all | Publish all

SQL pools

The serverless SQL pool, Built-in, is immediately available for

+ New Refresh

Filter by name

Showing 1-1 of 1 items (1 Serverless, 0 Dedicated)

Name

Built-in

New dedicated SQL pool

Basics * Additional settings * Tags Review + create

Create a dedicated SQL pool with your preferred configurations. Complete the Basics tab then go to the Additional settings tab to configure your pool. [Learn more](#)

Dedicated SQL pool details

Name your dedicated SQL pool and choose its initial settings.

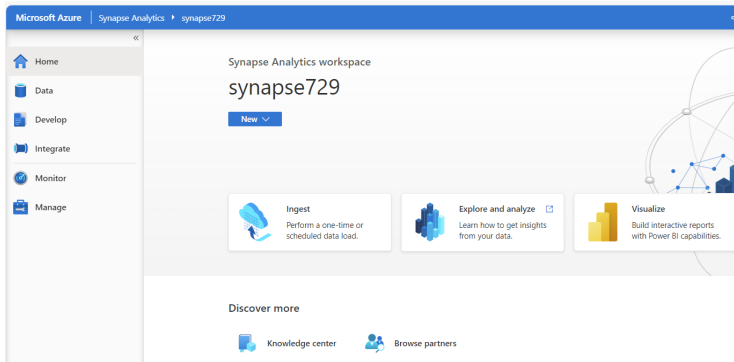
Dedicated SQL pool name *

Performance level ☐ Built-in

Estimated price ☐ Est. cost per hour 1.20 USD [View pricing details](#)

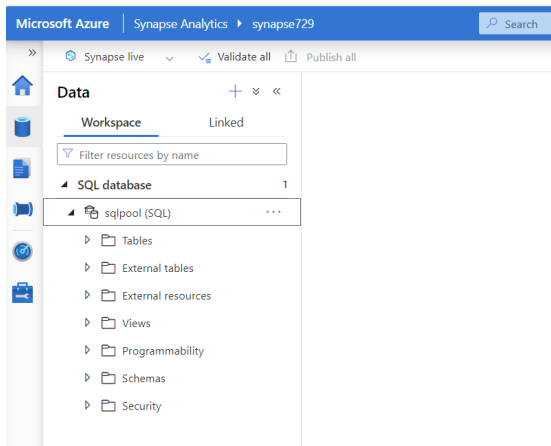
Review + create Next: Additional settings >

- Go to the home page of the Synapse workspace. The Synapse Studio dashboard page will be as shown in the image below.

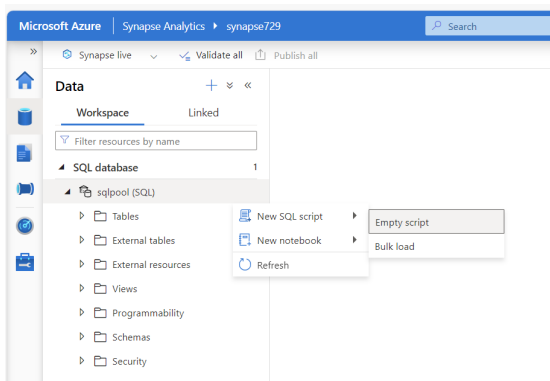


Lab

- 12 Click on the “Data” tab of the menu section on the left side of the screen.
- 13 In the Data menu, click on “SQL database” and click on the section for the custom SQL pool.



- 14 Click on the “...” option to the right of the custom sql pool’s name and create a new script by choosing “New SQL script” and “Empty script”.



Lab

- 15 Create a new table by typing the code given below.

```
CREATE TABLE [dbo].[Orders]
(
  [OrderID] varchar(10),
  [Quantity] int,
  [UnitPrice] decimal(5,2),
  [DiscountCategory] varchar(10)
)
```


- 16 Click on the “Run” option to execute this code.

The screenshot displays the Synapse live SQL editor interface. On the left, the 'Data' pane shows a tree view of resources under the 'Workspace' tab, including 'SQL database' and 'sqlpool (SQL)'. The main editor area shows a SQL script titled 'SQL script 1' with the following code:


```
1 CREATE TABLE [dbo].[Orders]
2 (
3 [OrderID] varchar(10),
4 [Quantity] int,
5 [UnitPrice] decimal(5,2),
6 [DiscountCategory] varchar(10)
7 )
8
```

Below the code editor, the 'Messages' tab is active, showing the execution status: '12:10:58 PM Started executing query at Line 1 (0 record affected)'. The 'Connect to' dropdown menu is set to 'sqlpool'.


Lab



Step 3. Search for Event Hubs to create an event hubs namespace

- 17 Search for “Event Hubs” in the search bar and click on the “Create event hubs namespace” button.



Event Hubs



 Create  View

Description

An Event Hubs namespace is a management container for event hubs which provides DNS-integrated network endpoints and a range of access control and network integration management.

Free training from Microsoft

[Enable reliable messaging for Big Data applicat...](#)
8 units · 56 min

Lab

- 18 Enter the below details and click on the “Review + create” button.

[Home](#) > [Event Hubs](#) >

Create Namespace

Event Hubs

[Basics](#) [Advanced](#) [Networking](#) [Tags](#) [Review + create](#)

Project Details

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

Subscription *

Resource group *
[Create new](#)

Instance Details

Enter required settings for this namespace, including a price tier and configuring the number of units (capacity).

Namespace name * ✓
.servicebus.windows.net

Location *
The region selected supports Availability zones. Your namespace will have Availability Zones enabled. [Learn more](#).

Pricing tier *
[Browse the available plans and their features](#)

Throughput Units *

[Review + create](#) [< Previous](#) [Next: Advanced >](#)

Lab

- 19 Validate the configuration and click on the “Create” button.

[Home](#) > [Event Hubs](#) >



Create Namespace ...

Event Hubs



Validation succeeded.

Basics

Advanced

Networking

Tags

Review + create

Event Hubs Namespace
by Microsoft

Basics

Namespace name	xzhang2eventhub1
Subscription	AzureSubscription
Resource group	xzhang2
Location	East US
Pricing tier	Basic
Throughput Units	1
Availability Zones (Zone Redundancy)	Enabled


Networking

Connectivity method	Public access
---------------------	---------------








Security

Minimum TLS version	1.2
Local Authentication	Enabled

- Note: The deployment process will take a few minutes.

Home > **xzhang2eventhub1** | Overview  ...

Deployment

Search    Delete  Cancel  Redeploy  Download  Refresh

Overview

Inputs

Outputs

Template

*** Deployment is in progress

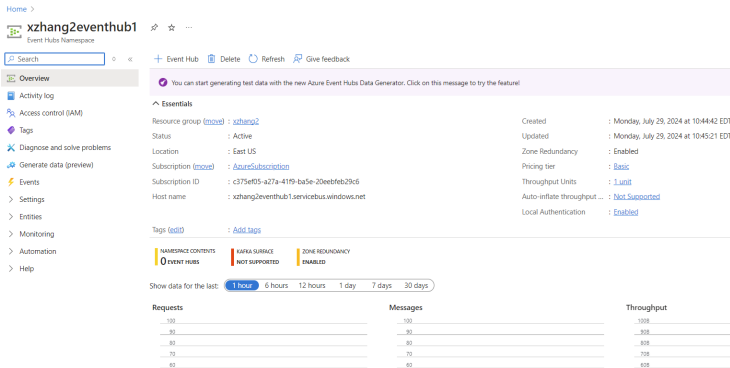
Deployment name : xzhang2eventhub1
Subscription : AzureSubscription
Resource group : xzhang2

Start time : 7/20/2024, 10:44:41 AM
Correlation ID : c62863c4-6c32-4af6-89a7-1eda2ed7b772

Deployment details

Resource	Type	Status	Operation details
xzhang2eventhub1	Event Hubs Namespace	Created	Operation details

20 The Event Hubs dashboard will be as shown in the image below.



21 Here, click on “+ Event Hub”.

Home >

xzhang2eventhub1 Event Hubs Namespace

Search

+ Event Hub Delete Refresh Give feedback

Overview

- Activity log
- Access control (IAM)
- Tags
- Diagnose and solve problems
- Generate data (preview)
- Events
- Settings
- Entities
- Monitoring
- Automation
- Help

Essentials

Resource group (move) : xzhang2

Status : Active

Location : East US

Subscription (move) : AzureSubscription

Subscription ID : c375ef05-a27a-41f9-ba5e-20eebf29c6

Host name : xzhang2eventhub1.servicebus.windows.net

Tags (edit) : Add tags

Created Updated Zone Red Pricing tie Throughput Auto-infla Local Aut

Namespaces

0 EVENT HUBS

KAFKA SURFACE NOT SUPPORTED

ZONE REDUNDANCY ENABLED

Show data for the last: 1 hour 6 hours 12 hours 1 day 7 days 30 days

- 22 Add the name and retention period shown in the image and click on the “Create” button.

[Home](#) > [xzhang2eventhub1](#) >

Create Event Hub

Event Hubs

Basics Capture Review + create

Event Hub Details

Enter required settings for this event hub, including partition count and message retention.

Name * ⓘ

demohub ✓

Partition count ⓘ



1

Retention

Configure retention settings for this Event Hub. [Learn more](#)

Cleanup policy ⓘ

Delete ✓

Retention time (hrs) * ⓘ

7 ✓

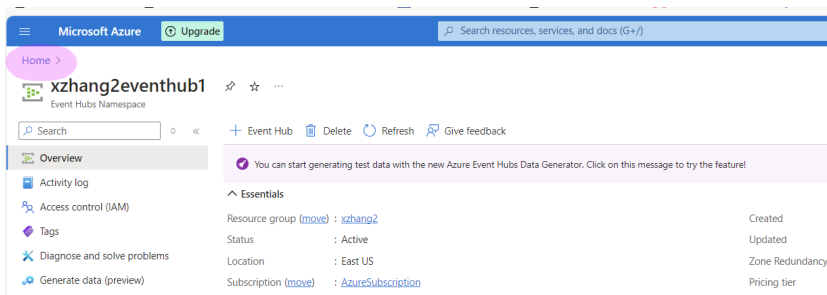
min. 1 hour, max. 24 hours (1day)

Review + create

< Previous

Next: Capture >

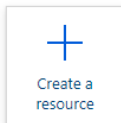
- Click on the “Home” option in the top-left corner of the screen.



Step 4. Create a Stream Analytics job

- 24 Click on the “Create a resource” option on the home page.

Azure services



Quickstart
Center



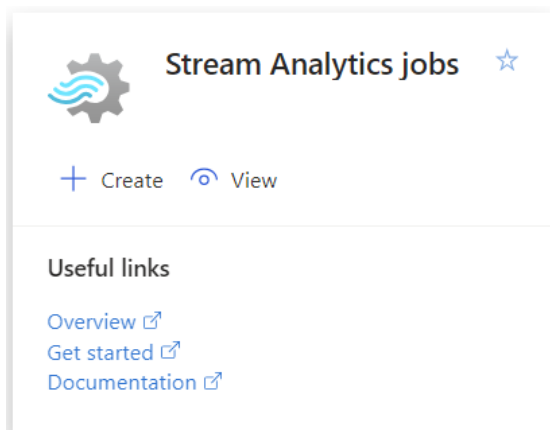
Virtual
machines



App Services

Lab

- 25 Search for “Stream Analytics jobs” in the search bar and click on the “Create stream analytics job” button.



26 Enter the below details and click on the “Create” button.

The screenshot shows the 'New Stream Analytics job' page in the Microsoft Azure portal. The page is divided into several sections for configuration:

- Project details:** Includes fields for 'Subscription' (set to 'AzureSubscription') and 'Resource group' (set to 'xzhang2'). A 'Create new' link is visible below the resource group field.
- Instance details:** Includes fields for 'Name' (set to 'jobstream1') and 'Region' (set to '(US) East US'). The 'Hosting environment' is set to 'Cloud' with a radio button.
- Streaming unit details:** Includes a description of Streaming Units (SUs) and a slider for 'Streaming units' set to '1'. A link 'For more options, visit here' is provided.

At the bottom of the page, there are three buttons: 'Previous', 'Next', and 'Review + create'.

- Note: The deployment process will take a few minutes.

The screenshot displays the Microsoft Azure portal interface. At the top, there's a navigation bar with the Microsoft Azure logo, an 'Upgrade' button, a search bar, and a 'Copilot' button. Below the navigation bar, the breadcrumb trail shows 'Home > StreamAnalyticsJob | Overview'. The main content area is titled 'StreamAnalyticsJob | Overview' and includes a 'Deployment' sub-header. A search bar and action buttons (Delete, Cancel, Redeploy, Download, Refresh) are present. On the left, a sidebar lists navigation options: Overview (selected), Inputs, Outputs, and Template. The main content area shows a status message '*** Deployment is in progress' with a blue progress indicator. Below this, deployment metadata is listed: Deployment name (StreamAnalyticsJob), Subscription (AzureSubscription), and Resource group (xzhang2). To the right, the Start time (7/29/2024, 10:32:40 AM) and Correlation ID (e673a5f9-7f9e-448d-bb67-dd1c188c1dcf) are displayed. A section titled 'Deployment details' is expanded, showing a table with columns: Resource, Type, Status, and Operation details. The table is currently empty, with a message 'There are no resources to display.' at the bottom. A 'Give feedback' link is located at the bottom left of the main content area.

Microsoft Azure Upgrade Search resources, services, and docs (G+/) Copilot

Home > StreamAnalyticsJob | Overview

Deployment

Search Delete Cancel Redeploy Download Refresh

Overview

Inputs

Outputs

Template

*** Deployment is in progress

Deployment name : StreamAnalyticsJob
Subscription : AzureSubscription
Resource group : xzhang2

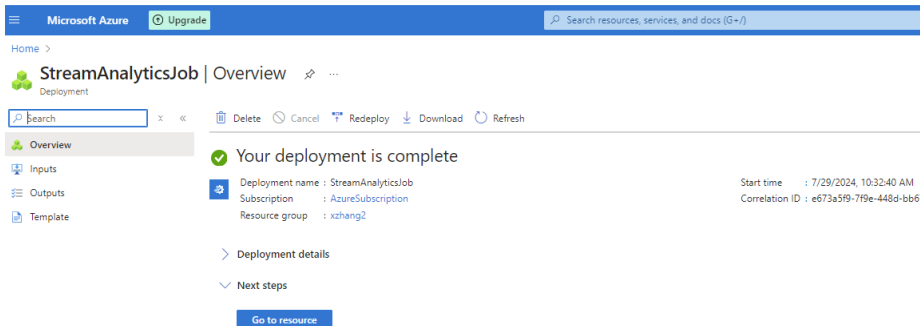
Start time : 7/29/2024, 10:32:40 AM
Correlation ID : e673a5f9-7f9e-448d-bb67-dd1c188c1dcf

Deployment details

Resource	Type	Status	Operation details
There are no resources to display.			

Give feedback

- 27 Once the deployment is completed, click on the “Go to resource” button.



The screenshot shows the Microsoft Azure portal interface. At the top, there's a blue header with the Microsoft Azure logo and an 'Upgrade' button. Below the header, a search bar is present. The main content area is titled 'StreamAnalyticsJob | Overview' with a 'Deployment' sub-header. A left sidebar contains navigation links: Overview (selected), Inputs, Outputs, and Template. The main content area displays a green checkmark icon and the text 'Your deployment is complete'. Below this, deployment details are listed: Deployment name: StreamAnalyticsJob, Subscription: AzureSubscription, and Resource group: xzhang2. To the right of these details, the Start time is 7/29/2024, 10:32:40 AM, and the Correlation ID is e673a5f9-7f9e-448d-bb6. At the bottom of the main content area, there are expandable sections for 'Deployment details' and 'Next steps', and a prominent blue 'Go to resource' button.

28 The Stream Analytics job dashboard will be as shown in the image below.

Home > StreamAnalyticsJob | Overview >

jobstream1 ☆ ...
Stream Analytics job

Search 🔍 ⌵ ⌵ ▶ Start job 🗑 Delete ➡ Move ⌵ 🔄 Refresh 😊 Share feedback

- Overview
- Activity log
- Access control (IAM)
- Tags
- Diagnose and solve problems
- Job topology
- Settings
- Developer tools
- Monitoring
- Automation
- Help

⌵ Essentials

Resource group (move)	: xzhango2	Created	: Monday, July 29, 2024 10:31
Location	: East US	Started	:
Status	: Created	Output watermark	:
Subscription (move)	: AzureSubscription	Cluster	: Shared
Subscription ID	: c375ef05-a27a-41f9-ba5e-20eebf29c6	Hosting environment	: Cloud
Pricing plan	: StandardV2 (manage)	Virtual Network	: Disabled
Tags (edit)	: Add tags		

[Get started](#) Properties Monitoring Tutorials

Which best describes your needs?

- 29 Now, click on the “Home” option in the top-left corner of the screen.

Home > StreamAnalyticsJob | Overview >

jobstream1 ☆ ...
Stream Analytics job

◊ ◀ ▶ Start job Delete → Move ▾ ↺ Refresh 😊 Share feedback

Overview

- Activity log
- Access control (IAM)
- Tags
- Diagnose and solve problems
- > Job topology
- > Settings

Created

^ Essentials

Resource group (move)	: xzhang2
Location	: East US
Status	: Created
Subscription (move)	: AzureSubscription
Subscription ID	: c375ef05-a27a-41f9-ba5e-20eebfeb29c6

- Click on the “Inputs” option under the Job topology category provided in the left section.

Home >

jobstream1 ☆ ...

Stream Analytics job

Search

Start job Delete Move Refresh Share feedback

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Job topology

Inputs ☆

Functions

Query

Outputs

No-code editor (preview)

Settings

Developer tools

Monitoring

Automation

Help

Created

Essentials

Resource group (move) : [xzhang2](#)

Location : East US

Status : Created

Subscription (move) : [AzureSubscription](#)

Subscription ID : c375ef05-a27a-41f9-ba5e-20eebfeb29c6

Pricing plan : StandardV2 (manage)

Tags (edit) : [Add tags](#)

Get started Properties Monitoring Tutorials

Which best d

I want to build my job using a c

<>

Lab

- 31 Click on the “+ Add input” button and click on the “Event Hub” option.

The screenshot shows the Microsoft Azure portal interface for a Stream Analytics job named 'jobstream1'. The top navigation bar includes the 'Microsoft Azure' logo and an 'Upgrade' button. Below the navigation bar, the breadcrumb 'Home > jobstream1' is visible. The main heading is 'jobstream1 | Inputs', with a star icon and a menu icon. Below the heading, there is a search bar and two buttons: '+ Add input' and 'Refresh'. The left sidebar contains a list of navigation options: Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Job topology, Inputs (selected), Functions, Query, Outputs, No-code editor (preview), Settings, Developer tools, Monitoring, Automation, and Help. The 'Add input' dropdown menu is open, showing two sections: 'Stream input' and 'Reference input'. Under 'Stream input', the options are Blob storage/ADLS Gen2, Event Hub (highlighted), IoT Hub, and Kafka (preview). Under 'Reference input', the options are Blob storage/ADLS Gen2 and SQL Database.

Lab

- 32 Enter the details shown in the images below in the first half of the Event Hub form:
- Click on the drop-down option for Authentication mode and select the “Connection string” option.

Event Hub

New input

Input alias *

demohub ✓

☐ Provide Event Hub settings manually

☒ Select Event Hub from your subscriptions

Subscription

AzureSubscription ▼

Event Hub namespace * ⓘ

xzhang2eventhub1 ▼

Event Hub name * ⓘ

☐ Create new ☒ Use existing

demohub ▼

Event Hub consumer group * ⓘ

☐ Create new ☒ Use existing

\$Default ▼

Authentication mode

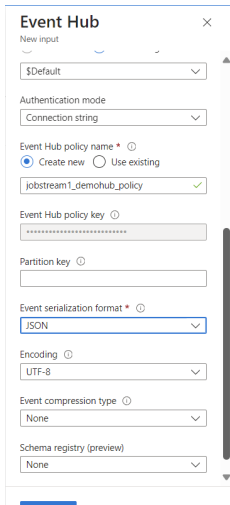
Connection string ▼

Event Hub policy name * ⓘ

☒ Create new ☐ Use existing

Lab

- 33 Ensure the details, which are filled in automatically, match the one given in the image below. Click on the “Save” button.
- The connection will be automatically tested.



The image shows a screenshot of the 'Event Hub' configuration form in the Azure portal. The form is titled 'Event Hub' and has a 'New input' label. It contains several fields and options for configuring an Event Hub. The 'Authentication mode' is set to 'Connection string'. The 'Event Hub policy name' is set to 'jobstream1_demohub_policy' with a green checkmark. The 'Event Hub policy key' is masked with asterisks. The 'Partition key' is empty. The 'Event serialization format' is set to 'JSON'. The 'Encoding' is set to 'UTF-8'. The 'Event compression type' is set to 'None'. The 'Schema registry (preview)' is set to 'None'. A 'Save' button is visible at the bottom left of the form.

Event Hub ✕

New input

\$Default

Authentication mode

Connection string

Event Hub policy name * ⓘ

☒ Create new ☐ Use existing

jobstream1_demohub_policy ✓

Event Hub policy key ⓘ

Partition key ⓘ

Event serialization format * ⓘ

JSON

Encoding ⓘ

UTF-8

Event compression type ⓘ

None

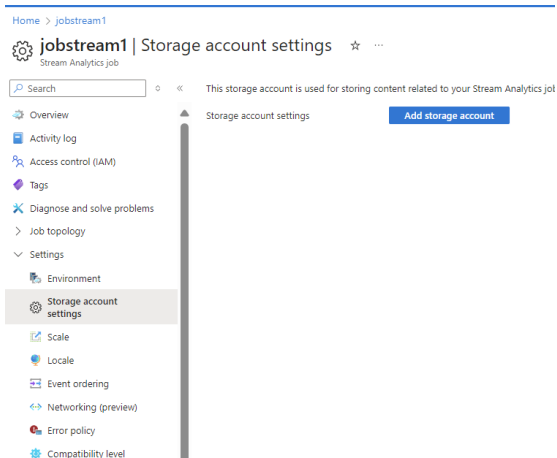
Schema registry (preview)

None

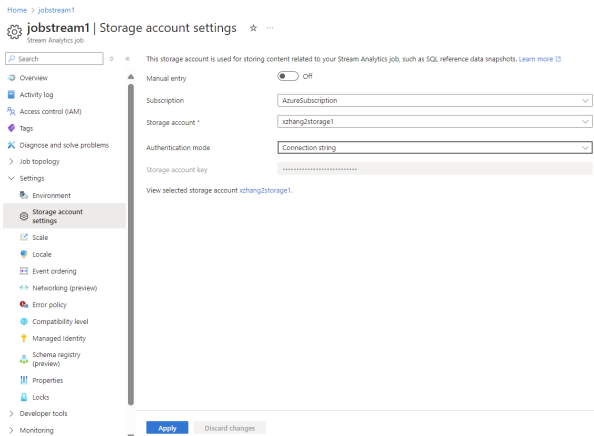
Save

Lab

- 34 Click on the “Storage account settings” option under the Settings on the left side of your screen.



- 35 Click on the “Add storage account” button.
- 36 Select the storage account that is already created and click on the “Apply” button at the bottom. Add other details as shown in the image given below.



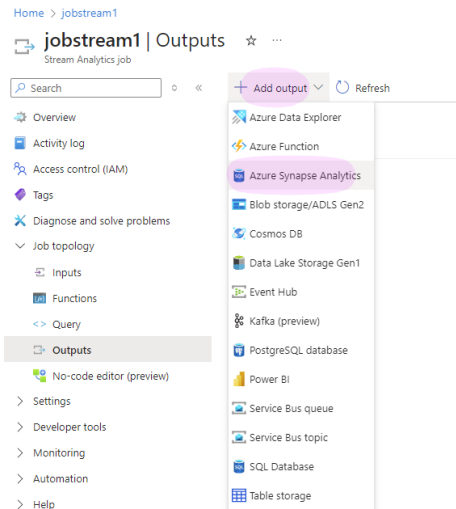
Lab

- 37 Click on the “Outputs” option under Job topology.

The screenshot shows the AWS Glue console interface for a job named 'jobstream1'. The breadcrumb navigation at the top reads 'Home > jobstream1'. Below this, the page title is 'jobstream1 | Outputs' with a star icon and a menu icon. Underneath the title is the text 'Stream Analytics job'. A search bar is present with the placeholder text 'Search'. To the right of the search bar are icons for 'Add output' (a plus sign) and 'Refresh' (a circular arrow). A horizontal line separates the header from the main content area. On the left side, there is a vertical navigation menu with the following items: 'Overview' (gear icon), 'Activity log' (document icon), 'Access control (IAM)' (person icon), 'Tags' (tag icon), 'Diagnose and solve problems' (wrench icon), 'Job topology' (dropdown arrow icon), 'Inputs' (document icon), 'Functions' (code icon), 'Query' (code icon), 'Outputs' (document icon, highlighted with a pink oval), 'No-code editor (preview)' (document icon), 'Settings' (dropdown arrow icon), 'Developer tools' (dropdown arrow icon), 'Monitoring' (dropdown arrow icon), 'Automation' (dropdown arrow icon), and 'Help' (dropdown arrow icon). On the right side, there is a section titled 'Alias' with an upward arrow icon.

Lab

- 38 Click on the “+ Add output” option and select “Azure Synapse Analytics”.



Lab

- 39 Enter the details provided in the image below and click on the “Save” button.
- Click on the drop-down option under Authentication mod and select “Connection string”.
- Enter the SQL login credentials created and provide the name for the table (Orders) created in the database.

The screenshot shows the 'Azure Synapse Analytics' 'New output' configuration window. It contains the following fields and options:

- Output alias ***: sqlpool
- Authentication options**:
 - ☐ Provide Azure Synapse Analytics settings manually
 - ☒ Select Azure Synapse Analytics from your subscriptions
- Subscription**: AzureSubscription
- Database ***: sqlpool (synapse729)
- Authentication mode**: SQL server authentication
- Username ***: sqladminuser
- Password ***: (masked with asterisks)
- Table ***: Orders
- Save** button

- 40 Click on Save. The connection will be automatically tested.

ts ☆ ...

×

+ Add output ▾ ↻ Refresh

Alias ↑	T...	A...	R...
sqlpool	A...	Conr	

Notifications

×

More events in the activity log →

Dismiss all ▾

Successful connection test

×

Connection to output 'sqlpool' succeeded.

a few seconds ago

Added output

×

Added output 'sqlpool' to Stream Analytics job 'jobstream1'.

a minute ago

Lab

- 41 Click on the “Query” option under Job topology.
- 42 The Query section page will appear as shown in the image below.

```
/*
Here are links to help you get started with Stream Analytics Query I
Common query patterns - https://go.microsoft.com/fwLink/?LinkID=6191
Query language - https://docs.microsoft.com/stream-analytics-query/c
*/
SELECT
    *
INTO
    [sqlpool]
FROM
    [demohub]
```

Home > jobstream1

<> jobstream1 | Query ☆ ...

Stream Analytics job

Search

- Overview
- Activity log
- Access control (IAM)
- Tags
- Diagnose and solve problems
- Job topology
 - Inputs
 - Functions
 - <> Query**
 - Outputs
 - No-code editor (preview)
- Settings
- Developer tools

Start job Open in VS Code Diagnostic settings Refresh Query language docs Share feedback Tutorial

- Inputs (1)
 - demohub
- Outputs (1)
 - sqlpool
- Functions (0)

Test query Save query Discard changes

```

1 /*
2 Here are links to help you get started with Stream Analytics Query Language:
3 Common query patterns - https://go.microsoft.com/fwlink/?linkID=619153
4 Query language - https://docs.microsoft.com/stream-analytics-query/query-language-elements-azure-stream-analytics
5 */
6 SELECT
7     *
8 INTO
9     [sqlpool]
10 FROM
11     [demohub]
    
```

Input preview Test results Job simulation (preview)

While sampling data, no data was received from 'I' partitions.

Table Raw Refresh Select time range Upload sample input Send events Download sample data

- 43 Query the output stream just created by using the code above.
 - 44 You can click on the “Save query” option to execute this code.
- Note: We are not executing the query as there is no data present at the input stream at this stage.

home > jobstream1

<> jobstream1 | Query ☆ ...
Stream Analytics job

Search ◊ « ▶ Start job 📄 Open in VS Code 🛠 Diagnostic settings 🔄 Refresh 📖 Query language docs ▾ 😊 Share feedback 📖 Tutorial

Overview

Activity log

Access control (IAM)

Tags

✓ Diagnose and solve problems

Test query

There is no data from input 'demohub'. Please make sure the input source has data and then try again.

OK

sqlpool

6 SELECT

45 We may upload a sample data

Home > StreamAnalyticsJob | Overview > jobstream1

<> jobstream1 | Query ☆ ...

Stream Analytics job

Search

Start job Open in VS Code Diagnostic settings Refresh Query language docs Share feedback Tutorial

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Job topology

Inputs

Functions

Query

Outputs

No-code editor (preview)

Settings

Developer tools

Monitoring

Automation

Help

Inputs (1)

demohub

Outputs (1)

sqlpool

Functions (0)

Test query Save query Discard changes

```
1 /*
2 Here are links to help you get started with Stream Analytics Query Lang
3 common query patterns - https://go.microsoft.com/fwlink/?linkID=619153
4 Query language - https://docs.microsoft.com/stream-analytics-query/quer
5 */
6 SELECT
7 *
8 INTO
9 [sqlpool]
10 FROM
11 [demohub]
```

Input preview Test results Job simulation (preview)

While sampling data, no data was received from "1" partitions.

Table Raw Refresh Select time range Upload sample input Send even

Upload sample data from file *

sample.csv

You can upload a file with sample input to test your query against. Stream Analytics supports processing input in JSON, CSV and Avro formats with UTF-8 encoding natively.

OK

- 46 Then test the query again

Test query Save query Discard changes

```

1  /*
2  Here are links to help you get started with Stream Analytics Query Language:
3  Common query patterns - https://go.microsoft.com/fwlink/?LinkID=619153
4  Query language - https://docs.microsoft.com/stream-analytics-query/query-language-elements-azure-stream-analytics
5  */
6  SELECT
7      *
8  INTO
9      [sqlpool]
10 FROM
11     [demohub]

```

Input preview Test results Job simulation (preview)

Download results

OrderID	Quantity	UnitPrice	DiscountCategory
string	string	string	string
"12"	"36"	"1120"	"5"
"57"	"42"	"1274.23"	"3"
"60"	"36"	"914.66"	"5"
"77"	"1"	"357.4"	"1"

Showing 15 rows from 'sqlpool'. Ln 1, Col 1

- 47 Go to the “Overview” option provided at the top of the section on the left side of your screen.

Home > **jobstream1** ☆ ...
Stream Analytics job

Search 0 << ▶ Start job Delete → Move ↕ Refresh ☺ Share feedback

Overview
Activity log
Access control (IAM)
Tags
Diagnose and solve problems
Job topology
Inputs
Functions
Query
Outputs
No-code editor (preview)
Settings
Developer tools
Monitoring
Automation
Help

Created

Essentials

Resource group (move)	: xzhang2	Created	: Monday, July 23, 2024 10:33 AM
Location	: East US	Started	:
Status	: Created	Output watermark	:
Subscription (move)	: AzureSubscription	Cluster	: Shared
Subscription ID	: c375ef05-a27a-41f9-ba5e-20eebf6b29c6	Hosting environment	: Cloud
Pricing plan	: StandardV2 (manage)	Virtual Network	: Disabled

Tags [\(edit\)](#) : [Add tags](#)

Get started **Properties** Monitoring Tutorials

Inputs (1)
demohub

<> **Query**
[Edit query](#)

Outputs (1)
sqlpool

Functions (0)
[Add function](#)

Settings

Environment	: Standard (multi-tenant)
Storage account settings	: xzhang2storage1
Scale	: 1 streaming units
Out of order tolerance	: 0 seconds
Error policy	: Retry
Managed identity	: 4ceec579-4163-48e0-942f-771

Monitoring

- Alert rules
- Diagnostic settings

- 48 Click on the “Start job” button.

Start job

×

jobstream1

① You have not configured the diagnostic settings for this job yet.
[Add diagnostic settings in the diagnostic settings pane.](#)

Streaming units ①

1

Environment ①

Standard

Job output start time ①





☒ Now

☐ Custom

Start

49 Set the Diagnostic settings


Diagnostic setting

 Save  Discard  Delete  Feedback

will occur. [Learn more about the different log categories and contents of those logs](#)

Diagnostic setting name
DiagnosticSetting1

Logs

Category groups 

☒ allLogs

Categories

☒ 執行

☒ 製作


Metrics

☒ AllMetrics

Destination details

☐ Send to Log Analytics workspace

☒ Archive to a storage account

 Showing all storage accounts including classic storage accounts

Location
East US

Subscription

Storage account *

Lab

- 90 Click on the “Start” button. It will take time to get the job started.

Start job

jobstream1

Streaming units ⓘ

1

Environment ⓘ

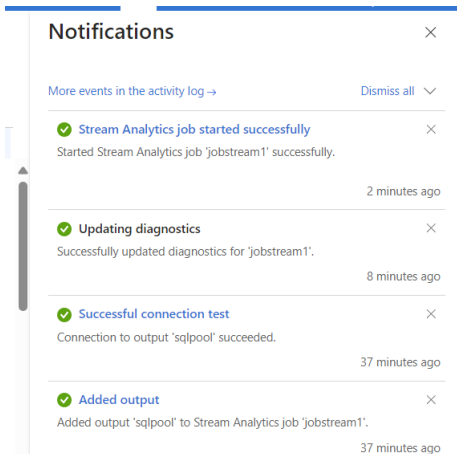
Standard

Job output start time ⓘ

☒ Now

☐ Custom

Start



- 51 Once done, you are ready to process real-time data from Event Hubs with Stream Analytics.

- 52 Go to Event Hubs Namespace (with Event Hubs Instance demohub) to generate real-time data

The screenshot displays the Microsoft Azure portal interface for an Event Hubs Namespace named 'demohub (xzhang2eventhub1/demohub)'. The left sidebar shows the navigation menu with 'Overview' selected. A purple banner at the top of the main content area reads: 'You can start generating test data with the new Azure Event Hubs Data Generator. Click on this message to try the feature!'. Below this, the 'Essentials' section provides key details: Resource group (xzhang2), Location (East US), Subscription ID (c375e059-a27e-41f9-ba5e-32aeb4b23bd), Partition count (1), Status (Active), Namespace (xzhang2eventhub1), Created (Monday, July 29, 2024 at 10:50:37 EDT), Updated (Monday, July 29, 2024 at 13:53:17 EDT), and Cleanup policy (Delete). The 'Actions' section contains six tiles: 'Capture events' (Use Capture to save your events to persistent storage), 'Process data' (Process data instantly with Azure Stream Analytics), 'Connect' (Authenticate with connection strings and SAS policies), 'Checkpoint' (Create consumer groups to checkpoint your events), 'Generate data (preview)' (Emit data to your event hub just with a few clicks), and 'Analyze data (preview)' (Ingest to Data Explorer for near real-time analysis). At the bottom, a status bar shows 'Event Hubs Namespace' with 1 consumer group, 'Event Hubs data' as ACTIVE, 'Cleanup policy' as DELETE, and a 'Partition count' of 1. On the right, a 'Show data for the last' dropdown is set to '1 hour'.

Lab

- 53 Click on the data generator, you will see

```
[
  {
    "key1": "value1",
    "key2": "value2",
    "key3": "value3",
    "nestedKey": {
      "nestedKey1": "nestedValue1"
    },
    "arrayKey": [
      "arrayValue1",
      "arrayValue2"
    ]
  }
]
```

Lab

- 54 We update the code to ensure that the schema of the data from the Event Hub matches the schema of the target table in Synapse SQL.

```
[
  {
    "OrderID": "1",
    "Quantity": 5,
    "UnitPrice": 120,
    "DiscountCategory": "Category 1"
  },
  {
    "OrderID": "2",
    "Quantity": 10,
    "UnitPrice": 200,
    "DiscountCategory": "Category 2"
  },
  {
    "OrderID": "3",
    "Quantity": 15,
    "UnitPrice": 300,
    "DiscountCategory": "Category 3"
  },
  {
    "OrderID": "4",
    "Quantity": 20,
    "UnitPrice": 400,
    "DiscountCategory": "Category 4"
  },
  {
    "OrderID": "5",
    "Quantity": 25,
    "UnitPrice": 500,
    "DiscountCategory": "Category 5"
  }
]
```

- 95 You may click **Select time range** to sample data from the event hub.
- Further, we see three more columns

Test query

Save query

Discard changes

```

1  /*
2  Here are links to help you get started with Stream Analytics Query Language:
3  Common query patterns - https://go.microsoft.com/fwlink/?LinkID=619153
4  Query language - https://docs.microsoft.com/stream-analytics-query/query-language-elements-azure-stream-analytics
5  */
6  SELECT
7  *
8  INTO
9  [sqlpool]
10 FROM
11 [demohub]

```

Input preview

Test results

Job simulation (preview)

Showing sample events from 'demohub'.

Table

Raw

Refresh

Select time range

Upload sample input

Send events

Download sample data

OrderID	Quantity	UnitPrice	DiscountCategory	EventProcessedUtcTime	PartitionId	EventEnqueuedUtcTime
string	string	string	string	datetime	bigint	datetime
"5"	"10"	"120"	"Category 5"	"2024-07-30T15:38:47...."	0	"2024-07-30T15:10:03.498000002"
"5"	"10"	"120"	"Category 5"	"2024-07-30T15:38:47...."	0	"2024-07-30T15:16:27.633000002"
"1"	"5"	"120"	"Category 1"	"2024-07-30T15:38:47...."	0	"2024-07-30T15:36:04.508000002"
"2"	"10"	"200"	"Category 2"	"2024-07-30T15:38:47...."	0	"2024-07-30T15:36:04.508000002"
"3"	"15"	"300"	"Category 3"	"2024-07-30T15:38:47...."	0	"2024-07-30T15:36:04.508000002"

Lab

96 So we need stop the job and update the query in the Stream Analytics job as

```
SELECT
    OrderID, Quantity, UnitPrice, DiscountCategory
INTO
    [sqlpool]
FROM
    [demohub];
```

Start job Open in VS Code Diagnostic settings Refresh Query language docs Share feedback Tutorial Job ready

Test query Save query Discard changes

```
1 SELECT
2 OrderID, Quantity, UnitPrice, DiscountCategory
3 INTO
4 [sqlpool]
5 FROM
6 [demohub]
```

Input preview Test results Job simulation (preview)


Showing sample events from 'demohub'.

Table Raw Refresh Select time range Upload sample input Send events

OrderID	Quantity	UnitPrice	DiscountCategory	EventProcessedUtcTime	Partition
string	string	string	string	datetime	bigint
"5"	"10"	"120"	"Category 5"	"2024-07-30T18:03:22...."	0
"5"	"10"	"120"	"Category 5"	"2024-07-30T18:03:22...."	0

57 Start the job again, and generate data again.

Home > demohub (xzhang2eventhub1/demohub)


 **demohub (xzhang2eventhub1/demohub)** | Generate data (preview) ☆ ...
Event Hubs Instance

◊ ◀

Generate Data in quick steps to stream into Azure Event Hubs

 Overview

 Access control (IAM)


 Diagnose and solve problems

> Settings


> Entities

▼ Features

 Capture

 **Generate data (preview)**

 Process data

 Analyze data (preview)

▼ Automation

 CLI / PS

 Tasks (preview)

 Export template

> Help

Select Dataset *

Custom payload ▼

Select Content-Type *


JSON ▼

Repeat Send *

10

Send

Cancel

 The content in the Enter payload section is treated as one event. For more information, [click here](#)

Enter payload

```
1  {
2      "OrderID": "1",
3      "Quantity": "5",
4      "UnitPrice": "120",
5      "DiscountCategory": "Category 1"
6  },
7
8  {
9      "OrderID": "2",
10     "Quantity": "10",
11     "UnitPrice": "200",
12     "DiscountCategory": "Category 2"
13 },
14 {
15     "OrderID": "3",
16     "Quantity": "15",
17     "UnitPrice": "300",
18     "DiscountCategory": "Category 3"
19 },
20 {
21     "OrderID": "4",
22     "Quantity": "20",
23     "UnitPrice": "400".
```

58 then check the Synapse database

The screenshot shows the Microsoft Azure Synapse Analytics workspace 'synapse729'. On the left, the 'Data' pane shows a linked 'SQL database' named 'sqlpool (SQL)'. The main editor shows 'SQL script 1' with the query: `select * from Orders`. The 'Run' button is active. Below the script, the 'Results' pane shows a table view of the query output.

OrderID	Quantity	UnitPrice	DiscountCategory
4	20	400.00	Category 4
3	15	300.00	Category 3
2	10	200.00	Category 2
2	10	200.00	Category 2
1	5	120.00	Category 1
3	15	300.00	Category 3
4	20	400.00	Category 4
1	5	120.00	Category 1
1	5	120.00	Category 1
5	25	500.00	Category 5
5	25	500.00	Category 5
5	25	500.00	Category 5

- 59 You may stream inputs from any application such as .NET application to send data to the **Event Hubs Namespace**. You need the connection authentication.

The screenshot displays the Azure portal interface for managing Event Hubs. The left sidebar shows the navigation menu with 'Shared access policies' selected. The main content area shows the 'demohub (xzhang2eventhub1/demohub) | Shared access policies' page. A table lists the policies, with 'jobstream1_demohub_policy' selected. The right-hand pane shows the details for this policy, including the 'SAS Policy: jobstream1_demohub_policy' title, action buttons (Save, Discard, Delete, Regenerate Primary Key), and checkboxes for 'Manage', 'Send', and 'Listen'. The 'Primary key' and 'Secondary key' sections show the respective keys. The 'Connection string-primary key' section is highlighted with a pink box, showing the connection string: 'Endpoint=sb://xzhang2eventhub1.servicebus.windows.net/SharedAccessKeyName=job...'. The 'Connection string-secondary key' section shows a similar connection string. The 'SAS Policy ARM ID' section shows the ARM ID: '/subscriptions/c375e105-a27a-41f9-ba5e-20eebfeb29c6/resourceGroups/xzhang2/prov...'.

License



This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-nc-sa/4.0/).