

Chapter 1: Getting Started with ADF

All services > Microsoft.StorageAccount-20201012115127 >

adfcookbookstorage Storage account

Search (Cmd+ /) Open in Explorer Move Refresh Delete Feedback

Overview Classic alerts in Azure Monitor is announced to retire in 2021, it is recommended that you upgrade your classic alert rules to retain alerting functionality with the new alerting platform. For more information, see Continue alerting with ARM storage accounts.

Activity log

Tags

Diagnose and solve problems

Access Control (IAM)

Data transfer

Events

Storage Explorer (preview)

Essentials

Resource group (change) : ADFCookbook Status : Primary: Available, Secondary: Available Location : East US, West US Subscription (change) : Free Trial Subscription ID : d388216e-67b8-41b3-9832-d6511a1a70be Tags (change) : Click here to add tags

Performance/Access tier : Standard/Hot Replication : Read-access geo-redundant storage (RA-GRS) Account kind : StorageV2 (general purpose v2)

Containers Scalable, cost-effective storage for unstructured data Learn more

File shares Serverless SMB and NFS file shares Learn more

Tables Tabular data storage Learn more

Settings

Access keys Geo-replication CORS Configuration

Data Factory Publish all Validate all Refresh

Factory Resources

Filter resources by name +

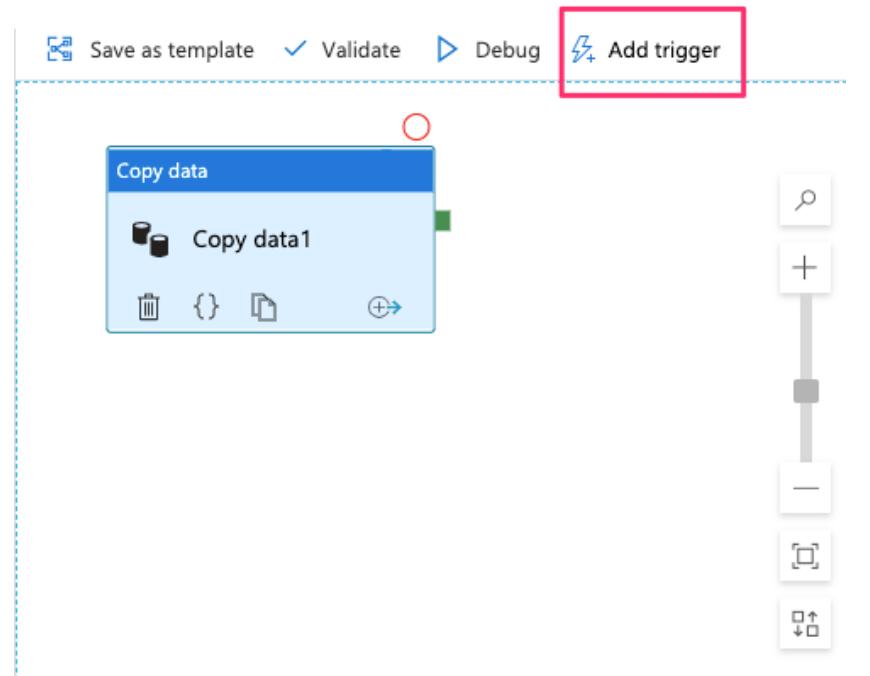
Pipelines 0

Datasets 0

Data flows

New dataset New folder

The image shows two screenshots from the Azure portal. The top screenshot is of a Storage Account named 'adfcookbookstorage'. It displays basic account details like resource group, location, and access tier. Below the details are three service cards: 'Containers' (selected and highlighted with a red box), 'File shares', and 'Tables'. The bottom screenshot is of the Data Factory interface. It shows 'Factory Resources' with sections for Pipelines, Datasets (selected and highlighted with a red box), and Data flows. Under Datasets, there are 'New dataset' and 'New folder' buttons.



The screenshot shows the 'Factory Resources' blade in the Azure Data Factory portal. On the left, there's a tree view of resources: Pipelines, Datasets, and Data flows. Under 'Pipelines', the item 'ADF-cookbook-pipeline1' is selected and highlighted with a red box. A context menu is open over this item, also highlighted with a red box. The menu includes options: 'Actions' (which is currently selected and highlighted with a red box), 'love & transform', 'Open', 'Monitor', 'Clone', 'Move to', 'Download support files' (which is also highlighted with a red box), and 'Delete'. On the right side of the blade, there are sections for 'SalesOrdersDataset' and 'Activities'.

Copy Data tool

- Properties**
One time copy
- Source**
Azure Blob Storage
 - ✓ Connection
 - ✓ Dataset
- Destination**
Azure Blob Storage
 - ✓ Connection
 - ✓ Dataset
- Settings**
- Summary**
- Deployment**

Azure Blob Storage → Azure Blob Storage

Deployment complete

- Validate copy runtime environment ✓
- Creating datasets ✓
- Creating pipelines ✓
- Running pipelines ✓

Datasets and pipelines have been created. You can now monitor and edit the copy pipelines or click finish to close Copy Data Tool.

[Edit pipeline](#) [Monitor](#)

Data Factory [Publish all 4](#) [Validate all](#) [Refresh](#) [Discard all](#) [Data flow debug](#) [ARM template](#)

Factory Resources

- [Filter resources by name](#)
- Pipelines** 3
 - ADF-cookbook-pipeline1
 - CDP-copy-job
 - DistinctRows**
- Datasets** 6
 - DestinationDataset_0rj
 - folderout**
 - MoviesD2**
 - SalesOrdersDataset
 - SalesOrdersDatasetOutput
 - SourceDataset_0rj
- Data flows** 1
 - DistinctRows**

Template gallery **DistinctRows**

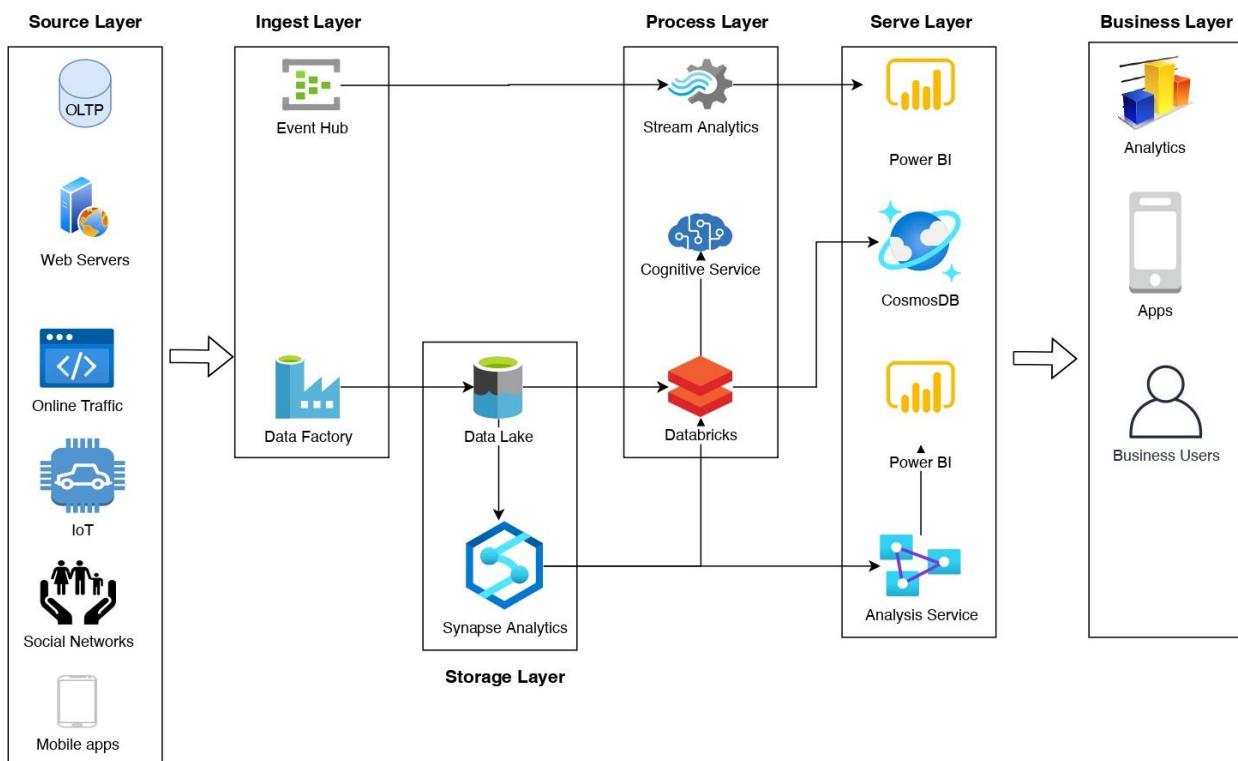
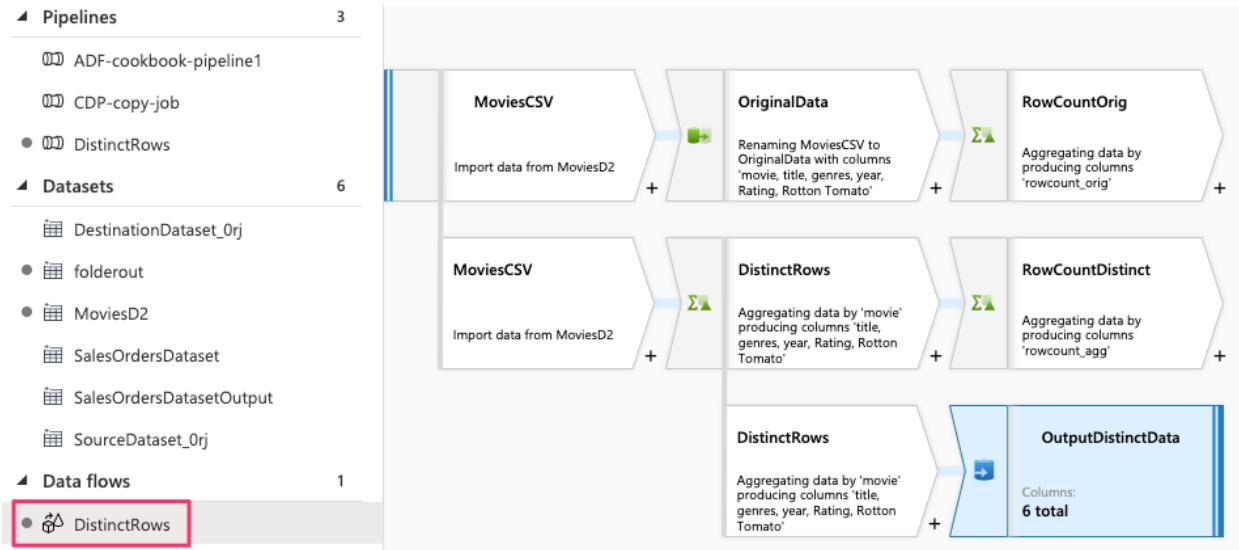
Activities

- [Move & transform](#)
- [Azure Data Explorer](#)
- [Azure Function](#)
- [Batch Service](#)
- [Databricks](#)
- [Data Lake Analytics](#)
- [General](#)
- [HDInsight](#)
- [Iteration & conditionals](#)
- [Machine Learning](#)

Data flow

DistinctRows

[Delete](#) [Edit](#) [Add](#)



Microsoft Azure data

Home > Quickstart Center Microsoft Azure

Get started Take an online course

Start a project

Learn about popular Azure services and create your first Azure project. If you're already familiar with Azure, try a new service below. [Otherwise, see All services.](#)

Create a web app
Build and deploy web apps that can scale

[Start >](#)

Deploy a virtual machine
Run your workloads in the cloud and reduce the redundancy and maintenance of physical hardware

[Start >](#)

Deploy and run a container-based app
Build and run your container-based applications

[Start >](#)

Set up a database
Explore options for managing relational or nonrelational databases in the cloud

[Start >](#)

Start a data analytics project
Put machine learning and artificial intelligence to work on your apps

[Start >](#)

Store, back up, or archive data
Extend data storage to the cloud and leverage it for disaster recovery

[Start >](#)

Build, deploy, and operate a serverless app
Focus on coding within an event-driven architecture, while Azure handles infrastructure-related

[Start >](#)

[All services >](#)

Azure Synapse Analytics (formerly SQL DW) Default Directory

[+ Add](#) [Edit columns](#) [Refresh](#) | [Assign tags](#) [Delete](#)

i Try our new Azure SQL resource browser! This experience offers a unified view of all your SQL Server resources in Azure as well as improved sorting and filtering. Click here to go to the new experience. X

Subscriptions: Free Trial

[Filter by name...](#) [All resource groups](#) [All locations](#) [All tags](#) [No grouping](#)

0 items

Name ↑↓	Status	Replication role	Server	Pricing tier	Location ↑↓	Subscription ↑
---------	--------	------------------	--------	--------------	-------------	----------------



No Synapse SQL pools (data warehouses) to display

Try changing your filters if you don't see what you're looking for.

[Create Synapse SQL pools \(data warehouses\)](#)

All services

 Search Analytics

Overview

ANALYTICS (14)

Categories

All

General

Compute

Networking

Storage

Web

Mobile

Containers

Databases

Analytics

 Azure Synapse Analytics (formerly SQL DW)

 HDInsight clusters

 Power BI Embedded

 Data Lake Analytics

 Event Hubs

 Log Analytics workspaces

 Azure Data Explorer Clusters

 Azure Databricks

 Data factories

 Stream Analytics jobs

 Analysis Services

 Event Hubs Clusters

 Data Lake Storage Gen1

 Power Platform

PREVIEW

Chapter 2: Orchestration and Control Flow

The screenshot shows the Azure Data Factory Studio interface. At the top, there are three tabs: "AzureSQLTables", "CsvData", and "pl_orchestration_reci...". The "AzureSQLTables" tab is active, indicated by a green icon and the text "Azure SQL Database" and "AzureSQLTables". Below the tabs, there are three tabs: "Connection" (selected), "Schema", and "Parameters". Under "Connection", the "Linked service" dropdown is set to "AzureSQLDatabase". To its right are buttons for "Test connection" (with a gear icon), "Open" (with a pencil icon), and "New" (with a plus icon). The "Table" dropdown is set to "dbo" and has a ". @dataset().tableName" placeholder next to it. A "Preview data" button with a magnifying glass icon is to the right. A "Edit" checkbox is checked below the table dropdown. On the left, under "Add dynamic content", there is a text input field containing "@dataset().filename". Below it are "Clear contents" and "Filter..." buttons. A note says "Use expressions, functions or refer to system variables." Under "Parameters", there is a list with "filename" highlighted with a red box. On the right, the "Factory Resources" sidebar shows "Pipelines" (0), "Datasets" (2), and "Data flows" (0). The "AzureSQLTables" dataset is selected, indicated by a blue background.

AzureSQLTables CsvData pl_orchestration_recie... ...

Save as template Validate Debug Add trigger { } ⚙️ ...

Activities « «

Search activities

Move & transform

- Copy data
- Data flow

Azure Data Explorer

Azure Function

Batch Service

Databricks

Data Lake Analytics

General

HDInsight

Iteration & conditionals

Machine Learning

Properties

General

Name *

Description

Concurrency

Annotations + New

Copy data

Copy From Blob to Azure SQL

General **Source** **Sink** **Mapping** **Settings** **User properties**

Name * Learn more ⓘ

Description

Timeout ⓘ

Retry ⓘ

Retry interval ⓘ

Secure output ⓘ

Secure input ⓘ

Parameters Variables Output ▼

Pipeline run ID: 2e44ad24-445e-4256-81f1-daffb8fcc665 [@] ↻ ⓘ View debug run consumption

NAME	TYPE	RUN START	DURATION	STATUS	INTEGRATION RUNTIME
Copy From Blob ↗ ↘ ⏺	Copy	2020-07-18T20:13:17.504	00:00:06	✓ Succeeded	DefaultIntegrationRuntime (West US 2)

Query 1

Run Cancel query Save query Export data as Show only Editor

```
1 select * from [dbo].[Country]
```

Results **Messages**

Search to filter items...

Name	Code
Aruba	AW
Antigua and Barbuda	AG
United Arab Emirates	AE
Afghanistan	AF
Algeria	DZ
Azerbaijan	AZ
Albania	AL
Armenia	AM
Angola	AO
American Samoa	AS
Argentina	AR

Add dynamic content

Clear contents

Filter... +

Use [expressions](#), [functions](#) or refer to [system variables](#).

▲ System variables

- Data factory name
Name of the data factory the pipeline run is running within
- Pipeline Name
Name of the pipeline
- Pipeline run ID
ID of the specific pipeline run
- Pipeline trigger ID
ID of the trigger that invokes the pipeline
- Pipeline trigger name
Name of the trigger that invokes the pipeline
- Pipeline trigger time
Time when the trigger that invoked the pipeline. The trigger time is the actual fired time, ...
- Pipeline trigger type
Type of the trigger that invoked the pipeline (Manual, Scheduler)

▲ Functions

- ▼ [Expand all](#)
- ▷ Collection Functions
- ▷ Conversion Functions
- ▷ Date Functions
- ▷ Logical Functions
- ▷ Math Functions
- ▷ String Functions

The screenshot shows the configuration interface for the 'Get Metadata' action in Microsoft Power Automate. At the top, there's a preview window showing a single step named 'Get Metadata1'. Below this, the main configuration area has three tabs: 'General', 'Dataset' (which is selected), and 'User properties'.

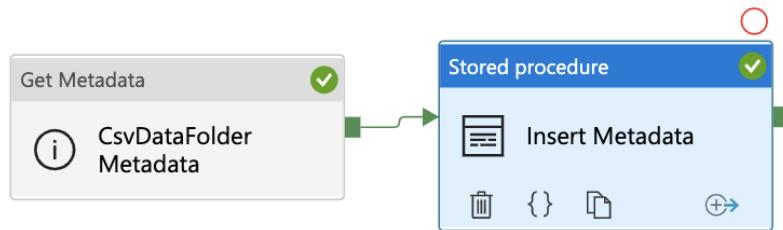
The 'Dataset' tab contains the following fields:

- Dataset ***: A dropdown menu set to 'CsvDataFolder' with options to 'Open' or 'New'.
- Start time (UTC)** and **End time (UTC)**: Two input fields for specifying time ranges.
- Filter by last modified**: An input field for filtering items based on their last modification date.
- Skip line count**: An input field for skipping a specified number of lines at the beginning of the file.

Below these fields is a 'Field list' section with buttons for '+ New' and 'Delete'.

Underneath the 'Field list' is a section titled 'ARGUMENT' with a checkbox. A dropdown menu is open, showing the following filter options:

- Item name
- Filter...
- Add dynamic content
- Child Items
- Exists
- Item name
- Item type
- Last modified



General Settings User properties

Linked service * ⓘ Test connection Open + New

Stored procedure name * ⓘ Edit ⓘ

▲ Stored procedure parameters ⓘ

← Import + New | Delete

<input type="checkbox"/>	NAME	TYPE	VALUE
	FileName	String	<code>@activity('CsvDataFolder Metadata').output.itemName</code>
	ModifiedAt	DateTime	<code>@convertFromUtc(activity('CsvDataFolder Metadata').output.lastModified, 'Pacific Standard Time')</code>
	UpdatedAt	DateTime	<code>@convertFromUtc(utcnow(), 'Pacific Standard Time')</code>

Metadata options

You can specify the following metadata types in the Get Metadata activity field list to retrieve the corresponding information:

Metadata type	Description
itemName	Name of the file or folder.
itemType	Type of the file or folder. Returned value is <code>File</code> or <code>Folder</code> .
size	Size of the file, in bytes. Applicable only to files.
created	Created datetime of the file or folder.
lastModified	Last modified datetime of the file or folder.
childItems	List of subfolders and files in the given folder. Applicable only to folders. Returned value is a list of the name and type of each child item.
contentMD5	MD5 of the file. Applicable only to files.
structure	Data structure of the file or relational database table. Returned value is a list of column names and column types.
columnCount	Number of columns in the file or relational table.
exists	Whether a file, folder, or table exists. Note that if <code>exists</code> is specified in the Get Metadata field list, the activity won't fail even if the file, folder, or table doesn't exist. Instead, <code>exists: false</code> is returned in the output.

General **Settings** User properties

Items `@activity('Get Metadata').output.childItems`

Condition `@endswith(item().name, '.csv')`

Parameters Variables Output

Pipeline run ID: `67b07a1e-f6ee-442f-95d8-04e9e8eaeae8`   

[View debug run consumption](#)

NAME	TYPE	RUN START	DURATION	STATUS	INTEGRATION RUNTIME	RUN ID
FilterOnCSV	Filter	2020-07-18T21:53:15.6354968Z	00:00:01	Succeeded	Unknown	2e4cf683-a620-4e2c-ab35-5776ce154...
Get Metadata	 GetMetadata	2020-07-18T21:53:12.3914818Z	00:00:03	Succeeded	DefaultIntegrationRuntime (West US)	5bc0a693-94c9-4221-a34b-cd74921e2...

Output

```
{  
    "itemName": "data",  
    "lastModified": "2020-07-11T02:36:47Z",  
    "childItems": [  
        {  
            "name": "Airlines-2020-07-18T20:30:47.2924367Z.backup",  
            "type": "File"  
        },  
        {  
            "name": "Country-2020-07-18T21:51:23.0941216Z.backup",  
            "type": "File"  
        },  
        {  
            "name": "airlines.csv",  
            "type": "File"  
        },  
        {  
            "name": "countries.csv",  
            "type": "File"  
        }  
    ],  
    "effectiveIntegrationRuntime": "DefaultIntegrationRuntime (West US)",  
    "executionDuration": 0,  
    "durationInQueue": {  
        "integrationRuntimeQueue": 0  
    },  
    "billingReference": {  
        "activityType": "PipelineActivity",  
        "billableDuration": [  
            {  
                "meterType": "AzureIR",  
                "duration": 0.016666666666666666,  
                "unit": "Hours"  
            }  
        ]  
    }  
}
```

General Dataset User properties

Dataset *

 CsvData

 Open  New

► Dataset properties 

NAME	VALUE	TYPE
filename	@item().name	string

Start time (UTC)

End time (UTC)



Filter by last modified

Skip line count

Field list

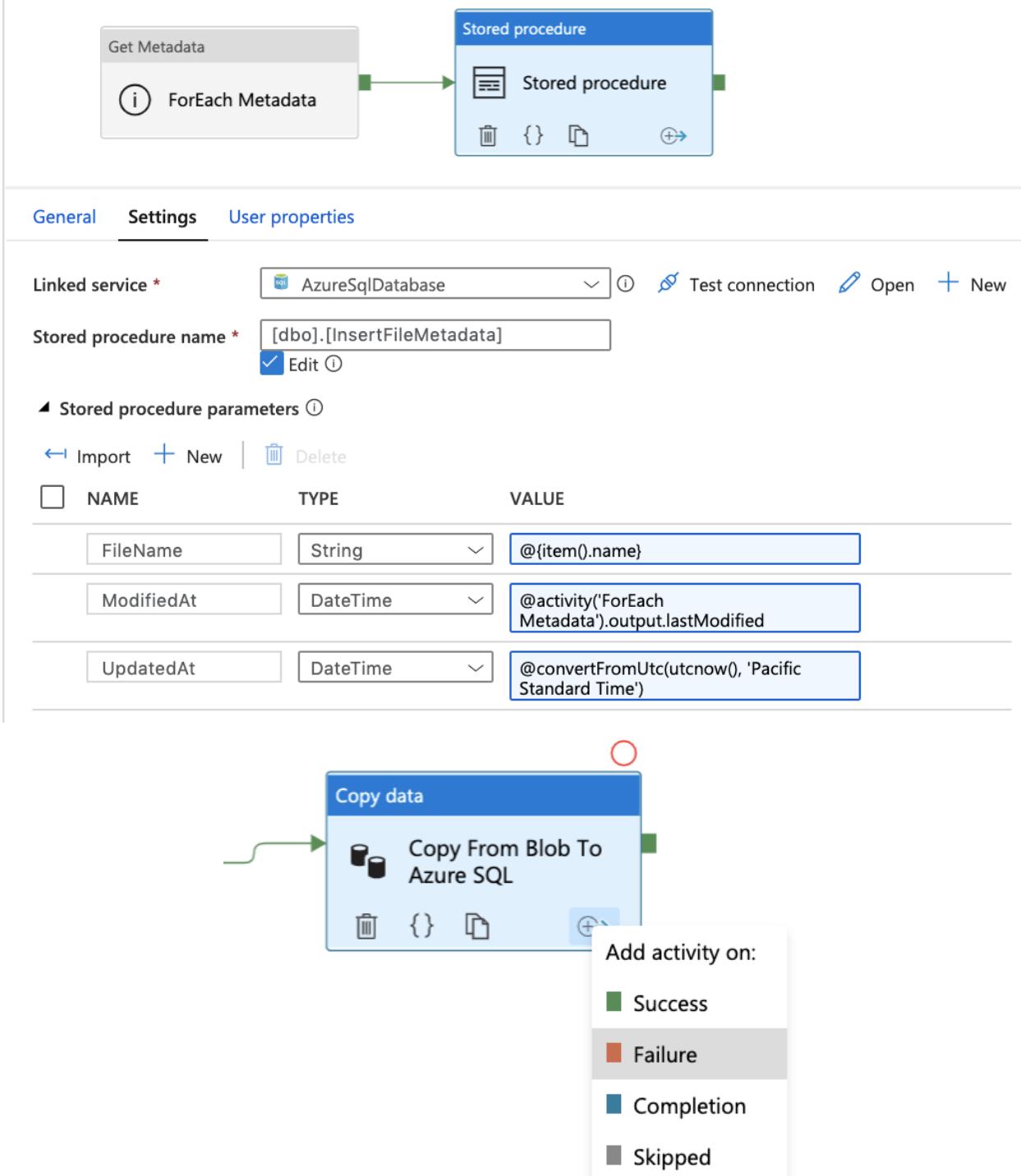
 New  Delete

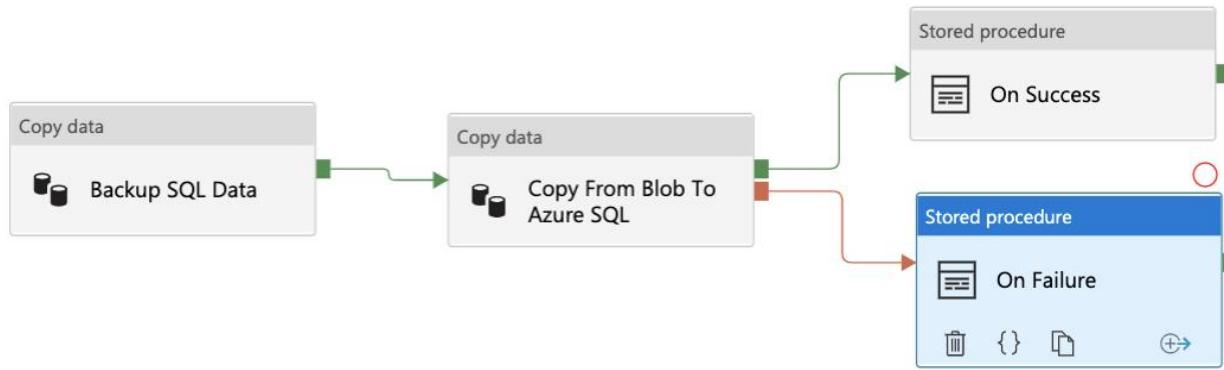
ARGUMENT

Item name

Last modified

pl_orchestration_recipe_3_clone > ForEach

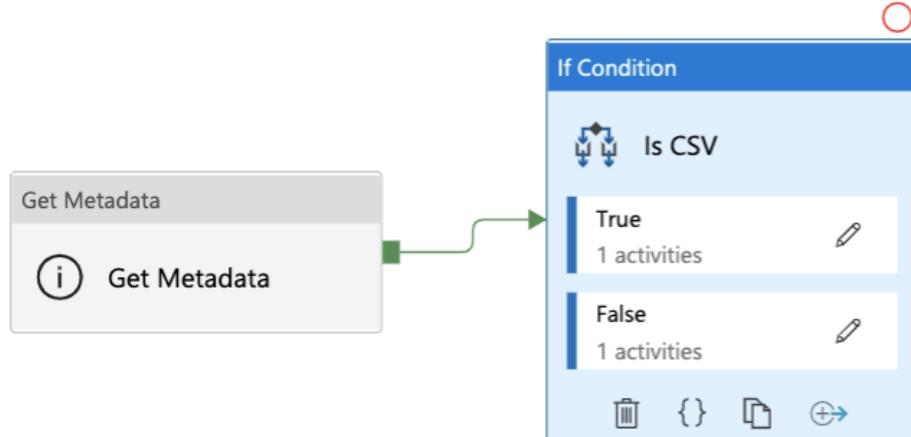




Results Messages

Search to filter items...

PipelineID	RunID	Status	UpdatedAt
pl_orchestration_recipe_4	cf1846ef-ee2f-4450-92bf-06f67b4b9854	Success	2020-07-20T00:24:47.5770000
pl_orchestration_recipe_4	5e9783c3-f8d8-40dd-bf58-31d2ff8d712f	Failure	2020-07-20T00:26:54.3700000



General Activities (2) User properties

Expression `@not(endswith(activity('Get Metadata').output.itemName, 'csv'))`

CASE	ACTIVITY
True	1 Activity
False	1 Activity



When a HTTP request is received

HTTP POST URL URL will be generated after save (i) ...

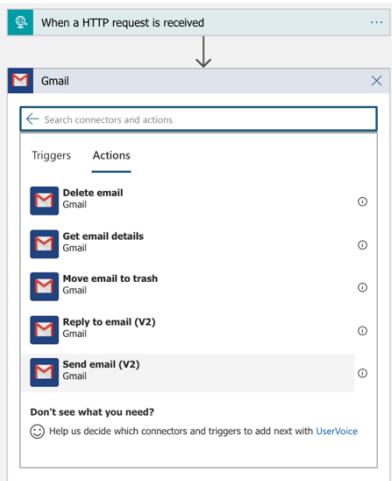
Request Body JSON Schema

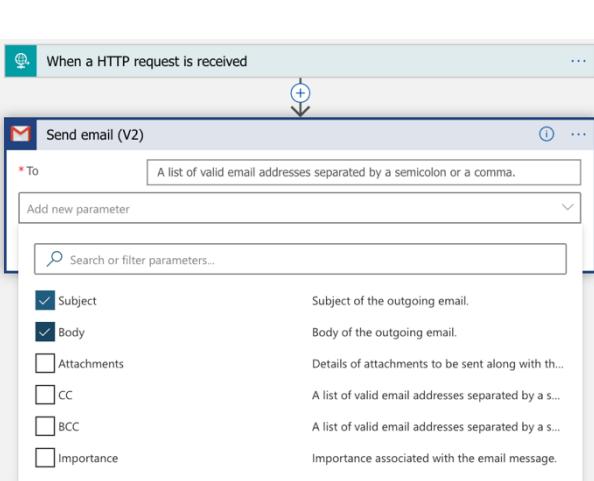
```
{
  "type": "object",
  "properties": {
    "subject": {
      "type": "string"
    },
    "messageBody": {
      "type": "string"
    }
  }
}
```

[Use sample payload to generate schema](#)

Add new parameter ▼

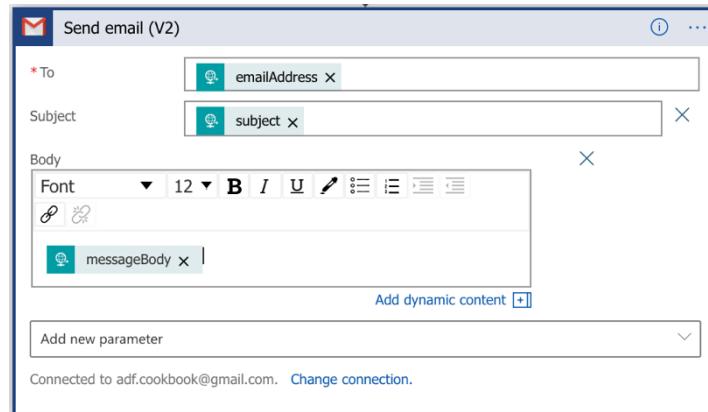
[+ New step](#)





The 'Send email (V2)' action configuration screen shows the following fields:

- To:** A list of valid email addresses separated by a semicolon or a comma.
- Add new parameter**: A dropdown menu.
- Search or filter parameters...**: A search bar.
- Subject**: Subject of the outgoing email.
- Body**: Body of the outgoing email.
- Attachments**: Details of attachments to be sent along with the email.
- CC**: A list of valid email addresses separated by a semicolon or a comma.
- BCC**: A list of valid email addresses separated by a semicolon or a comma.
- Importance**: Importance associated with the email message.



○

Lookup

Get Email Recipients

{} ↻ +→

General Settings User properties

Source dataset * Open New Preview data

Dataset properties

NAME	VALUE	TYPE
tableName	EmailRecipients	string

Use query Table Query Stored procedure

Query *

Query timeout (minutes)

Isolation level

First row only

Web 

Send Email

General Settings User properties

URL *

Method *

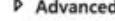
Headers *  

<input type="checkbox"/> NAME	VALUE
<input type="text" value="Content-Type"/>	<input type="text" value="application/json"/>

Body

Datasets 

Linked services 

Integration runtime * ⓘ 

ADF-Email-LogicApp

Logic app

Search (Cmd+ /)

Run Trigger Refresh Edit Delete Disable Update Schema Clone Export

Changes due to Google's data security and privacy policies now limit Gmail connector communication to only Google-approved apps and services. Non-compliant logic apps will be disabled. Click to learn how to ensure compliance. →

Resource group (change) : ADF-Orchestration-RG Definition : 1 trigger, 1 action
Location : West US Status : Enabled
Subscription (change) : Azure subscription 1 Runs last 24 hours : 6 successful, 0 failed
Subscription ID : e1c43362-444d-41ca-a320-85f93e5b2aa6 Integration Account : ---

Summary

Trigger Actions

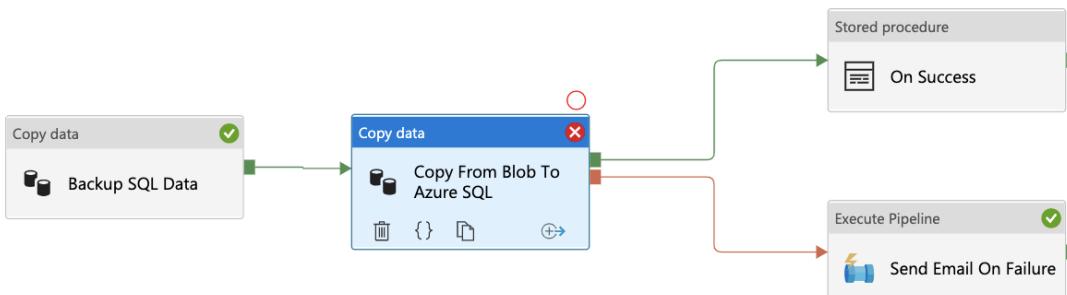
REQUEST
manual COUNT
1 action
[View in Logic Apps designer](#)

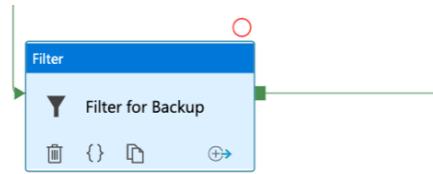
FREQUENCY

EVALUATION
Evaluated 6 times, fired 6 times in the last 24 hours
[See trigger history](#)

Runs history

Status	Start time	Identifier	Duration	Static Results
Succeeded	7/18/2020, 5:11 PM	08586064894049081890770626936CU35	31.56 Seconds	
Succeeded	7/18/2020, 5:11 PM	08586064894151596368779224498CU06	32.14 Seconds	
Succeeded	7/18/2020, 4:22 PM	08586064923220639686264962527CU06	998 Milliseconds	
Succeeded	7/18/2020, 4:22 PM	08586064923314914328149457756CU12	1.42 Seconds	
Succeeded	7/17/2020, 6:07 PM	08586065724566063492715210789CU100	1.83 Seconds	
Succeeded	7/17/2020, 6:07 PM	085860657245672546228462414676CU14	1.88 Seconds	

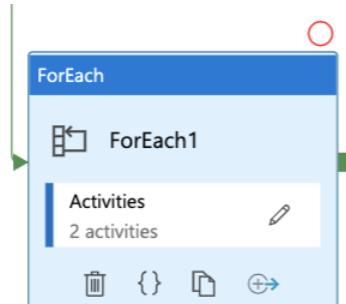




General Settings User properties

Items @activity('Get Metadata').output

Condition @endswith(item().name, 'backup')

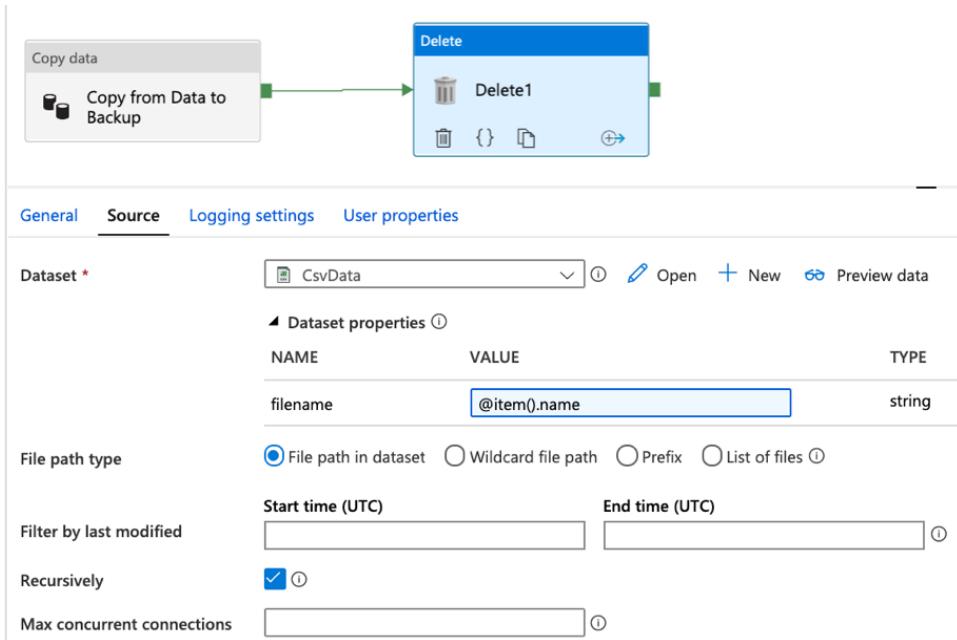


General Settings Activities (2) User properties

Sequential

Batch count ⓘ

Items @activity('Filter for Backup').output.Value



New trigger

Name *
trigger_blob_added

Description

Type *
 Schedule Tumbling window Event

Account selection method *
 From Azure subscription Enter manually

Azure subscription
Azure subscription 1

Storage account name *
adforchestrationstorage

Container name *
data

Blob path begins with

Blob path ends with

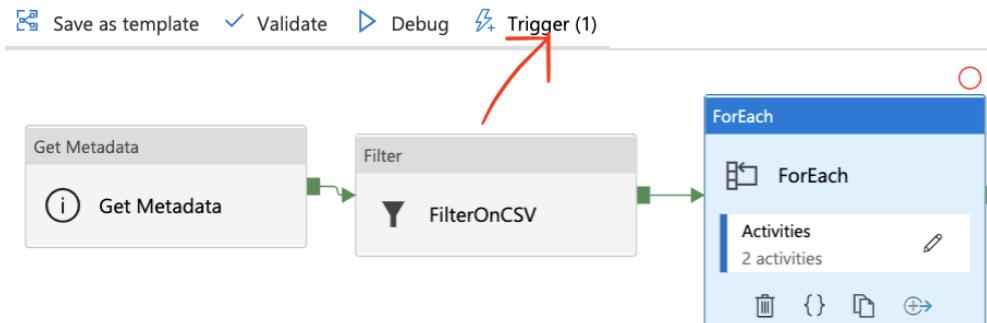
Event *
 Blob created Blob deleted

Ignore empty blobs *
 Yes No

Annotations
+ New

Activated *
 Yes No

Continue **Cancel**



adf-orchestration-sql-server | Firewalls and virtual networks



SQL server

<<

[Save](#)[Discard](#)[+ Add client IP](#)[Quick start](#)[Failover groups](#)[Manage Backups](#)[Active Directory admin](#)[SQL databases](#)[SQL elastic pools](#)[Deleted databases](#)[Import/Export history](#)[DTU quota](#)[Properties](#)[Locks](#)

Security

[Auditing](#)[Firewalls and virtual networks](#)[Private endpoint connections](#)[Security Center](#)[Transparent data encryption](#)

Intelligent Performance

[Automatic tuning](#)[Recommendations](#)Deny public network access [\(i\)](#)[Yes](#) [No](#)[Click here to create a new private endpoint.](#)
[Create Private Endpoint](#)Minimum TLS Version [\(i\)](#)[> 1.0](#) [> 1.1](#) [> 1.2](#)Connection Policy [\(i\)](#)[Default](#) [Proxy](#) [Redirect](#)Allow Azure services and resources to access this server [\(i\)](#)[Yes](#) [No](#)

Client IP address

Rule name	Start IP	End IP	...
ClientIPAddress_2020-7-10...	172.16.0.1	172.16.0.255	...

Virtual networks

[+ Add existing virtual network](#) [+ Create new virtual network](#)

Rule name	Virtual network	Subnet
No vnet rules for this server.		

New linked service

Data store Compute

Search

All Azure Database File Generic protocol NoSQL Services and apps

 Amazon Marketplace Web Service	 Amazon Redshift	 Amazon S3
 Apache Impala	 Azure Blob Storage	 Azure Cosmos DB (MongoDB API)
		

Continue Cancel

New linked service (Azure Blob Storage)

Name *

OrchestrationAzureBlobStorage1

Description

Connect via integration runtime *



AutoResolveIntegrationRuntime



Authentication method

Account key



Connection string

Azure Key Vault

Account selection method



From Azure subscription Enter manually

Azure subscription



Azure subscription 1

Storage account name *

adforchestrationstorage



Additional connection properties

+ New

Test connection



To linked service To file path

Annotations

+ New

Create

Back



Test connection

Cancel

New linked service (Azure SQL Database)

Name *

Description

Connect via integration runtime *



Account selection method



From Azure subscription Enter manually

Azure subscription

Server name *

Database name *



Authentication type *

User name *

Password *



Additional connection properties

+ New

Connection successful



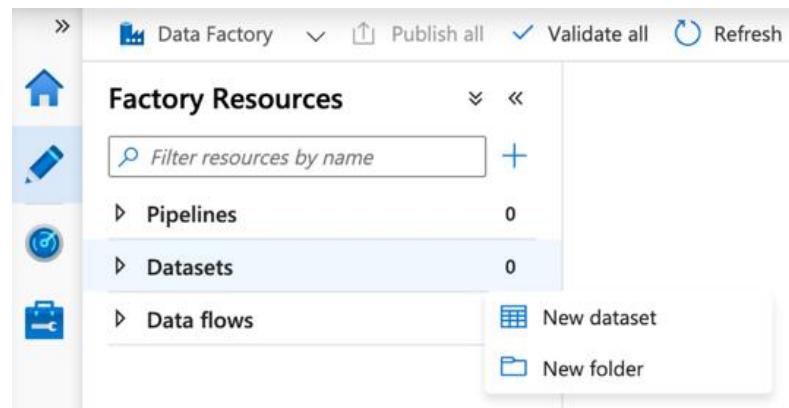
» Data Factory Publish all Validate all Refresh

Factory Resources

Filter resources by name +

- Pipelines 0
- Datasets 0
- Data flows 0

New dataset New folder



Set properties

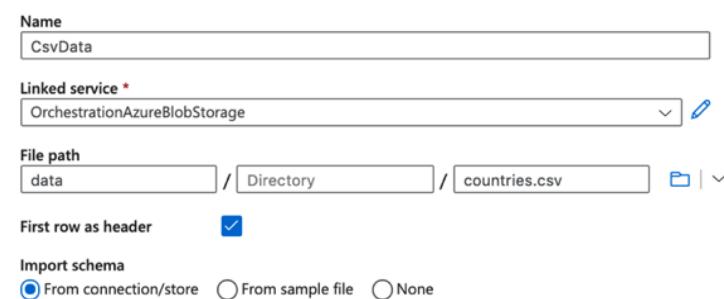
Name

Linked service * edit

File path / / edit

First row as header

Import schema From connection/store From sample file None



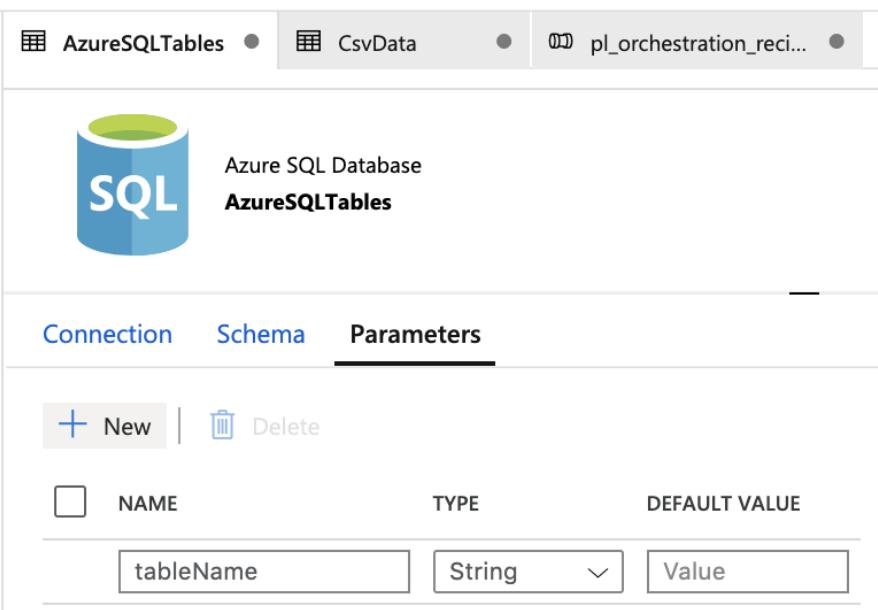
AzureSQLTables CsvData pl_orchestration_rec...

 Azure SQL Database
AzureSQLTables

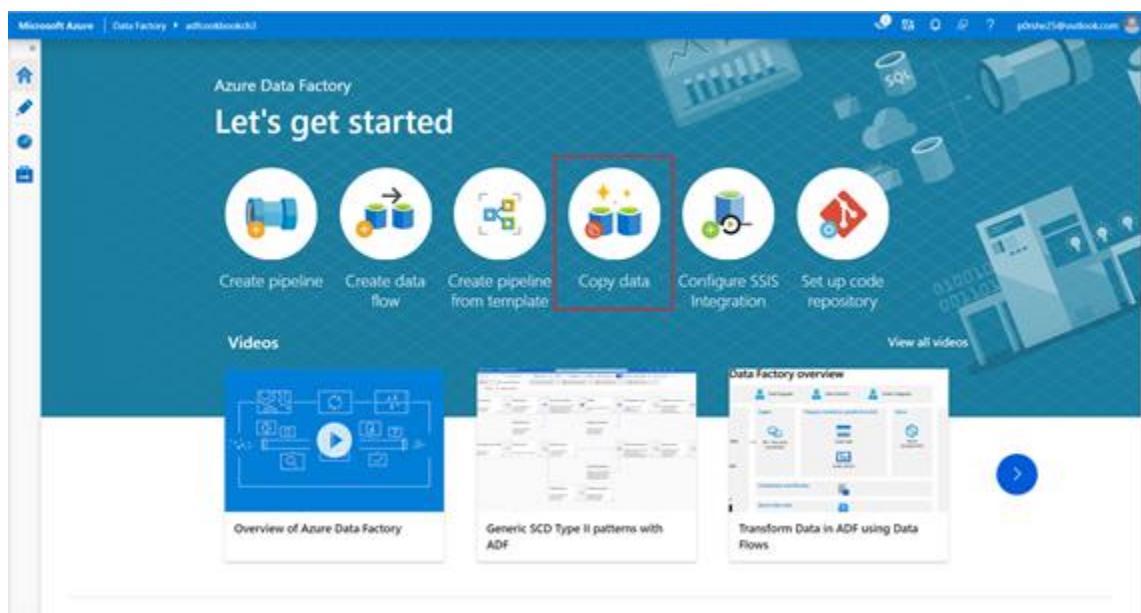
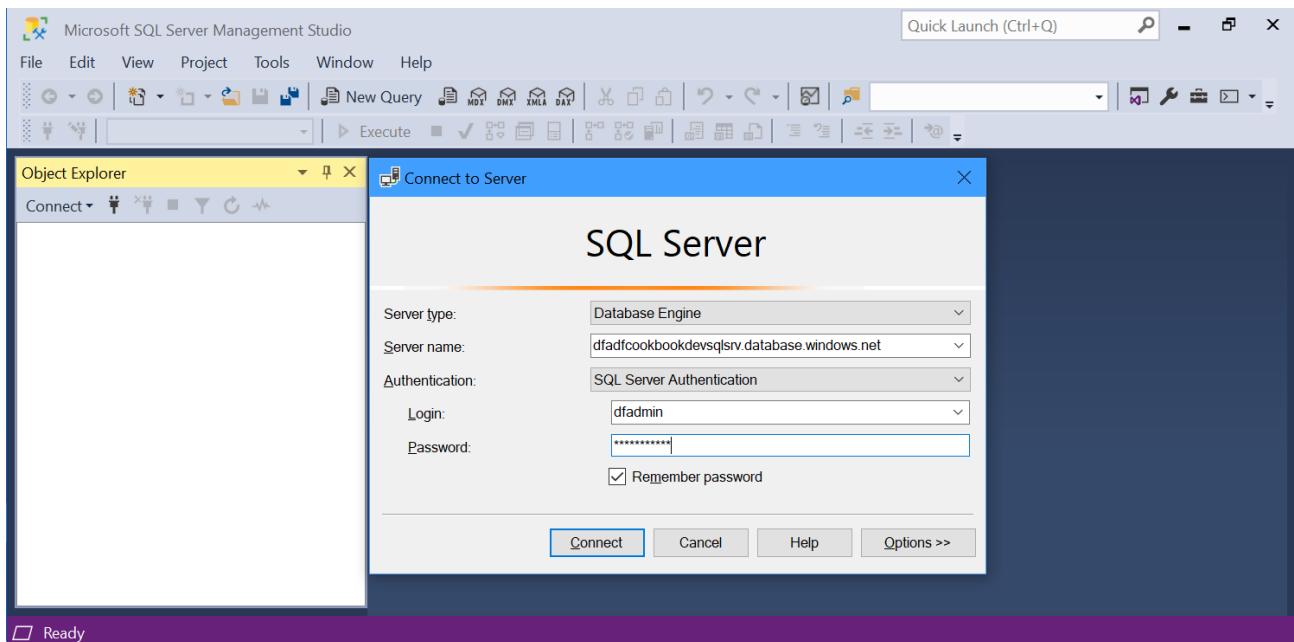
Connection Schema Parameters

+ New | Delete

<input type="checkbox"/>	NAME	TYPE	DEFAULT VALUE
<input type="checkbox"/>	tableName	String	Value



Chapter 3: Setting Up a Cloud Data Warehouse



Copy data

1 Properties
One time copy

2 Source

- Connection
- Dataset

3 Destination

- Connection
- Dataset

4 Settings

5 Summary

6 Deployment

Choose the input file or folder
Select a source file or folder to be copied to the destination data store.

File or folder *

If the identity you use to access the data store only has permission to subdirectory instead of the entire account, specify the path to browse.

Binary copy

Recursively

Max concurrent connections

Start time (UTC)

End time (UTC)

Filter by last modified

Copy data

1 Properties
One time copy

2 Source

- Connection
- Dataset

3 Destination

- Connection
- Dataset

4 Settings

5 Summary

6 Deployment

File format settings

File format

Column delimiter

Row delimiter

Name	IATA code	ICAO code
Aerospatiale (Nord) 262	ND2	N262
Aerospatiale (Sud Aviation) Se.210 Caravelle	CRV	S210

Microsoft Azure | adfcookbookch3adf

Copy data

Properties
One time copy

Source
Azure Blob Storage

Connection

Dataset

Destination

Connection

Dataset

Settings

Summary

Deployment

Destination data store
Specify the destination data store

All Azure Database

All

AzureBlobStorage

Connection string Azure Key Vault

Account selection method
 From Azure subscription Enter manually

Fully qualified domain name * dfadfcbookdevsq|srv.database.windows.net

Database name * adfcookbookch1devsqldb

Authentication type * SQL authentication

User name * adfuser

Password Azure Key Vault

Password *

Connection successful

Test connection Cancel

Previous Next

Apply

SQLQuery2.sql - dfadfcbookdevsq|srv.database.windows.net.adfcookbookch1devsqldb (adfuser (122)) - Microsoft SQL Server Management Studio Quick Launch

File Edit View Project Tools Window Help

Object Explorer

Connect Databases adfcookbookch1devsqldb Tables External Tables dbo.Airlines dbo.Countries dbo.Planes

SQLQuery2.sql - dfa...ldb (adfuser (122)) SQLQuery1.sql - dfa...ldb (adfuser (113))

SELECT TOP 100 * FROM Planes

100 %

Results Messages

	Name	IATA_code	ICAO_code
1	Aerospatiale (Nord) 262	ND2	N262
2	Aerospatiale (Sud Aviation) Se.210 Caravelle	CRV	S210
3	Aerospatiale SN 601 Corvette	NDC	S601
4	Aerospatiale/Alenia ATR 42-300	AT4	AT43
5	Aerospatiale/Alenia ATR 42-500	AT5	AT45

Copy data

Properties
One time copy

Source
Azure Blob Storage

Connection

Dataset

Destination
Azure Synapse Analytics (formerly SQL DW)

Connection

Dataset

Settings

Summary

Deployment

Settings
More options for data movement

Performance settings

Enable staging

Staging settings

Staging account linked service AzureBlobStorageLS Connection successful

Storage Path flightsblobstorage

Enable Compression

Advanced settings

Copy method PolyBase Copy command (Preview) Bulk insert

Allow PolyBase

Reject type Value

Previous **Next**

Microsoft Azure | adfcookbookch3adf | Search resources

Data Factory | Publish all 1 | Validate all | Refresh | Discard all | Data flow debug | ARM template | ...

Factory Resources | AutoStartDWpipeline

Pipelines: 2
 ADFcookbookCopyToSynapse
 AutoStartDWpipeline

Datasets: 2

Data flows: 0

Activities: web

General: Web, WebHook

Properties

General

Name * startDW

Description

Concurrency

Annotations

Timeout 7.00:00:00

Microsoft Azure Upgrade Search resources, services, and docs (G+) ⚡ ? p0rshe25@yandex.ru DEFAULT DIRECTORY

Home > adfcookbookch1devsql (dfadfcbookdevsqlsrv/adfcookbookch1devsql)

Synapse SQL pool (data warehouse)

Search (Ctrl+ /) Pause Scale Restore + New restore point Delete

Overview Activity log Tags Diagnose and solve problems Workload management (previ... Settings Maintenance schedule Quick start Geo-backup policy Connection strings Properties Locks Export template

Azure Synapse Analytics brings new workload management capabilities. Click here to learn more.

Welcome to Azure Synapse Analytics (formerly known as Azure SQL Data Warehouse). Learn more.

Resource group (change)
adf-cookbook-ch1-rg

Status Online
Location East US
Subscription (change)
Azure subscription 1
Subscription ID 69efd1df-7ca2-48e6-ba8a-3057c8d3536f
Tags (change)
Click here to add tags

Server name dfadfcbookdevsqlsrv.database.windows.net
Connection strings Show database connection strings
Performance level Gen2: DW100
Maintenance schedule Sun 20:00 UTC (4h) / Wed 12:00 UTC (8h)
Geo-backup policy Enabled

<https://portal.azure.com/#>

Microsoft Azure | adfcookbookch3adf

Search resources Publish all 1 Validate all Refresh Discard all Data flow debug ARM template

Factory Resources Pipelines Datasets Data flows

Filter resources by name

AutoStartDWpipeline

Save as template Validate Debug Add trigger

General Settings User properties

Method * POST

Headers + New

Body {}

Add dataset reference Add linked service reference

Integration runtime * AutoResolveIntegrationRuntime

Advanced

Authentication None Basic MSI ClientCertificate

Resource * https://management.core.windows.net

Properties General Name * AutoStartDWpipeline Description Concurrency Annotations Name

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

Home > dfadfcbookdevsqlsr | Access control (IAM)

SQL server

Search (Ctrl+)

Add Download role assignments Edit columns Refresh

Check access Role assignments Roles Deny assignments Classic

Check access Review the level of access a user, group, service principal, or managed identity has to this resource. Learn more

Find Azure AD user, group, or service principal

Search by name or email address

Add a ro Grant access role to a user identity.

Select Search by name

No results to display.

Add

View rol View the identities access at

View

Selected members:

adfcbookch3adf /subscriptions/69ef1df-7ca2-48e6-b... Remove

Save Discard

This screenshot shows the Microsoft Azure Access control (IAM) interface for a specific resource named 'dfadfcbookdevsqlsr'. The left sidebar includes links for Overview, Activity log, Tags, Diagnose and solve problems, Settings (Quick start, Failover groups, Manage Backups, Active Directory admin, SQL databases, SQL elastic pools, Deleted databases), and Help. The main area displays the 'Check access' tab, which allows users to search for Azure AD users, groups, or service principals. A modal window titled 'Add role assignment' is open, showing the 'Contributor' role assigned to the 'Data Factory' under the 'Subscription' 'Azure subscription 1'. The 'Selected members' section lists the 'adfcbookch3adf' service principal. The bottom right of the modal contains 'Save' and 'Discard' buttons.

Microsoft Azure | adfcbookch3adf

Search resources

Data Factory Publish all Validate all Refresh Discard all Data flow debug ARM template

Factory Resources

Pipelines: ADFcookbookCopyToSynapse, AutoStartDW, AutoPauseDW

Datasets: 2

Data flows: 0

Activities

Move & transform, Azure Data Explorer, Azure Function, Batch Service, Databricks, Data Lake Analytics

General, HDInsight, Iteration & conditionals, Machine Learning

Web: PauseDW

Save as template, Validate, Debug, Add trigger

Properties

General

Name: AutoPauseDW

Description

Concurrency

Annotations

Body

This screenshot shows the Microsoft Azure Data Factory interface. The left sidebar lists 'Factory Resources' with sections for Pipelines (containing 'ADFCookbookCopyToSynapse', 'AutoStartDW', and 'AutoPauseDW'), Datasets (2 items), and Data flows (0 items). The main area shows the 'Activities' section with options like Move & transform, Azure Data Explorer, Azure Function, etc. A specific 'Web' activity named 'PauseDW' is selected, shown in a detailed view. The 'Properties' panel on the right shows the 'General' tab with fields for Name (set to 'AutoPauseDW'), Description, Concurrency, Annotations, and Body (empty). Other tabs include 'Settings' and 'User properties'.

Microsoft Azure [Upgrade](#) [Search resources, services, and docs \(G+/\)](#) [p0rshe25@yandex.ru](#) [DEFAULT DIRECTORY](#)

Home > New > Azure Synapse Analytics (workspaces preview) >

Create Synapse workspace

Project details

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

Subscription * [Azure subscription 1](#)
The Synapse resource provider needs to be registered to this subscription.
[Click here to register.](#)

Resource group * [adf-cookbook-ch1-rg](#)
[Create new](#)

Workspace details

Name your workspace, select a location, and choose a primary Data Lake Storage Gen2 file system to serve as the default location for logs and job output.

Workspace name *

Region * [East US](#)

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

Microsoft Azure [Upgrade](#) [Search resources, services, and docs \(G+/\)](#) [p0rshe25@yandex.ru](#) [DEFAULT DIRECTORY](#)

Home >

Create Synapse workspace

Workspace details

Name your workspace, select a location, and choose a primary Data Lake Storage Gen2 file system to serve as the default location for logs and job output.

Workspace name *

Region * [East US](#)

Select Data Lake Storage Gen2 * From subscription Manually via URL

Account name * [\(New\) adfcookbookch3adls](#)
[Create new](#)

File system name * [\(New\) adfcookbookch3fs](#)
[Create new](#)
 Assign myself the Storage Blob Data Contributor role on the Data Lake Storage Gen2 account 'adfcookbookch3adls'.

We will automatically grant the workspace identity data access to the specified Data Lake Storage Gen2 account, using the [Storage Blob Data Contributor](#) role.

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

Microsoft Azure [Upgrade](#) p0rshe25@yandex.ru DEFAULT DIRECTORY

Home > Create Synapse workspace

* Basics * **Security + networking** Tags Summary

Configure security options and networking settings for your workspace.

SQL administrator credentials

Provide credentials that can be used for administrator access to the workspace's SQL pools. If you don't provide a password, one will be automatically generated. You can change the password later.

Admin username *

Password

Confirm password

Workspace managed identity

Choose whether you'd like your workspace's managed identity to be automatically granted CONTROL permission (inherits all permissions). This applies when using SQL pools and SQL on-demand within the current workspace.

Grant CONTROL to the workspace's managed identity on all SQL pools and SQL on-demand.

Review + create < Previous Next: Tags >

Microsoft Azure [Upgrade](#) p0rshe25@yandex.ru DEFAULT DIRECTORY

Home > Microsoft.Azure.SynapseAnalytics-20200802073933 | Overview

Deployment

Search (Ctrl+/) Delete Cancel Redeploy Refresh

We'd love your feedback!

Your deployment is complete

Deployment name: Microsoft.Azure.SynapseAnalytics-20200802073... Start time: 8/2/2020, 7:52:07 AM
Subscription: Azure subscription 1 Correlation ID: 036ec35c-c291-4c53-a523-4958de74ed3
Resource group: adf-cookbook-ch1-rg

Deployment details (Download) Next steps

Go to resource

7:56 AM Deployment succeeded Deployment 'Microsoft.Azure.SynapseAnalytics-20200802073933' to resource group 'adf-cookbook-ch1-rg' was successful.

Go to resource Pin to dashboard

Security Center Secure your apps [Go to Azure security](#)

Free Microsoft tools [Start learning today](#)

Work with an expert Azure experts are who can help make and be your first [Find an Azure expert](#)

Microsoft Azure [Upgrade](#)       ? p0rshe25@yandex.ru  DEFAULT DIRECTORY

Home > Microsoft.Azure.SynapseAnalytics-20200802073933 | Overview >

adfcookbooksynapse | SQL pools

Synapse workspace

Search (Ctrl+/
New Refresh Assign tags Delete

Overview Activity log Access control (IAM) Tags

Settings

SQL Active Directory admin Properties Locks

Synapse resources

SQL pools Apache Spark pools

Security

Microsoft Azure [Upgrade](#)      ? p0rshe25@yandex.ru  DEFAULT DIRECTORY

Home >

adfcookbookch1devsql (dfadfcbookdevsqlsr/adfcookbookch1devsql)

Synapse SQL pool (data warehouse)

Search (Ctrl+/
Pause Scale Restore New restore point Delete

Overview Activity log Tags Diagnose and solve problems Workload management (previ...
Maintenance schedule Quick start Geo-backup policy Connection strings Properties Locks Export template

Azure Synapse Analytics brings new workload management capabilities. Click here to learn more. →
Welcome to Azure Synapse Analytics (formerly known as Azure SQL Data Warehouse). Learn more.

Resource group (change)
adf-cookbook-ch1-rg
Status
Online
Location
East US
Subscription (change)
Azure subscription 1
Subscription ID
69ef1df-7ca2-48e6-ba8a-3057c8d3536f
Tags (change)
Click here to add tags

Server name
dfadfcbookdevsqlsr.
Connection strings
Show database connectio...
Performance level
Gen2: DW100c
Maintenance schedule
Sun 20:00 UTC (4h) / Wed
Geo-backup policy
Enabled

Notifications (1) Features (4) Tasks (6)

Microsoft Azure Upgrade Search resources, services, and docs (G+) ? p0rshe25@yandex.ru DEFAULT DIRECTORY

Home > adfcookbooksynapse | SQL pools >

Create SQL pool

Synapse

* Basics * Additional settings Tags Review + create

Customize additional configuration parameters including collation & data source.

Data source

Start with a blank SQL pool, restore from a backup or leverage a restore point to populate your new SQL pool.

Use existing data *

None Backup Restore point

i Leverage a restore point from any SQL pool in this subscription to recover or copy from a previous state. [Learn more](#)

Server or Workspace * ① dfadfcookbookdevsqlsr (East US)

SQL pool name * ① adfcookbookch1devsql1db

Restore point type * ①

Automatic restore points User-defined restore points

Restore point * synapserestorepoint (2020-08-02T05:19:54Z)

Review + create < Previous Next : Tags >

Microsoft Azure Upgrade Search resources, services, and docs (G+) ? p0rshe25@yandex.ru DEFAULT DIRECTORY

Home >

adfcookbooksynapse | SQL pools

Synapse workspace

Name	Type	Status	Size	...
SQL on-demand	SQL on-demand	N/A	Auto	...
adfcookbooksql	SQL pool	Online	DW100c	...

Settings

- SQL Active Directory admin
- Properties
- Locks

Synapse resources

- SQL pools Selected
- Apache Spark pools

Security

- Firewalls
- Managed identities
- Private endpoint connections ...

Monitoring

Microsoft Azure | Synapse Analytics > adfcookbooksynapse

Search resources

Home Data Develop Orchestrate Monitor Manage

Data Workspace Linked

Databases 1

adfcookbooksql (SQL pool)

Tables External tables External resources Views Programmability Schemas Security

SQL script 1

Run Undo Publish Query plan Connect to adfcookbooksql

1 CREATE TABLE [dbo].[Table]
2 (coll int NOT NULL
3)
4 WITH
5 (DISTROBUTION = HASH (coll),
6)
7

New SQL script Refresh

Microsoft Azure | Synapse Analytics > adfcookbooksynapse

Search resources

Home Data Develop Orchestrate Monitor Manage

Data Workspace Linked

Databases 1

adfcookbooksql (SQL pool)

Tables

dbo.Airlines dbo.Countries dbo.Planes dbo.Routes

External tables External resources Views Programmability Schemas Security

SQL script 1

Run Undo Publish Query plan Connect to adfcookbooksql

1 CREATE TABLE [dbo].[Routes]
2 (Airline VARCHAR(10) NOT NULL,
3 Airline_ID VARCHAR(20) NULL,
4 Source_airport VARCHAR(20) NOT NULL,
5 Source_airport_ID VARCHAR(50) NULL,
6 Destination_airport VARCHAR(20) NOT NULL,
7 .nation_airport_ID VARCHAR(50) NULL,
8 share VARCHAR(10) NULL,
9 .int NOT NULL,
10 equipment VARCHAR(50) NULL
11)
12)
13 WITH
14 (

Results Messages

11:41:42 PM Started executing query at Line 1

Microsoft Azure | Synapse Analytics > adfcookbooksynapse

Search resources

flightscontainer

Data

Storage accounts

adfcookbooksynapse (Primary - a...
adfcookbookch3fs (Primary)
flightscontainer

Datasets

routes.csv

NAME LAST MODIFIED CONTENT TYPE

airlines.csv 8/2/2020, 12:32:16 PM

countries.csv 8/2/2020, 12:32:30 PM

data_4e7712aa-7db0-4180-a6d1-8f05dc14c5ff_3b51df33-ec55-4aec-b58e-1c42f41934d3.pa... 8/3/2020, 12:43:39 PM

planes.csv 8/2/2020, 12:32:41 PM

routes.csv preview

New SQL script Bulk load

New notebook

New data flow

New dataset

Copy ABFSS path

Manage access...

Rename...

Download

Delete

Showing 1 to 3 of 3 items

<https://web.azuresynapse.net/assets/external/storage/Scripts/0.2.7/StorageExplorer/FrameIndex.html?&enabledFeature=D&id=a365TabContentDiv95dd0282-47ba-4b6d-8cd8-9cc17a5a1f1e&implPath=%2Fassets%2Fexterna...>

Microsoft Azure | Synapse Analytics > adfcookbooksynapse

Search resources

flightscontainer

Data

Storage accounts

adfcookbooksynapse (Primary - a...
adfcookbookch3fs (Primary)
flightscontainer

Datasets

routes.csv

Bulk load

Source storage location

Select the storage location where the files containing the data is staged. Azure Data Lake Storage (ADLS) Gen2 and Azure Blob Storage are supported. [Learn more](#)

Storage account

adfcookbooksynapse-WorkspaceDefaultStorage

Input file or folder

flightscontainer/routes.csv

Continue Cancel

Microsoft Azure | Synapse Analytics > adfcookbooksynapse

Search resources

Preview data

Linked service: adfcookbooksynapse-WorkspaceDefaultStorage

Object: /routes.csv

Prop_0	Prop_1	Prop_2	Prop_3	Prop_4	Prop_5
Airline	Airline_ID	Source_airport	Source_airport_ID	Destination_airport	Destination_airpo
2B	410	AER	2965	KZN	2990
2B	410	ASF	2966	KZN	2990
2B	410	ASF	2966	MRV	2962
2B	410	CEK	2968	KZN	2990
2B	410	CEK	2968	OVB	4078

Showing 1 to 5 of 5 cached items

Bulk load

Source data file format settings

Specify the format and layout of your data. [Learn more](#)

Source storage

Preview data

File type

Field terminator

Edit

Row terminator

Edit

First row

Field quote

Edit

Date format

Microsoft Azure | Synapse Analytics > adfcookbooksynapse

Search resources

Data

Workspace

Storage accounts

Datasets

Showing 1 to 6 of 6 cached items

Bulk load

Select target SQL pool

Specify the target location for your load including the SQL pool, table, and column mapping. [Learn more](#)

Select SQL pool *

Select target table

Existing table Create New

Existing target table

Configure column mapping

Checking this box will automatically load into the SQL pool table when opening the script.

Microsoft Azure | Synapse Analytics > adfcookbooksynapse

Search resources

Publish all Validate all Refresh Discard all

Orchestrate + <>

CopyData Activities

Validate Debug Add trigger

Properties

General

Name * CopyData

Description

Concurrency

Annotations

Timeout 7:00:00

This screenshot shows the Microsoft Azure Synapse Analytics pipeline editor. On the left, the 'Orchestrate' sidebar is open, showing 'Pipelines' with two items: 'CopyToParquet' and 'CopyData'. The 'CopyData' pipeline is selected. The main workspace displays a single 'Copy data' activity. The 'Properties' panel on the right is configured for this activity, with the 'Name' field set to 'CopyData'. Other settings like 'Concurrency' and 'Timeout' are also visible.

Microsoft Azure | Synapse Analytics > adfcookbooksynapse

Search resources

Publish all Validate all Refresh Discard all

Orchestrate + <>

CopyData AzureSynapseAnalytics... ●

Properties

General

Name * AzureSynapseAnalyticsTable2

Description

Annotations

Name

Connection Schema Parameters

Linked service * adfcookbooksynapse-WorkspaceDef.✓ Connection successful

Test connection Open New

Linked service properties

NAME VALUE

DBName adfcookbooksql

Table dbo . Routes

Edit Preview data

This screenshot shows the Microsoft Azure Synapse Analytics dataset editor. The 'Connection' tab is selected, displaying a 'Linked service' named 'adfcookbooksynapse-WorkspaceDef.' with a green checkmark indicating a successful connection. Below this, 'Linked service properties' are listed, showing 'DBName' set to 'adfcookbooksql'. The 'Schema' tab is also visible, showing a table named 'dbo' with a 'Routes' column. The 'Properties' panel on the right is configured for this dataset, with the 'Name' field set to 'AzureSynapseAnalyticsTable2'.

Microsoft Azure | Synapse Analytics > adfcookbooksynapse

Search resources

Publish all Validate all Refresh Discard all

Orchestrate

Activities

CopyData

AzureSynapseAnalyti...

Copy data

CopyData

Parquet1

flightscontainer / Directory / File

Import schema

From connection/store From sample file None

OK Back Cancel

Set properties

Choose a name for your dataset. This name can be updated at any time until it is published.

Name: Parquet1

Linked service: adfcookbooksynapse-WorkspaceDefaultStorage

File path: flightscontainer / Directory / File

Import schema: From connection/store From sample file None

Advanced

General Source

Sink dataset: New

Microsoft Azure | Synapse Analytics > adfcookbooksynapse

Search resources

Publishing 3 Validate all Refresh Discard all

Orchestrate

Activities

CopyData

AzureSynapseAnalyti...

Copy data

CopyData

CopyData (New)

AzureSynapseAnalyticsTab... (New)

Parquet1 (New)

Name: CopyData

Description

Timeout: 7.00:00:00

OK Back Cancel

Publish all

You are about to publish all pending changes to the live environment. [Learn more](#)

Pending changes (3)

NAME	CHANGE	EXISTING
▲ Pipelines		
CopyData	(New)	-
▲ Datasets		
AzureSynapseAnalyticsTab...	(New)	-
Parquet1	(New)	-

Microsoft Azure | Synapse Analytics > adfcookbooksynapse

Pipeline runs

Time : Last 24 hours (8/3/20 8:29 AM -...) Time zone : Moscow, St. Petersburg, Volgog... Runs : Latest runs List Gantt Filter

CopyData Add filter

All status Rerun Cancel Refresh Edit columns

Showing 1 - 1 items

Pipeline name	Run start	Duration	Triggered by	Status	Parameters	Annotations
CopyData	8/4/20, 9:32:34 AM	00:00:11	Manual trigger	Succeeded		

Microsoft Azure | Synapse Analytics > adfcookbooksynapse

Search resources

Publish all Validate all Refresh Discard all

Data Workspace Linked

Filter resources by name

Storage accounts 1

- adfcookbooksynapse (Primary - a...
 - adfcookbookch3fs (Primary)
 - flightscontainer

Datasets 4

- AirlinesParquet
- AzureSynapseAnalyticsTable1
- AzureSynapseAnalyticsTable2
- Parquet1

CopyData AzureSynapseAnalyti... flightscontainer

New SQL script New notebook New data flow New dataset Upload Download New folder More

flightscontainer

NAME	LAST MODIFIED	CONTENT TYPE	SIZE
airlines.csv	8/2/2020, 12:32:16 PM		320.5 KB
countries.csv	8/2/2020, 12:32:30 PM		4.6 KB
dbo.Routes.parquet	8/4/2020, 9:32:44 AM		1.1 MB
planes.csv	8/2/2020, 12:32:41 PM		7.0 KB
routes.csv	8/2/2020, 11:48:18 PM		2.3 MB

Showing 1 to 5 of 5 cached items

Microsoft Azure | Synapse Analytics > adfcookbooksynapse

Search resources

Publish all Validate all Refresh Discard all

Data Workspace Linked

Filter resources by name

Storage accounts 1

- adfcookbooksynapse (Primary - a...
 - adfcookbookch3fs (Primary)
 - flightscontainer

Datasets 4

- AirlinesParquet
- AzureSynapseAnalyticsTable1
- AzureSynapseAnalyticsTable2
- Parquet1

CopyData AzureSynapseAnalyti... flightscontainer

New SQL script New notebook New data flow New dataset Upload Download New folder More

flightscontainer

NAME	LAST MODIFIED	CONTENT TYPE	SIZE
airlines.csv	8/2/2020, 12:32:16 PM		320.5 KB
countries.csv	8/2/2020, 12:32:30 PM		4.6 KB
dbo.Routes.parquet		Select TOP 100 rows	1.1 MB
planes.csv		New SQL script	7.0 KB
routes.csv		New notebook	2.3 MB

New data flow New dataset Copy ABFSS path Manage access... Rename... Download Delete Properties...

<https://web.azuresynapse.net/assets/external/storage/Scripts/0.2.7/StorageExplorer/FrameIndex.html?&enabledFeature=D&id=a365TabContentDiv5d75e762-4c6b-49d3-8c3a-70cd14f4cb20&implPath=%2Fassets%2Fexternal..>

Microsoft Azure | Synapse Analytics > adfcookbooksynapse

Search resources

Publish all 1 Validate all Refresh Discard all

Data Workspace Linked

Filter resources by name

Storage accounts 1

adfcookbooksynapse (Primary - a...
adfcookbookch3fs (Primary)
flightscontainer

Datasets 4

AirlinesParquet
AzureSynapseAnalyticsTable1
AzureSynapseAnalyticsTable2
Parquet1

CopyData AzureSynapseAnalytic... flightscontainer SQL script 1

Run Undo Publish Query plan Connect to SQL on-demand Use database master

```

1 SELECT
2 TOP 100 *
3 FROM
4 OPENROWSET(
5 BULK 'https://adfcookbookch3adls.dfs.core.windows.net/flightscontainer dbo.Routes.parquet',
6 FORMAT='PARQUET'
7 ) AS [r];
8

```

Results Messages

View Table Chart Export results

Search

Airline	Airline_ID	Source_airport	Source_airport...	Destination_air...	Destination_air...	Codeshare	Stops
2B	410	DME	4029	NBC	6969	NULL	0
AS	439	SEA	3577	PHX	3462	NULL	0
ET	2220	COO	245	ADD	1107	NULL	0

00:00:05 Query executed successfully.

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

p0rshe25@yandex.ru DEFAULT DIRECTORY

Home >

adfcookbooksynapse

Synapse workspace

Search (Ctrl+/)

New SQL pool New Apache Spark pool Refresh Reset SQL admin password Delete Launch Synapse Studio

Overview

Resource group (change) adf-cookbook-ch1-rg

Status Succeeded

Location East US

Subscription (change) Azure subscription 1

Subscription ID 69efd1df-7ca2-48e6-ba8a-3057c8d3536f

Managed virtual network No

Managed Identity object ID aa4da46e-e054-42fd-a8b2-00d1d844d016

Workspace web URL <https://web.azuresynthesize.net?workspace=%2bsubscriptions%2f69ef...>

Tags (change) Click here to add tags

Firewalls Show firewall settings

Primary ADLS Gen2 account URL <https://adfcookbookch3adls.dfs.core.windows.net>

Primary ADLS Gen2 file system adfcookbookch3fs

SQL admin username synapseadmin

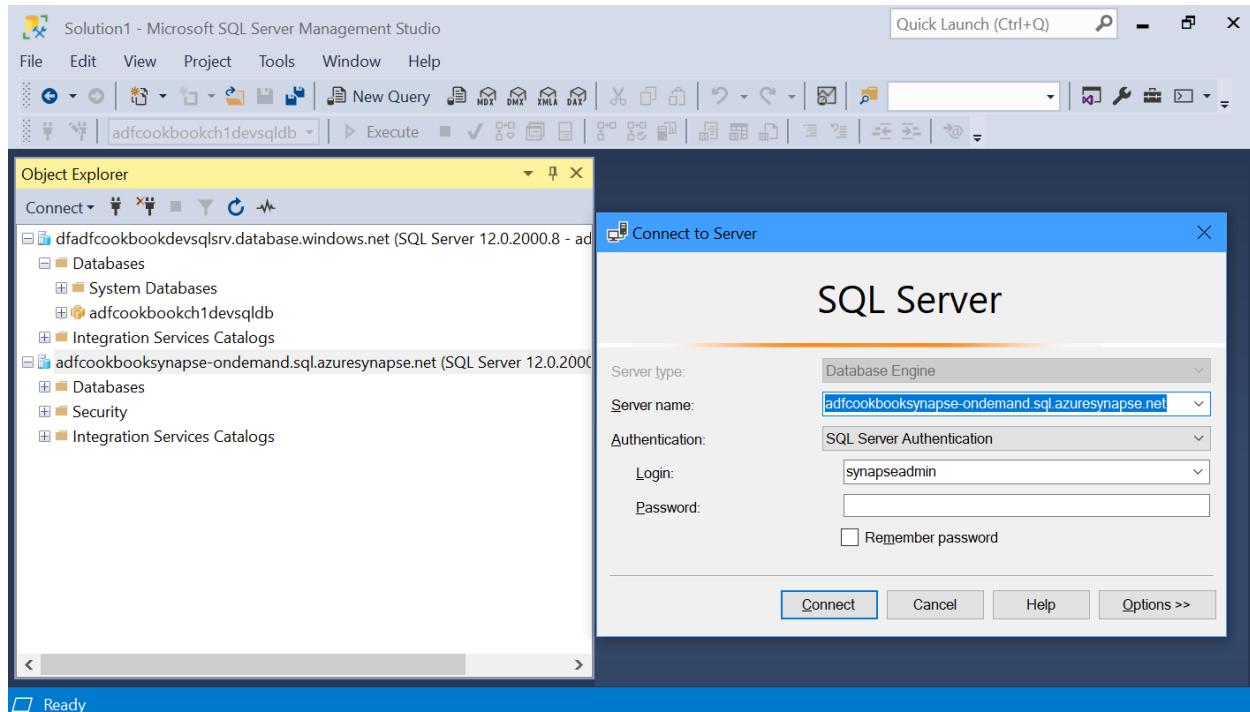
SQL Active Directory admin live.com#p0rshe25@yandex.ru

SQL endpoint adfcookbooksynapse.sql.azuresynapse.net

SQL on-demand endpoint adfcookbooksynapse-ondemand.sql.azuresynapse.net

Development endpoint <https://adfcookbooksynapse.dev.azuresynapse.net>

Copy to clipboard



Screenshot of Microsoft SQL Server Management Studio showing a query window and results.

The query window displays the following T-SQL code:

```
SELECT
    TOP 100 *
FROM
    OPENROWSET(
        BULK 'https://adfcookbookch3adls.dfs.core.windows.net/flightscontainer/dbo.Routes.parquet'
        FORMAT='PARQUET'
    ) AS [r];
```

The results pane shows a table with 13 rows of flight route data:

	Airline	Airline_ID	Source_airport	Source_airport_ID	Destination_airport	Destination_airport_ID	Codeshare	Stops	Equipment
1	2B	410	DME	4029	NBC	6969	NULL	0	CR2
2	AS	439	SEA	3577	PHX	3462	NULL	0	73G 734 73J 73H
3	ET	2220	COO	245	ADD	1107	NULL	0	73W 752
4	KC	1006	IST	1701	SCO	4367	NULL	0	320
5	QR	4091	LAS	3877	JFK	3797	Y	0	757
6	US	5265	DFW	3670	ICN	3930	NULL	0	777
7	2B	410	KZN	2900	AER	2965	NULL	0	CR2
8	AS	439	SEA	3577	SAT	3621	NULL	0	73H 73J
9	ET	2220	DAR	1177	ZNZ	1186	NULL	0	73W
10	KC	1006	KUL	3304	ALA	2908	NULL	0	757
11	QR	4091	MAD	1229	DOH	11051	NULL	0	77W
12	US	5265	DFW	3670	LAS	3877	NULL	0	738 M80 M83 757
13	2B	410	LED	2948	KZN	2990	NULL	0	CR2

Message bar at the bottom: `Query executed successfully.`

Microsoft Azure Search resources, services, and docs (G+) 🔍 ⚡ 🛡️ 🌐 🛡️ ? 😊 dfoshin@mail.ru DEFAULT DIRECTORY

Dashboard > New > Azure Synapse Analytics (formerly SQL DW) >

Azure Synapse Analytics

Microsoft

Welcome to Azure Synapse Analytics (formerly known as Azure SQL Data Warehouse). Learn more.

* Basics * Additional settings Tags Review + create

Create a 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](#)

Project details

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

Subscription * 無料試用版

Resource group * df-adfcookbook-dev-ch1

Create new

Review + create **Next : Additional settings >**

Microsoft Azure Search resources, services, and docs (G+) 🔍 ⚡ 🛡️ 🌐 🛡️ ? 😊 dfoshin@mail.ru DEFAULT DIRECTORY

Dashboard > New > Azure Synapse Analytics (formerly SQL DW) >

Azure Synapse Analytics

Microsoft

Welcome to Azure Synapse Analytics (formerly known as Azure SQL Data Warehouse). Learn more.

SQL pool details

Enter required settings for this SQL pool, including picking a logical server and configuring the performance level.

SQL pool name * Enter SQL pool name

Server dfadfcookbookdevsqlsrv (East US)

Create new

Performance level * Gen2

DW400c

Select performance level

Review + create **Next : Additional settings >**

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

Home > adfcookbookch1devsqlldb (dfadfcbookdevsqlsrv/adfcookbookch1devsqlldb) X

Synapse SQL pool (data warehouse)

Search (Ctrl+ /) Pause Scale Restore + New restore point Delete

Overview Activity log Tags Diagnose and solve problems Workload management (previ... Settings Maintenance schedule Quick start Geo-backup policy Connection strings Properties Locks Export template

Azure Synapse Analytics brings new workload management capabilities. Click here to learn more.

Welcome to Azure Synapse Analytics (formerly known as Azure SQL Data Warehouse). Learn more.

Resource group (change)
adf-cookbook-ch1-rg

Status
Online

Location
East US

Subscription (change)
Azure subscription 1

Subscription ID
69efd1df-7ca2-48e6-ba8a-3057c8d3536f

Server name
dfadfcbookdevsqlsrv.database.windows.net

Connection strings
Show database connection strings

Performance level
Gen2: DW100c

Maintenance schedule
Sun 20:00 UTC (4h) / Wed 12:00 UTC (8h)

Geo-backup policy
Enabled

Tags (change)
Click here to add tags

<https://portal.azure.com/#>

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

Home > adfcookbookch1devsqlldb (dfadfcbookdevsqlsrv/adfcookbookch1devsqlldb) | Firewalls and virtu...

Synapse SQL pool (data warehouse)

Search (Ctrl+ /) Save Discard + Add client IP

You are setting the Minimal TLS Version property for all SQL Database and SQL Data Warehouse databases associated with the server. Any login attempts from clients using TLS version less than the Minimal TLS Version shall be rejected.

Connection Policy Default Proxy Redirect

Allow Azure services and resources to access this server Yes No

Connections from the IPs specified below provides access to all the databases in dfadfcbookdevsqlsrv.

Client IP address 165.225.66.75

Rule name	Start IP	End IP	...
ClientIPAddress_2020-07-19...	176.59.35.81	176.59.35.81	...
ClientIPAddress_2020-07-19....	165.225.66.74	165.225.66.74	...

Locks Export template

Advanced data security Auditing Firewalls and virtual networks Transparent data encryption

View streaming jobs Load Data Query editor (preview) Build dashboards + reports Model and cache data

Chapter 4: Working with Azure Data Lake

The screenshot shows the Microsoft Azure Storage Explorer interface. A container named 'movielens' is selected. Inside, there is a blob named 'ratings.csv'. The properties for 'ratings.csv' are displayed, including its URL (<https://adfdatalake22...>), last modified date (7/28/2020, 5:02:00 PM), creation time (7/28/2020, 5:02:00 PM), type (Block blob), size (2.37 MiB), and access tier (Cool (Inferred)). Other files in the container include 'links.csv', 'movies.csv', 'README.txt', and 'tags.csv'.

The screenshot shows the Microsoft Azure Data Factory pipeline configuration. A pipeline named 'CopyPipeline_Chicago' is selected. On the left, the 'Factory Resources' sidebar shows 'Pipelines' (1), 'Datasets' (2), and 'Data flows' (0). On the right, the 'Set properties' dialog is open for the 'ml_movies' activity. The activity is configured to use the 'ADF_DataLake' linked service and copy data from the 'movielens' dataset to a 'movies.csv' file. The 'File path' is set to 'movielens / Directory / movies.csv'. The 'Import schema' section has 'From connection/store' selected. Buttons at the bottom of the dialog are 'OK', 'Back', and 'Cancel'.

Microsoft Azure | Data Factory > ADFbook

Data Factory > pipeline1 > ml_preprocess

Validate

mlratings

Columns: 4 total

Add Source

Source settings

Output stream name: mlratings

Source type: Dataset

Dataset: ml_ratings

Options:

- Allow schema drift
- Infer drifted column types
- Validate schema

Skip line count: []

Sampling:

- Enable
- Disable

Rows limit: 100

Description

```
graph TD; Start(( )) --> Pipeline1[pipeline1]; Pipeline1 --> Step1[ml_preprocess]; Step1 --> MlRatings[mlratings]; MlRatings --> End(( ));
```

Microsoft Azure | Data Factory > ADFbook

Data Factory > pipeline1 > ml_ratings

Validate

mlratings

Columns: 4 total

Add Source

Source settings

Projection

Projection

Define default format

Column name: userid

Type: 12s short

Format: Specify format

Column name: movieid

Type: 12s integer

Format: Specify format

Column name: rating

Type: 12 double

Format: Specify format

Column name: timestamp

Type: 12s integer

Format: Specify format

Description

Factory Resources

Pipelines

- CopyPipeline_Chicago
- pipeline1

Datasets

- Chicago_ADF_DataLake
- ml_links
- ml_movies
- ml_ratings
- ml_tags
- SourceDataset_5g6
- wrangling_sink

Data flows

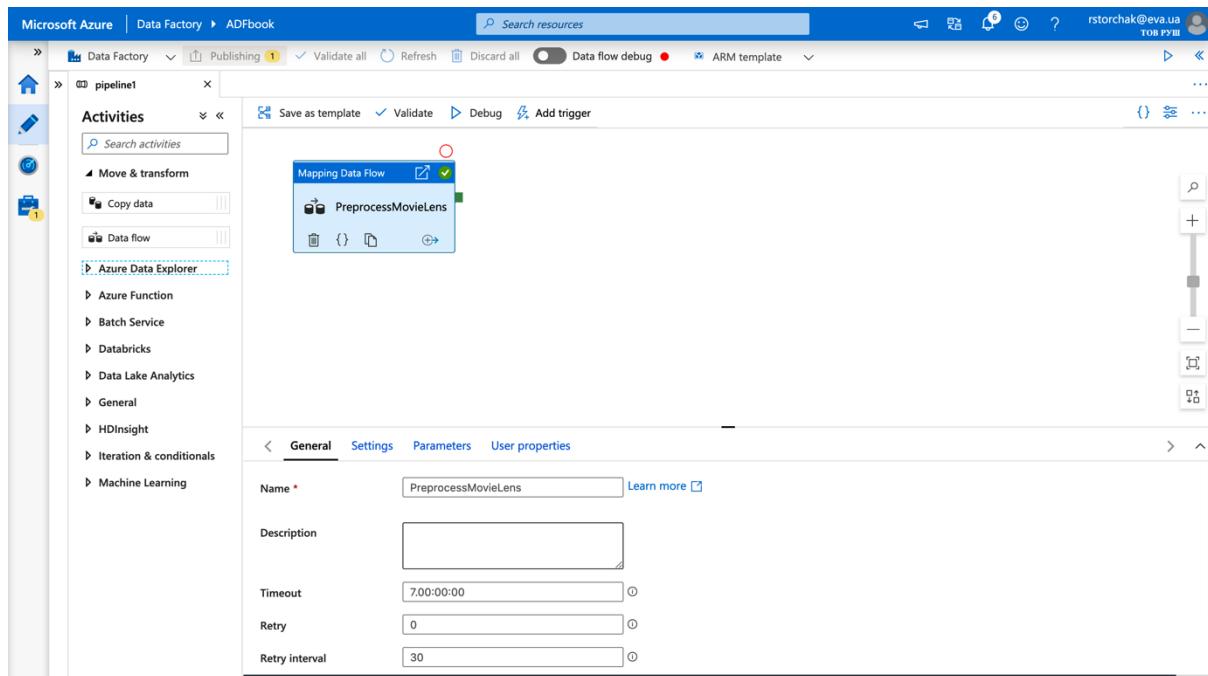
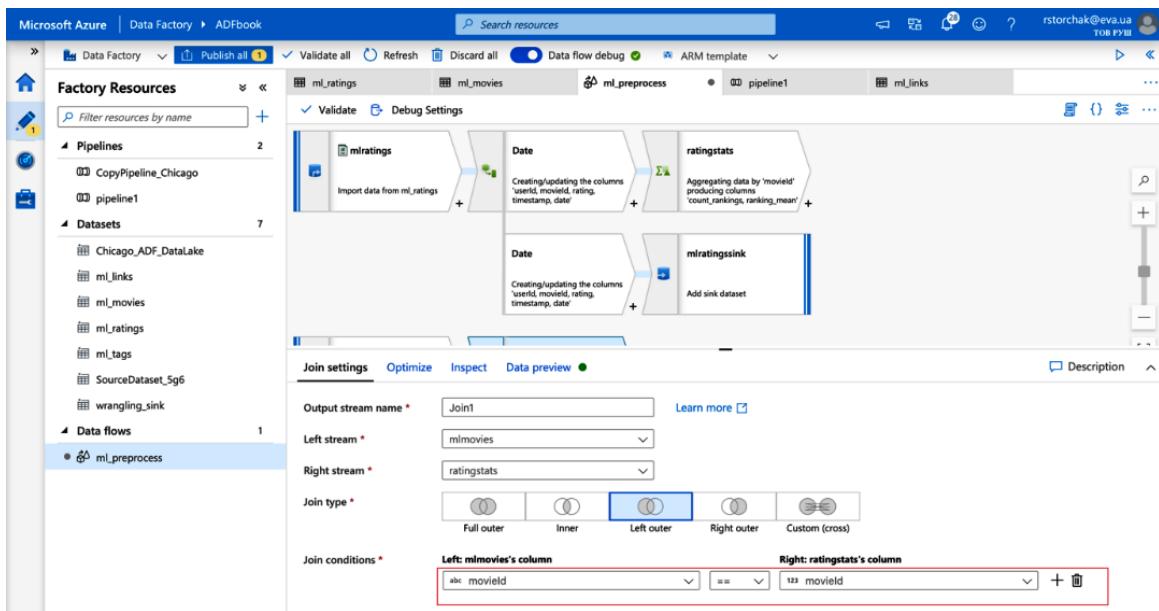
- ml_preprocess

```
graph TD; Start(( )) --> Pipeline1[pipeline1]; Pipeline1 --> Step1[ml_ratings]; Step1 --> ProjectionTab[Projection]; ProjectionTab --> End(( ));
```

The screenshot shows the Microsoft Azure Data Factory interface. The top navigation bar includes 'Microsoft Azure', 'Data Factory', 'ADFbook', 'Search resources', and user information 'rstorchak@eva.ua TOR PUBL'. The left sidebar lists 'Factory Resources' under 'Pipelines' (CopyPipeline_Chicago, pipeline1), 'Datasets' (Chicago_ADF_DataLake, ml_links, ml_movies, ml_ratings, ml_tags, SourceDataset_5g6, wrangling_sink), and 'Data flows' (ml_preprocess). The main workspace displays a pipeline named 'ml_preprocess' with two stages: 'ml_ratings' (source) and 'ml_movies' (sink). The 'ml_ratings' stage shows a table with 4 total rows. The 'ml_movies' stage shows an import from 'ml_ratings'. A 'Multiple inputs/outputs' panel is open, listing options like Join, Conditional Split, Exists, Union, and Lookup. The 'Data preview' tab is selected, showing a table with columns 'userId' (12s), 'rating' (12), and 'timestamp' (12s). The data preview shows 1000 rows with various values.

userId	rating	timestamp
6	4.0	845554024
5	4.0	847434801
7	4.0	1106636718
6	3.0	845555024
5	3.0	847434802
6	4.0	845556181

The screenshot shows the Microsoft Azure Data Factory interface. The top navigation bar includes 'Microsoft Azure', 'Data Factory', 'ADFbook', 'Search resources', and various account and notification icons. The left sidebar, titled 'Factory Resources', lists 'Pipelines' (CopyPipeline_Chicago, pipeline1), 'Datasets' (Chicago_ADF_DataLake, ml_links, ml_movies, ml_ratings, ml_tags, SourceDataset_5g6, wrangling_sink), and 'Data flows' (ml_preprocess). The main workspace displays the 'ml_preprocess' data flow. The data flow diagram consists of three main stages: 1) 'ml_ratings' (Import data from ml_ratings) which joins with 'Date' (Creating/updating the columns 'userid, movieid, rating, timestamp, date'); 2) 'Date' which joins with 'ratingstats' (Column: 3 total); 3) 'mlmovies' (Import data from ml_movies) which joins with 'ratingstats'. Below the diagram, the 'Aggregate settings' tab is selected, showing 'Output stream name' as 'ratingstats' and 'Incoming stream' as 'Date'. The 'Group by' dropdown is set to 'movielid' and the 'Aggregates' section contains two rows: 'count_rankings' with expression 'count()' and 'ranking_mean' with expression 'mean(rating)'.



Microsoft Azure | Data Factory > ADFbook

Data Factory Publishing Validate all Refresh Discard all Data flow debug ARM template

Activities Search activities Move & transform Copy data Data flow Azure Data Explorer Azure Function Batch Service Databricks Data Lake Analytics General HDInsight Iteration & conditionals Machine Learning

Save as template Validate Debug Add trigger

Mapping Data Flow PreprocessMovieLens

General Settings Parameters User properties

Data flow ml_preprocess Open New

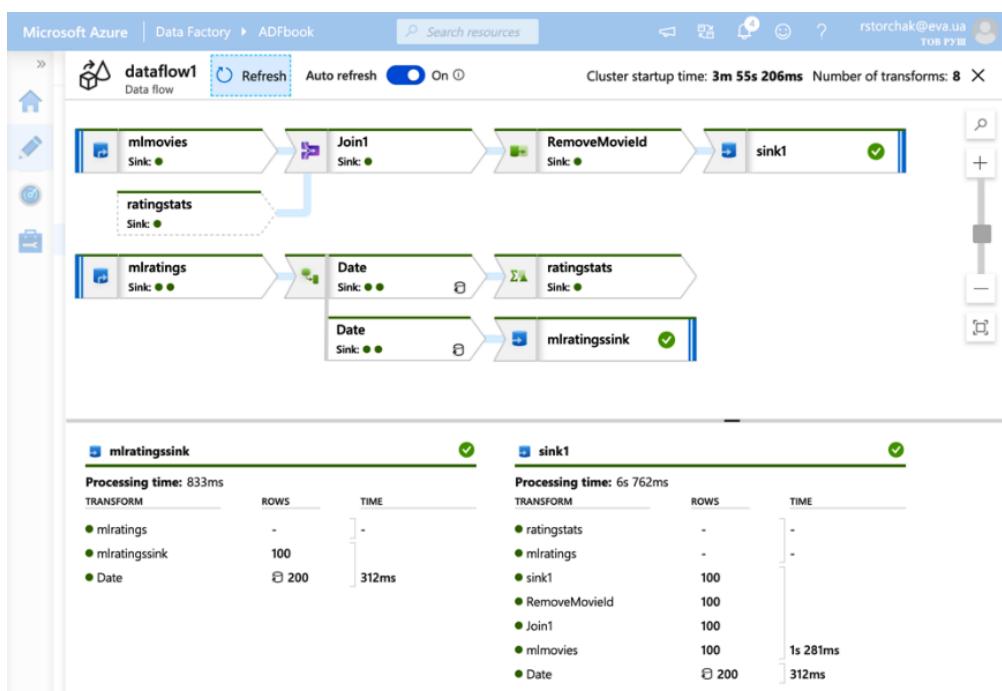
Run on (Azure IR) AutoResolveIntegrationRuntime

Compute type General purpose

Core count 4 (+ 4 Driver cores)

PolyBase

This screenshot shows the Azure Data Factory pipeline editor. On the left, a sidebar lists various activity types: Move & transform, Copy data, Data flow, Azure Data Explorer, Azure Function, Batch Service, Databricks, Data Lake Analytics, General, HDInsight, Iteration & conditionals, and Machine Learning. The main workspace displays a mapping data flow named 'PreprocessMovieLens'. Below the workspace, there are tabs for General, Settings, Parameters, and User properties. Under the Settings tab, the data flow is set to 'ml_preprocess', run on 'AutoResolveIntegrationRuntime', and use 'General purpose' compute type with 4 cores. The PolyBase setting is also visible.



Microsoft Azure  Search resources, services, and docs (G+/)     ?  
Home >

Create a resource group

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

Resource group - A container that holds related resources for an Azure solution. The resource group can include all the resources for the solution, or only those resources that you want to manage as a group. You decide how you want to allocate resources to resource groups based on what makes the most sense for your organization. [Learn more](#)

Project details

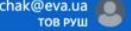
Subscription *  Azure subscription 1

Resource group *  ADF-test

Resource details

Region *  (US) East US

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

Microsoft Azure  Search resources, services, and docs (G+/)    ?   
Home >

Create storage account

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

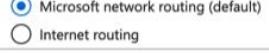
Network connectivity

You can connect to your storage account either publicly, via public IP addresses or service endpoints, or privately, using a private endpoint.

Connectivity method * Public endpoint (all networks) Public endpoint (selected networks) Private endpoint
All networks will be able to access this storage account. [Learn more about connectivity methods](#)

Network routing

Determine how to route your traffic as it travels from the source to its Azure endpoint. Microsoft network routing is recommended for most customers.

Routing preference *  Microsoft network routing (default) Internet routing

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

Microsoft Azure Search resources, services, and docs (G+) Dashboard Storage accounts Create storage account X

Basics Networking Data protection Advanced Tags Review + create

Recovery

Turn on point-in-time restore for containers
Use point-in-time restore to restore one or more containers to an earlier state. If point-in-time restore is enabled, then versioning, change feed, and blob soft delete must also be enabled. [Learn more](#)

Turn on soft delete for blobs
Soft delete enables you to recover blobs that were previously marked for deletion, including blobs that were overwritten. [Learn more](#)

Turn on soft delete for containers
Soft delete enables you to recover containers that were previously marked for deletion. [Learn more](#)
Sign up is required on a per-subscription basis to use container soft delete. [Sign up for container soft delete](#)

Turn on soft delete for file shares
Soft delete enables you to recover file shares that were previously marked for deletion. [Learn more](#)

Tracking

Turn on versioning for blobs
Use versioning to automatically maintain previous versions of your blobs for recovery and restoration. [Learn more](#)

Turn on blob change feed
Keep track of create, modification, and delete changes to blobs in your account. [Learn more](#)

Review + create < Previous Next : Advanced >

Dashboard > Storage accounts > Create storage account

Basics Networking Data protection **Advanced** Tags Review + create

Security

Secure transfer required Enabled Disabled

Minimum TLS version

Infrastructure encryption Enabled Disabled
Sign up is currently required to enable infrastructure encryption on a per-subscription basis. [Sign up for infrastructure encryption](#)

Blob storage

Allow Blob public access Enabled Disabled

Blob access tier (default) Hot Cool

NFS v3 Enabled Disabled
Sign up is currently required to utilize the NFS v3 feature on a per-subscription basis. [Sign up for NFS v3](#)

Data Lake Storage Gen2

Hierarchical namespace Enabled Disabled

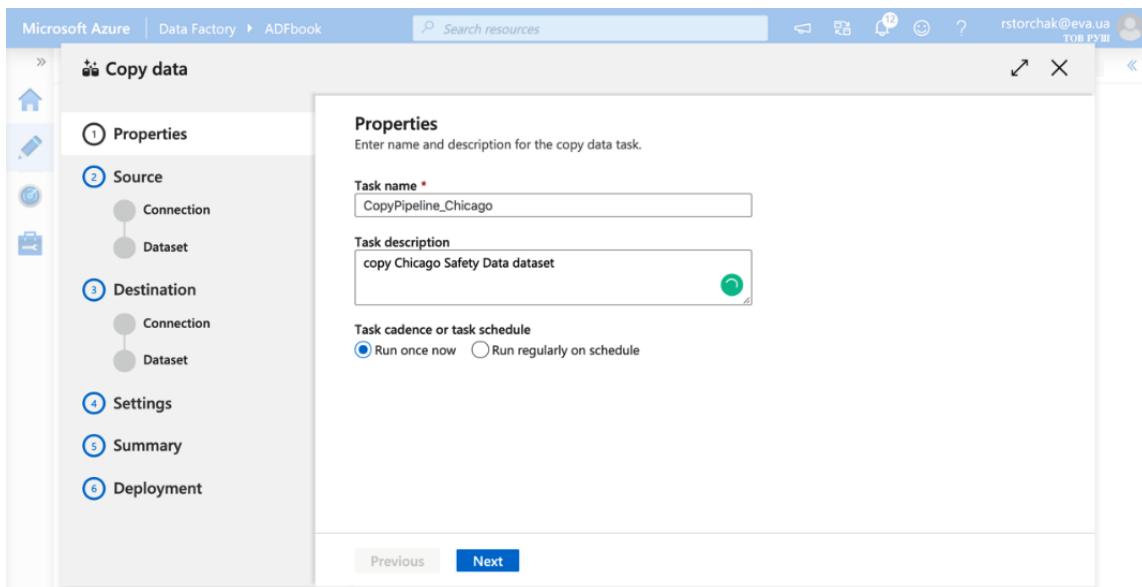
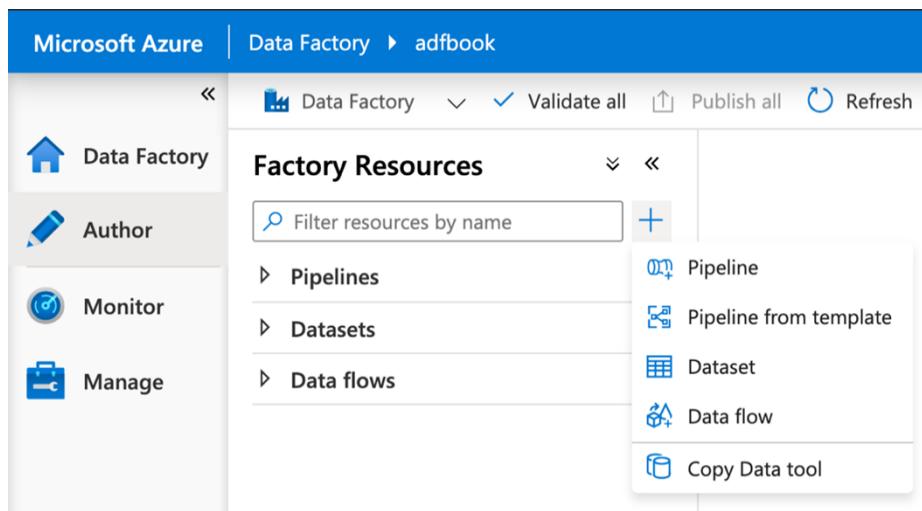
Azure Files

Large file shares Enabled Disabled
The current combination of storage account kind, performance, replication and location does not support large file shares.

Tables and Queues

Customer-managed keys support Enabled Disabled
Sign up is currently required to enable customer-managed keys support for tables and queues on a per-subscription basis. [Sign up for CMK support](#)

Review + create < Previous Next : Tags >



Microsoft Azure | Data Factory > ADFbook

Search resources

New linked service

Source data Specify the source

All Azure Database File Generic protocol NoSQL Services and apps

Copy data

Properties One time copy

Source Connection Dataset

Destination Connection Dataset

Settings Summary Deployment

Azure Blob Storage

Azure Cosmos DB (MongoDB API)

Azure Cosmos DB (SQL API)

Azure Data Explorer (Kusto)

Azure Data Lake Storage Gen1

Azure Data Lake Storage Gen2

Azure Database for MariaDB

Azure Database for MySQL

Azure Database for PostgreSQL

Azure File Storage

Azure SQL Database

Azure SQL Database Managed Instance

Azure Synapse Analytics (formerly SQL DW)

Azure Table Storage

Cancel

Previous Continue

Microsoft Azure | Data Factory > ADFbook

Search resources

New linked service (Azure Blob Storage)

Source data Specify the source

All Azure

Filter by name

AzureB1

DataLak

Name * OpenData_Chicago_AzureBlobStorage

Description Linked Service to Azure Open Data Chicago Safety Data

Connect via integration runtime * AutoResolveIntegrationRuntime

Authentication method SAS URI

SAS URL * https://azureopendatastorage.blob.core.windows.net/citydatacontainer/Safety/Release/city=Chic

SAS token Azure Key Vault

SAS token sample: ?sv=<storage services version>&st=<start time>&se=<expire time>&sr=<resource>

Test connection To linked service To file path

Annotations + New

Advanced

Connection successful Test connection Cancel

Copy data

Properties One time copy

Source Connection Dataset

Destination Connection Dataset

Settings Summary Deployment

Previous Create Back

This screenshot shows the 'New linked service (Azure Blob Storage)' configuration page in the Microsoft Azure Data Factory interface. The 'Source data' section is set to 'Azure'. The 'Name' field is filled with 'OpenData_Chicago_AzureBlobStorage'. The 'Description' field contains 'Linked Service to Azure Open Data Chicago Safety Data'. The 'Connect via integration runtime' dropdown is set to 'AutoResolveIntegrationRuntime'. The 'Authentication method' dropdown is set to 'SAS URI', and the 'SAS URL' field contains the URL 'https://azureopendatastorage.blob.core.windows.net/citydatacontainer/Safety/Release/city=Chic'. Below this, there are tabs for 'SAS token' and 'Azure Key Vault', with 'SAS token' selected. The 'Test connection' button is visible at the bottom right. The left sidebar shows a navigation tree with 'Copy data' selected, followed by 'Properties', 'Source', 'Destination', 'Settings', 'Summary', and 'Deployment'.

Microsoft Azure | Data Factory > ADFbook

Search resources

Choose the input file or folder

Select a source file or folder to be copied to the destination data store.

File or folder * citydatacontainer/Safety/Release/city=Chicago

Browse

If the identity you use to access the data store only has permission to subdirectory instead of the entire account, specify the path to browse.

Binary copy

Recursively

Max concurrent connections

Start time (UTC) End time (UTC)

Filter by last modified

Previous Next

This screenshot shows the 'Choose the input file or folder' configuration page in the Microsoft Azure Data Factory interface. The 'File or folder' field is set to 'citydatacontainer/Safety/Release/city=Chicago'. The 'Binary copy' checkbox is unchecked, and the 'Recursively' checkbox is checked. There are fields for 'Max concurrent connections', 'Start time (UTC)', 'End time (UTC)', and 'Filter by last modified'. At the bottom are 'Previous' and 'Next' buttons. The left sidebar shows a navigation tree with 'Copy data' selected, followed by 'Properties', 'Source', 'Destination', 'Settings', 'Summary', and 'Deployment'.

Copy data

Properties
One time copy

Source
Connection
Dataset

Destination
Connection
Dataset

Settings

Summary

Deployment

File format settings

File format: Parquet format

Compression type: snappy

Additional columns: + New

Schema

Column name	Type
dataType	UTF8
dataSubtype	UTF8
dateTime	INT96
category	UTF8
subcategory	UTF8
status	UTF8
address	UTF8
latitude	DOUBLE
longitude	DOUBLE
source	UTF8
extendedProperties	UTF8

Previous **Next**

Copy data

Properties
One time copy

Source
Azure Blob Storage
Connection
Dataset

Destination
Connection
Dataset

Settings

Summary

Deployment

Edit linked service (Azure Data Lake Storage Gen2)

Destination data store
Specify the destination data store for

All Azure Database File

Filter by name

- ADF_DataLake
- ChicagoSafetyDataLink
- OpenData_Chicago_Azu

Name *: ADF_DataLake

Description: Azure Data Lake gen2 storage

Connect via integration runtime *: AutoResolveIntegrationRuntime

Authentication method: Account key

Account selection method: From Azure subscription (radio button) Enter manually (radio button) selected

URL *: https://adfdatalake22.dfs.core.windows.net

Storage account key: Azure Key Vault

Storage account key *: *****

Test connection: To linked service (radio button) To file path (radio button)

Annotations: + New

Advanced

Previous **Next** **Apply** **Test connection** **Cancel**

Copy data

Choose the output file or folder
Specify a folder that will contain output files or a specific output file in the destination data store.

Properties
One time copy

Source
Azure Blob Storage

Connection

Dataset

Destination

Connection

Dataset

Settings

Summary

Deployment

Folder path chicago [Browse](#)

If the identity you use to access the data store only has permission to subdirectory instead of the entire account, specify the path to browse.

File name

Copy behavior Flatten hierarchy [\(i\)](#)

Max concurrent connections [\(i\)](#)

Block size (MB) [\(i\)](#)

[Previous](#) [Next](#)

Copy data

Properties
One time copy

Source
Azure Blob Storage

Connection

Dataset

Destination

Connection

Dataset

Settings

Summary

Deployment

File format settings

File format Parquet format [\(i\)](#)

Compression type snappy [\(i\)](#)

[Previous](#) [Next](#)

Copy data

Properties
One time copy

Source
Azure Blob Storage
 Connection
 Dataset

Destination
Azure Data Lake Storage Gen2
 Connection
 Dataset

Settings

Summary

Deployment

Settings
More options for data movement

Fault tolerance: Skip incompatible rows

Enable logging:

Logging settings

Storage connection name: ADF_DataLake

Folder path: chicago/logs

Performance settings

Enable staging:

Advanced settings

Previous **Next**

Copy data

Properties
One time copy

Source
Azure Blob Storage
 Connection
 Dataset

Destination
Azure Data Lake Storage Gen2
 Connection
 Dataset

Settings

Summary

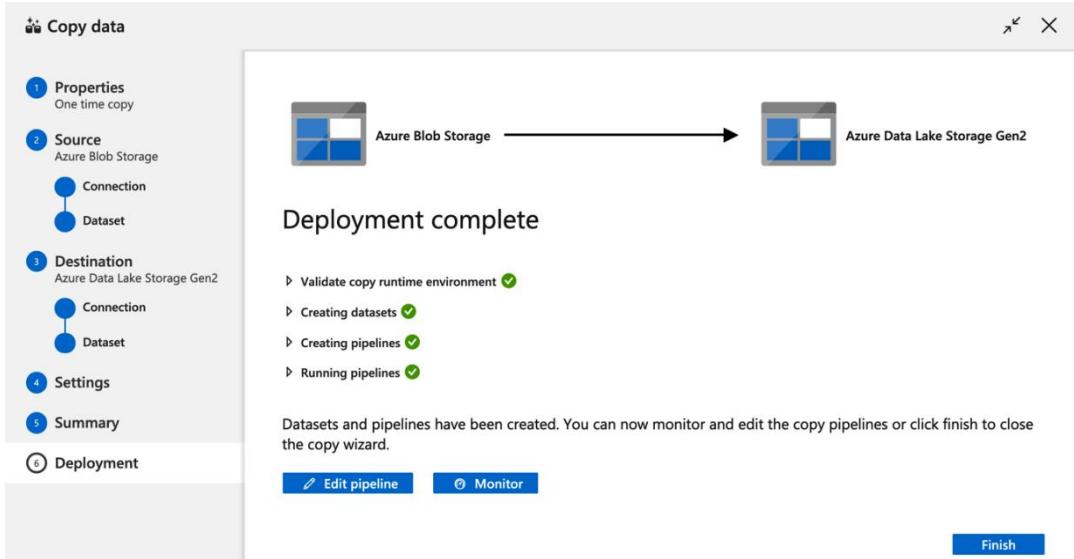
Deployment

Summary
You are running pipeline to copy data from Azure Blob Storage to Azure Data Lake Storage Gen2.

Properties

Task name	CopyPipeline_Chicago	Edit
Task description	copy Chicago Safety Data dataset	
Source		
Connection name	OpenData_Chicago_AzureBlobStorage	
Dataset name	SourceDataset_Sg6	
Compression codec	snappy	
Folder path	Safety/Release/city=Chicago	
Container	citydatacontainer	
Destination		
Connection name	ADF_DataLake	Edit
Dataset name	DestinationDataset_Sg6	
Compression codec	snappy	
Copy settings		
Timeout	7.00:00:00	
Retry	0	
Retry interval	30	
Secure output	false	
Secure input	false	

Previous **Next**



The screenshot shows the Microsoft Azure storage blob upload interface. The left pane displays the 'movielens' container listing. The right pane is titled 'Upload blob' with the path 'movielens/'. It contains fields for 'Files' (with a list box containing 'links.csv', 'movies.csv', 'ratings.csv', and 'READ...'), an unchecked 'Overwrite if files already exist' checkbox, and an 'Advanced' section. The 'Advanced' section includes settings for 'Authentication type' (set to 'Account key'), 'Blob type' (set to 'Block blob'), 'Block size' (set to '4 MB'), 'Access tier' (set to 'Cool (Inferred)'), and an 'Upload to folder' input field. At the bottom is an 'Upload' button.

Chapter 5: Working with Big Data – HDInsight and Databricks

Create HDInsight cluster

Primary storage

Select or create a storage account that will be the default location for cluster logs and other output.

Primary storage type * Azure Data Lake Storage Gen2

Primary storage account * adfadlv2

Filesystem * afdhdihadoop-2020-09-14t15-18-43-322z

Identity

Select a user-assigned managed identity to represent the cluster for Azure Data Lake Gen2 Storage account access. Only identities with access to the selected storage account are listed. Assign the managed identity to the 'Storage Blob Data Owner' role on the storage account. [Learn more](#)

User-assigned managed identity * managed_identity_adf

Create HDInsight cluster

This configuration will use 40 of 40 available cores in the East US region.
[View cores usage](#)

+ Add application

Node type	Node size	Number of ...	Estimated cost/h...
Head node	D12 v2 (4 Cores, 28 GB RAM), 0.37 USD/hour	2	0.75 USD
Worker node	D4 v2 (8 Cores, 28 GB RAM), 0.59 USD/hour	4	2.36 USD

Enable autoscale
[Learn More](#)

Total estimated cost/hour 3.11 USD

Home > HDInsight clusters >

afdhdihadoop X

HDInsight cluster

Search (Cmd+ /) Move Delete Refresh

- Overview
- Activity log
- Access control (IAM)
- Tags
- Diagnose and solve problems
- Quick start
- Tools

Settings

- Cluster size
- Quota limits
- SSH + Cluster login

● Applying changes...

Essentials

Resource group (change) ADF	Learn more Documentation
Status HDInsight configuration	Cluster type, HDI version Hadoop 2.7 (HDI 3.6)
Location East US	URL https://afdhdihadoop.azurehdinsight.net
Subscription (change) Azure subscription 1	Cluster ID 42cde877aeef6450ab65a1465d7ae1298
Subscription ID 1e70af7e-a8d1-4e0d-9a95-e6b389f01eb4	Getting started Quickstart
Tags (change) Click here to add tags	

External Identities Switch tenant Delete tenant Create a tenant What's new Preview features ...

External Identities Roles and administrators Administrative units (Preview) Enterprise applications Devices App registrations Identity Governance Application proxy Licenses Azure AD Connect

Tenant information

Your role
Global administrator [More info](#)

License
Azure AD Free

Tenant ID
ac76bf9e-678c-4ce2-bdb8-4a76... [Copy](#)

Primary domain
kshpundragmail.onmicrosoft.com

Azure AD Connect

Status
Not enabled

Last sync
Sync has never run

Register an application

* Name

The user-facing display name for this application (this can be changed later).

Supported account types

Who can use this application or access this API?

Accounts in this organizational directory only (Каталог по умолчанию only - Single tenant)

Accounts in any organizational directory (Any Azure AD directory - Multitenant)

Accounts in any organizational directory (Any Azure AD directory - Multitenant) and personal Microsoft accounts (e.g. Skype, Xbox)

Personal Microsoft accounts only

[By proceeding, you agree to the Microsoft Platform Policies](#)

Register

Home >

Subscriptions

Каталог по умолчанию

+ Add

View list of subscriptions for which you have role-based access control (RBAC) permissions to manage Azure resources. To view subscriptions for which you have billing access, [click here](#)

Showing subscriptions in Каталог по умолчанию directory. Don't see a subscription? [Switch directories](#)

My role	Status
8 selected	3 selected

Apply

Showing 1 of 1 subscriptions Show only subscriptions selected in the global subscriptions filter

Search to filter items...

Subscription name	Subscription ID	My role	Current cost	Status
Azure subscription 1	1e70af7e-a8d1-4e0d-9a95-e6b389f01eb4	Owner	\$173.11	Active

Microsoft Azure | portal.azure.com/?l=en-en-us#lkshpundragmail.onmicrosoft.com/resource/subscriptions/1e70af7e-a8d1-4e0d-9a95-e6b389f01eb4/users

Home > Subscriptions > Azure subscription 1

Azure subscription 1 | Access control (IAM)

Subscription

Search (Cmd+/)

Add Download role assignment

- Overview
- Activity log
- Access control (IAM)**
- Tags
- Diagnose and solve problems
- Security
- Events

Cost Management

- Cost analysis
- Cost alerts
- Budgets

Check access

View my level of access to this resource [View my access](#)

Check access

Review the level of access a user, group, principal, or managed identity has to this resource. [Learn more](#)

Find

Azure AD user, group, or service principal

Search by name or email address

Add role assignment

Role

Contributor

Assign access to

Azure AD user, group, or service principal

Select

adfbook

adfbook

Selected members:

adfbook Remove

Save **Discard**

» Data Factory Publish all 1 Validate all Refresh Discard all Data flow debug ARM template ▶ ⌂

» HDI_Hive

Activities Search activities

- Move & transform
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning

Save as template Validate Debug Add trigger ⌂ ⌂ ⌂

Properties

General Related

Name * HDI_Hive

Description Runs Hive queries on HDI Hadoop cluster

Concurrency 1

Annotations + New

Parameters Variables Output

This screenshot shows the Azure Data Factory designer interface. A new pipeline named 'HDI_Hive' is being created. The left sidebar lists various activity types under 'Activities'. The main workspace is currently empty. The properties panel on the right is configured for a general activity, with the name set to 'HDI_Hive' and a description indicating it runs Hive queries on an HDI Hadoop cluster. The concurrency is set to 1.

» Data Factory Publish all 1 Validate all Refresh Discard all Data flow debug ARM template ▶ ⌂

» HDI_Hive

Activities Search activities

- Move & transform
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Hive
- MapReduce
- Pig
- Spark

Save as template Validate Debug Add trigger ⌂ ⌂ ⌂

Hive

HiveJob

General HDI Cluster 1 Script 2 User properties

Name * HiveJob

Description Runs Hive script that builds DB

Timeout 7.00:00:00

Retry 0

Retry interval 30

Secure output

Secure input

This screenshot shows the Azure Data Factory designer interface. A new HDInsight activity named 'HiveJob' is being created within the 'HDI_Hive' pipeline. The left sidebar shows the 'HDInsight' section with options for Hive, MapReduce, Pig, and Spark. The main workspace contains a single 'Hive' activity node. The properties panel on the right is configured for the 'Hive' activity, with the name set to 'HiveJob' and a description indicating it runs a Hive script that builds a database. The timeout is set to 7.00:00:00, and the retry count is 0.

New linked service

Name *
HDI_Hive_linkedservice

Description
Linked service for HDI cluster with Hadoop

Type *
 Bring your own HDInsight On-demand HDInsight

Connect via integration runtime *
AutoResolveIntegrationRuntime

Azure Storage linked service *
Isadlv2

HDI Cluster
Cluster type * hadoop
Cluster size * 4
Time to live * 00:05:00

Script

Service principal ID *

Service principal key **Azure Key Vault**
Service principal key *

adfbook 

Search (Cmd+/)   

Overview   

Quickstart

Integration assistant | Preview

Manage

-  [Branding](#)
-  [Authentication](#)
-  [Certificates & secrets](#)
-  [Token configuration](#)
-  [API permissions](#)
-  [Expose an API](#)
-  [App roles | Preview](#)

Essentials

Display name adfbook	Supported account types My organization only
Application (client) ID d7955ea1-55b3-41c1-83c2-6e159f3d7f27	Redirect URLs Add a Redirect URI 
Directory (tenant) ID ac76bf9e-678c-4ce2-bdb8-4a7619dc1eb0	Application ID URI Add an Application ID URI
Object ID ed6e1846-dbd1-4463-b0fd-a67c125aeefc	Managed application in local directory adfbook

Important Starting June 30th, 2020 we will no longer add any new features to Azure Active Directory Authentication Library (ADAL) and Azure AD Graph. We will continue to provide technical support and security updates but we will no longer provide feature updates. Applications will need to be upgraded to Microsoft Authentication Library (MSAL) and Microsoft Graph. [Learn more](#)

New linked service

Name *
HDI_Hive_linkedservice

Description
Linked service for HDI cluster with Hadoop

Type *
 On-demand HDInsight

Connect via integration runtime *
AutoResolveIntegrationRuntime

Azure Storage linked service *
Isadv2

Cluster type *
hadoop

Cluster size *
4

Time to live *
00:05:00

Service principal ID *
d7955ea1-55b3-41c1-83c2-6e159f3d7f27

Service principal key **Azure Key Vault**

Service principal key *
.....

Tenant *
ac76bf9e-678c-4ce2-bdb8-4a7619dc1eb0

Version
3.6

Cluster name prefix

Resource group selection method
 From Azure subscription Enter manually

Subscription
Azure subscription 1 (1e70af7e-a8d1-4e0d-9a95-e6b389f01eb4)

Select resource group *
ADF

Additional storage linked service
Select...

Hcatalog linked service
Select...

OS type

Script action ⓘ

Advanced properties

Annotations
+ New

Name

Advanced ⓘ

Create **Cancel**

General HDI Cluster **Script** User properties

Details

Script linked service * **Isadlv2** ⓘ [Test connection](#) [Edit](#) [New](#)

File path * **dataLens/hive.hql** ⓘ [Browse storage](#) [Browse local](#) [Add/Edit script](#) [Preview script](#)

Advanced

Details		Refresh
Name	Value	
Duration	00:07:14	
Integration runtime	DefaultIntegrationRuntime (East US)	
Log location	abfss://adfjobs@adfdalv2.dfs.core.windows.net/HiveQueryJobs/f1c65e62-c14b-4d2fa74c-986105d7a13b/15_09_2020_07_38_49_335>Status	
Compute information	https://hi-742f6925-b633-4d7c-ac9d-3c1049017c47.azurehdinsight.net:443/	
Job ID	job_1600198594805_0001	
Execution progress	Succeeded	
YarnApplicationUri	https://hi-742f6925-b633-4d7c-ac9d-3c1049017c47.azurehdinsight.net/yarnui/hn/cluster/app/application_16001985948...	
JobStatusDetailsFromHDI	application_1600198594805_0001	
Duration in queue	{"integrationRuntimeQueue":10}	
billingReference	{"activityType":"ExternalActivity","billableDuration":[]}]}	

[Home](#) > [Azure Databricks](#) >

Azure Databricks Service

X

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

Subscription *	Azure subscription 1
Resource group *	ADF Create new

Instance Details

Workspace name *	adf_databricks
Location *	East US
Pricing Tier *	Standard (Apache Spark, Secure with Azure AD)

[Review + Create](#)

[Next : Networking >](#)

» Data Factory Publish all 1 Validate all Refresh

Connections

- Linked services
- Integration runtimes

Source control

- Git configuration
- Parameterization template

Author

- Triggers
- Global parameters
- Security
- Customer managed key
- Managed private endpoints

Linked services

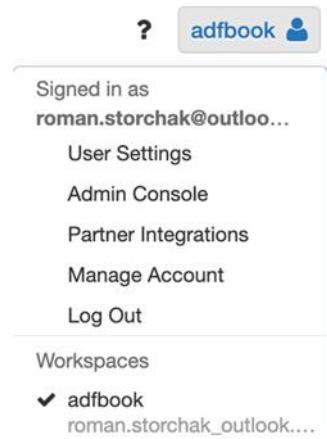
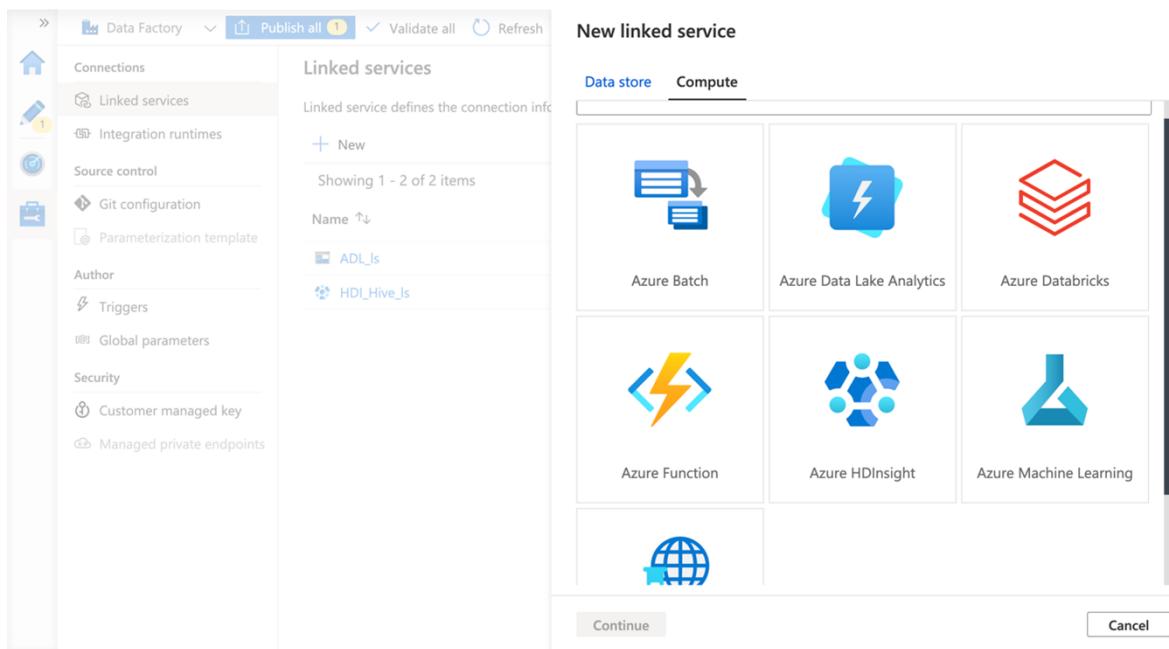
Linked service defines the connection info

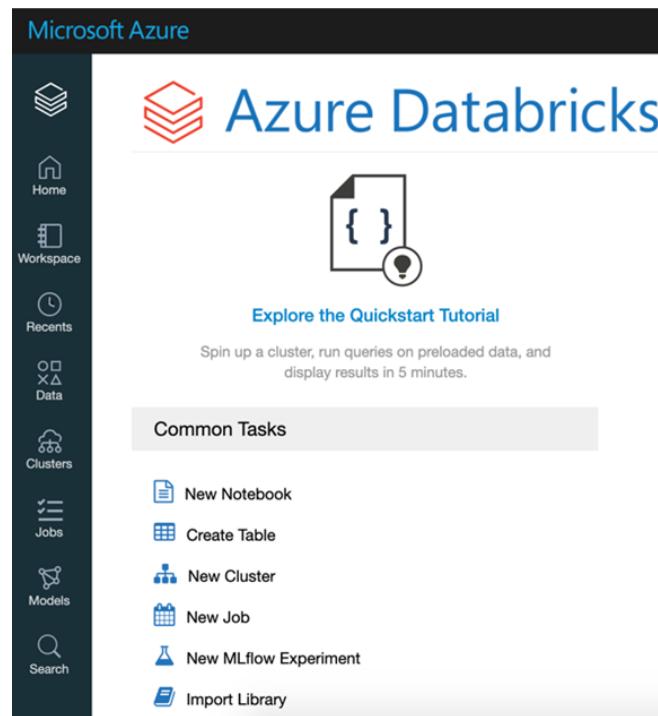
+ New Showing 1 - 2 of 2 items

Name ↑

ADL_ls	HDI_Hive_ls	
Azure Batch	Azure Data Lake Analytics	Azure Databricks
Azure Function	Azure HDInsight	Azure Machine Learning

Continue Cancel





Create Cluster

New Cluster

Cancel Create Cluster 0 Workers: 0.0 GB Memory, 0 Cores, 0 DBU
1 Driver: 8.0 GB Memory, 4 Cores, 0.5 DBU

Cluster Name

Cluster Mode

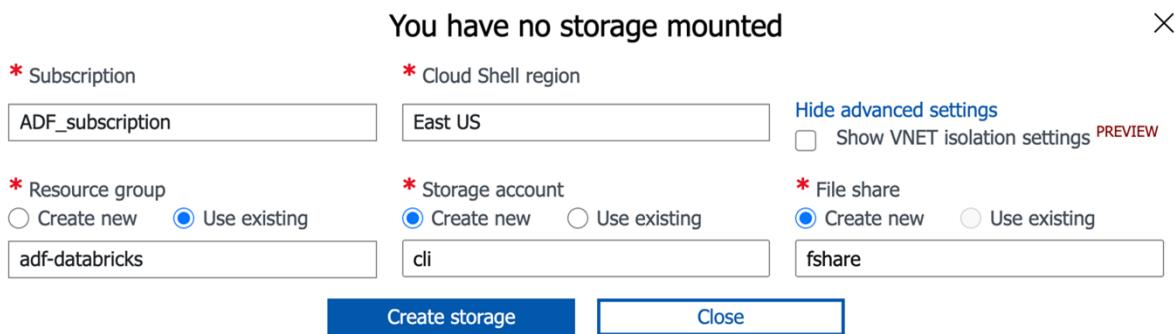
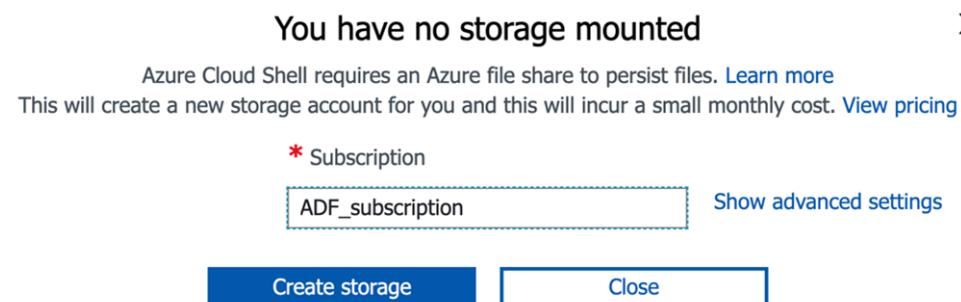
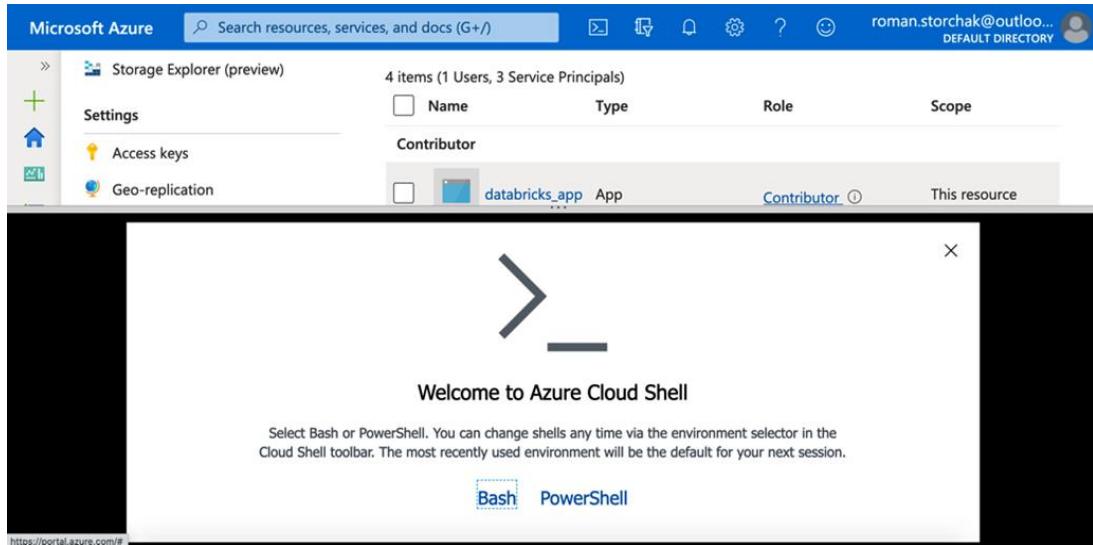
Pool

Databricks Runtime Version [Learn more](#)

New This Runtime version supports only Python 3.

Autopilot Options
 Terminate after minutes of inactivity

Node Type
 8.0 GB Memory, 4 Cores, 0.5 DBU

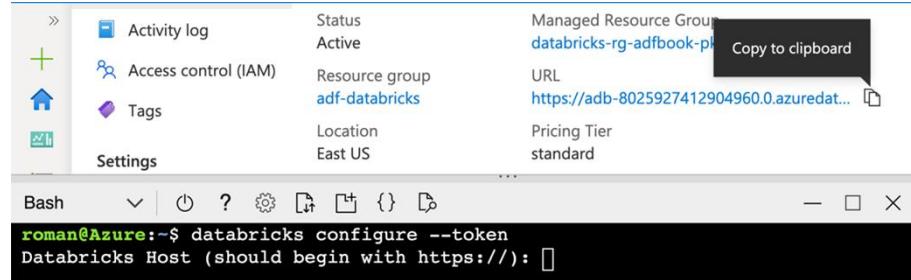


```
Bash ⌂ ? ⓘ { } ⌂
Requesting a Cloud Shell.Succeeded.
Connecting terminal...

Welcome to Azure Cloud Shell

Type "az" to use Azure CLI
Type "help" to learn about Cloud Shell

roman@Azure:~$
```



```
Bash    ⌂ ? ⌂ { } ⌂ X
roman@Azure:~$ databricks configure --token
Databricks Host (should begin with https://): [ ]
```

```
Bash    ⌂ ? ⌂ { } ⌂ X
3. /usr/bin/vim.basic

/usr/bin/select-editor: 32: /usr/bin/select-editor: gettext: not found
1-3 [1]: 1
Use "fg" to return to nano.

[1]+  Stopped                  databricks secrets put --scope adf-book --key adls
roman@Azure:~$ databricks secrets put --scope adf-book --key adls
roman@Azure:~$ databricks secrets list --scope adf-book
Key name      Last updated
-----
adls          1603208672111
roman@Azure:~$ [ ]
```

Microsoft Azure

PORTAL roman.storchak@outlook.com

test_mini (Python)

mini

Cmd 1

```
1 configs = {"fs.azure.account.auth.type": "OAuth",
2             "fs.azure.account.oauth.provider.type": "org.apache.hadoop.fs.azurebfs.oauth2.ClientCredsTokenProvider",
3             "fs.azure.account.oauth2.client.id": "26d8de10-8398-4572-9bf9-a8b571cec726",
4             "fs.azure.account.oauth2.client.secret": dbutils.secrets.get(scope="adf-book",key="adls"),
5             "fs.azure.account.oauth2.client.endpoint": "https://login.microsoftonline.com/7f2a6054-c522-4449-ba7e-
25559476ef47/oauth2/token"}
6
7 dbutils.fs.mount(
8     source = "abfss://datalens@adfbbook.dfs.core.windows.net/",
9     mount_point = "/mnt/adfs",
10    extra_configs = configs)

> (1) Spark Jobs
Out[1]: True

Command took 25.75 seconds -- by roman.storchak@outlook.com at 20.10.2020, 19:23:56 on mini
```

Cmd 2

```
1 dbutils.fs.ls("/mnt/adfs/")

Out[12]: [FileInfo(path='dbfs:/mnt/adfs/README.txt', name='README.txt', size=8342),
 FileInfo(path='dbfs:/mnt/adfs/links.csv', name='links.csv', size=197979),
 FileInfo(path='dbfs:/mnt/adfs/movies.csv', name='movies.csv', size=494431),
 FileInfo(path='dbfs:/mnt/adfs/ratings.csv', name='ratings.csv', size=2483723),
 FileInfo(path='dbfs:/mnt/adfs/tags.csv', name='tags.csv', size=118660)]
```

Command took 0.21 seconds -- by roman.storchak@outlook.com at 20.10.2020, 19:24:23 on mini

Clusters / mini

mini !

Configuration Notebooks (0) Libraries Event Log Spark UI Driver Logs Metrics Apps Spark Cluster UI - Master ▾

Cluster Mode ? Single Node

Databricks Runtime Version 7.0 (includes Apache Spark 3.0.0, Scala 2.12)

New This Runtime version supports only Python 3.

Autopilot Options Terminate after **10** minutes of inactivity ?

Node Type Standard_F4s 8.0 GB Memory, 4 Cores, 0.5 DBU

► Advanced Options

Clusters / mini

mini !

Configuration Notebooks (1) Libraries Event Log **Spark UI** Driver Logs Metrics Apps Spark Cluster UI - Master ▾

Jobs Stages Storage Environment Executors SQL JDBC/ODBC Server Structured Streaming

Spark Jobs (7)

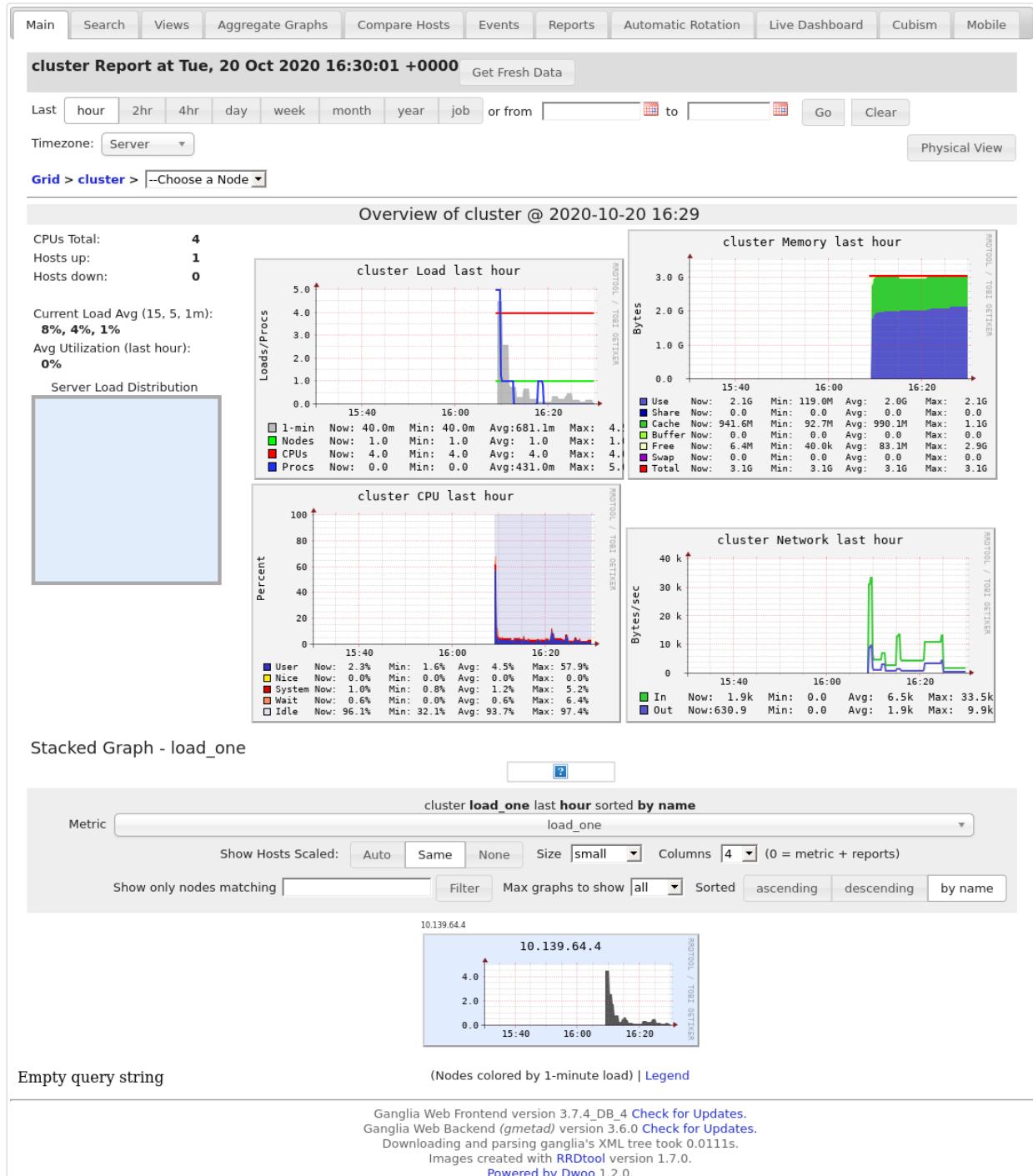
User: root
Total Uptime: 1.6 min
Scheduling Mode: FAIR
Completed Jobs: 7

► Event Timeline
▼ Completed Jobs (7)

Page: **1** 1 Pages. Jump to . Show items in a page.

Job Id (Job Group) ▾	Description	Submitted	Duration	Stages: Succeeded/Total	Tasks (for all stages): Succeeded/Total
6 (5533287454040499978_8904114122325394392_433a621d18c34c1d8337ab66226092e3)	from pyspark.ml.evaluation import RegressionEva... treeAggregate at Statistics.scala:58	2020/10/20 17:42:01	12 s	6/6 (7 skipped)	434/434 (43)
5 (5533287454040499978_8045897224296953182_ed85514bacfa40cfbf89ec0c7ddf9da5)	from pyspark.ml.recommendation import ALS als ... count at ALS.scala:1048	2020/10/20 17:41:59	0.2 s	1/1 (6 skipped)	10/10 (33)
4 (5533287454040499978_8045897224296953182_ed85514bacfa40cfbf89ec0c7ddf9da5)	from pyspark.ml.recommendation import ALS als ... count at ALS.scala:1043	2020/10/20 17:41:56	3 s	5/5 (3 skipped)	50/50 (3)
3 (5533287454040499978_8045897224296953182_ed85514bacfa40cfbf89ec0c7ddf9da5)	from pyspark.ml.recommendation import ALS als ... count at ALS.scala:946	2020/10/20 17:41:55	0.4 s	2/2 (1 skipped)	11/11 (1)
2 (5533287454040499978_8045897224296953182_ed85514bacfa40cfbf89ec0c7ddf9da5)	from pyspark.ml.recommendation import ALS als ... count at ALS.scala:939	2020/10/20 17:41:53	2 s	3/3	12/12
1 (5533287454040499978_8045897224296953182_ed85514bacfa40cfbf89ec0c7ddf9da5)	from pyspark.ml.recommendation import ALS als ... isEmpty at ALS.scala:926	2020/10/20 17:41:52	0.6 s	1/1	1/1
0 (5533287454040499978_7930159292396040565_979ec5974a884df18e7859cda59b971f)	ratings_df_schema = "userId integer, movieId in... head at <command-2760279187566901>:-4	2020/10/20 17:41:46	5 s	1/1	1/1

Page: **1** 1 Pages. Jump to . Show items in a page.



dataLens
Container

Search (Cmd+ /) << Upload Add Directory Refresh | Rename Delete | ...

Overview Access Control (IAM)

Authentication method: Access key ([Switch to Azure AD User Account](#))
Location: dataLens / models

Search blobs by prefix (case-sensitive)

Name	Modified	Access tier
<input type="checkbox"/> [..]		
<input type="checkbox"/> itemFactors		
<input type="checkbox"/> metadata		
<input type="checkbox"/> userFactors		

Home > New > User Assigned Managed Identity >

Create User Assigned Managed Identity

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

Subscription * Azure subscription 1

Resource group * ADF [Create new](#)

Instance details

Region * East US

Name * managed_identity_adf

Review + create < Previous Next : Tags >

Home > Storage accounts > adfadlv2

adfadlv2 | Access Control (IAM)

Storage account

Search (Cmd+ /) Add Download role assignments Edit columns Refresh Remove ...

Overview Activity log Tags Diagnose and solve problems Access Control (IAM) Data transfer Events Storage Explorer (preview)

Access Control (IAM)

Add role assignment Add co-administrator Roles Deny assignments Classic administrators

My access
View my level of access to this resource [View my access](#)

Check access
Review the level of access a user, group, service principal, or managed identity has to this resource. [Learn more](#)

Find Azure AD user, group, or service principal Search by name or email address

Add a role assignment
Grant access to resources at this scope by assigning a role to a user, group, service principal, or managed identity.

Add Learn more

View role assignments
View the users, groups, service principals and identities that have role assignments granting access at this scope.

Home > Storage accounts > adfadlv2

adfadlv2 | Access Control (IAM)

Storage account

Search (Cmd+ /) Add Download role assignments

Overview Activity log Tags Diagnose and solve problems Access Control (IAM) Data transfer Events Storage Explorer (preview)

Access Control (IAM)

Check access Role assignments

My access
View my level of access to this resource [View my access](#)

Check access
Review the level of access a user, group, service principal, or managed identity has to this resource. [Learn more](#)

Find Azure AD user, group, or service principal Search by name or email address

Add role assignment

Role Storage Blob Data Owner

Assign access to User assigned managed identity

Subscription * Azure subscription 1

Select Search by name

No results to display.

Selected members:

managed_identity_adf /subscriptions/1e70af7e-a8d1-4e0d-9a... Remove

Save Discard

Create HDInsight cluster

Subscription *

Resource group *

Cluster details
Name your cluster, pick a region, and choose a cluster type and version. [Learn more](#)

Cluster name *

Region *

Cluster type *

Version *

Cluster credentials
Enter new credentials that will be used to administer or access the cluster.

Cluster login username *

Cluster login password *

Confirm cluster login password *

Secure Shell (SSH) username *

Use cluster login password for SSH

Chapter 6: Integration with MS SSIS

Dashboard > SQL databases > movielens (adfbooksq/movielens) >

Firewall settings

adfbooksq (SQL server)

Save Discard + Add client IP

Info You are setting the Minimal TLS Version property for all SQL Database and SQL Data Warehouse databases associated with the server. Any login attempts from clients using TLS version less than the Minimal TLS Version shall be rejected.

Connection Policy [\(i\)](#) Default Proxy Redirect

Allow Azure services and resources to access this server Yes No

Info Connections from the IPs specified below provides access to all the databases in adfbooksq.

Client IP address	185.209.58.148		
Rule name	Start IP	End IP	...
adf-range	20.42.2.1	20.42.2.255	...
ClientIPAddress_2020-10-22...	185.209.58.148	185.209.58.148	...

Microsoft Azure | adfbook

Connections Linked services

Linked service defines the connection info

+ New Filter by keyword

Showing 1 - 2 of 2 items

Name ↑

ADLS_linkedservice AzureDatabricks_ADF

New linked service (Azure SQL Database)

Name * SQLLinkedService

Description

Connect via integration runtime * AutoResolveIntegrationRuntime

Connection string Azure Key Vault

Account selection method

From Azure subscription Enter manually

Azure subscription ADF_subscription (245d4f8f-e6c4-439c-b774-d373fba7b7ae)

Server name * adfbooksq1

Database name * movielens

Authentication type * SQL authentication

User name * roman

Password Azure Key Vault

Password *

Additional connection properties

+ New

Annotations

+ New

Name

Parameters Advanced

✓ Connection successful

Create Back Test connection Cancel

The screenshot shows the 'Linked services' blade in the Microsoft Azure Data Factory interface. On the left, there's a sidebar with various options like Connections, Linked services, Integration runtimes, Source control, etc. The main area shows a list of existing linked services: 'ADLS_linkedservice' and 'AzureDatabricks_ADF'. A 'New' button is visible. The right side is a form for creating a new linked service for an Azure SQL Database. It asks for a 'Name' (SQLLinkedService), a 'Description' (with a 'G' icon), and a 'Connect via integration runtime' (AutoResolveIntegrationRuntime). Under 'Account selection method', the 'From Azure subscription' option is selected. The 'Azure subscription' dropdown shows 'ADF_subscription (245d4f8f-e6c4-439c-b774-d373fba7b7ae)'. The 'Server name' is set to 'adfbooksq1' and the 'Database name' to 'movielens'. The 'Authentication type' is 'SQL authentication' with 'User name' 'roman' and 'Password' (redacted). There's also a 'Password' section with 'Azure Key Vault'. At the bottom, there are buttons for 'Create', 'Back', 'Test connection', and 'Cancel'. A success message '✓ Connection successful' is displayed above the 'Create' button.

Configure your new project

Integration Services Project (Azure-Enabled)

Project name

preprocess_movieLens

Location

C:\Users\roman\source\repos



Solution name

preprocess_movieLens

Place solution and project in the same directory

Back

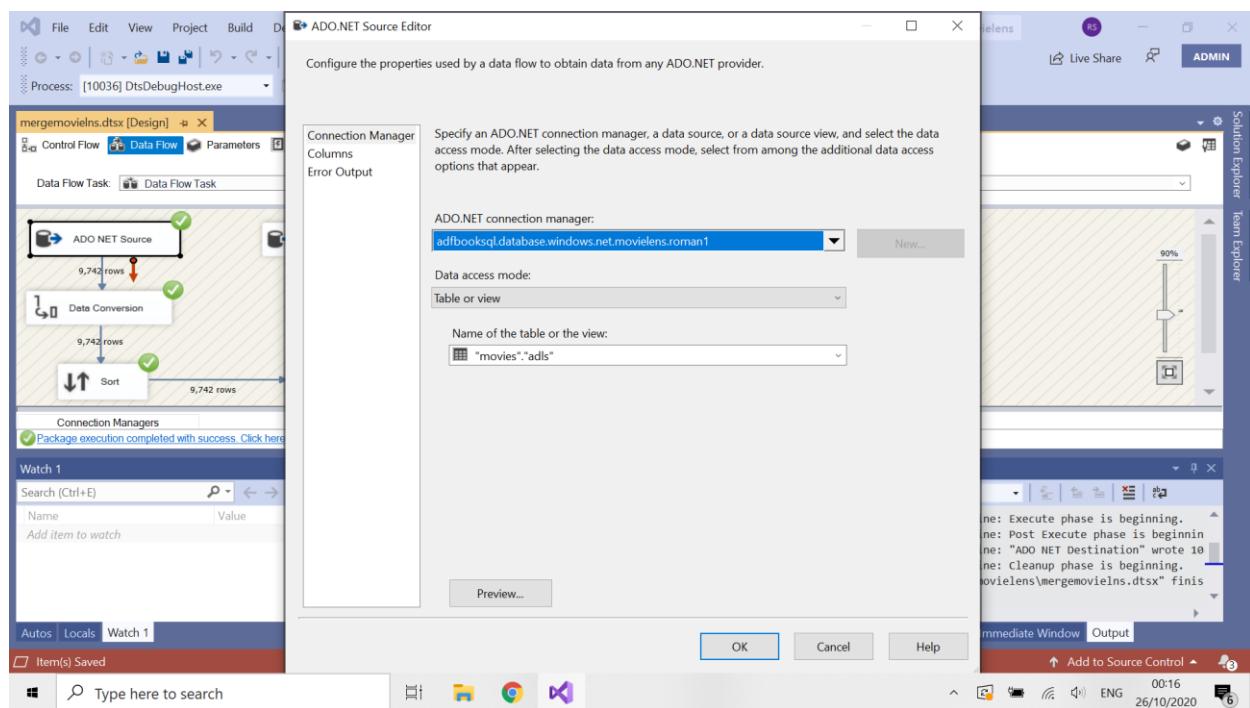
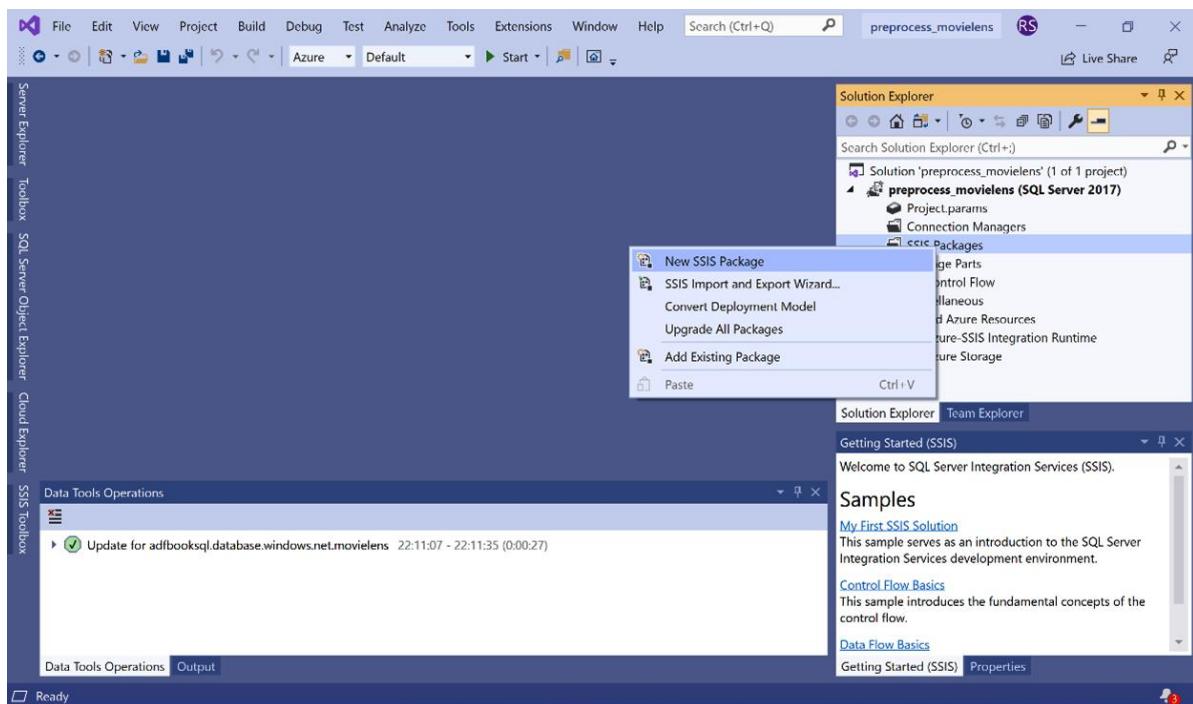
Create

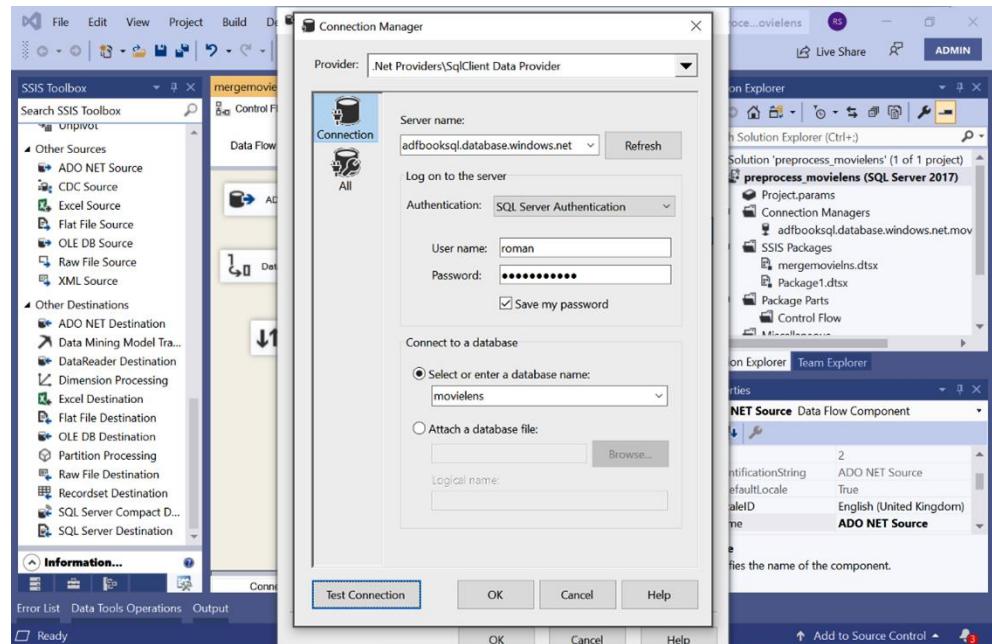
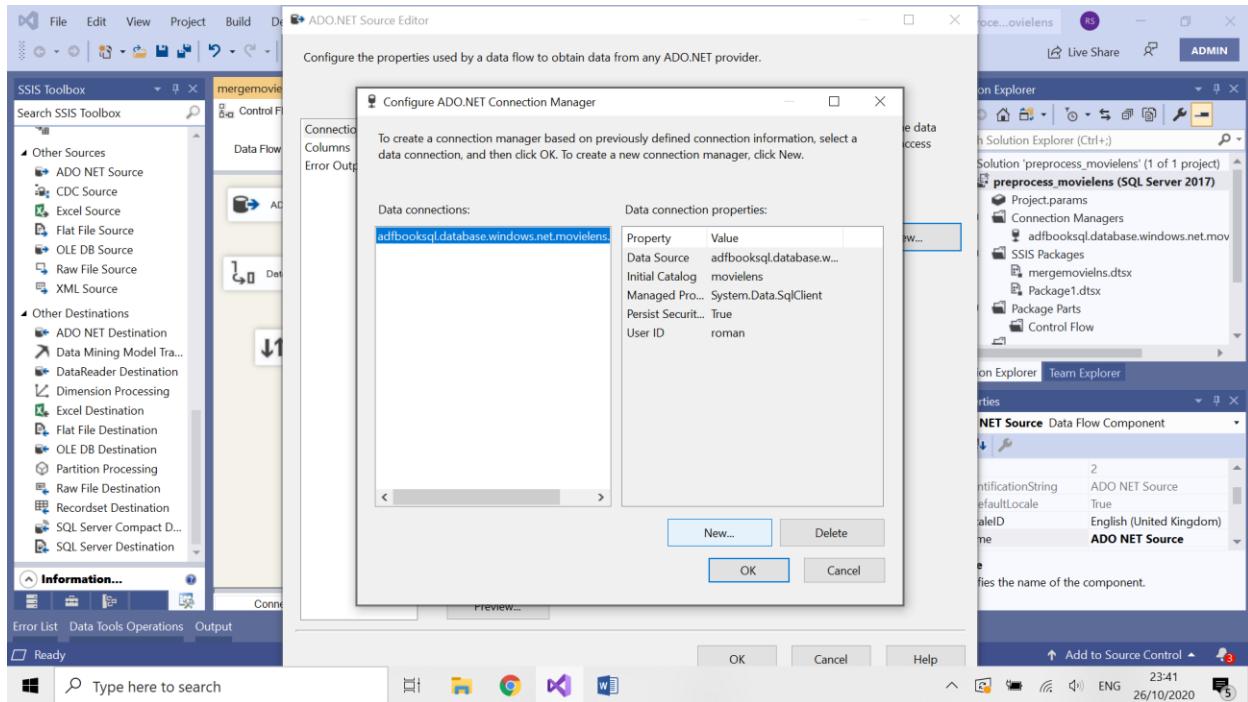
File Edit View Project Build Debug Test Analyze Tools Extensions Window Help Search (Ctrl+Q) RS Live Share

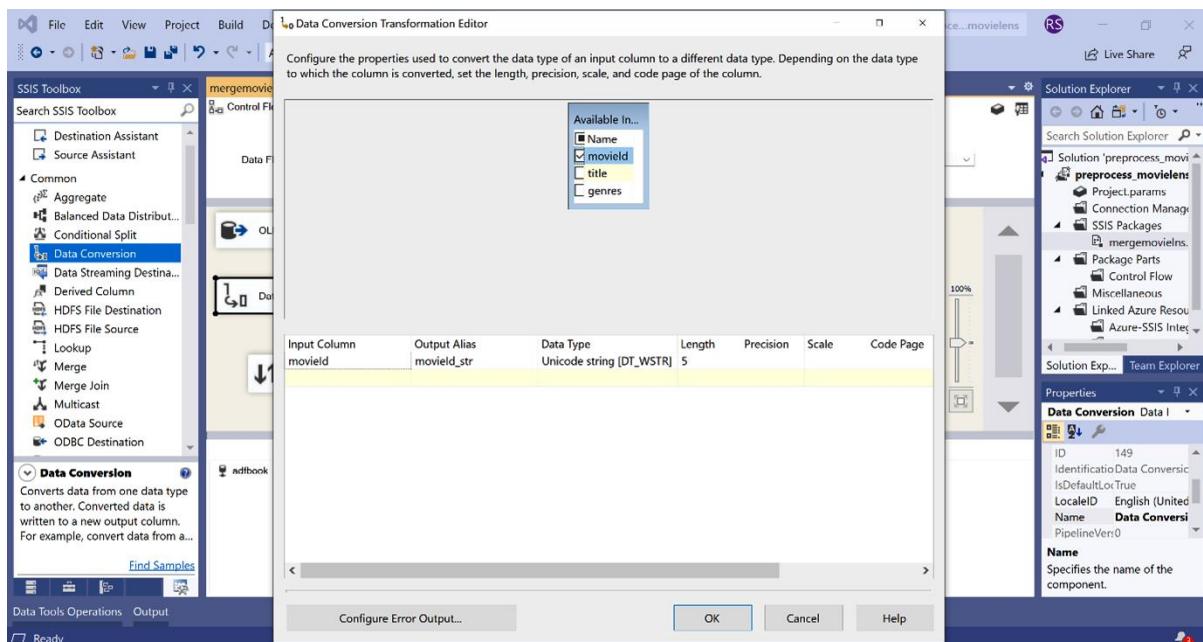
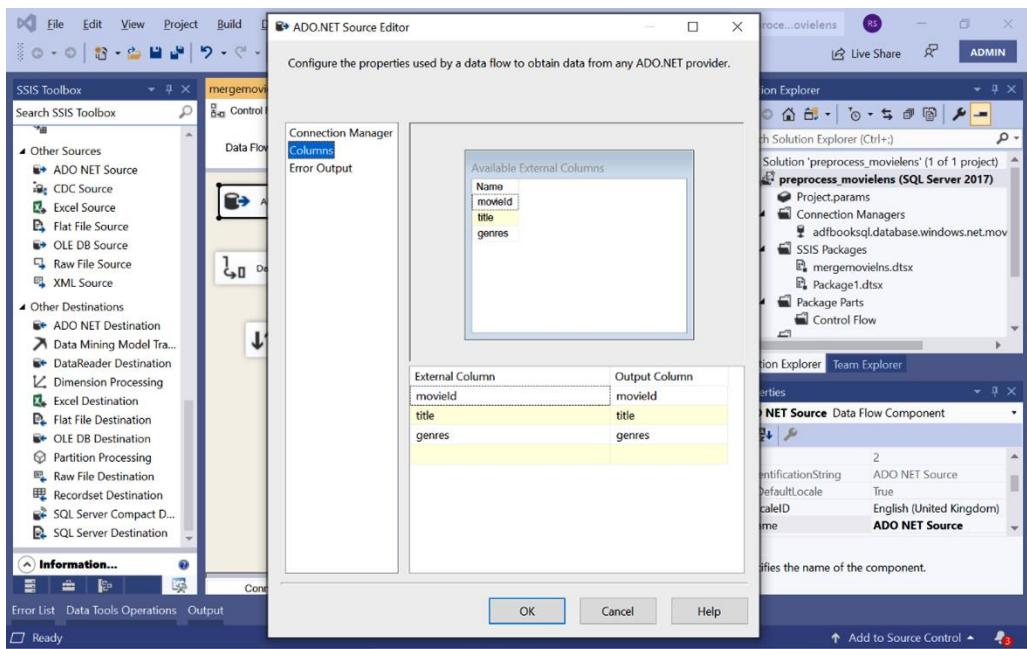
Server Explorer Toolbox SQL Server Object Explorer Cloud Explorer SSIS Toolbox

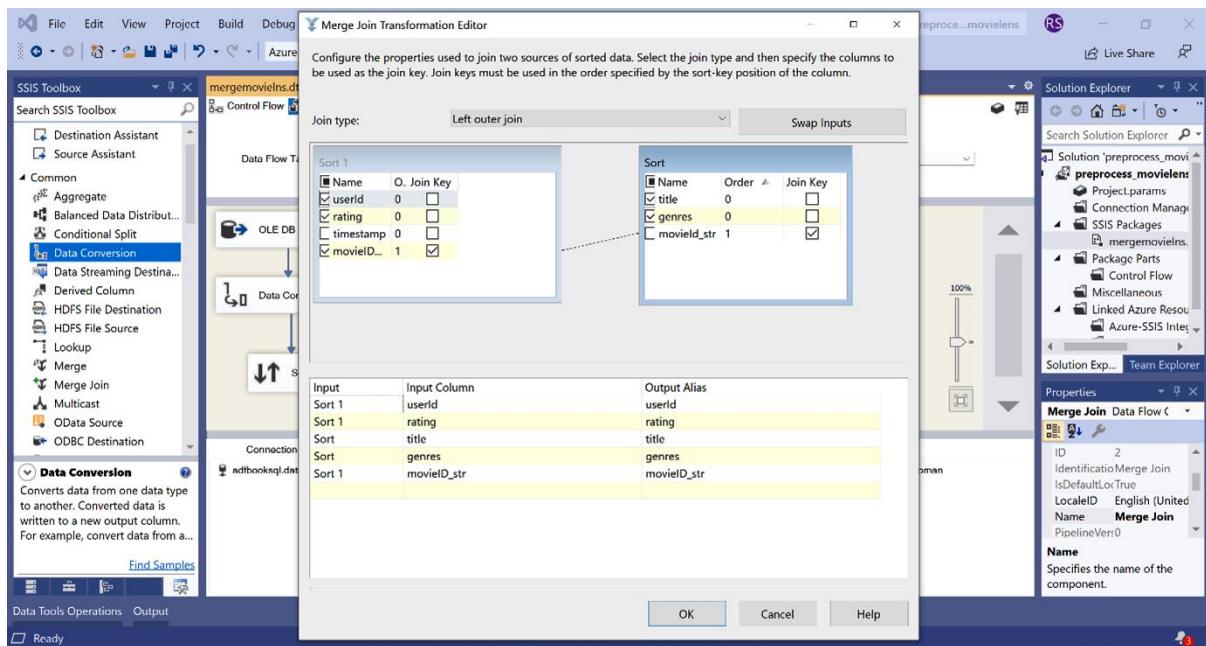
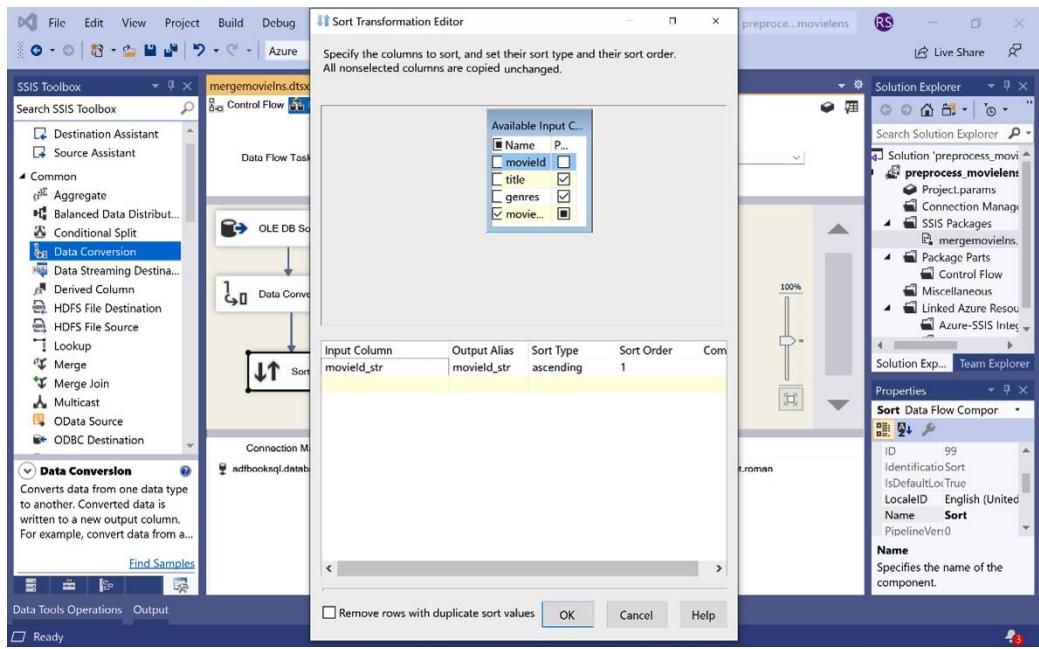
Ready Type here to search

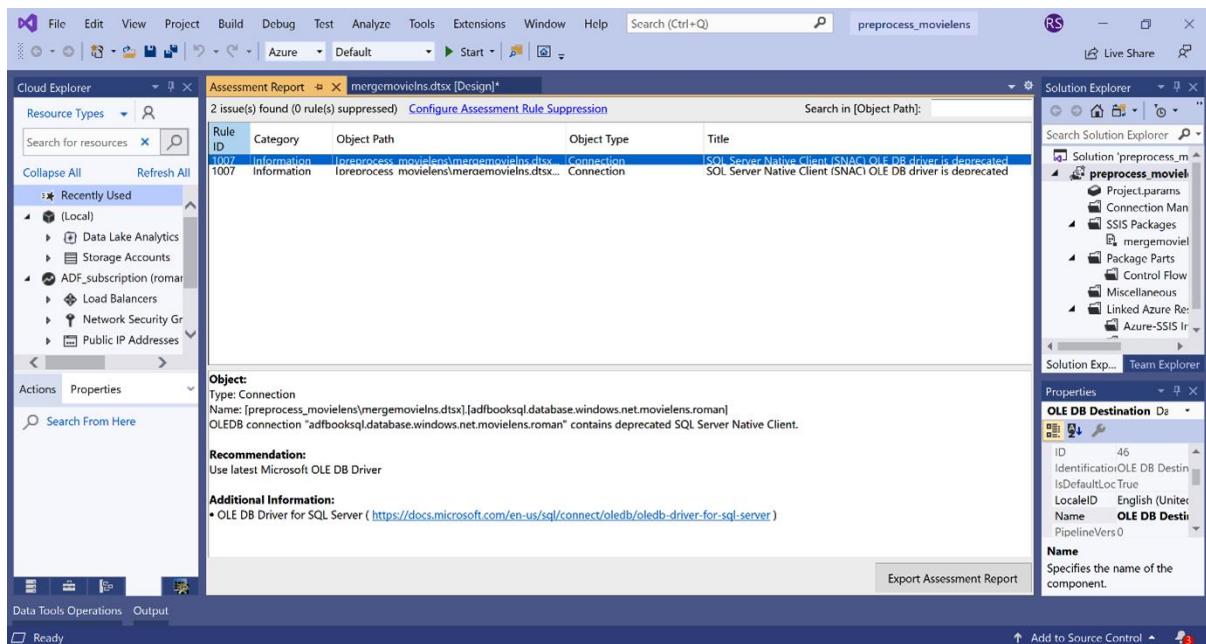
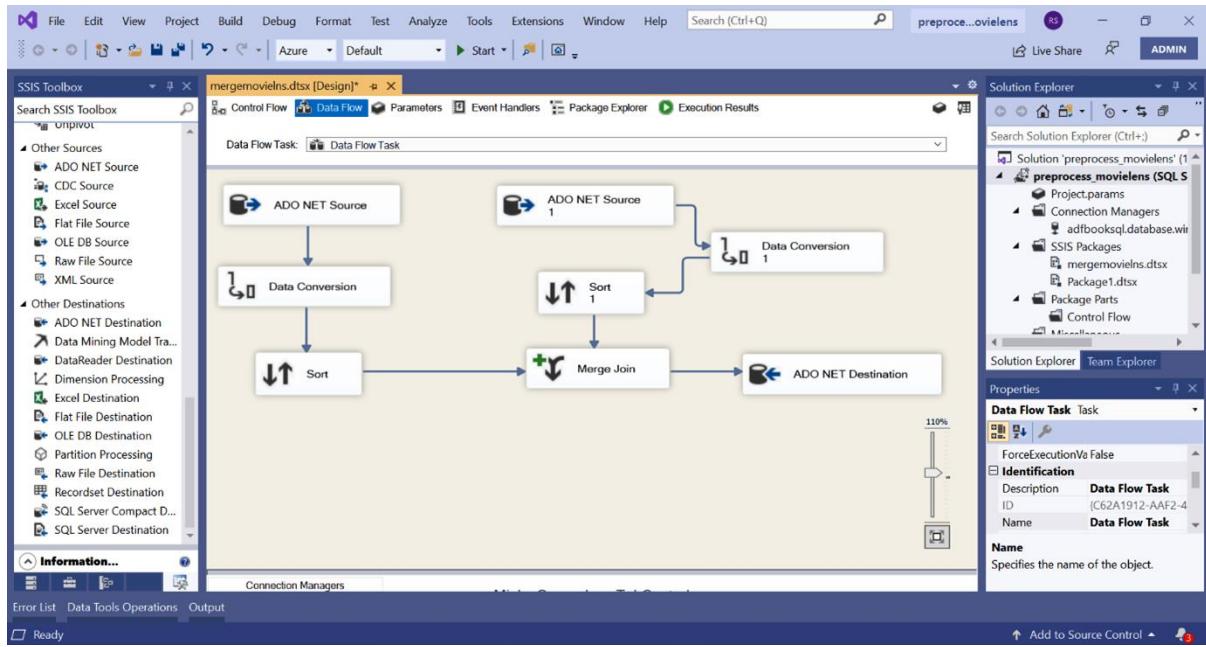
17:12 23/10/2020 ENG 8

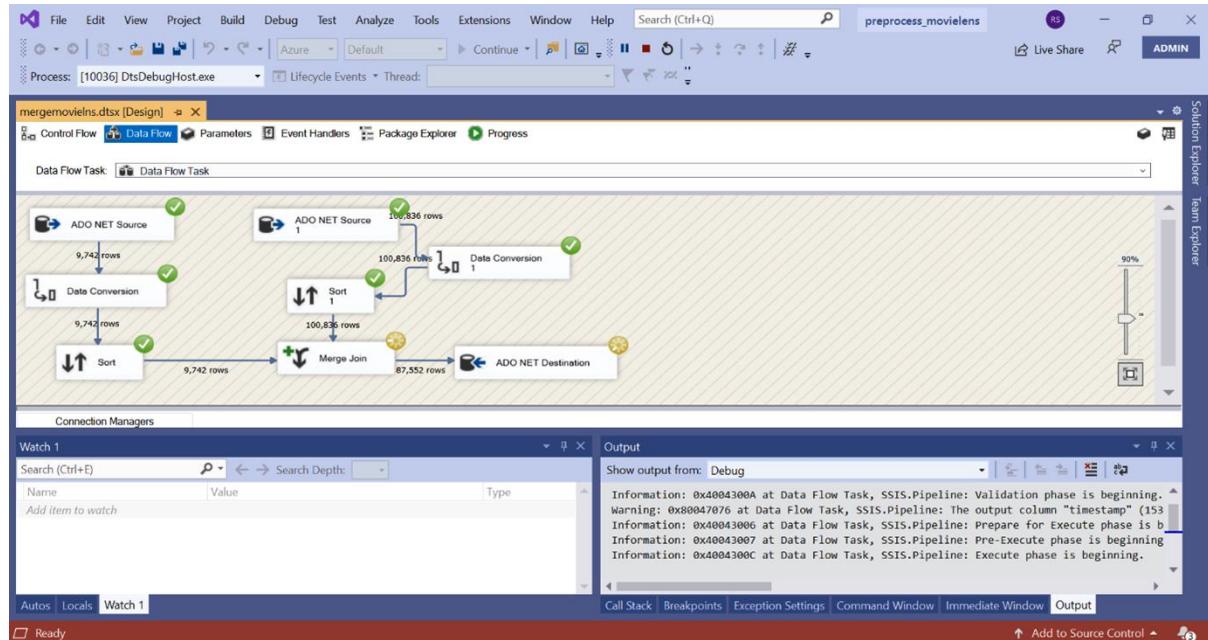
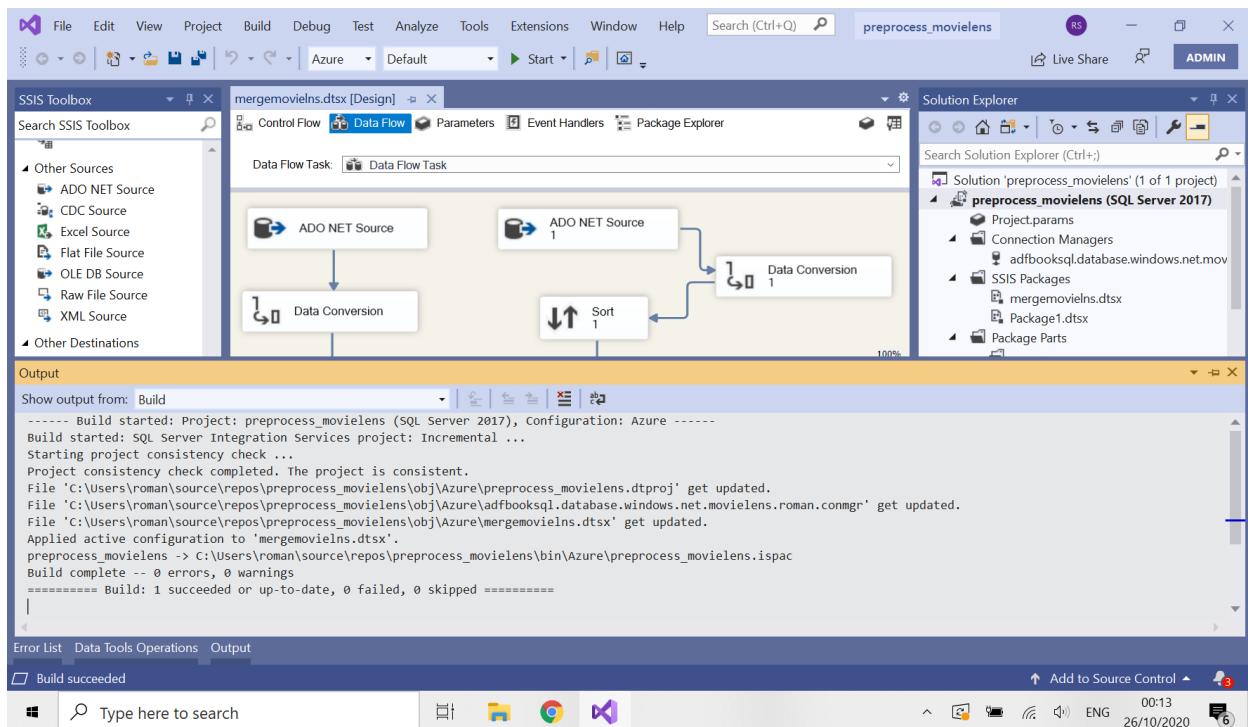












Integration runtimes

The integration runtime (IR) is the compute infrastructure to provide the following data integration capabilities across different network environment. [Learn more](#)

[+ New](#) [Refresh](#)

[Filter by keyword](#)

Showing 1 - 2 of 2 items

Name ↑↓	Type ↑↓	Sub-type ↑↓	Status ↑↓	Region ↑↓	Version ↑↓
AutoResolve... Azure	Public	Running	Auto Resolve	---	
Azure-SSIS	---	Stopped	East US	---	

Edit integration runtime

General settings

Name *

ssis-ir-open

(i)

Description

ssis integration runtime

(i)

Location *

East US

(i)

Node size *

D1_v2 (1 Core(s), 3584 MB)

(i)

Node number *

1

(i)

Edition/license *

Standard

(i)

(i)

Save money

Save with a license you already own. Already have a SQL Server license?

Yes

No

By selecting "yes", I confirm I have a SQL Server license with Software Assurance to apply this
[Azure Hybrid Benefit for SQL Server](#).

Please be aware that the cost estimate for running your Azure-SSIS Integration Runtime is
 $(1 * \text{US\$ 0.592})/\text{hour} = \text{US\$ 0.592}/\text{hour}$, see [here](#) for current prices.

[Continue](#)

[Cancel](#)

Edit integration runtime

Deployment settings

Create SSIS catalog (SSISDB) hosted by Azure SQL Database server/Managed Instance to store your projects/packages/environments/execution logs
(See more info [here](#))

Catalog database server endpoint *

adfbooksqldatabase.windows.net

Use AAD authentication with the managed identity for your Data Factory
(See how to enable it [here](#))

Admin username *

roman

Admin password *

Catalog database service tier *

Basic

Create package stores to manage your packages that are deployed into file system/Azure Files/SQL Server database (MSDB) hosted by Azure SQL Database Managed Instance
(See more info [here](#))

Continue

Back

Cancel

Edit integration runtime

Advanced settings

Maximum parallel executions per node *

1

Customize your Azure-SSIS Integration Runtime with additional system configurations/component installations
(See more info [here](#))

Select a VNet for your Azure-SSIS Integration Runtime to join, allow ADF to create certain network resources, and optionally bring your own static public IP addresses
(See more info [here](#))

Set up Self-Hosted Integration Runtime as a proxy for your Azure-SSIS Integration Runtime
(See more info [here](#))

Continue

Back

Cancel

Microsoft Azure | Data Factory > adfbook

Data Factory Author Monitor Manage

Connections
Linked services
Integration runtimes
Source control
Git configuration
Parameterization template
Author
Triggers
Global parameters
Security
Customer managed key
Managed private endpoints

Integration runtimes

The integration runtime (IR) is the compute infrastructure to provide the following data integration capabilities across different network environment. [Learn more](#)

+ New Refresh

Filter by keyword

Showing 1 - 2 of 2 items

Name	Type	Sub-type	Status	Region	Version
AutoResolvein...	Azure	Public	Running	Auto Resolve	---
ssis-ir-open	Azure-SSIS	---	Stopped	East US	---

File Edit View Project B

SSIS Toolbox Search SSIS Toolbox

- Other Sources
 - ADO.NET Source
 - CDC Source
 - Excel Source
 - Flat File Source
 - OLE DB Source
 - Raw File Source
 - XML Source
- Other Destinations
 - ADO.NET Destination
 - Data Mining Model Tra...
 - DataReader Destination
 - Dimension Processing
 - Excel Destination
 - Flat File Destination
 - OLE DB Destination
 - Partition Processing
 - Raw File Destination
 - Recordset Destination
 - SQL Server Compact D...
 - SQL Server Destination

Information... Error List Data Tools Operations Output Ready

Integration Services Deployment Wizard

Introduction

Deploy Integration Services project or packages.

This wizard deploys Integration Services project or packages to an Integration Services catalog (SSISDB) hosted by SQL Server/Azure SQL Database server/Azure SQL Managed Instance.

There are five steps to completing this wizard:

- Select to deploy a whole project or individual packages incrementally.
- Select the Integration Services project or packages that you want to deploy.
- Select the destination.
- Review your selections.
- Deploy the project or packages.

Click Next to continue.

Do not show this page again

< Previous Next > Deploy Cancel Add to Source Control

Solution Explorer Solution 'preprocess_movieLens' (1)

- preprocess_movieLens (SQL S)
 - Project.params
 - Connection Managers
 - adfbksql.database.wir
 - SSIS Packages
 - mergemovielns.dtsx
 - Package1.dtsx
 - Package Parts
 - Control Flow

Properties Data Flow Task Task

ForceExecutionValue False

Identification

Description Data Flow Task
(C62A1912-AAF2-4)

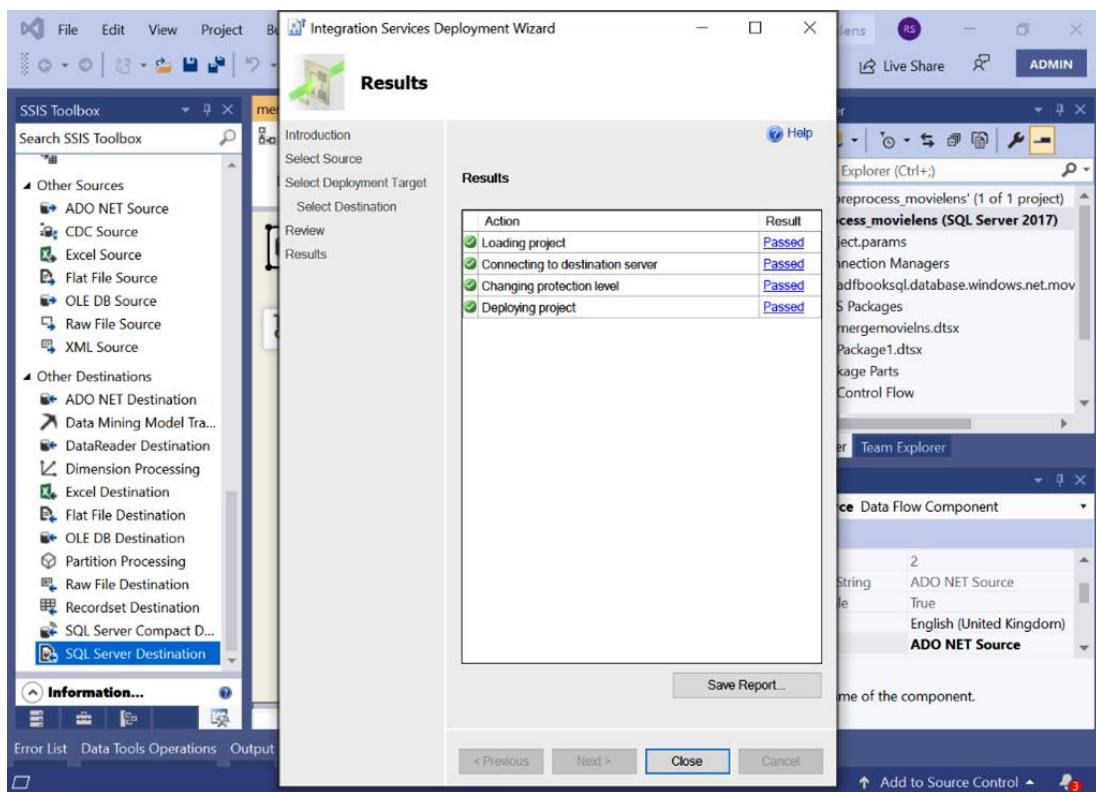
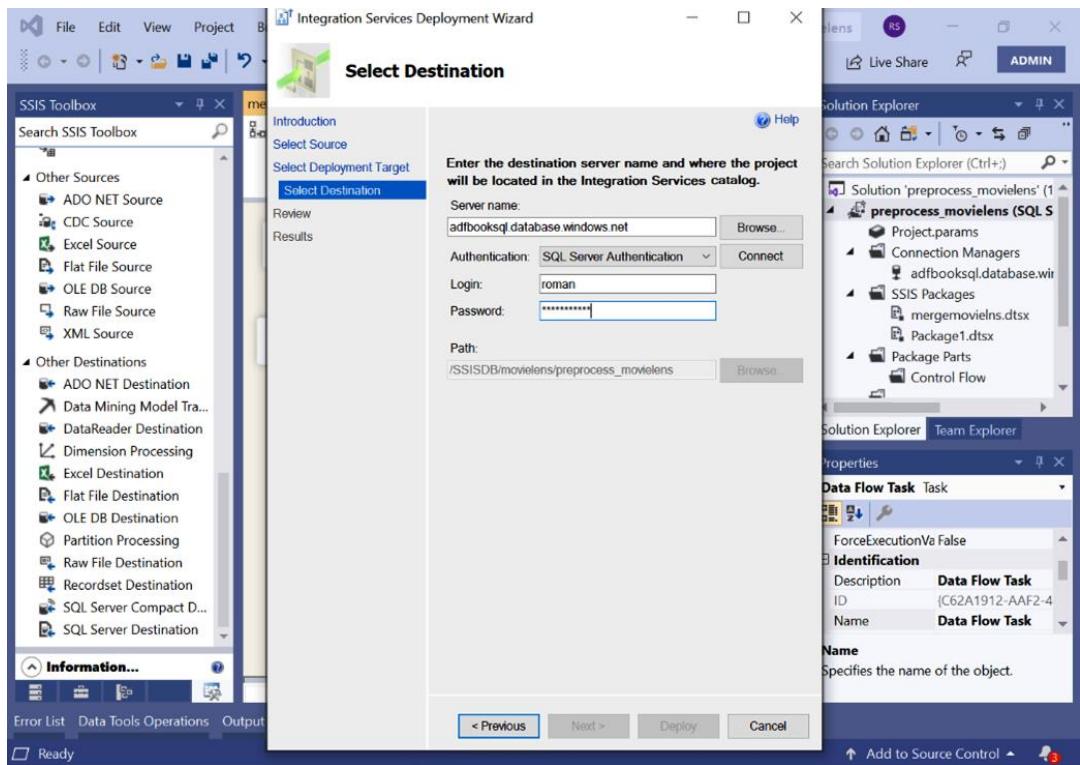
ID C62A1912-AAF2-4

Name Data Flow Task

Name Specifies the name of the object.

Live Share ADMIN

27/10/2020 00:03



Microsoft Azure | Data Factory > adfbook

Activities

Execute SSIS package Movilens

General

Name: Execute SSIS package Movilens

Description: Runs SSIS package

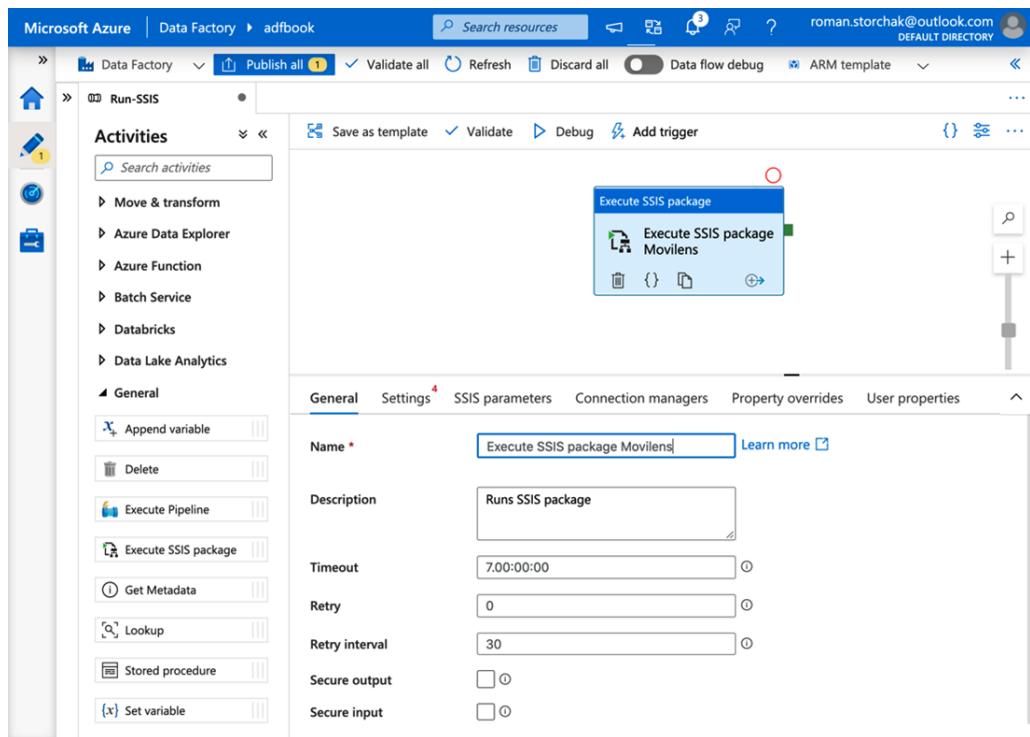
Timeout: 7:00:00:00

Retry: 0

Retry interval: 30

Secure output:

Secure input:



Microsoft Azure | Data Factory > adfbook

Activities

Execute SSIS package Movilens

Settings

Azure-SSIS IR: ssis-ir-open

Windows authentication: 32-bit runtime:

Package location: SSISDB

Folder: movielens Refresh:

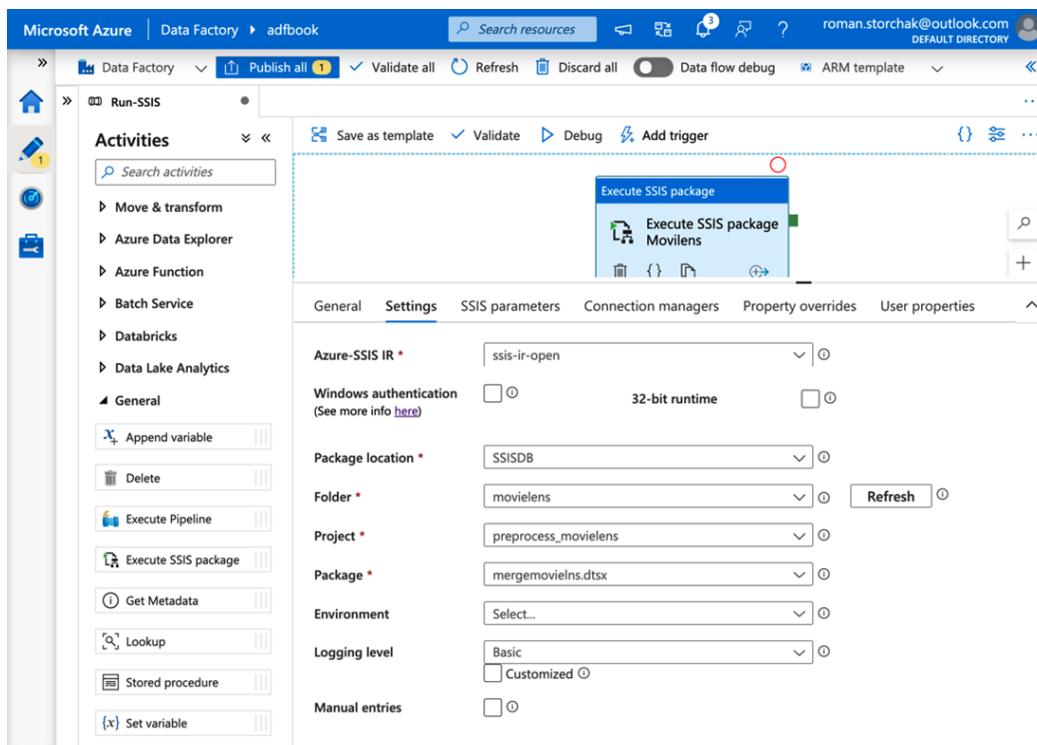
Project: preprocess_movelens

Package: mergemovielns.dtsx

Environment: Select...

Logging level: Basic Customized

Manual entries:



Screenshot of the Azure Data Factory SSIS package configuration interface.

Execute SSIS package Movilens

General Settings

SSIS parameters

Connection managers

Property overrides

User properties

Azure Key Vault

SCOPE	NAME	PROPERTY	TYPE	SENSITIVE	VALUE	AZURE KEY VAULT
Project	adfbooksqldatabi	ConnectByProxy	Boolean	<input type="checkbox"/>	False	<input type="checkbox"/>
		ConnectionString	String	<input type="checkbox"/>	Data Source=adfbooksqldatabas...	<input type="checkbox"/>
		InitialCatalog	String	<input type="checkbox"/>	movielens	<input type="checkbox"/>
		Password	String	<input checked="" type="checkbox"/>	*****	<input type="checkbox"/>
		RetainSameConne	Boolean	<input type="checkbox"/>	False	<input type="checkbox"/>
		ServerName	String	<input type="checkbox"/>	adfbooksqldatabase.windows.net	<input type="checkbox"/>
		UserName	String	<input type="checkbox"/>	roman	<input type="checkbox"/>
Package	adfbooksqldatabi	ConnectByProxy	Boolean	<input type="checkbox"/>	False	<input type="checkbox"/>
		ConnectUsingMar	Boolean	<input type="checkbox"/>	False	<input type="checkbox"/>
		ConnectionString	String	<input type="checkbox"/>	Data Source=adfbooksqldatabas...	<input type="checkbox"/>
		InitialCatalog	String	<input type="checkbox"/>	movielens	<input type="checkbox"/>
		Password	String	<input checked="" type="checkbox"/>	*****	<input type="checkbox"/>
		RetainSameConne	Boolean	<input type="checkbox"/>	False	<input type="checkbox"/>
		ServerName	String	<input type="checkbox"/>	adfbooksqldatabase.windows.net	<input type="checkbox"/>
		UserName	String	<input type="checkbox"/>	roman	<input type="checkbox"/>

Microsoft Azure | Data Factory > adfbook

All pipeline runs > Run-SSIS - Activity runs

Run-SSIS

List Gantt

Rerun Rerun from activity Execute SSIS package Movilens Rerun from failed activity Refresh

Activity runs

Pipeline run ID 88048ec8-8c62-41e8-a157-e021b290df40

All status ▾

Showing 1 - 1 of 1 items

Activity name	Activity type	Run start ↑	Duration	Status	Integration runtime
Execute SSIS package Movilens	ExecuteSSISPack	10/26/20, 8:50:32 PM	00:02:20	✓ Succeeded	ssis-ir-open (East US)

Microsoft Azure Search resources, services, and docs (G+) Dashboard > SQL databases X

Default Directory

+ Add Reservations Edit columns Refresh | Assign tags Delete

Try our new Azure SQL resource browser! This experience offers a unified view of all your SQL Server resources in Azure as well as improved sorting and filtering. Click here to go to the new experience.

Subscriptions: ADF_subscription – Don't see a subscription? [Open Directory + Subscription settings](#)

Filter by name... All resource groups All locations All tags No grouping

0 items

Name ↑↓	Status	Replication role	Server	Pricing tier	Location
---------	--------	------------------	--------	--------------	----------

Microsoft Azure Search resources, services, and docs (G+) Database details Microsoft

Enter required settings for this database, including picking a logical server and configuring the compute resources

Database name * movielens

Server * Select a server Create new The value must not be empty.

Want to use SQL elastic pool? * Yes No

Compute + storage * Please select a server first. [Configure database](#)

Server name * adfbooksqldatabase.windows.net

Server admin login * roman

Password * *****

Confirm password * *****

Location * (US) East US

Review + create **Next : Networking >** **OK**

Configure

[Feedback](#)

Basic

For less demanding workloads

Standard

For workloads with typical performance requirements

Premium

For IO-intensive workloads.

DTUs [What is a DTU?](#)

5 (Basic)

Data max size



Microsoft Azure Search resources, services, and docs (G+) 1 roman.storchak@outlook.com DEFAULT DIRECTORY

Dashboard > SQL databases > **Create SQL Database** X

Microsoft

Basics Networking Additional settings Tags Review + create

Configure network access and connectivity for your server. The configuration selected below will apply to the selected server 'adfbooksq1' and all databases it manages. [Learn more](#)

Network connectivity

Choose an option for configuring connectivity to your server via public endpoint or private endpoint. Choosing no access creates with defaults and you can configure connection method after server creation. [Learn more](#)

Connectivity method * ①

No access
 Public endpoint
 Private endpoint

Review + create < Previous Next : Additional settings >

Microsoft Azure Search resources, services, and docs (G+) Dashboard SQL databases roman.storchak@outlook.com DEFAULT DIRECTORY

Create SQL Database

Microsoft

Data source

Start with a blank database, restore from a backup or select sample data to populate your new database.

Use existing data * None Backup Sample

Database collation

Database collation defines the rules that sort and compare data, and cannot be changed after database creation. The default database collation is SQL_Latin1_General_CI_AS. [Learn more](#)

Collation * SQL_Latin1_General_CI_AS [Find a collation](#)

Azure Defender for SQL

Protect your data using Azure Defender for SQL, a unified security package including vulnerability assessment and advanced threat protection for your server. [Learn more](#)

Get started with a 30 day free trial period, and then 15 USD/server/month.

Enable Azure Defender for SQL * Start free trial Not now

Review + create < Previous Next : Tags >

New linked service (Azure S)

Name * SQLlinkedService

Description

Connect via integration runtime * AutoResolveIntegrationRuntime

Connection string Azure Key Vault

Account selection method From Azure subscription Enter

Azure subscription ADF_subscription (245d4f8f-e6c4-439

Server name *

Connection failed

Cannot connect to SQL Database: 'adfbbooksql.database.windows.net', Database: 'movielens', User: 'roman'. Check the linked service configuration is correct, and make sure the SQL Database firewall allows the integration runtime to access.

Cannot open server 'adfbbooksql' requested by the login. Client with IP address '20.42.2.58' is not allowed to access the server. To enable access, use the Windows Azure Management Portal or run sp_set_firewall_rule on the master database to create a firewall rule for this IP address or address range. It may take up to five minutes for this change to take effect.

SqlErrorNumber=40615,Class=14,State=1, Activity ID: 2b6dc2b0-a43b-44c7-bab7-b594fffd95a.

This may happen if your data source only allows secured connections. If that's the case, please use a VNet integration runtime.

How helpful or unhelpful was this error message? ★ ★ ★ ★ ★

Connection failed. More Test connection Cancel

Chapter 7: Data Migration – Azure Data Factory and Other Cloud Services

Choose the output file or folder
Specify a folder that will contain output files or a specific output file in the destination data store.

Folder path * [Browse](#)

If the identity you use to access the data store only has permission to subdirectory instead of the entire account, specify the path to browse.

File name

Copy behavior ⓘ

Max concurrent connections

Block size (MB) ⓘ

Copy Data

Properties
One time copy

Source
Amazon S3

Connection
Dataset

Destination
Azure Blob Storage

Connection
Dataset

Settings

Summary

Deployment

Summary
You are running pipeline to copy data from Amazon S3 to Azure Blob Storage.

Amazon S3 → Azure Blob Storage

Properties

Task name

Task description

Source

Connection name AmazonS3Bucket

Dataset name SourceDataset_8z2

Column delimiter ,

Escape character \

Quote char "

First row as header true

Destination

Connection name AzureBlobFromS3

Dataset name DestinationDataset_8z2

Column delimiter ,

Escape character \

Quote char "

[Previous](#) [Next >](#)



Binary
CommonCrawlS3DataSet

[Connection](#) [Parameters](#)

Linked service * [Test connection](#) [Edit](#) [New](#)

File path * / / [Browse](#) | [...](#)

Compression type



Binary
AzureCommonCrawl

[Connection](#) [Parameters](#)

Linked service * [Test connection](#) [Edit](#) [New](#)

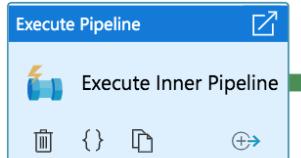
File path * / / [Browse](#) | [...](#)

Compression type

[Parameters](#) [Variables](#) [Output](#)

[+ New](#) | [Delete](#)

<input type="checkbox"/>	NAME	TYPE	DEFAULT VALUE
	<input type="text" value="Path"/>	<input type="text" value="String"/>	<input type="text" value="Value"/>



General Settings User properties

Invoked pipeline * [Open](#) [New](#)

Wait on completion

Parameters

Name	Type	Value
Path	string	<input type="text" value="@{item().Path}"/> Add dynamic content [Alt+P]

<input type="checkbox"/> Pipeline name	Run start ↗↓	Run end	Duration	Triggered by	Status	Parameters	Annotations	Error
<input type="checkbox"/> pl_commoncrawl_data_inner	8/29/20, 3:08:15 PM	--	00:02:23	6c0918e1-5250-4150-83	In progress	@1		
<input type="checkbox"/> pl_commoncrawl_data_inner	8/29/20, 3:08:15 PM	--	00:02:23	86d64fda-aeba-49bb-ab	In progress	@2		
<input type="checkbox"/> pl_commoncrawl_data_outer	8/29/20, 3:08:10 PM	8/29/20, 3:08:18 PM	00:00:08	Manual trigger	Succeeded			

Settings

PROJECT ACCESS

INTEROPERABILITY

The Interoperability API allows Google Cloud Storage to interoperate with tools written for other cloud storage systems. This enables you to [run migrations to Cloud Storage](#) and to authenticate both user and service accounts using keyed-hash message authentication codes (HMAC). [Learn more](#)

Request endpoint

Make sure the request endpoint in the tools or libraries you use with other cloud storage systems (e.g., Amazon S3) uses the Cloud Storage URI.

Storage URI

<https://storage.googleapis.com>



Service account HMAC

Use access keys with your organization's Cloud Platform service accounts when you don't want to tie HMAC authentication to specific user accounts. Recommended for production workloads. [Learn more](#)

- Each service account can use up to five keys.
- Note that keys must be deactivated before they can be deleted.
- Grant your service accounts the required permissions for their intended operations – typically this is the IAM Storage Object Admin role.

Access keys for service accounts

This project doesn't have any service account HMAC keys.

[+ CREATE A KEY FOR A SERVICE ACCOUNT](#)

Select a service account

Search service accounts

Email	Name	Keys
adf-service-account@tranquil-apogee-285814.iam.gserviceaccount.c...	adf-service-accou...	3

CANCEL

[CREATE KEY](#)

[CREATE NEW ACCOUNT](#)

New linked service (Google Cloud Storage (S3 API))

Name *

Description

Connect via integration runtime *



Access Key ID *

Secret Access Key

Azure Key Vault

Secret Access Key *



Service URL



Annotations

+ New

Name

Advanced ⓘ

✓ Connection successful

Create

Back

🔗 Test connection

Cancel

Connection Schema Parameters

Linked service * AzureBlobStorageLinkedService Test connection Edit + New

File path * datafromgoogle / Directory / File

Compression type none

Column delimiter Comma (,) Edit

Row delimiter Auto detect (\r,\n, or \r\n) Edit

Encoding Default(UTF-8)

Escape character Backslash (\) Edit

Quote character Double quote ("") Edit

First row as header

Null value

Add members to "BigQueryToAzure"

Add members, roles to "BigQueryToAzure" project

Enter one or more members below. Then select a role for these members to grant them access to your resources. Multiple roles allowed. [Learn more](#)

New members

adf.cookbook@gmail.com



Role

BigQuery Job User

Access to run jobs

Condition

[Add condition](#)



[+ ADD ANOTHER ROLE](#)

Send notification email

This email will inform members that you've granted them access to this role for "BigQueryToAzure"

SAVE

CANCEL

New linked service (Google BigQuery)

Name *

Description

Connect via integration runtime *



Project ID *

Additional project IDs

Request access to Google Drive

True False

Authentication type *



Client ID

Client secret

Azure Key Vault

Client secret

Refresh Token

Azure Key Vault

Refresh Token

Annotations

✓ Connection successful

Create

Back

🔗 Test connection

Cancel

Connection Parameters

✓ Connection successful

Linked service *

GoogleBigQueryToAzure2

🔗 Test connection

>Edit



New

Table

bqquery-public-data

github_repos.sample_files

🔗 Preview data

Add dynamic content [Alt+P]

Edit

Create a new app on the DBX Platform



ADF Cookbook ▾

1. Choose an API

Scoped access **New**

Select the level of access your app needs to Dropbox data. [Learn more](#)

Dropbox Legacy API (Deprecated) - This uses the legacy app permissions model. Not recommended. [Learn more](#)

Dropbox Business API (Deprecated) - This uses the legacy app permissions model. Not recommended. [Learn more](#)



2. Choose the type of access you need

[Learn more about access types](#)

App folder – Access to a single folder created specifically for your app.

Full Dropbox – Access to all files and folders in a user's Dropbox.

3. Name your app

UploadData

Create app

Account Info

Permissions that allow your app to view and manage Dropbox account info

- account_info.read Read Dropbox account information (Required)
- account_info.write Create, modify, and delete Dropbox account information

Files and folders

Permissions that allow your app to view and manage files and folders

- files.content.read Read file data
- files.content.write Create, modify, and delete file data
- files.metadata.read Read file metadata
- files.metadata.write Create, modify, and delete file metadata
- files.permanent_delete Permanently delete files. This scope needs to be authorized by a team admin during OAuth flow

Redirect URIs

https:// (http allowed for localhost)

Add

Allow public clients (Implicit Grant & PKCE) i

Allow

Generated access token i

Generate

Access token expiration i

Short-lived

data																																		
Container																																		
Search (Cmd+Shift+F)																																		
Upload		Change access level Refresh Delete Change tier Acquire lease Break lease View snapshots Create snapshot																																
Overview		Authentication method: Access key (Switch to Azure AD User Account) Location: data																																
Access Control (IAM)		Search blobs by prefix (case-sensitive) <input type="checkbox"/> Show deleted blobs																																
Settings		<table><thead><tr><th>Name</th><th>Modified</th><th>Access tier</th><th>Blob type</th><th>Size</th><th>Lease state</th><th>...</th></tr></thead><tbody><tr><td><input type="checkbox"/> airlines.csv</td><td>8/23/2020, 4:37:57 PM</td><td>Hot (Inferred)</td><td>Block blob</td><td>228.37 KiB</td><td>Available</td><td>...</td></tr><tr><td><input type="checkbox"/> airports.csv</td><td>8/23/2020, 4:37:58 PM</td><td>Hot (Inferred)</td><td>Block blob</td><td>1.08 MiB</td><td>Available</td><td>...</td></tr><tr><td><input type="checkbox"/> upload_2_dropbox.py</td><td>8/23/2020, 4:37:36 PM</td><td>Hot (Inferred)</td><td>Block blob</td><td>2.5 KiB</td><td>Available</td><td>...</td></tr></tbody></table>					Name	Modified	Access tier	Blob type	Size	Lease state	...	<input type="checkbox"/> airlines.csv	8/23/2020, 4:37:57 PM	Hot (Inferred)	Block blob	228.37 KiB	Available	...	<input type="checkbox"/> airports.csv	8/23/2020, 4:37:58 PM	Hot (Inferred)	Block blob	1.08 MiB	Available	...	<input type="checkbox"/> upload_2_dropbox.py	8/23/2020, 4:37:36 PM	Hot (Inferred)	Block blob	2.5 KiB	Available	...
Name	Modified	Access tier	Blob type	Size	Lease state	...																												
<input type="checkbox"/> airlines.csv	8/23/2020, 4:37:57 PM	Hot (Inferred)	Block blob	228.37 KiB	Available	...																												
<input type="checkbox"/> airports.csv	8/23/2020, 4:37:58 PM	Hot (Inferred)	Block blob	1.08 MiB	Available	...																												
<input type="checkbox"/> upload_2_dropbox.py	8/23/2020, 4:37:36 PM	Hot (Inferred)	Block blob	2.5 KiB	Available	...																												

Custom deployment

Deploy from a custom template

Learn about template deployment

 [Read the docs](#) 

 [Build your own template in the editor](#)

Common templates

 [Create a Linux virtual machine](#)

 [Create a Windows virtual machine](#)

 [Create a web app](#)

 [Create a SQL database](#)

Load a GitHub quickstart template

Select a template (disclaimer) 

[Home](#) >
Custom deployment
Deploy from a custom template

TEMPLATE

 Customized template
2 resources [Edit tem...](#) [Edit para...](#) [Learn m...](#)

BASICS

Subscription * [Azure subscription 1](#)
Resource group * [Create new](#)
Location * [\(US\) West US](#)

SETTINGS

Batch
Accounts_adfdatamigration [adfdatamigration](#)
Pools_upload_2_dropbox_p [^4Tm0LnZ&U188T](#)

TERMS AND CONDITIONS

[Azure Marketplace Terms](#) | [Azure Marketplace](#)

By clicking "Purchase," I (a) agree to the applicable legal terms associated with the offering; (b) authorize Microsoft to charge or bill my current payment method for the fees associated with the offering(s), including applicable taxes, with the same billing

I agree to the terms and conditions stated above

adfdatamigration | Pools

Batch account

Search (Cmd+)

+ Add + Add (JSON editor) Columns Refresh

Overview Activity log Access control (IAM) Tags

Settings Quick start Properties Quotas Identity Encryption Storage account Keys Locks Export template

Features Applications Pools Jobs Job schedules Certificates

Monitoring Alerts Metrics Diagnostic settings Logs Advisor recommend...

Support +

Queries: All pools

Advanced query

Search for pool by ID or display name

Pool ID	Dedicated nodes	Low-priority nodes	Current vCPUs	VM size	Allocation state
upload-2-dropbox	1	0	0	basic_a1	Steady

upload-2-dropbox ✎

Pool

Search (Cmd+ /) < Refresh + Add job Scale Delete View node

Overview

General

Properties

Nodes

Settings

Certificates

Start task

Application packages

Scale

Metadata

Current vCPUs : 1
Dedicated nodes : 1
Low-priority nodes : 0
Auto scale : false
Metadata : Click here to view Metadata

Batch account : adfdatamigration
Pool ID : upload-2-dropbox
Operating System : microsoft-dsvm dsvm-win-2019 server-2019 (latest)
VM size : basic_a1
Allocation state : Steady

Nodes heat map



adfdatamigration | Keys ✎

Batch account

Search (Cmd+ /) Regenerate primary Regenerate secondary

Overview

Activity log

Access control (IAM)

Tags

Settings

Quick start

Properties

Quotas

Identity

Encryption

Storage account

Keys

Locks

Export template

Features

Applications

Pools

Jobs

Job schedules

Certificates

Monitoring

Alerts

Metrics

Diagnostic settings

BATCH ACCOUNT CREDENTIALS

Batch account	adfdatamigration
URL	https://adfdatamigration.eastus.batch.azure.com
Primary access key	UqCepDDbkpz0/B9tLRsiPQqak/aNC0NPZGnLG1mic509DMHwT/27CmwiySl6p78bG5Mv925cW/n9XYYmEfAB7A==
Secondary access key	C0zp5dO0he0ndh9Q5gagZ1z7d4X+GuZ0jqIzG5LQQ/fcdvEGPd7HVCxz3H2rcdJ1JjEtuzBUIEKXszbUkoOlwg==

New linked service

Data store Compute

Search

Azure Batch	Azure Data Lake Analytics	Azure Databricks
Azure Function	Azure HDInsight	Azure Machine Learning

Continue

Cancel

New linked service (Azure Batch)

Name *

Description

Connect via integration runtime *



Account name *

 Access key Azure Key Vault

Access key *

Batch URL *

Pool name *

Storage linked service name *

Annotations

New

Name

Advanced

Connection successful

Test connection

Custom

Azure Storage to Dropbox

General Azure Batch Settings User properties

Command * python upload_2_dropbox.py

Resource linked service * BatchServiceStorageLinkedService Test connection Edit New

Folder path * data Browse storage Browse local

Reference objects Add reference object

Retention time in days

Advanced

Extended properties New

Home > Batch accounts > adfdatamigration | Pools >

upload-2-dropbox

Pool

Search (Cmd+ /) Refresh Add job Scale Delete View node

Overview

General

Properties

Nodes

Nodes

Current vCPUs : 1
Dedicated nodes : 1
Low-priority nodes : 0
Auto scale : false
Metadata : Click here to view Metadata

Settings

Certificates

Start task

Application packages

Scale

Metadata

Nodes heat map

tvmmps_2db8332fbfea12082df0a4ac76dee8b6450f1fec3393849444fd9a3aafe1f98_d

Search (Cmd+/) <> Refresh Reboot Reimage || Disable Connect Delete node Upload batch logs

Overview

General

- Properties
- Files
- Recent tasks
- Start task info
- Certificate references

Users

- Add user account
- Update user account
- Remove user account

Batch account : adfdatamigration
Pool ID : upload-2-dropbox
VM size : basic_a1
Operating System : microsoft-dsvm dsvm-win-2019 server-2019 (latest)
Total tasks run : 2 (0 Running tasks, 1 Total tasks succeeded)

Essentials

Path://

Location: root /

File name	Creation
applications	
fsmounts	
shared	
startup	
volatile	
workitems	

Location: root / workitems / adfv2-upload-2-dropbox / job-1 / fffc6250-cc65-4d98-8b7c-7164b3a2adfc

File name	Creation time	Last modified	Size	Content type	File mode
[.]					
wd					
stderr.txt	Aug 23, 16:45:14	Aug 23, 16:45:24	972 Bytes	text/plain	
stdout.txt	Aug 23, 16:45:14	Aug 23, 16:45:24	163 Bytes	text/plain	

New linked service (Amazon S3)

i If the identity you use to access the data store only has permission to subdirectory instead of the entire account, specify the path to test connection.

Name *

Description

Connect via integration runtime *

(i)

▼

Access Key ID *

Secret Access Key

Azure Key Vault

Secret Access Key *



Service URL

(i)

Test connection

To linked service To file path

Annotations

+ New

Name

Advanced (i)

✓ Connection successful

Create

Back

🔗 Test connection

Cancel

File format settings

File format ⓘ

Column delimiter
 ⓘ
 Edit

Row delimiter
 ⓘ
 Edit

Skip line count ⓘ

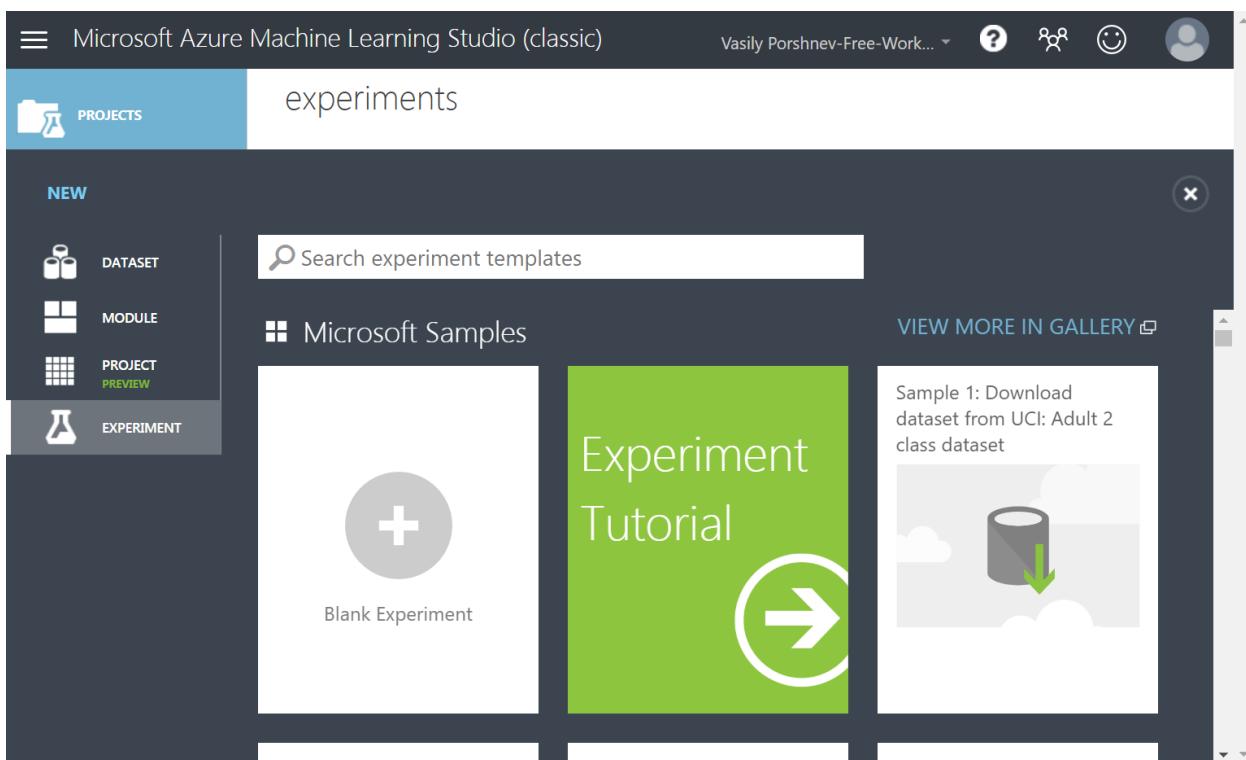
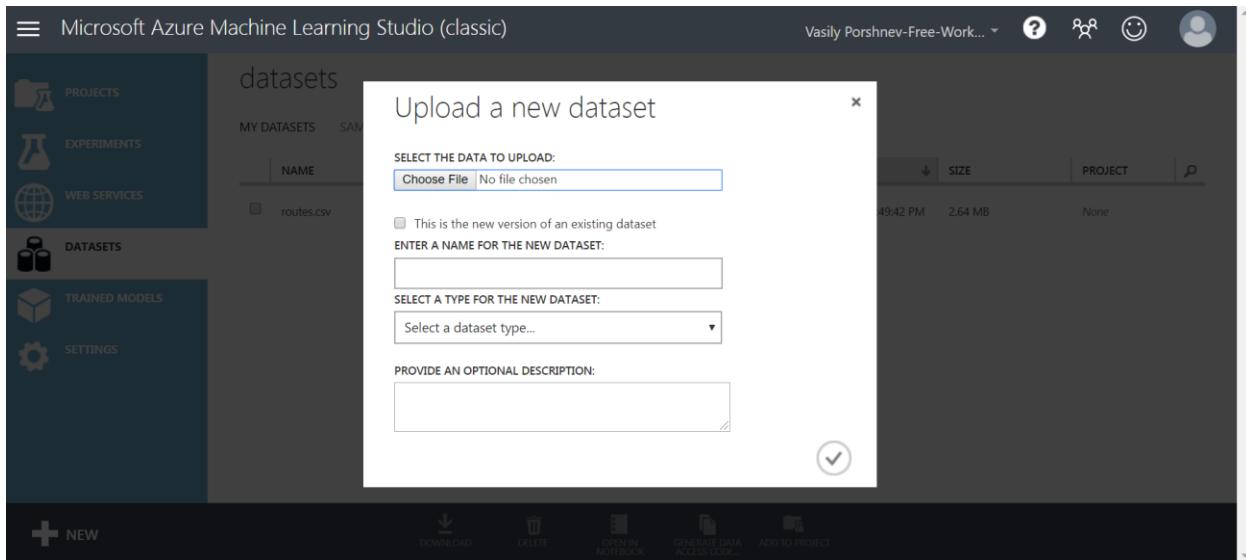
First row as header ⓘ

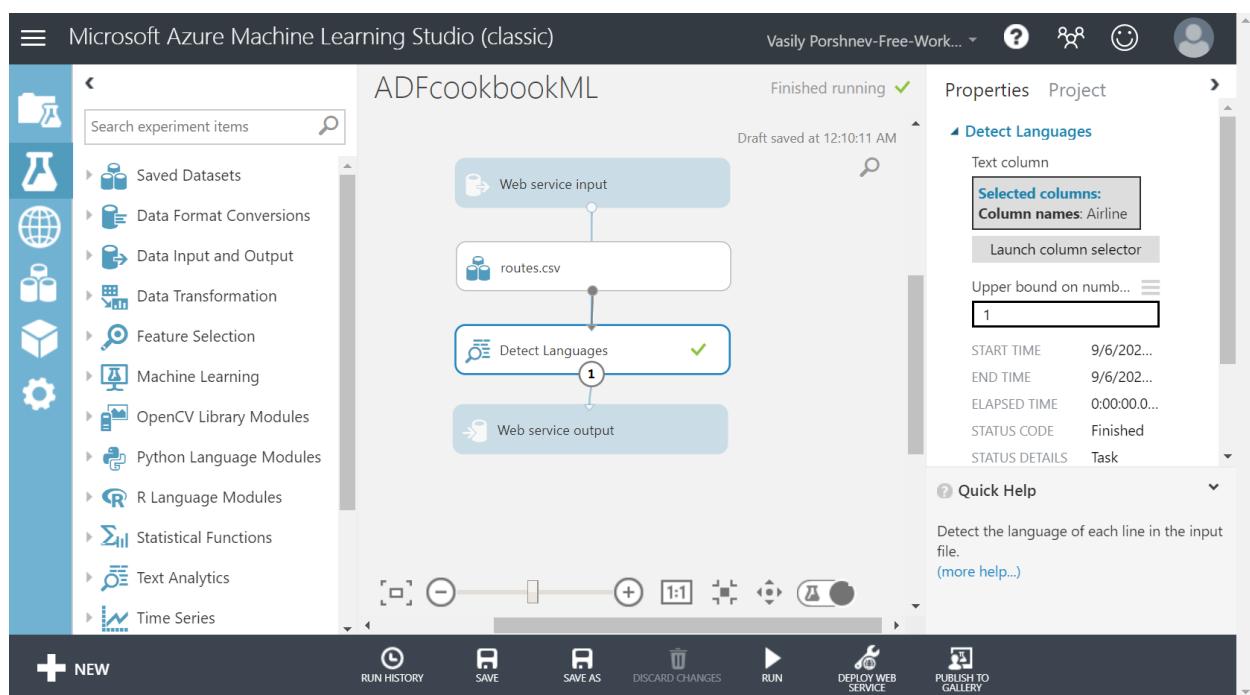
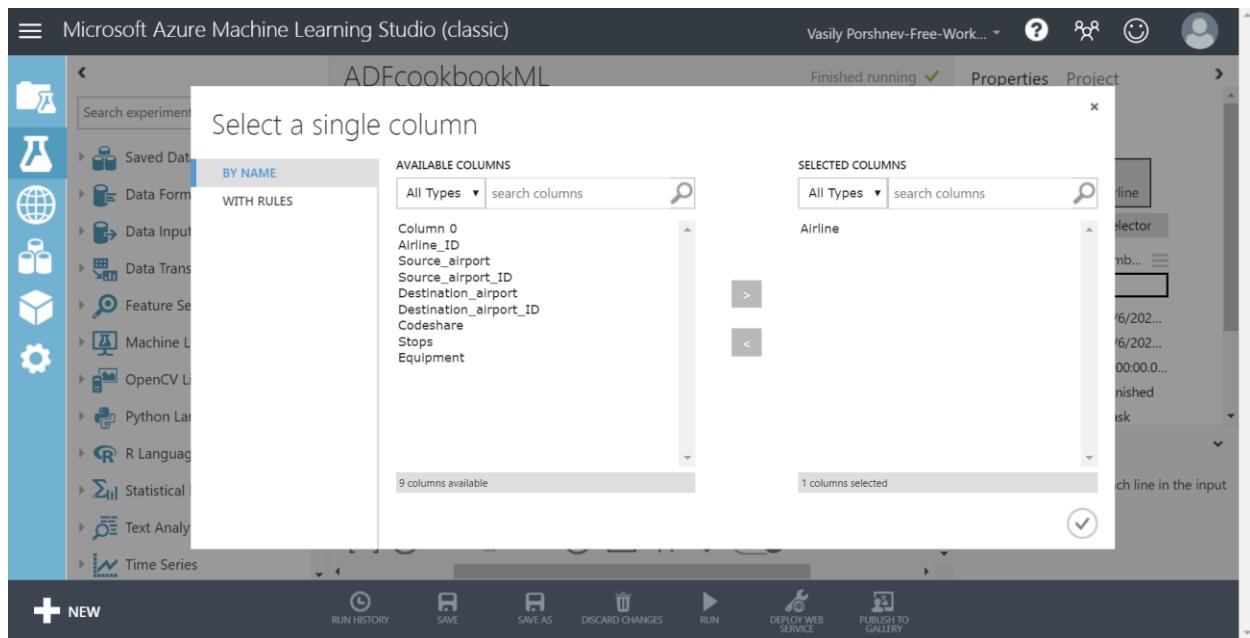
► Advanced

Compression type
 ⓘ

Additional columns ⓘ
+ New

Chapter 8: Working with Azure Services Integration





Microsoft Azure Machine Learning Studio (classic) Vasily Porshnev-Free-Work...    

adfcookbookml

DASHBOARD CONFIGURATION

General New Web Services Experience [preview](#)

Published experiment

[View snapshot](#) [View latest](#)

Description

No description provided for this web service.

API key

```
q+UvuEo8E48BKwRMU5aU0DhprT7VBf8h+ltHalVH7UCw7lEooTf4RpokBpLmpbplgPgNqaVcelR8h00oZPEPow==
```

Default Endpoint

API HELP PAGE TEST APPS LAST UPDATED 

REQUEST/RESPONSE	Test preview	Excel 2013 or later Excel 2010 or earlier work	9/6/2020 11:59:49 PM
BATCH EXECUTION	Test preview	Excel 2013 or later workbook	9/6/2020 11:59:49 PM

 NEW  DELETE

Batch Execution API Documentation for ADFcookbookML

Updated: 09/06/2020 20:59

No description provided for this web service.

- [Previous version of this API](#)
- [Submit a Job](#)
- [Start a Job](#)
- [Get Job Status](#)
- [Delete a Job](#)
- [Web App Template for BES](#)
- [BES SDK](#)
- [Sample Code](#)
- [API Swagger Document](#) 
- [Endpoint Management Swagger Document](#) 

Submit (but not start) a Batch Execution job

Request

Method	Request URI	HTTP Version
POST	https://ussouthcentral.services.azureml.net/workspaces/56eaa3840c6b47068ced50b6517ba76b/services/190503e3f5f24be5a68a9c2860dbe2de/jobs?api-version=2.0	HTTP/1.1

Microsoft Azure | adfcookbookch8adf

Factory Resources

Pipelines: Delete_blob_LA, ADFcookbookMLpipe, Pipeline_with_Function

Datasets: 1

Data flows: 0

Activities

Move & transform, Azure Data Explorer, Azure Function, Batch Service, Databricks, Data Lake Analytics, General, HDInsight, Iteration & conditionals, Machine Learning

Machine Learning Batch Execution1

New linked service (Azure Machine Learning Studio (classic))

Name *: AzureMILLS

Description:

Connect via integration runtime *: AutoResolveIntegrationRuntime

Endpoint *: https://ussouthcentral.services.azureml.net/workspaces/56eaa3840c6b47068ced50b6517ba76b/ser

API key: Azure Key Vault

API key *: [REDACTED]

Enable update resource: Disable update resource:

Annotations: + New

Name:

Advanced: Connection successful Test connection

Create

Microsoft Azure | adfcookbookch8adf

Factory Resources

Pipelines: Delete_blob_LA, ADFcookbookMLpipe, Pipeline_with_Function

Datasets: 1

Data flows: 0

Activities

Move & transform, Azure Data Explorer, Azure Function, Batch Service, Databricks, Data Lake Analytics, General, HDInsight, Iteration & conditionals, Machine Learning

Machine Learning Batch Execution1

Properties

General

Name *: ADFcookbookMLpipe

Description:

Concurrency:

Annotations: + New

Name:

Settings

Web service inputs

+ New | Delete

INPUT KEY: input1, INPUT STORAGE: AzureBlobStorageLS, flightscontainer/routes.csv, Test connection, Edit, Browse

Web service outputs

+ New | Delete

OUTPUT KEY: output1, OUTPUTS STORAGE: AzureBlobStorageLS, flightscontainer/outputML.csv, Test connection, Edit, Browse

Microsoft Azure [Upgrade](#) [Search resources, services, and docs \(G+/\)](#)

Dashboard > All resources > adfcookbookch8adls >

flightscontainer

Container

Search (Ctrl+ /) [Upload](#) [Change access level](#) [Refresh](#) [Delete](#) [Change tier](#) [Acquire lease](#) [Buckets](#)

Authentication method: Access key ([Switch to Azure AD User Account](#))
Location: flightscontainer

Search blobs by prefix (case-sensitive) Show deleted blobs

Name	Modified	Access tier	Blob type	Size	Lease state	...
airlines.csv	9/5/2020, 7:16:26 PM	Hot (Inferred)	Block blob	320.49 KiB	Available	...
countries.csv	9/5/2020, 7:16:24 PM	Hot (Inferred)	Block blob	4.6 KiB	Available	...
outputML.csv	9/7/2020, 12:04:35 AM	Hot (Inferred)	Block blob	3.71 MiB	Available	...
routes.csv	9/7/2020, 12:04:09 AM	Hot (Inferred)	Block blob	2.64 MiB	Available	...

Microsoft Azure [Upgrade](#) [Search resources, services, and docs \(G+/\)](#)

Home > All resources > New > Azure Databricks >

Azure Databricks Service

*Basics 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 * [\(i\)](#)

Resource group * [\(i\)](#) [Create new](#)

Instance Details

Workspace name *

Location *

Pricing Tier * [\(i\)](#)

[Review + Create](#) [Next : Networking >](#)

Azure
Databricks

Home



Workspace



Recents



Data



Clusters



Jobs



Models



Create Cluster



adfcookbookadb

New Cluster

Cancel

Create Cluster

2-8 Workers: 28.0-112.0 GB Memory, 8-32 Cores, 1.5-6 DBU
1 Driver: 14.0 GB Memory, 4 Cores, 0.75 DBU UI | [JSON](#)

Cluster Name

adfcookbookcluster

Cluster Mode

Standard

Pool

None

Databricks Runtime Version

[Learn more](#)

Runtime: 7.0 (Scala 2.12, Spark 3.0.0)

New This Runtime version supports only Python 3.

Autopilot Options

 Enable autoscaling Terminate after minutes of inactivity

Worker Type

Min Workers Max Workers

Standard_DS3_v2

14.0 GB Memory, 4 Cores, 0.75 DBU

2

8

Driver Type



The screenshot shows the Microsoft Azure Databricks workspace interface. On the left is a vertical sidebar with icons for Azure Databricks, Home, Workspace, Recents, Data, Clusters, Jobs, and Models. The 'Workspace' icon is selected. The main area shows a 'Workspace' sidebar with options: Shared, Users, adfcookbooknb (which is highlighted), and mount. A context menu is open over the 'adfcookbooknb' item, showing 'Create' as the active option. The 'Create' menu has sub-options: Notebook, Library, Folder, and MLflow Experiment. The URL at the bottom of the page is <https://adb-1486914505250902.2.azure.databricks.net/?o=1486914505250902#>.

Microsoft Azure

PORTAL vasily.porshnev@outlook.com

adfcookbooknb (Python)

adfcookbookcluster

```
1 print('ADFcookbook')
```

ADFcookbook

Command took 0.12 seconds -- by vasily.porshnev@outlook.com at 9/7/2020, 4:37:26 PM on adfcookbookcluster

Shift+Enter to run shortcuts

Azure Databricks

- Home
- Workspace
- Recents
- Data
- Clusters
- Jobs
- Models
- Search

Microsoft Azure | adfcookbookch8adf

Factory Resources

Pipelines: ADFcookbookMLpipe, Delete_blob_LA, pipeline1, Pipeline_with_Function

Datasets: 1

Data flows: 0

Activities

- Move & transform
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Notebook
- Jar
- Python
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning

Save as template Validate Add trigger

Notebook

New linked service (Azure Databricks)

Name * AzureDatabricks

Description

Connect via integration runtime * AutoResolveIntegrationRuntime

Account selection method * From Azure subscription

Azure subscription * Azure subscription 1 (ec19a3db-92c1-49ef-8692-15bfce1202c8)

Databricks workspace * adfcookbookadb

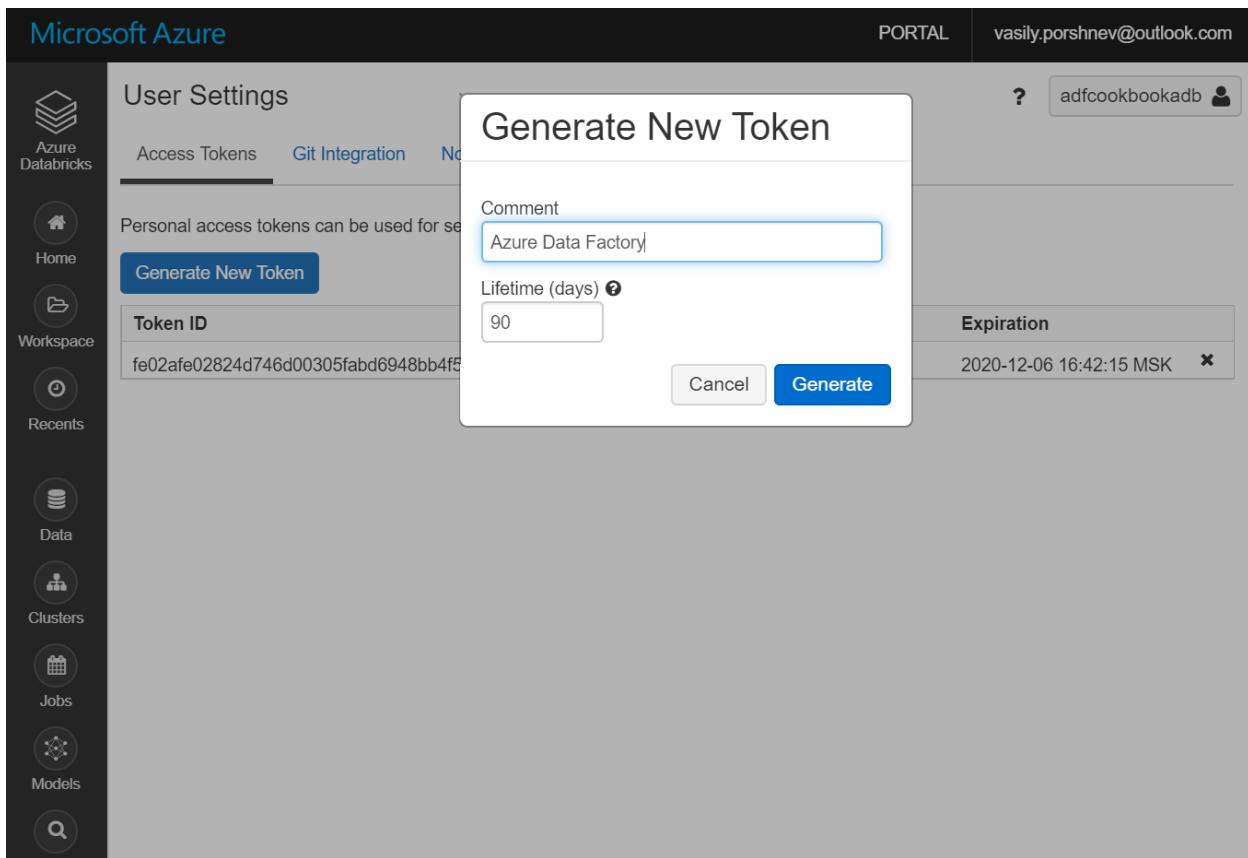
Select cluster
 New job cluster
 Existing interactive cluster
 Existing instance pool

Databrick Workspace URL * https://adb-1486914505250902.azuredatabricks.net

Access token * Access token *

Choose from existing clusters * adfcookbookcluster

Create Connection successful Test connection Cancel



The screenshot shows the Azure Data Factory pipeline editor. The left sidebar lists "Factory Resources" including "Pipelines" (with items like "ADFcookbookMLpipe", "Delete_blob_LA", and "pipeline1" selected), "Datasets", and "Data flows". The main workspace displays a pipeline named "pipeline1" with activities: "Move & transform", "Azure Data Explorer", "Azure Function", "Batch Service", and "Databricks" (with sub-options for "Notebook", "Jar", and "Python"). A "Notebook" activity named "Notebook1" is highlighted. The "Properties" pane on the right shows the pipeline's name as "pipeline1", its description, concurrency settings, annotations, and a "General" tab where the "Notebook path" is set to "/adfcookbooknb".

Microsoft Azure | adfcookbookch8adf

Validate all Refresh Discard all Data flow debug ARM template

ADFcookbookDB

Activities

Pipelines 4

Datasets 1

Data flows 0

Notebook Notebook1

Properties

Name: ADFcookbookDB

Description:

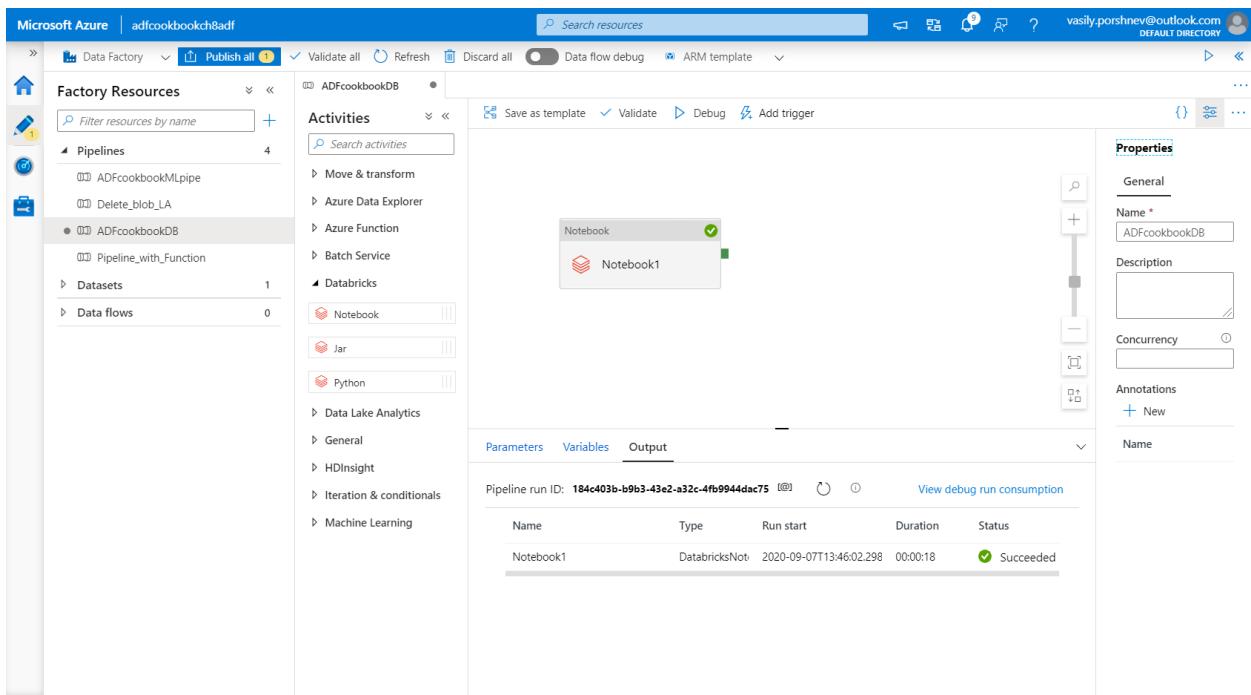
Concurrency:

Annotations

Parameters Variables Output

Pipeline run ID: 184c403b-b9b3-43e2-a32c-4fb9944dac75 View debug run consumption

Name	Type	Run start	Duration	Status
Notebook1	DatabricksNotebook	2020-09-07T13:46:02.298	00:00:18	Succeeded



General Azure Databricks Settings User properties

LIBRARY TYPE

LIBRARY CONFIGURATION

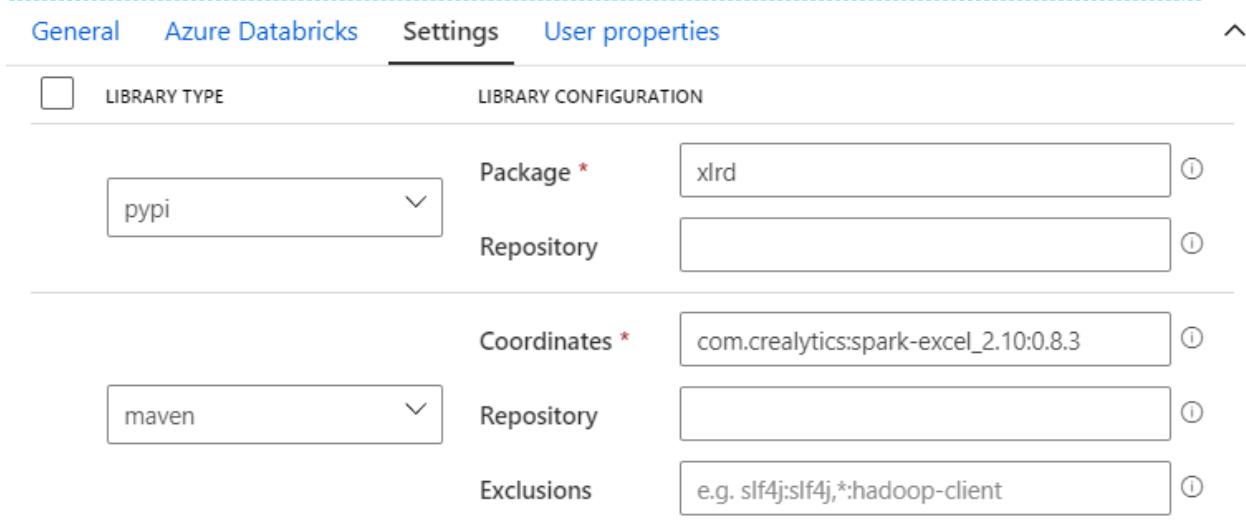
Package * xldr

Repository

Coordinates * com.crealytics:spark-excel_2.10:0.8.3

Repository

Exclusions e.g. slf4j:slf4j-*:hadoop-client



Microsoft Azure [Upgrade](#) Search resources, services, and docs (G+/)

Dashboard > New > Logic App >

Logic App

* Basics 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

Logic App name *

Select the location Region Integration Service Environment

Location *

[Review + create](#) < Previous : Basics [Next : Tags >](#) [Download a template for automation](#)

Microsoft Azure [Upgrade](#) Search resources, services, and docs (G+/)

Dashboard > ADFcookbookLA >

Logic Apps Designer

[Save](#) [Discard](#) [Run](#) [Designer](#) [Code view](#) [Parameters](#) [Templates](#) [Connectors](#) [Help](#)

HTTP POST URL

Request Body JSON Schema

Use sample payload to generate schema

Add new parameter

+ New step

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

Dashboard > ADFcookbookLA >

Logic Apps Designer

Save Discard Run Designer Code view Parameters Templates Connectors Help

Choose an action

Search: azure data factory

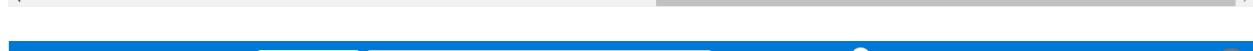
For You All Built-in Standard Enterprise Custom

Azure Data Factory

Triggers Actions

Cancel a pipeline run Azure Data Factory

Create a pipeline run Azure Data Factory



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

Dashboard > ADFcookbookLA >

Logic Apps Designer

Save Discard Run Designer Code view Parameters Templates Connectors Help

When a HTTP request is received

Create a pipeline run

Subscription: Azure subscription 1

Resource Group: adf-cookbook-ch1-rg

Data Factory Name: adfcookbookch3adf

Data Factory Pipeline Name: ADCookbookCopyToSynapse

Add new parameter

Connected to live.com#p0rshe25@yandex.ru. Change connection.

+ New step



Microsoft Azure [Upgrade](#) Search resources, services, and docs (G+)

Dashboard > ADFcookbookLA > Logic Apps Designer

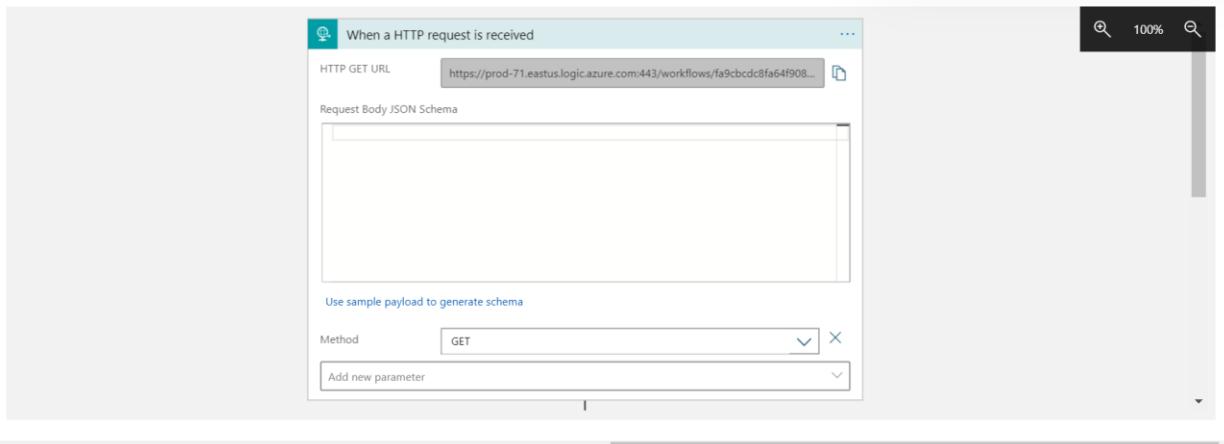
Save Discard Run Designer Code view Parameters Templates Connectors Help

When a HTTP request is received

HTTP GET URL https://prod-71.eastus.logic.azure.com:443/workflows/fa9cbcdc8fa64f908...
Request Body JSON Schema

Use sample payload to generate schema

Method GET Add new parameter



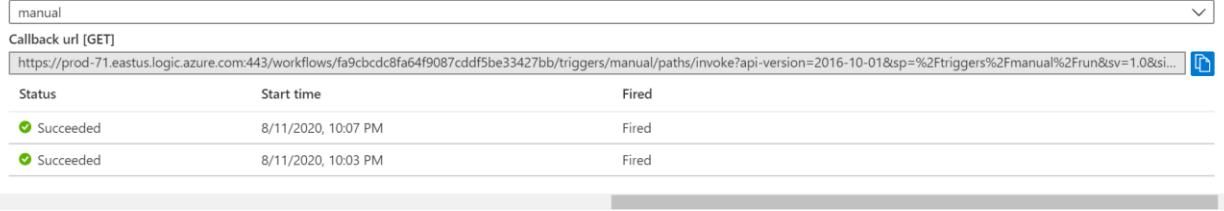
Microsoft Azure [Upgrade](#) Search resources, services, and docs (G+)

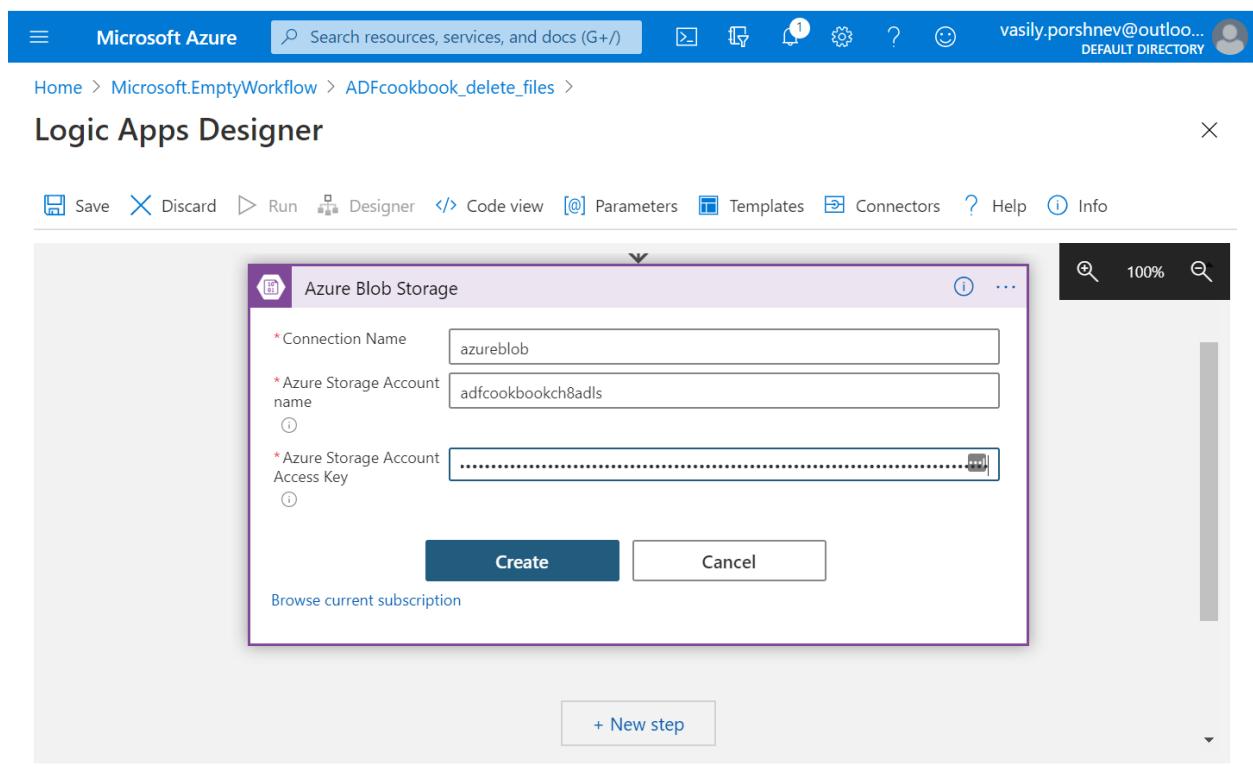
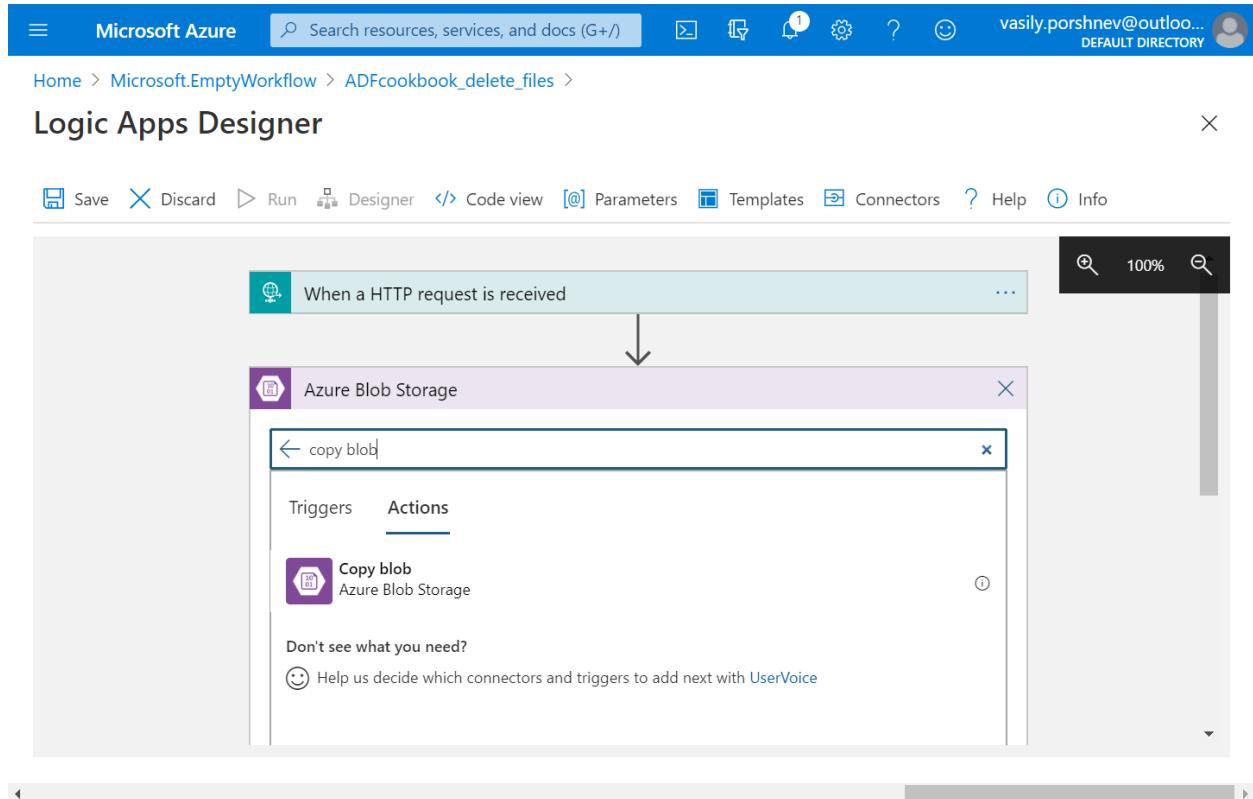
Dashboard > ADFcookbookLA > manual

All manual

Callback url [GET]
https://prod-71.eastus.logic.azure.com:443/workflows/fa9cbcdc8fa64f9087cddf5be33427bb/triggers/manual/paths/invoke?api-version=2016-10-01&sp=%2Ftriggers%2Fmanual%2Frun&sv=1.0&si...

Status	Start time	Fired
Succeeded	8/11/2020, 10:07 PM	Fired
Succeeded	8/11/2020, 10:03 PM	Fired





Microsoft Azure Upgrade Search resources, services, and docs (G+/-) vasily.porshnev@outlook.com DEFAULT DIRECTORY

Home > Microsoft.StorageAccount-20200905185908 > adfcookbookch8adls

adfcookbookch8adls | Access keys

Storage account

Search (Ctrl+ /)

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Data transfer

Events

Storage Explorer (preview)

Settings

Access keys

Geo-replication

CORS

Configuration

Use access keys to authenticate your applications when making requests to this Azure storage account. Store your access keys securely - for example, using Azure Key Vault - and don't share them. We recommend regenerating your access keys regularly. You are provided two access keys so that you can maintain connections using one key while regenerating the other.

When you regenerate your access keys, you must update any Azure resources and applications that access this storage account to use the new keys. This action will not interrupt access to disks from your virtual machines. [Learn more about regenerating storage access keys](#)

Storage account name: adfcookbookch8adls

Hide keys

key1

Key: EdrFo66CnxmRUC0THdNr37bCqbeYfR/3jgOB9dsfLePlb1ZHQpElKku+mKGjP3NNFM68+SRUnE/FTJfE7Qv8zw==

Connection string: DefaultEndpointsProtocol=https;AccountName=adfcookbookch8adls;AccountKey=EdrFo66CnxmRUC0THdNr37bCqbeYfR/3jgOB9dsfLePlb1ZHQpEl...

key2

Key: 7MBwhBxSQu4N6AaveI7NJXNQG6p0tL6+GH75HQ/BXgaum6FFT6NsfrFrN8yZf+tLCVW+FVaq3kbE4TuRc7lgA==

Microsoft Azure Search resources, services, and docs (G+/-) vasily.porshnev@outlook.com DEFAULT DIRECTORY

Home > Microsoft.EmptyWorkflow > ADFcookbook_delete_files >

Logic Apps Designer

Save Discard Run Designer Code view Parameters Templates Connectors Help Info

Copy blob

*Source url: Container FileName

*Destination blob path: archive/FileName

Overwrite?: No

Connected to azureblob. [Change connection](#).

+ New step

Dynamic content Expression

Search dynamic content

When a HTTP request is received See more

Container

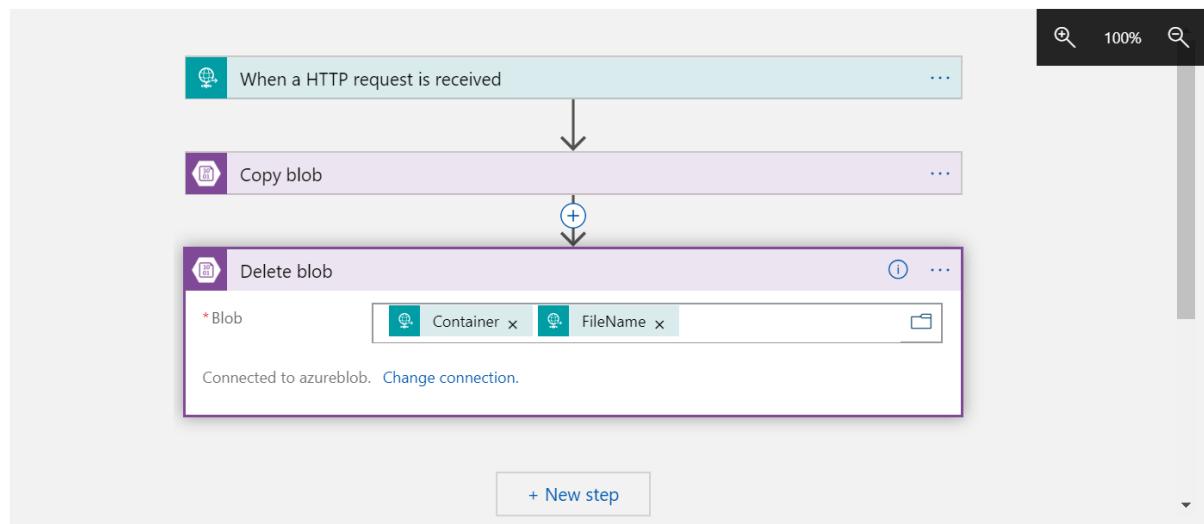
FileName



Home > Microsoft.EmptyWorkflow > ADFcookbook_delete_files >

Logic Apps Designer

Save Discard Run Designer Code view Parameters Templates Connectors Help Info



The screenshot shows the Microsoft Azure Data Factory pipeline editor. On the left, the 'Factory Resources' sidebar lists Pipelines (1), Datasets (1), and Data flows (0). The 'Activities' section shows a pipeline named 'Delete_blob_LA' with one activity: 'Get Metadata1'. The main area displays the 'Get Metadata1' activity configuration. The 'Dataset' tab is selected, showing 'flights_blob.ds' as the dataset. Other tabs include General, User properties, and ARGUMENT. To the right, the 'Properties' panel shows the pipeline's name as 'Delete_blob_LA'.

Microsoft Azure | adfcookbookch8adf

Validate all Refresh Discard all Data flow debug ARM template

Search resources

Factory Resources

Pipelines 1 Delete_blob_LA

Datasets 1 flights_blob_ds

Data flows 0

Activities

General

- Append variable
- Delete
- Execute Pipeline
- Execute SSIS package
- Get Metadata
- Lookup
- Stored procedure
- Set variable
- Validation
- Web
- WebHook
- Wait

HDInsight

Iteration & conditionals

Machine Learning

flights_blob_ds

Save as template Validate Debug Add trigger

Get Metadata1

Web1

Properties

General

Name * Delete_blob_LA

Description

Concurrency

Annotations

+ New

Name

URL * https://prod-43.eastus.logic.azure.com:443

Method * POST

Headers + New

Body

```
{ "Container": "@{pipeline().parameters.Container}", "FileName": "@activity('Get Metadata1').output.itemName" }
```

Datasets + Add dataset reference

General Settings User properties

Microsoft Azure | adfcookbookch8adf

Validate all Refresh Discard all Data flow debug ARM template

Search resources

Factory Resources

Pipelines 1 Delete_blob_LA

Datasets 1 flights_blob_ds

Data flows 0

Activities

General

- Append variable
- Delete
- Execute Pipeline
- Execute SSIS package
- Get Metadata
- Lookup
- Stored procedure
- Set variable
- Validation
- Web
- WebHook
- Wait

HDInsight

Iteration & conditionals

Machine Learning

flights_blob_ds

Save as template Validate Debug Add trigger

Get Metadata1

Web1

Properties

General

Name * Delete_blob_LA

Description

Concurrency

Annotations

+ New

Name

Parameters Variables Output

+ New | Delete

NAME	TYPE	DEFAULT VALUE
Container	String	flightscontainer/

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

Home > Microsoft.EmptyWorkflow > ADFcookbook_delete_files

ADFcookbook_delete_files

Logic app

Search (Ctrl+ /) Overview Activity log Access control (IAM) Tags Diagnose and solve problems Development Tools Logic app designer Logic app code view Versions API connections Quick start guides Release notes Settings

Run Trigger Refresh Edit Delete Disable Update Schema Clone Export

To improve traffic flow, we're adding new outbound IP addresses for Logic Apps. Review action needed if you're filtering IP addresses with firewall settings before 08/31/2020. Click to learn more.

REQUEST
manual

COUNT
2 actions
[View in Logic Apps designer](#)

FREQUENCY
Evaluated 2 times, fired 2 times in the last 24 hours
[See trigger history](#)

Runs history

Status	Start time	Identifier	Duration	Static Results
Succeeded	9/5/2020, 8:54 PM	0858602278414889012932210898...	22.81 Seconds	

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

Home > Function App >

Create Function App

Select a 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

Function App name * .azurewebsites.net

Publish * Code Docker Container

Runtime stack *

Version *

Region *

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

Microsoft Azure [Upgrade](#) [Search resources, services, and docs \(G+\)](#)

Home > New > Function App >

Create Function App

Basics **Hosting** Monitoring Tags Review + create

Storage

When creating a function app, you must create or link to a general-purpose Azure Storage account that supports Blobs, Queue, and Table storage.

Storage account * [Create new](#)

Operating system

The Operating System has been recommended for you based on your selection of runtime stack.

Operating System * Windows Linux

Plan

The plan you choose dictates how your app scales, what features are enabled, and how it is priced. [Learn more](#)

Plan type * Consumption (Serverless)

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

Microsoft Azure [Upgrade](#) [Search resources, services, and docs \(G+\)](#)

Home > Function App >

Create Function App

Basics Hosting **Monitoring** Tags Review + create

Application Insights

Application Insights is a code-less attach to provide detailed observability in to your application. [Learn more](#)

Enable Application Insights * Yes No

Application Insights * [Create new](#)

Region

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

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

Home > adfcookbookfunctionapp

adfcookbookfunctionapp | Functions

Function App

Search (Ctrl+ /) Add Develop Locally Refresh

Overview Activity log Access control (IAM) Tags Diagnose and solve problems Security Events (preview)

Functions Functions App keys App files Proxies Deployment

New Function

Create a new function in this function app. Start by selecting a template below.

Templates **Details**

Search by template name

HTTP trigger A function that will be run whenever it receives an HTTP request, responding based on data in the body or query string

Timer trigger A function that will be run on a specified schedule

Azure Queue Storage trigger A function that will be run whenever a message is added to a specified Azure Storage queue

Azure Service Bus Queue trigger A function that will be run whenever a message is added to a specified Service Bus queue

Microsoft Azure | adfcookbookch8adf

Data Factory Publish all Validate all Refresh Discard all Data flow debug ARM template

Factory Resources Pipelines Datasets Data flows

Activities Pipeline_with_Function_1 Move & transform Azure Data Explorer Azure Function Databricks Data Lake Analytics General HDInsight Iteration & conditionals Machine Learning

Azure Function

New Linked service (Azure Function)

Name * AzureFunction1

Description

Connect via integration runtime * AutoResolveIntegrationRuntime

Azure Function App selection method From Azure subscription Enter manually

Azure subscription Azure subscription 1 (ec19a3db-92c1-49ef-8692-15bfce1202c8)

Azure Function App url * adfcookbookfunctionapp; https://adfcookbookfunctionapp.azurewebsites.net

Function Key Azure Key Vault

Annotations + New

Name

Advanced

Create Cancel

Microsoft Azure | adfcookbookch8adf

Validate all Refresh Discard all Data flow debug ARM template

Factory Resources Pipeline_with_Function... Activities

Pipelines 2 Delete_blob_LA Pipeline_with_Function Datasets 1 Data flows 0

Save as template Validate Debug Add trigger

Azure Function

Azure Function1

Properties

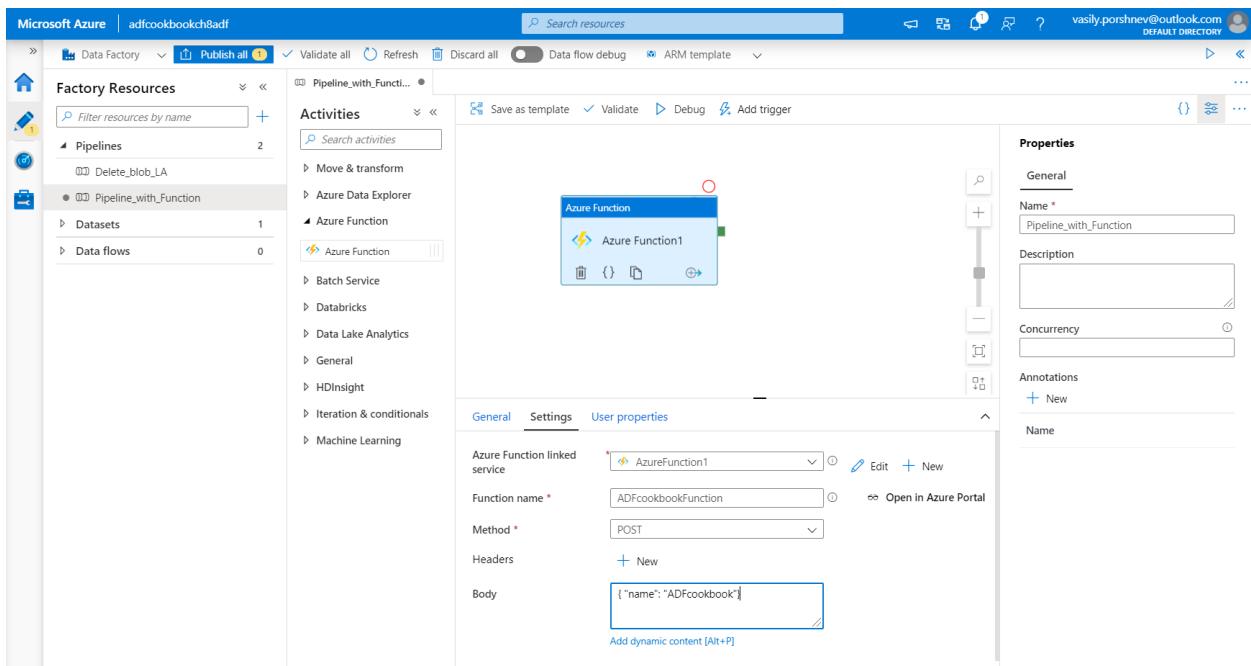
Name * Pipeline_with_Function

Description

Concurrency

Annotations

Name



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

vasily.porshnev@outlook.com DEFAULT DIRECTORY

Home > adfcookbookfunctionapp > ADFcookbookFunction

ADFcookbookFunction | Monitor

Function

Search (Ctrl+ /)

Overview Invocations Logs

Developer

Code + Test Integration Monitor Function Keys

Success Count 3 Last 30 Days Error Count 3 Last 30 Days

Invocation Traces

The twenty most recent function invocation traces. For more advanced analysis, run the query in Application Insights.

Run query in Application Insights Refresh

Filter invocations

Date (UTC)	Success	Result Code	Duration (ms)	Operation Id
2020-09-06 11:10:57.454	Success	200	7	3771a1492e0fb944afec4b67474
2020-09-06 11:07:52.915	Success	200	4	193e411af9281546b34e96ed72c



Dashboard > All resources > New > Machine Learning Studio (classic) Web Service >

Machine Learning Studio (classic) web service

Machine Learning Studio (classic) web service

Web Service Creation

If you have authored an experiment in Azure Machine Learning Studio (classic), you can create a web service from within the Studio.

The following tools help you create, manage and view Machine Learning Studio (classic) Web Services.

1. Create



Machine Learning Studio
(classic)

2. Manage



Web Service Portal

3. View



Azure Portal

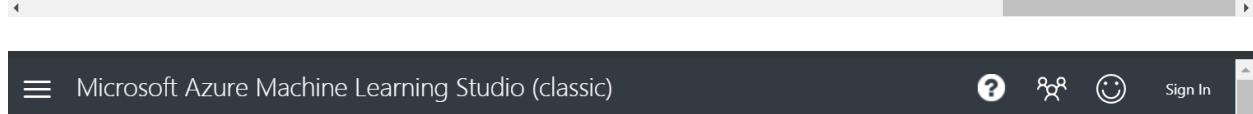
1. Create your web service and deploy from within Azure Machine Learning Studio (classic).

To create a Machine Learning Studio (classic) web service you need to have a standard Machine Learning Studio (classic) workspace.

[View your workspaces](#)

2. Manage and monitor your web service in the Web Service Management Portal.

3. View your web services in Azure Portal.



Welcome to Azure
Machine Learning
Studio (classic)

Try it for free

No [Azure subscription](#)? No credit card? No problem! Choose anonymous Guest Access, or sign in with your work or school account, or a Microsoft account.

[Sign In](#) 

Not an Azure ML Studio (classic) user?
[Sign up here](#)

[Pricing & FAQ](#)

By using this free version, you agree to be bound by the Microsoft Azure Website Terms of Use.

Announcements NEW!

[Azure Machine Learning Studio R Runtime](#)

[Mining Campaign Funds](#)

[Inside the Data Science VM](#)

Chapter 9: Managing Deployment Processes with Azure DevOps

The screenshot shows the 'Configure a repository' dialog box. At the top left, there's a 'Set up code repository' button with a dropdown arrow, which is highlighted with a red box. The main area contains several configuration fields:

- Repository type ***: A dropdown menu set to 'Azure DevOps Git'.
- Azure Active Directory**: A dropdown menu set to 'Default Directory (3b7b3863-42dc-4d1c-b559-6403c9b3a91d)'.
- Azure DevOps Account ***: A dropdown menu set to 'ADFCookbook-org'.
- Project name ***: A dropdown menu set to 'ADF DevOps'.
- Repository name ***: A dropdown menu with options 'Create new' (radio button) and 'Use existing' (radio button, selected). The 'Use existing' option is set to 'ADF DevOps'.
- Collaboration branch ***: A dropdown menu set to 'master'.
- Publish branch ***: A dropdown menu set to 'adf_publish'.
- Root folder ***: A text input field containing '/'.
- Import existing resource**: A checkbox checked with the text 'Import existing resources to repository'.
- Import resource into this branch ***: A radio button group where 'Use Collaboration' is selected.

At the bottom right of the dialog are 'Apply' and 'Cancel' buttons.

The screenshot shows the 'Factory Resources' blade in the Azure DevOps interface. The 'Pipelines' section is expanded, showing 1 item. A context menu is open over the 'master branch' entry, with the following options visible:

- Create pull request [Alt+P]
- + New branch [Alt+N]
- master branch (selected)
- adf_publish branch

Below the menu, the 'Filter...' search bar is highlighted with a blue box. The top navigation bar includes 'Search resources', 'Save all', 'Publish', 'Validate all', 'Refresh', 'Discard all', 'Data flow debug', 'ARM template', and user information.

Microsoft Azure | adfcookbookch8adf

Search resources

Save all Publish Validate all Refresh Discard all Data flow debug ARM template

Factory Resources <>

Pipelines 5

- ADFcookbookDB
- ADFcookbookMLpipe
- Delete_blob_LA
- DevOpsPipe
- Pipeline_with_Function

Datasets 1

Data flows 0

Templates 0

Activities <>

Wait

General

Wait Wait1

Properties

General Related

Name * DevOpsPipe

Description

Concurrency

Annotations + New

Parameters Variables Output

Microsoft Azure | adfcookbookch8adf

You have GIT enabled in your data factory. Publishing in 'Data Factory' mode is disabled. Please switch back to GIT mode to make further changes. [Read about our Best Practices.](#)

Search resources

Data Factory Publish all Validate all Refresh Discard all Data flow debug ARM template

Factory Resources <>

Pipelines 4

- ADFcookbookDB
- ADFcookbookMLpipe
- Delete_blob_LA
- Pipeline_with_Function

Datasets 1

Data flows 0

Select an item

Use the resource explorer to select or create a new item

Azure DevOps

/ ADF DevOps / Repos / Files / ADF DevOps

Search

ADF DevOps

ADF DevOps

dataset

factory

linkedService

pipeline

ADFcookbookDB.json

ADFcookbookMLpipe.json

Delete_blob_LA.json

DevOpsPipe.json

Pipeline_with_Function.json

ML readme.md

master

Branches Tags

New branch

```
5
6
7
8
9
10
11
12
13
14
15
16
17
{
  "name": "Wait1",
  "type": "Wait",
  "dependsOn": [],
  "userProperties": [
    {
      "typeProperties": {
        "waitTimeInSeconds": 1
      }
    }
  ],
  "annotations": []
}
```

Edit

ADF DevOps

Overview

Boards

Repos

Files

Commits

Pushes

Branches

Tags

Pull requests

Pipelines

Test Plans

Project settings

Microsoft Azure | adfcookbookh8adf

Search resources

master branch

Save all Publishing Validate all Refresh Discard all Data

Factory Resources

Filter resources by name

Pipelines

ADFcookbookDB

ADFcookbookMLpipe

Delete_blob_LA

DevOpsPipe

Pipeline_with_Function

Datasets

Data flows

Templates

Activities

Move & transform

Azure Data Explorer

Azure Function

Batch Service

Databricks

Data Lake Analytics

General

HDInsight

Iteration & conditionals

Machine Learning

DevOpsPipe

Saved Save as template Validate Debug

OK Cancel

Pending changes

The Publish branch is the branch in your repository where publishing related ARM templates are stored and updated.

Publish branch

Data Factory properties

Pipelines

DevOpsPipe (New)

Datasets

Data flows

Integration runtimes

Linked services

Triggers

Azure DevOps

ADF DevOps / Repos / Files / ADF DevOps

Search

ADF DevOps

Overview Boards Repos Files Commits Pushes Branches Tags Pull requests Pipelines Test Plans Project settings

ADF DevOps

adf_publish

linkedTemplates / ArmTemplate_master.json

Filter branches

Branches Tags

master Default

Mine

adf_publish

a pull request

New branch

contentVersion: "1.0.0.0",
parameters: {
factoryName: {
type: "string",
metadata: "Data Factory name",
defaultValue: "adfcookbookch8adf"
},
"AzureBlobStorageLSConnectionString": {
type: "secureString",
metadata: "Secure string for 'connectionString' of 'AzureBlobStorageL...
},
"AzureDatabricks1AccessToken": {
type: "secureString",
metadata: "Secure string for 'accessToken' of 'AzureDatabricks1'"
},
"AzureFunction1FunctionKey": {
type: "secureString",
metadata: "Secure string for 'functionKey' of 'AzureFunction1'"
},

Microsoft Azure | adfcookbookch8adf

Search resources

master branch

Save all Publish Validate all Refresh Discard all Data flow debug ARM template

Factory Resources

Pipelines

- ADFCookbookDB
- ADFCookbookMLpipe
- Delete_blob_LA
- DevOpsPipe
- Pipeline_with_Function

Datasets

Data flows

Templates

Create pull request [Alt+P] New branch [Alt+N]

Wait Wait1

Parameters Variables Output

Microsoft Azure | adfcookbookch8adf

Search resource: 1 3 ?

New branch [Alt+N]

vasily.porshnev@outlook.com DEFAULT DIRECTORY

Factory Resources

- Pipelines 5
 - ADFcookbookDB
 - ADFcookbookMLpipe
 - Delete_blob_LA
 - DevOpsPipe
 - Pipeline_with_Function
- Datasets 1
- Data flows 0
- Templates 0

DevOpsPipe Activities

- Move & transform
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning

Parameters Variables Output

Save Cancel

The screenshot shows the Azure Data Factory interface. On the left, there's a sidebar with 'Factory Resources' containing sections for Pipelines, Datasets, Data flows, and Templates. Under Pipelines, several pipelines are listed: ADFcookbookDB, ADFcookbookMLpipe, Delete_blob_LA, DevOpsPipe, and Pipeline_with_Function. The 'DevOpsPipe' pipeline is currently selected. On the right, a large panel titled 'Activities' lists various data processing and machine learning activities. At the bottom of this panel are tabs for 'Parameters', 'Variables', and 'Output'. Below these tabs are two buttons: 'Save' and 'Cancel'. The top navigation bar includes a search bar, notification icons, and user information.

Microsoft Azure | adfcookbookch8adf

Search resources

vasily.porshnev@outlook.com DEFAULT DIRECTORY

Factory Resources

- Pipelines 5
 - ADFcookbookDB
 - ADFcookbookMLpipe
 - Delete_blob_LA
 - DevOpsPipe
 - Pipeline_with_Function
- Datasets 2
 - Airlines_blob_ds
 - flights_blob.ds
- Data flows 0
- Templates 0

Airlines_blob_ds

Save

DelimitedText Airlines_blob_ds

Properties

General Related

Name * Airlines_blob_ds

Description

Annotations + New

Name

Connection Schema Parameters

Linked service * AzureBlobStorageLS Test connection Edit + New

File path * flightscontainer / Directory / airlines.csv Browse Preview data

Compression type none

Column delimiter Comma (,)

Row delimiter Auto detect (\r\n, or \n)

Encoding Default(UTF-8)

Escape character Backslash (\)

Quote character Double quote (")

First row as header

The screenshot shows the Azure Data Factory interface. On the left, there's a sidebar with 'Factory Resources' containing sections for Pipelines, Datasets, Data flows, and Templates. Under Datasets, two datasets are listed: 'Airlines_blob_ds' and 'flights_blob.ds'. The 'Airlines_blob_ds' dataset is currently selected. The main panel shows its configuration details. It's defined as a 'DelimitedText' dataset named 'Airlines_blob_ds'. The 'Connection' tab is selected, showing it's connected to 'AzureBlobStorageLS'. The 'File path' is set to 'flightscontainer / Directory / airlines.csv'. Other settings include 'Compression type' (none), 'Column delimiter' (Comma (,)), 'Row delimiter' (Auto detect (\r\n, or \n)), 'Encoding' (Default(UTF-8)), 'Escape character' (Backslash (\)), 'Quote character' (Double quote (")), and 'First row as header' (unchecked). To the right, there's a 'Properties' panel with tabs for 'General' and 'Related', and fields for 'Name' (set to 'Airlines_blob_ds') and 'Description'.

Microsoft Azure | adfcookbookch8adf

Factory Resources

Pipelines: ADFCookbookDB, ADFCookbookMLpipe, Delete_blob_LA, DevOpsPipe, Pipeline_with_Function

Datasets: Airlines_blob_ds (selected), flights_blob_ds

Data flows: 0

Templates: 0

Save: Save, Publish, Validate all, Refresh, Discard all, Data flow debug, ARM template

Search resources: Search bar

Properties: General, Related

Name: Airlines_blob.ds

Description:

Annotations: + New

OK button

Airlines_blob.ds (CSV) file preview

Connection: Linked service (AzureBlobStorageLS)

Schema: DelimitedText

Parameters: File path (flightscontainer), Compression type (none), Column delimiter (Comma ()), Row delimiter (Auto detect (\r\n, or \n)), Encoding (Default(UTF-8)), Escape character (Backslash (\)), Quote character (Double quote (")), First row as header (checkbox)

Message: Publish is only allowed from collaboration ('master') branch. Merge the changes to 'master'. OK

Microsoft Azure | adfcookbookch8adf

Factory Resources

Pipelines: ADFCookbookDB, ADFCookbookMLpipe, Delete_blob_LA, DevOpsPipe, Pipeline_with_Function

Datasets: Airlines_blob_ds (selected), flights_blob_ds

Data flows: 0

Templates: 0

Save: Save all, Publish, Validate all, Refresh, Discard all, Data flow debug, ARM template

Search resources: Search bar

Properties: General, Related

Name: Airlines_blob.ds

Description:

Annotations: + New

OK button

Airlines_blob.ds (CSV) file preview

Connection: Linked service (AzureBlobStorageLS)

Schema: DelimitedText

Parameters: File path (flightscontainer / Directory / airlines.csv), Compression type (none), Column delimiter (Comma ()), Row delimiter (Auto detect (\r\n, or \n)), Encoding (Default(UTF-8)), Escape character (Backslash (\)), Quote character (Double quote (")), First row as header (checkbox)

Branch selection dropdown: Filter..., Create pull request [Alt+P], master branch, adfcookbook_feature branch, adf_publish branch

Azure DevOps

adfcookbook / ADF DevOps / Repos / Pull requests / ADF DevOps

New pull request

adfcookbook_feature into master

Overview Files 1 Commits 1

Title: Adding dataset: Airlines_blob_ds

Description: Adding dataset: Airlines_blob_ds

Markdown supported. Link work items.

Adding dataset: Airlines_blob_ds

Reviewers: Add required reviewers

Search users and groups to add as reviewers

Work items to link:

Search work items by ID or title

Tags: dataset

Create

Project settings

This screenshot shows the 'New pull request' page in Azure DevOps. It displays the source branch 'adfcookbook_feature', target branch 'master', and a single file 'Airlines_blob.ds'. The 'Title' and 'Description' fields are filled with the same text: 'Adding dataset: Airlines_blob_ds'. There are sections for 'Reviewers' and 'Work items to link', both currently empty. The 'Tags' field contains the tag 'dataset'. At the bottom right is a prominent blue 'Create' button.

Azure DevOps

adfcookbook / ADF DevOps / Repos / Pull requests / ADF DevOps

New pull request

adfcookbook_feature into master

Overview Files 1 Commits 1

1 changed file

Inline

ADF DevOps

dataset

Airlines_blob.ds.json [+]

```
1 + {
2 +   "name": "Airlines_blob_ds",
3 +   "properties": {
4 +     "linkedServiceName": {
5 +       "referenceName": "AzureBlobStorageLS",
6 +       "type": "LinkedServiceReference"
7 +     },
8 +     "annotations": [],
9 +     "type": "DelimitedText",
10 +    "typeProperties": {
11 +      "location": {
12 +        "type": "AzureBlobStorageLocation",
13 +        "fileName": "airlines.csv",
14 +        "container": "flightscontainer"
15 +      },
16 +      "columnDelimiter": ",",
17 +      "escapeChar": "\\",
18 +      "quoteChar": "\\"
19 +    },
20 +    "schema": [
21 +      {
22 +        "type": "String"
23 +      }
24 +    ]
25 +  }
26 + }
```

Project settings

This screenshot shows the 'New pull request' page with the 'Files' tab selected, displaying the contents of the 'Airlines_blob.ds.json' file. The file is a JSON object with several properties, including 'name', 'properties' (containing a 'linkedServiceName' reference), and a 'typeProperties' section with a 'location' object pointing to an 'AzureBlobStorageLocation' with specific file and container names. The 'schema' section defines a single column with type 'String'. The code is presented in a diff-style format with '+' signs for additions and '-' signs for deletions.

Azure DevOps adfcookbook / ADF DevOps / Repos / Pull requests / ADF DevOps

Adding dataset: Airlines_blob_ds

Active | 1 | Vasily Porshnev adfcookbook_feature into master

Overview Files Updates Commits

No merge conflicts Last checked Just now

Description

Adding dataset: Airlines_blob_ds

Show everything (1) ▾

Add a comment...

Vasily Porshnev created the pull request Just now

Reviewers Add ▾

Required No required reviewers

Optional No optional reviewers

Tags + dataset

Work items + No work items

Welcome to the new pull request experience!

Learn more Dismiss

Learn more about all the new features and let us know what you think through this survey.

Azure DevOps adfcookbook / ADF DevOps / Repos / Pull requests / ADF DevOps

Adding dataset: Airlines_blob_ds

Active | 1 | Vasily Porshnev adfcookbook_feature into master

Overview Files Updates Commits

No merge conflicts Last checked 31m ago

Description

Adding dataset: Airlines_blob_ds

Show every

Add a comment...

Vasily Porshnev approved the pull request

Vasily Porshnev joined as a reviewer

Vasily Porshnev created the pull request

Complete pull request

Merge type Merge (no fast forward)

Post-completion options

Complete associated work items after merging

Delete adfcookbook_feature after merging

Customize merge commit message

Cancel Complete merge

Azure DevOps / ADF DevOps / Repos / Pull requests / ADF DevOps / Search / Boards / Overview / Files / Updates / Commits

Adding dataset: Airlines_blob_ds

Completed 11 VP Vasily Porshnev adfcookbook_feature into master

Overview Files Updates Commits

Vasily Porshnev completed this pull request 3m ago Cherry-pick Revert

Merged PR 1: Adding dataset: Airlines_blob_ds 3738f429 Vasily Porshnev Today at 2:25 PM

Show details

No merge conflicts Last checked 3m ago

Description

Adding dataset: Airlines_blob_ds

Reviewers Add Required No required reviewers

Optional Vasily Porshnev Approved

Tags dataset +

Work items No work items

Add a comment...

Vasily Porshnev completed the pull request 3m ago

Vasily Porshnev approved the pull request 10m ago

Azure DevOps / ADF DevOps / Repos / Files / ADF DevOps / Search / Boards / Overview / Files / Commits / Pushes / Branches / Tags / Pull requests / Pipelines / Test Plans / Artifacts / Project settings

ADF DevOps

master / Type to find a file or folder...

dataset

- Airlines_blob.ds.json
- flights_blob.ds.json

factory

linkedService

pipeline

readme.md

Set up build Clone ...

Filter branches Branches Tags

master Default Mine adf_publish

+ New branch linkedService Sep 16

pipeline Friday a5532c69 Adding pipeline: DevOpsPipe Vasily Porshnev

readme.md Sep 16 f3385d60 Initial commit. Vasily Porshnev

Initialized by Azure Data Factory!

Microsoft Azure | adfcookbookch8adf

Search resources

Airlines_blob.ds

Save all Publish Validate all Refresh Discard all Data

Factory Resources

Pipelines 5

- ADFCookbookDB
- ADFcookbookMLpipe
- Delete_blob_LA
- DevOpsPipe
- Pipeline_with_Function

Datasets 2

- Airlines_blob.ds
- flights_blob.ds

Data flows 0

Templates 0

CSV DelimitedText Airlines_blob.ds

Connection Schema Parameters

Linked service * AzureBlobStorageLS Test connection Edit

File path * flightscontainer / Directory / airlines.csv

Compression type none

Column delimiter Comma (,) Edit

Row delimiter Auto detect (\r,\n, or \v\n) Edit

Encoding Default(UTF-8)

Escape character Backslash (\) Edit

Quote character Double quote (") Edit

First row as header

Save

Select working branch

Current working branch 'adfcookbook_feature' was deleted, create a new one or select existing branch.

Working branch Create new Use existing

* master

The Pending changes dialog displays the following information:

- Publish branch:** master branch
- Data Factory properties:** Data Factory properties
- Pipelines:** Pipelines
- Datasets:** Datasets
- Airlines_blob.ds (New):** Airlines_blob.ds
- Data flows:** Data flows
- Integration runtimes:** Integration runtimes
- Linked services:** Linked services
- Triggers:** Triggers

The dialog also contains a note: "The Publish branch is the branch in your repository where publishing related ARM templates are stored and updated."

Microsoft Azure | adfcookbookch8adf

Search resources

Airlines_blob.ds

Save all Publishing Validate all Refresh Discard all Data

Factory Resources

Pipelines 5

- ADFCookbookDB
- ADFcookbookMLpipe
- Delete_blob_LA
- DevOpsPipe
- Pipeline_with_Function

Datasets 2

- Airlines_blob.ds
- flights_blob.ds

Data flows 0

Templates 0

CSV DelimitedText Airlines_blob.ds

Connection Schema Parameters

Linked service * AzureBlobStorageLS Test connection Edit

File path * flightscontainer / Directory / airlines.csv

Compression type none

Column delimiter Comma (,) Edit

Row delimiter Auto detect (\r,\n, or \v\n) Edit

Encoding Default(UTF-8)

Escape character Backslash (\) Edit

Quote character Double quote (") Edit

First row as header

OK Cancel

Pending changes

The Publish branch is the branch in your repository where publishing related ARM templates are stored and updated.

Publish branch

- Data Factory properties
- Pipelines
- Datasets
- Airlines_blob.ds (New)
- Data flows
- Integration runtimes
- Linked services
- Triggers

Microsoft Azure [Upgrade](#) Search resources, services, and docs (G+)

Home > All resources > New > Key Vault >

Create key vault

Project details

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

Subscription *	Free Trial
Resource group *	ADFCOOKBOOK-DEV-RG
	Create new

Instance details

Key vault name *	ADFCOOKBOOK-DEV-AKV
Region *	East US
Pricing tier *	Standard

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

Microsoft Azure [Upgrade](#) Search resources, services, and docs (G+)

Home > All resources > New > Key Vault > Create key vault >

Add access policy

Add access policy

Configure from template (optional)

Key permissions	0 selected
Secret permissions	2 selected
Certificate permissions	0 selected

Select principal *

ADFCOOKBOOK-DEV-ADF
Object ID: 039b5525-784b-4d67-9ec1-46d14dc470cf

Authorized application ○
None selected

[Add](#) [Select](#)

Principal

Select a principal

Select a directory member by entering object ID, name or email address

- AAD Request Verification Service - PROD
c728155f-7b2a-4502-a08b-b8af9b269319
- ADFCOOKBOOK-DEV-ADF
451c3060-7b90-444c-8756-ea6aa7e22d32
Selected
- ADFCOOKBOOK-UAT-ADF
1fa266fa-fe03-405a-92d3-ac6718223a6a

Selected items

- ADFCOOKBOOK-DEV-ADF
451c3060-7b90-444c-8756-ea6aa7e22d32

[Remove](#)

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

Home > adfcookbookdevadls

adfcookbookdevadls | Access keys

Storage account

Search (Ctrl+ /) Diagnose and solve problems Access Control (IAM) Data transfer Events Storage Explorer (preview)

Settings

Access keys (selected)

Geo-replication CORS Configuration Encryption Shared access signature Firewalls and virtual networks Private endpoint connections

Storage account name: adfcookbookdevadls

When you regenerate your access keys, you must update any Azure resources and applications that access this storage account to use the new keys. This action will not interrupt access to disks from your virtual machines. [Learn more about regenerating storage access keys](#)

Hide keys

key1

Key: RHds1vuiMNJHTpvMSL8t/MOY5tbaBYZWuGF/8nF6uL6WJY6mVzaH2q9EtR5Lpxw2JSY5x6o49TcCIxwVBdzFw==

Connection string: DefaultEndpointsProtocol=https;AccountName=adfcookbookdevadls;AccountKey=RHds1vuiMNJHTpvMSL8t/MOY5tbaBYZWuGF/8nF6uL6WJY6mVZ...

key2

Key: Mmpyf1RMT5KHynPSONYR4hDyjuWT6zQEseSY5XYbh9r8k2ENxqY9GORxiP8ard1Xjr3GycpSUwtxEybxt0kw==

Connection string: DefaultEndpointsProtocol=https;AccountName=adfcookbookdevadls;AccountKey=Mmpyf1RMT5KHynPSONYR4hDyjuWT6zQEseSY5XYbh9r8k2ENx...

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

Home > ADFCOOKBOOK-DEV-RG > ADFCOOKBOOK-DEV-AKV >

Create a secret

Upload options: Manual

Name *: adfcookbookdevadls-secret

Value *:

Content type (optional):

Set activation date?

Set expiration date?

Enabled? Yes No

Create

Microsoft Azure | ADFCOOKBOOK-DEV-ADF

Factory Resources

- Pipelines 1
- Datasets 0
- Data flows 0
- Templates 0

Activities

- Copy data
- Move & transform
- Azure Data Explorer
- Data Lake Analytics
- U-SQL
- General
- Get Metadata
- Validation
- Machine Learning
- Machine Learning Up...

Search resources

Save all Publish Validate all Refresh Discard all Data flow

New linked service (Azure Key Vault)

Name * AKV_DEV_LS

Description

Azure key vault selection method

From Azure subscription Enter manually

Azure subscription Free Trial (2a8f3ccf-550c-48ea-9b06-06c2f1e03472)

Azure key vault name * ADFCOOKBOOK-DEV-AKV

Edit key vault

Managed identity name: ADFCOOKBOOK-DEV-ADF
Managed identity object ID: 039b5525-784b-4d67-9ec1-46d14dc470cf
Grant Data Factory service managed identity access to your Azure Key Vault.

Annotations + New

Name

Advanced

Connection successful Test connection Cancel

Create

This screenshot shows the Azure Data Factory pipeline editor. On the left, the 'Factory Resources' sidebar lists Pipelines (1), Datasets (0), Data flows (0), and Templates (0). The 'Activities' section is expanded, showing Copy data, Move & transform, Azure Data Explorer, Data Lake Analytics, U-SQL, General, Get Metadata, Validation, and Machine Learning. A 'Copy data' activity is selected. On the right, a 'New linked service (Azure Key Vault)' dialog is open. It has fields for Name (AKV_DEV_LS), Description, and Azure key vault selection method (From Azure subscription). It also shows the Azure subscription (Free Trial) and Azure key vault name (ADFCOOKBOOK-DEV-AKV). At the bottom, there are 'Create', 'Test connection', and 'Cancel' buttons, with a message indicating a successful connection.

Microsoft Azure | ADFCOOKBOOK-DEV-ADF

Factory Resources

- Pipelines 1
- Datasets 0
- Data flows 0
- Templates 0

Activities

- Copy data
- Move & transform
- Azure Data Explorer
- Data Lake Analytics
- U-SQL
- General
- Get Metadata
- Validation
- Machine Learning
- Machine Learning Up...

Search resources

Save all Publish Validate all Refresh Discard all Data flow

New linked service (Azure Blob Storage)

Name * ADLS_DEV_LS

Description

Connect via integration runtime *

AutoResolveIntegrationRuntime

Authentication method Account key

Connection string Azure Key Vault

AKV linked service * AKV_DEV_LS

Secret name * adfcookbookdevads-secret

Secret version Use the latest version if left blank

Test connection To linked service To file path

Annotations + New

Name

Connection successful Test connection Cancel

Create

This screenshot shows the Azure Data Factory pipeline editor. The left sidebar and activities list are identical to the previous screenshot. The right side shows a 'New linked service (Azure Blob Storage)' dialog. It includes fields for Name (ADLS_DEV_LS), Description, Connect via integration runtime (AutoResolveIntegrationRuntime), Authentication method (Account key), Connection string (Azure Key Vault), AKV linked service (AKV_DEV_LS), Secret name (adfcookbookdevads-secret), and Secret version (Use the latest version if left blank). It also includes 'Test connection' and 'Annotations' sections. At the bottom, there are 'Create', 'Test connection', and 'Cancel' buttons, with a message indicating a successful connection.

Microsoft Azure | ADFCOOKBOOK-DEV-ADF

Azure DevOps GIT / master branch Save all Publish Validate all Refresh Discard all Data flow debug ARM template ...

Factory Resources

- Pipelines 1
 - ADFcookbook_pipeline
- Datasets 2
 - Input_DS
 - Output_DS
- Data flows 0
- Templates 0

Activities

- Move & transform
 - Copy data
 - Data flow
- Azure Data Explorer
- Azure Data Explorer C...
- Databricks
 - Notebook
 - Jar
 - Python
- Data Lake Analytics
 - U-SQL
- General
 - Get Metadata
 - Validation
 - Machine Learning
 - Machine Learning Up...

Properties

General

Name * ADFcookbook_pipeline

Description

Concurrency

Annotations + New

Name

Azure DevOps

ADF cookbook-org / ADF DevOps / Pipelines / Releases

All pipelines > New release pipeline

Pipeline Tasks Variables Retention Options History

Artifacts | + Add

+ Add an artifact

Schedule not set

Stages | + Add

+ Stage 1 Select a template

Select a template

Or start with an Empty job

Search

Featured

- Azure App Service deployment**
Deploy your application to Azure App Service. Choose from Web App on Windows, Linux, containers, Function Apps, or WebJobs.
- Deploy a Java app to Azure App Service**
Deploy a Java application to an Azure Web App.
- Deploy a Node.js app to Azure App Service**
Deploy a Node.js application to an Azure Web App.
- Deploy a PHP app to Azure App Service and Azure Database for MySQL**
Deploy a PHP application to an Azure Web App and database to Azure Database for MySQL.
- Deploy a Python app to Azure App Service and Azure database for MySQL**
Deploy a Python Django, Bottle, or Flask application to an Azure Web App and database to Azure Database for MySQL.
- Deploy to a Kubernetes cluster**
Deploy, configure, update your containerized applications to a Kubernetes cluster.
- IIS** IIS website and SQL database deployment
Deployment Group: Deploy ASP.NET or ASP.NET Core web applications to an IIS Website and SQL database on physical or virtual machines (VM).

Others

Apply

Azure DevOps

All pipelines > ADFcookbookReleasePipeline

Pipeline Tasks Variables Retention Options History

Artifacts | + Add Stages | + Add UAT

Properties Stage name: UAT Stage owner: Vasily Porshen

```
graph LR; subgraph Pipeline [All pipelines]; subgraph Stages [Stages]; UAT[UAT]; end; UAT --- StageProperties[Properties];
```

Azure DevOps

All pipelines > ADFcookbookReleasePipeline

Pipeline Tasks Variables Retention Options History

Artifacts | + Add Stages | + Add UAT

Source type: Azure Repo... (selected)

Project: ADF DevOps

Source (repository): ADF DevOps

Default branch: adf_publish

Default version: Latest from the default branch

Shallow fetch depth: (empty)

Source alias: _ADF DevOps

Add

```
graph LR; subgraph Pipeline [All pipelines]; subgraph Artifacts [Artifacts]; subgraph Stages [Stages]; UAT[UAT]; end; UAT --- SourceType[Source type]; SourceType --- Repo[Azure Repo...]; SourceType --- GitHub[GitHub]; SourceType --- TFVC[TFVC]; SourceType --- More[5 more artifact types]; SourceType --- Project[Project: ADF DevOps]; SourceType --- Source[Source (repository): ADF DevOps]; SourceType --- Branch[Default branch: adf_publish]; SourceType --- Version[Default version: Latest from the default branch]; SourceType --- ShallowDepth[Shallow fetch depth: (empty)]; SourceType --- Alias[Source alias: _ADF DevOps]; SourceType --- AddButton[Add];
```

Azure DevOps

All pipelines > ADFcookbookReleasePipeline

Pipeline Tasks Variables Retention Options History

UAT Deployment process

Agent job Run on agent

ARM Template deployment: Resource G... Some settings need attention

Add tasks Refresh

azure resource

Azure Monitor alerts (Deprecated)

Configure alerts on available metrics for an Azure resource (Deprecated)

ARM template deployment

Deploy an Azure Resource Manager (ARM) template to all the deployment scopes

Add

Marketplace

Delete an Azure Resource Group if it is empty

Build and Release extension that enable you to delete an Azure Resource Group if it does NOT contain any resource.

Azure Resource Group with Tags and Lock

Build and Release Task for the creation of a Azure Resource Group with specific tags and a delete lock

Azure Resource Group Tagging

Build and Release Task for performing Role Based Access Control on a Azure Resource Group

Azure VM Image Builder DevOps Task (Preview)

DevOps task for building images using Azure Image Builder Resource Provider. This is a preview version.

Azure Role Based Access Control

Project settings

Azure DevOps

All pipelines > ADFcookbookReleasePipeline

Pipeline Tasks Variables Retention Options History

UAT Deployment process

Agent job Run on agent

ARM Template deployment: Resource G... Some settings need attention

Display name *

ARM Template deployment: Resource Group scope

Azure Details

Deployment scope *

Resource Group

Azure Resource Manager connection *

Free Trial (2a8f3ccf-550c-48ea-9b06-06c2f1e03472)

Scoped to subscription 'Free Trial'

Subscription *

Free Trial (2a8f3ccf-550c-48ea-9b06-06c2f1e03472)

Action *

Create or update resource group

Resource group *

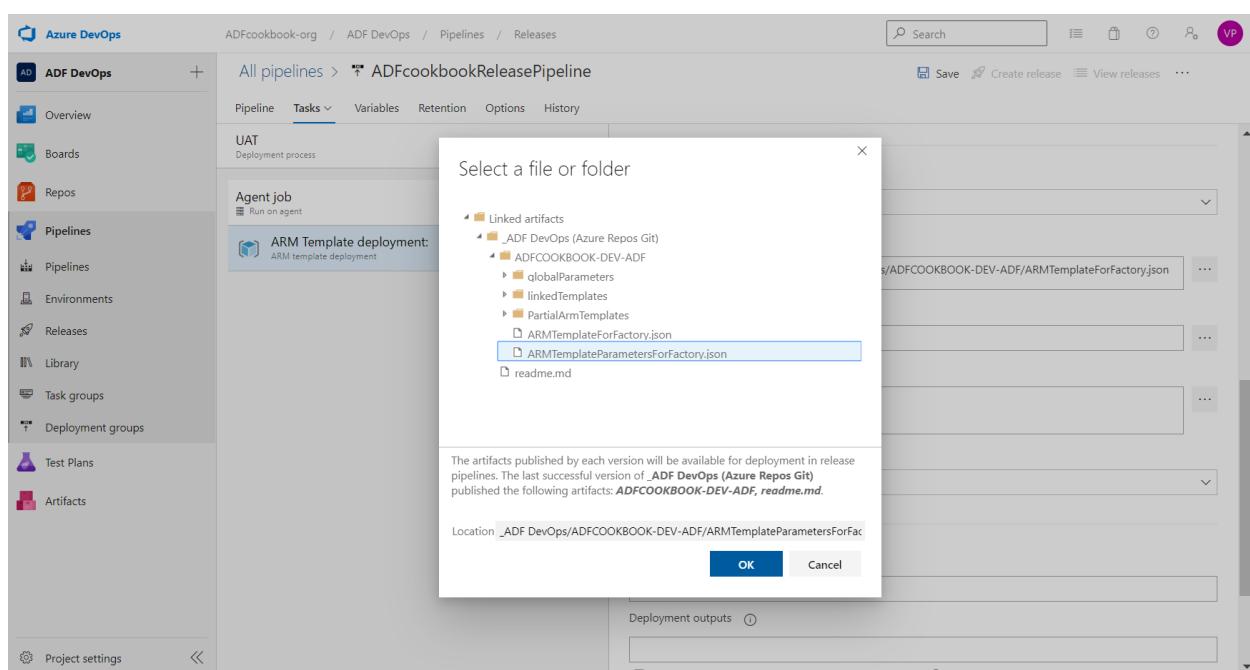
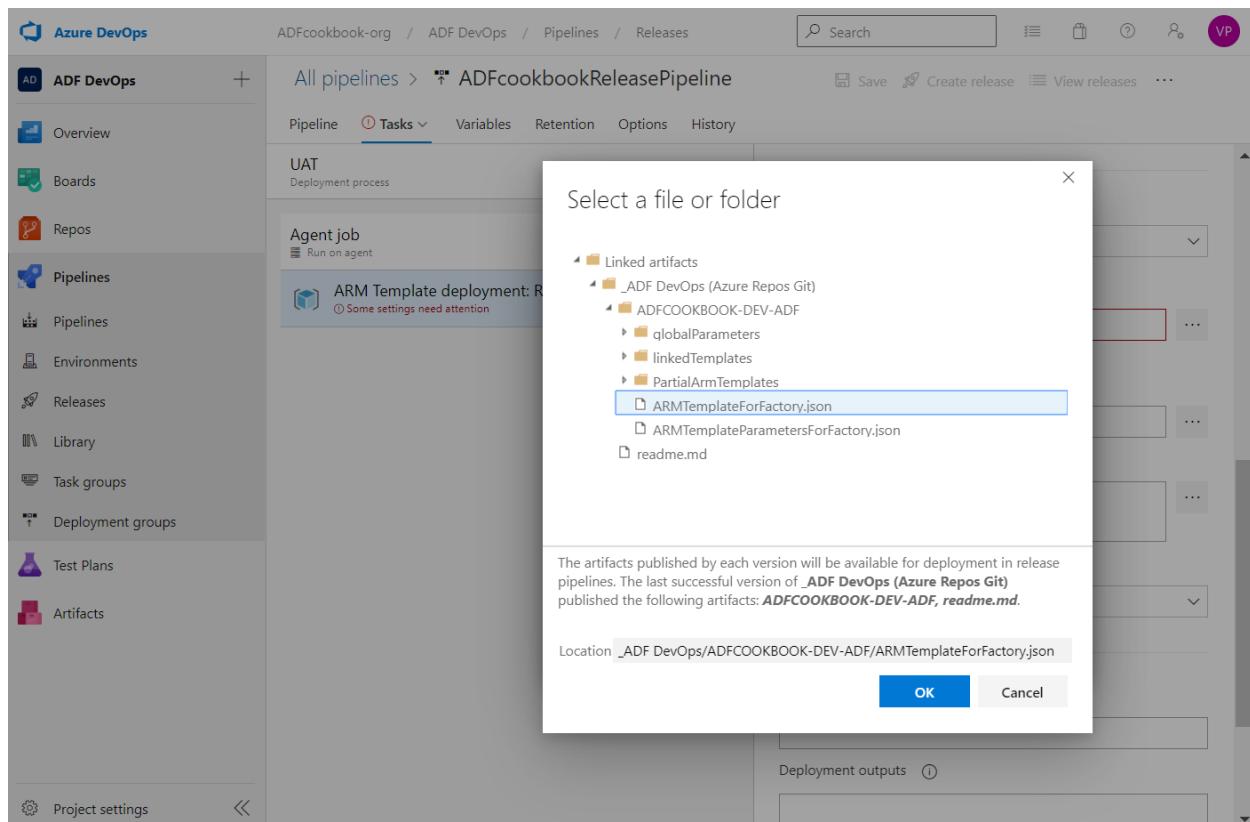
ADFCOOKBOOK-UAT-RG

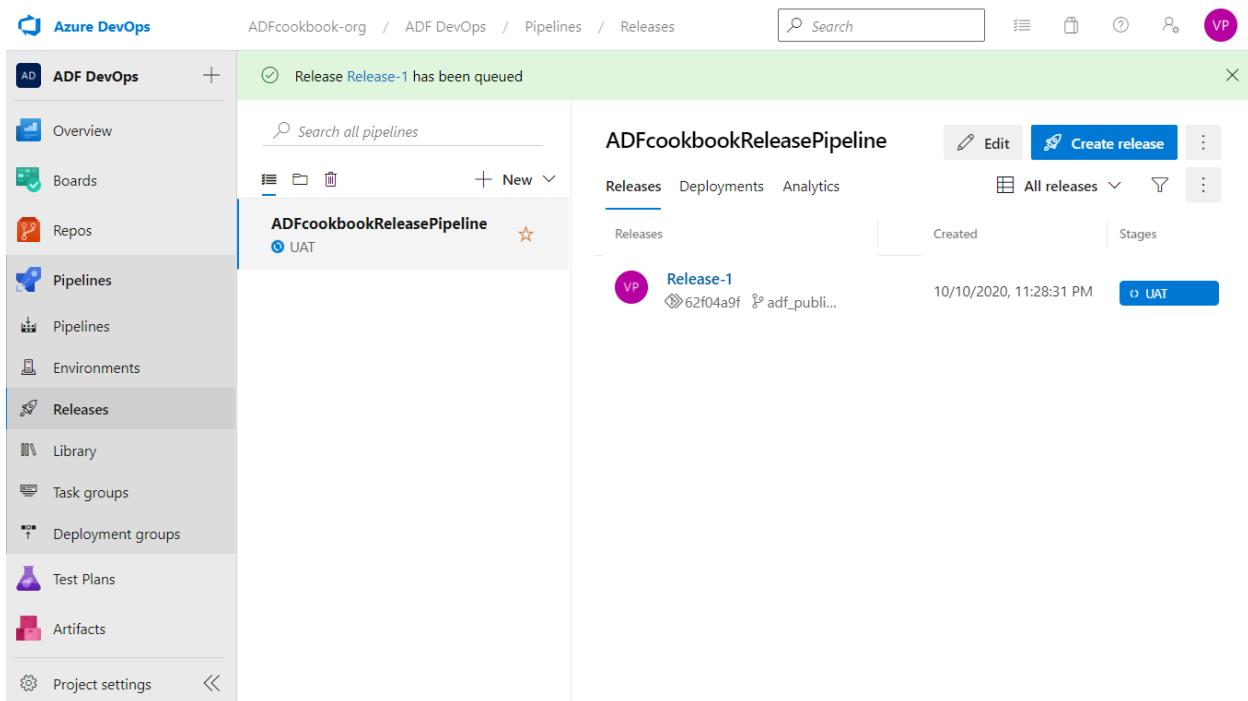
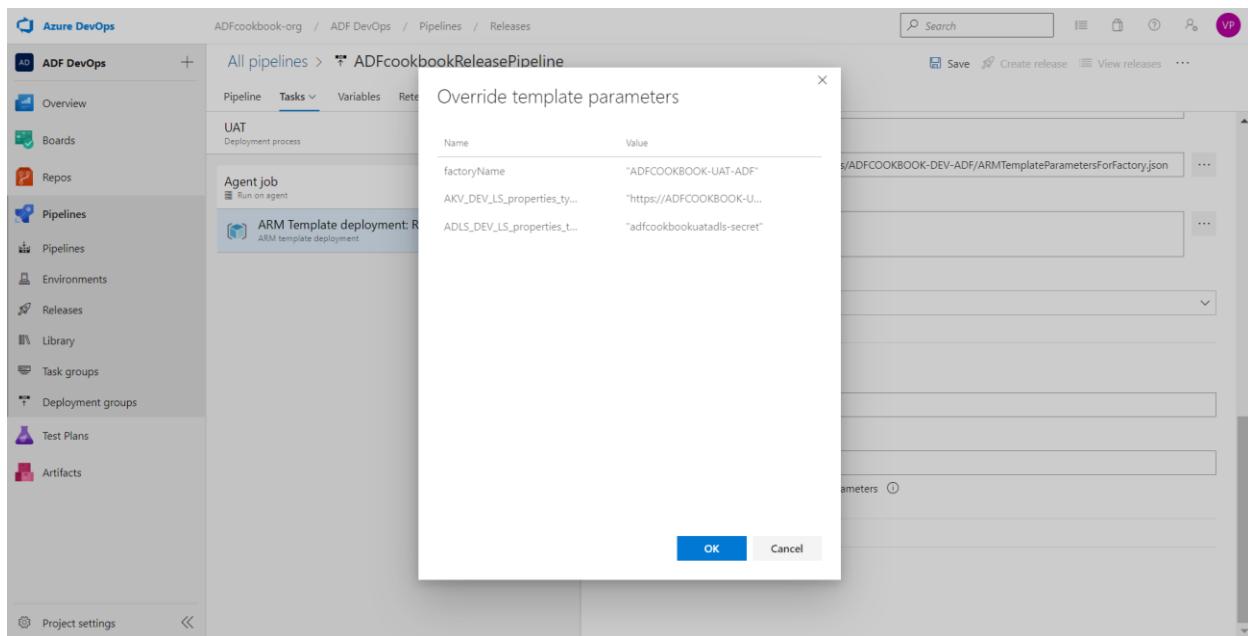
Location *

East US

Template

Project settings





Azure DevOps

ADFcookbook-org / ADF DevOps / Pipelines / Releases / ADFcookbookReleasePipeline / Release-1

ADF DevOps Overview Boards Repos Pipelines Pipelines Environments Releases Library Task groups Deployment groups Test Plans Artifacts Project settings

ADF cookbookReleasePipeline > Release-1

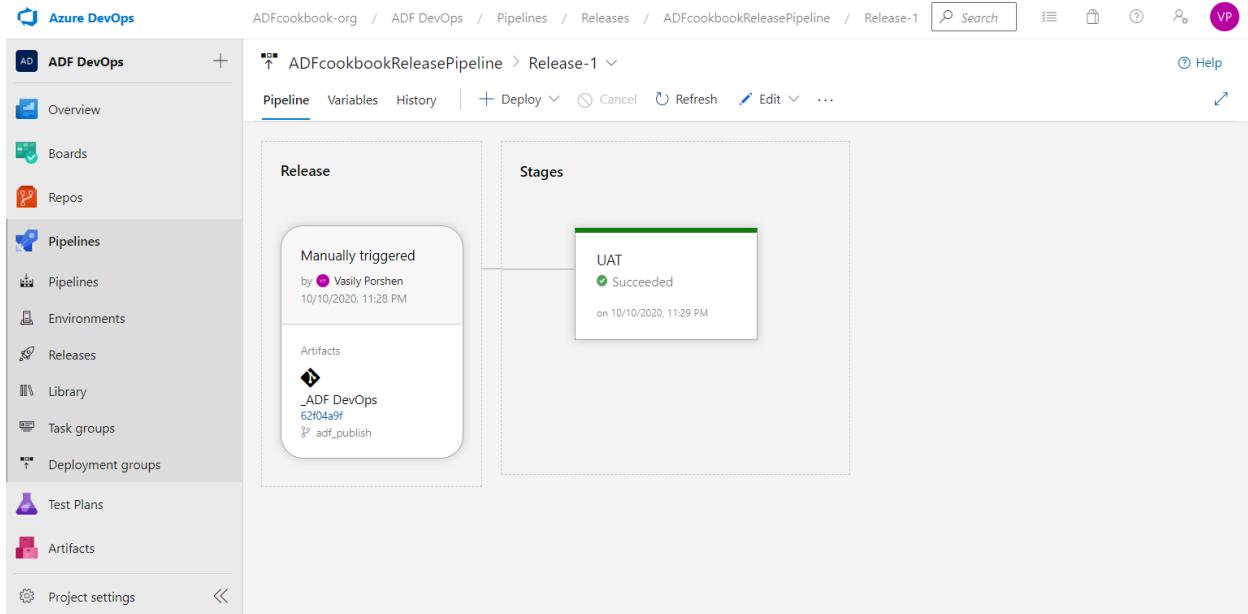
Pipeline Variables History Deploy Cancel Refresh Edit ...

Release Stages

Manually triggered by Vasily Porshen 10/10/2020, 11:28 PM

Artifacts _ADF DevOps 62f04a9f adf_publish

UAT Succeeded on 10/10/2020, 11:29 PM



Microsoft Azure | ADFCOOKBOOK-UAT-ADF

Data Factory Publish all Validate all Refresh Discard all Data flow debug ARM template

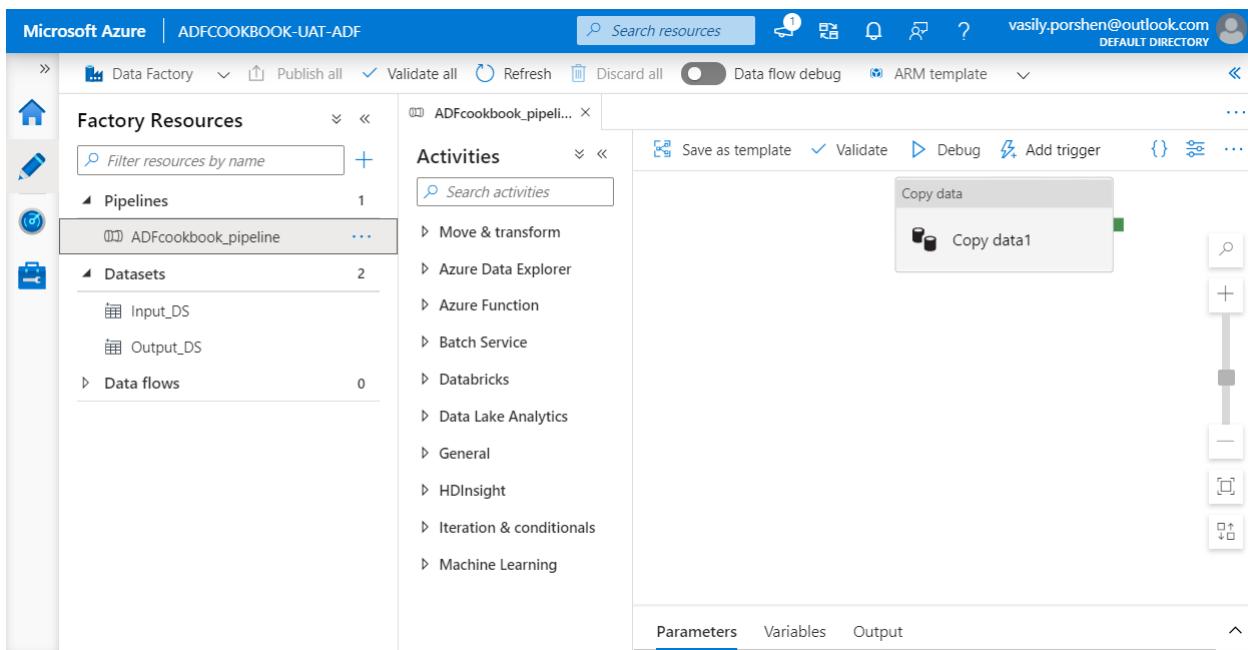
Factory Resources ADFcookbook_pipeline Activities

Pipelines ADCookbook_pipeline Datasets Input_DS Output_DS Data flows

Move & transform Azure Data Explorer Azure Function Batch Service Databricks Data Lake Analytics General HDInsight Iteration & conditionals Machine Learning

Save as template Validate Debug Add trigger

Copy data Copy data1



Azure DevOps

All pipelines > ADFcookbookReleasePipeline

Pipeline Tasks Variables Retention Options History

Artifacts | + Add Stages | + Add

ADF DevOps UAT 1 job, 1 task

Schedule not set

Continuous deployment trigger

Git: _ADF DevOps

Enabled

Creates a release every time a Git push occurs in the selected repository.

Branch filters

Type Branch

Include adf.publish

+ Add

Pull request trigger

Git: _ADF DevOps

Disabled

Enabling this will create a release every time a selected artifact is available as part of a pull request workflow

This screenshot shows the Azure DevOps pipeline configuration for the 'ADFcookbookReleasePipeline'. It features a 'Continuous deployment trigger' for the '_ADF DevOps' repository, which is enabled to create a release on every Git push. It also includes a 'Pull request trigger' for the same repository, which is currently disabled. The pipeline consists of an 'ADF DevOps' artifact and a single 'UAT' stage containing one job and one task.

Azure DevOps

All pipelines > ADFcookbookReleasePipeline

Pipeline Tasks Variables Retention Options History

Artifacts | + Add Stages | + Add

ADF DevOps UAT 1 job, 1 task

Schedule set

Scheduled release trigger

Create new releases at the specified times

Enabled

Mon through Fri at 3:00

Mon Tue Wed Thu Fri Sat Sun

03h 00m (UTC) Coordinated Universal Time

Only schedule releases if the source or pipeline has changed

+ Add a new time

This screenshot shows the same Azure DevOps pipeline configuration, but with a 'Scheduled release trigger' instead of a continuous deployment trigger. The trigger is set to run every day from Monday to Friday at 3:00 UTC. The pipeline structure remains the same, with an 'ADF DevOps' artifact and a 'UAT' stage.

 Microsoft Azure | Azure DevOps Services Documentation Support Pricing Blog Sign in

Home / Services / Azure DevOps

Azure DevOps

Plan smarter, collaborate better, and ship faster with a set of modern dev services.

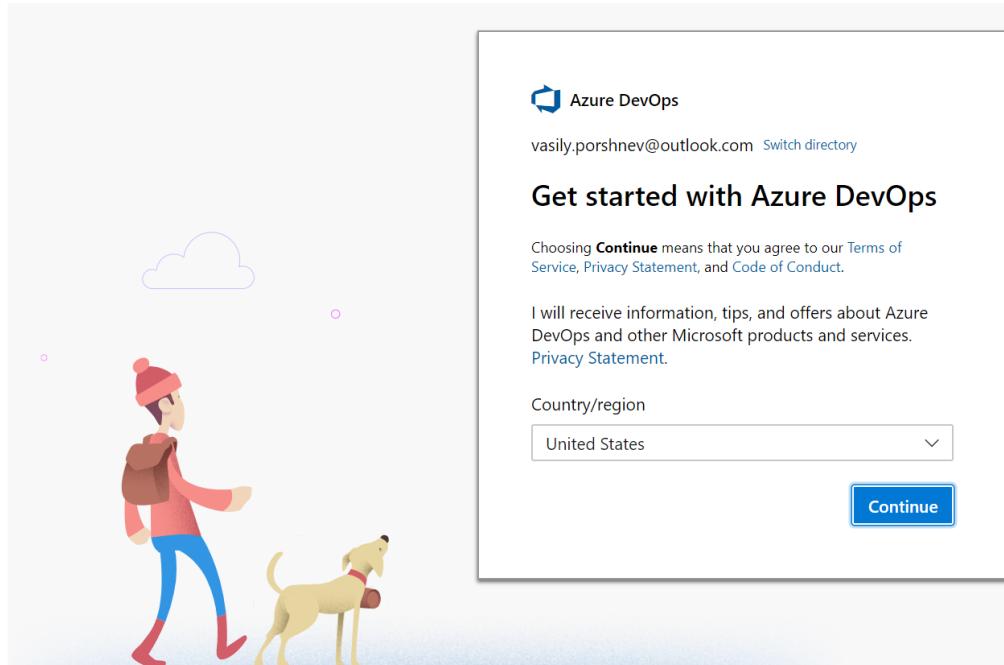
[Start free](#) [Start free with GitHub](#)

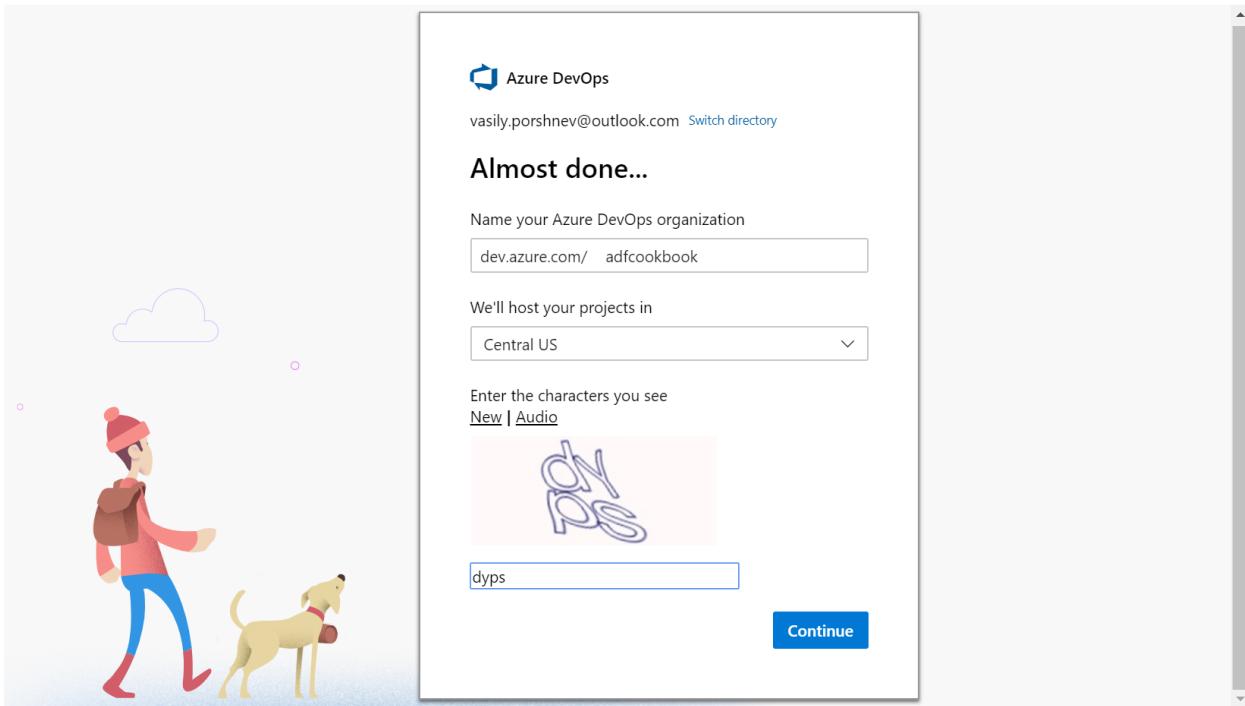
Already have an account? [Sign in to Azure DevOps >](#)



Azure DevOps Services Customer stories Documentation > Support > Pricing > Blog >

[Chat with Sales](#)





A screenshot of the Azure DevOps project creation interface. At the top, there's a navigation bar with the Azure DevOps logo, a search bar, and user profile icons. The main area shows the organization name "adfcookbook" and a title "Create a project to get started".
The "Project name *" field contains "ADF DevOps".
The "Visibility" section offers two options:

- Public**: Anyone on the internet can view the project. Certain features like TFVC are not supported.
- Private**: Only people you give access to will be able to view this project.

Azure DevOps

adfcookbook / Settings / Azure Active Directory

Organization
Settings
adfcookbook

Search Settings

General

- Overview
- Projects
- Users
- Billing
- Auditing
- Global notifications
- Usage
- Extensions
- Azure Active Directory**

Security

- Policies
- Permissions

Azure Active Directory

Connect your organization to Azure Active Directory. No content or history will be deleted.

Follow steps and learn more.

Connect directory

Azure Active Directory Connection

Connect your organization to a directory.

You are signed in as

 Vasily Porshnev
vasily.porshnev@outlook.com

Azure Active Directory

Default Directory

 Default Directory
Tenant Id: 22b33133-e461-429f-bf81-2a7ec718e759

Cancel **Connect**

Azure DevOps

adfcookbook / Settings / Azure Active Directory

Search

Organization
Settings
adfcookbook

Search Settings

General

- Overview
- Projects
- Users
- Billing
- Auditing
- Global notifications
- Usage
- Extensions
- Azure Active Directory**

Security

- Policies
- Permissions

Azure Active Directory

Your organization is connected to the **Default Directory** directory.

 **Default Directory**
vasilyporshnevoutlook.onmicrosoft.com
Tenant Id: 22b33133-e461-429f-bf81-2a7ec718e759

Check out [other frequently asked questions](#).

Disconnect directory

Download Azure DevOps organizations connected to **Default Directory** directory.

Download

Chapter 10: Monitoring and Troubleshooting Data Pipelines

The screenshot displays the Azure Data Factory Resource Monitor interface, specifically the 'Integration runtimes' and 'Resource monitor (details)' sections.

Annotations: A callout box highlights the 'Orchestration' checkbox under the 'Pipeline name' filter.

Integration runtimes:

Name	Type	Sub-type	Status	Region	Created on
AutoResolveIntegrationRuntime	Azure	Public	Running	Auto Resolve	---
CustomIntegrationRuntime	Azure	Public	Running	Auto Resolve	---
Self-Hosted-IR	Self-Hosted	---	Unavailable	More	12/10/20, 9:34:42 AM

Resource monitor (details):

STATUS Unavailable	TYPE Self-Hosted	SUB-TYPE Shared	VERSION 4.11.7521.1
RUNNING / REGISTERED NODE(S) 0 / 2	HIGH AVAILABILITY ENABLED False	LINKED COUNT 3	QUEUE LENGTH 0
AVERAGE QUEUE DURATION 0.00s			

Node Details:

Name	Status	Version	Available mem...	CPU utilization	Network (In/Out)	Concurrent job...	Role	Credential status	Message
ADFCookbookIRDev1	Unavailable	4.11.7521.1	---	---	---/---	---/---	Worker	In sync	More
ADFCookbookIRDev2	Unavailable	4.11.7521.1	---	---	---/---	---/---	Dispatcher/wor...	In sync	More

Custom deployment

Deploy from a custom template

Select a template **Basics** Review + create

Template



Customized template

2 resources



Edit template



Edit parameters

Deployment scope

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

Subscription *

Azure subscription 1

Resource group *

Create new

Parameters

Region *

West US

Data Factory Name

[concat('datafactory', uniqueString(resourceGroup().id))]

Location

[resourceGroup().location]

Storage Account Name

[concat('storage', uniqueString(resourceGroup().id))]

Blob Container

[concat('blob', uniqueString(resourceGroup().id))]

Review + create

< Previous

Next : Review + create >

Factory Resources

Filter resources by name



l_copy_from_mult...

pl

Pipelines



Actions

pl_copy_from_mult_sources

pl_orchestration_recipe_5_child

Activities

Search activities

▷ Move & transform

▷ Azure Data Explorer

Factory Resources < Use local template

Categories

- Copy
- Transform with data flow
- External activities
- SSIS
- Prep with data flow

Created by

- Microsoft
- My templates

Tag

Services used

Bulk Copy from Database to Azure Data Explorer
Use this template to copy large amount of data in bulk from database like SQL Server, Google...
by Microsoft

Bulk Copy from Database
Use this template to copy data in bulk from database using external control table to store partition list of source tables.
by Microsoft

Bulk Copy from Files to Database
Use this template to copy data in bulk from Azure Data Lake Storage Gen2 to Azure Synapse Analytics / Azure Sql Database.
If you want to copy data from a small number of...
by Microsoft

Copy and convert data from Office 365 into Common Data Model for Open Data...
Use this template to copy data from your Office 365 organization and convert it into Common Data...
by Microsoft

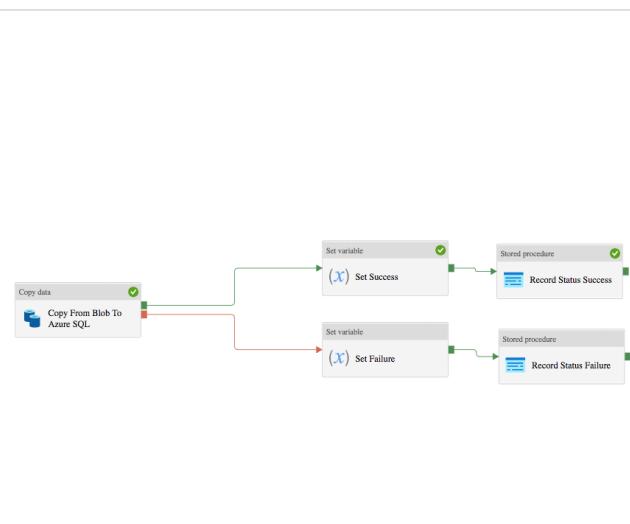
pl_troubleshooting

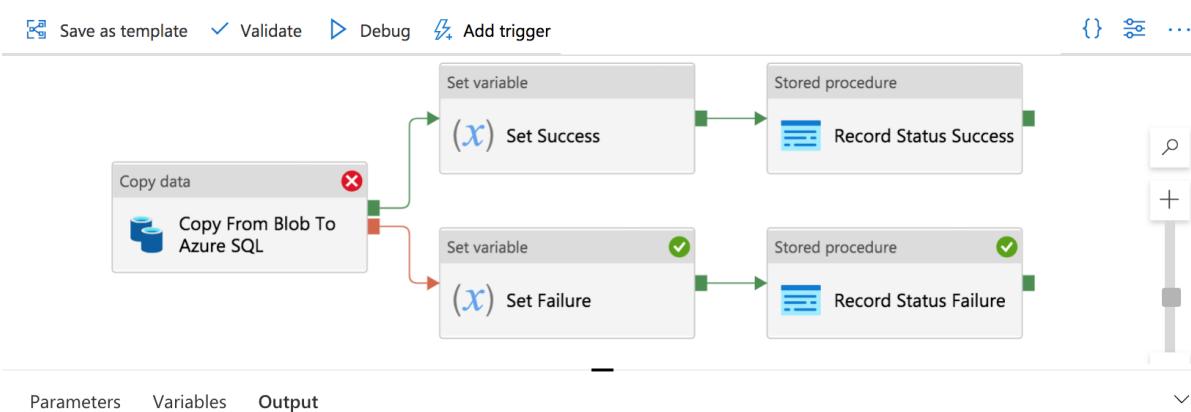
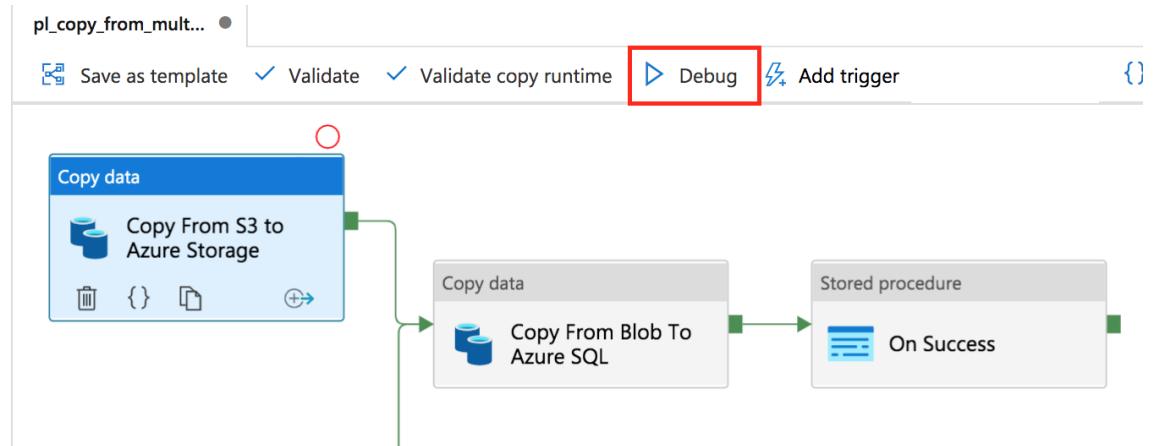
Description

User inputs ①

- CsvData** ①
DelimitedText
- AzureSQLTablesDataset** ①
Azure SQL Database
- Record Status Success** ①
Stored procedure
- Record Status Failure** ①
Stored procedure

Preview





Parameters Variables Output

Pipeline run ID: 42013279-08fe-4dec-a74c-e1335a0239ec [@] ⏪ ⏴ View debug run consumption

Name	Type	Run start	Duration	Status	Integration runtime
Record Status Failure	SqlServerStore	2020-12-10T20:58:38.542	00:00:02	✓ Succeeded	DefaultIntegrationRuntime
Set Failure	SetVariable	2020-12-10T20:58:38.042	00:00:01	✓ Succeeded	Unknown
Copy From Blob To Azure SQL	Copy	2020-12-10T20:58:31.277	00:00:06	✗ Failed	DefaultIntegrationRuntime

Parameters Variables Output

Pipeline run ID: 42013279-08fe-4dec-a74c-e1335a0239ec [@] ⏪ ⏴ View debug run consumption

Name	Type	Run start	Duration	Status	Integration runtime
Record Status Failure	SqlServerStore	2020-12-10T20:58:38.542	00:00:02	✓ Succeeded	DefaultIntegrationRuntime (West US 2)
Set Failure	SetVariable	2020-12-10T20:58:38.042	00:00:01	✓ Succeeded	Unknown
Copy From Blob To Azure SQL	Copy	2020-12-10T20:58:31.277	00:00:06	✗ Failed	DefaultIntegrationRuntime (West US 2)

Error



Troubleshoot activity failures



How helpful or unhelpful was this error message?



```
{  
    "errorCode": "2200",  
    "message":  
        "ErrorCode=UserErrorSourceBlobNotExist,'Type=Microsoft.DataTransfer.Common.Shared.HybridDeliveryException,Message=The  
        required Blob is missing. ContainerName: https://storagekg6kxqj7wtn2o.blob.core.windows.net/data, path:  
        airlines.csv.,Source=Microsoft.DataTransfer.ClientLibrary,'",  
    "failureType": "UserError",  
    "target": "Copy From Blob To Azure SQL",  
    "details": []  
}
```

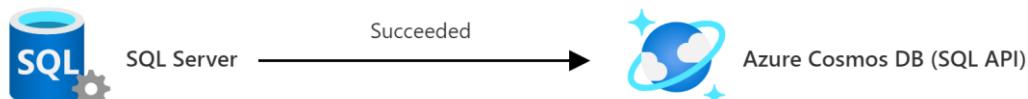
Name	Type	Run start	Duration	Status	Integration runtime	Run ID
Set Success	SetVariable	2020-10-05T03:13:23.050	00:00:01	✓ Succeeded	Unknown	008fd98e-a2
Copy From Blob	Copy	2020-10-05T03:13:15.152	00:00:07	✓ Succeeded	DefaultIntegrationRuntime (West US 2)	1c6d6eb2-2c
Backup SQL Data	Copy	2020-10-05T03:13:07.607	00:00:07	✓ Succeeded	DefaultIntegrationRuntime (East US 2)	d5aa0b73-62

Details



[Learn more on copy performance details from here.](#)

Activity run id:



Data read: 41.053 KB
Rows read: 295
Peak connections: 1

Data written: 68.709 KB
Rows written: 295
Peak connections: 4
Throughput: 4.561 KB/s

Copy duration 00:00:09

▲ SQL Server → Azure Cosmos DB (SQL API)

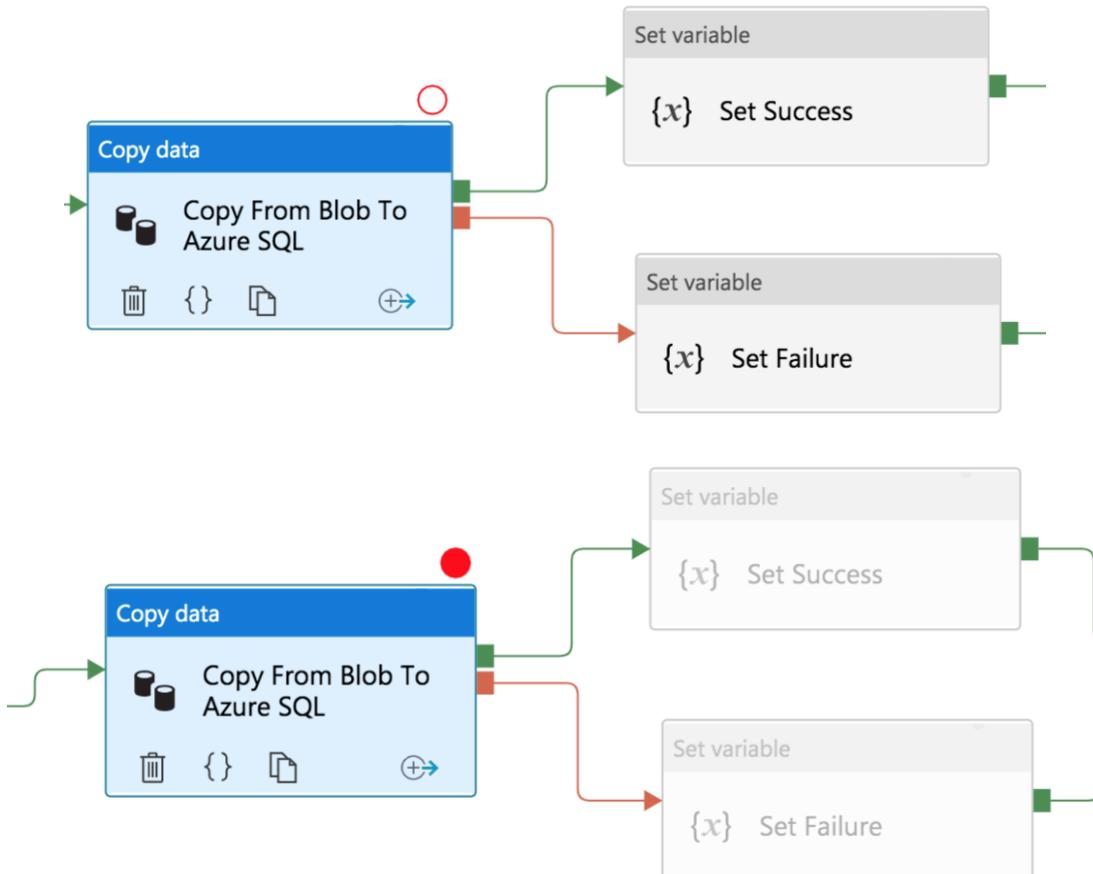
Start time 12/10/20, 4:01:24 AM

Used parallel copies 4

▲ Duration 00:00:09

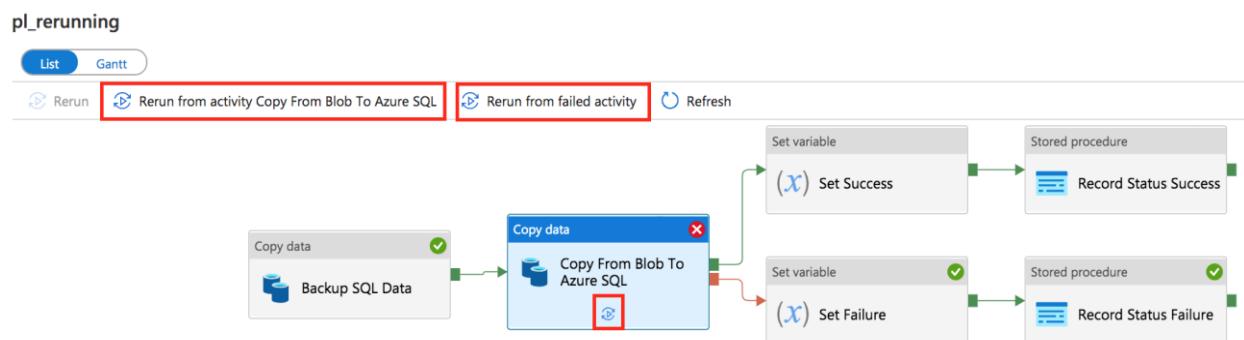
Details	Working duration	Total duration
✓ Queue ⓘ	00:00:00	00:00:07
✓ Transfer ⓘ	[Time to first byte ⓘ 00:00:00 Reading from source ⓘ 00:00:00 Writing to sink ⓘ 00:00:00]	00:00:01

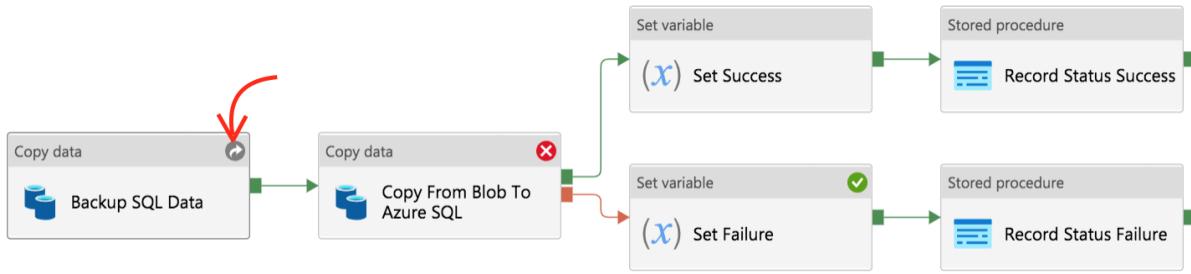
Data consistency verification ⓘ Not verified



Pipeline runs

Triggered						
<input type="button" value="Triggered"/>		<input type="button" value="Debug"/>	<input type="button" value="Rerun"/>	<input type="button" value="Cancel"/>	<input type="button" value="Refresh"/>	<input type="button" value="List"/>
<input type="text" value="Search by run ID or name"/>		<input type="text" value="Pacific Time (US & C... : Last 24 hours"/>		<input type="text" value="Pipeline name : All"/>		<input type="text" value="Status : All"/>
<input type="text" value="Runs : Latest runs"/>		<input type="button" value="Add filter"/>				
Showing 1 - 2 items						
<input type="checkbox"/> Pipeline name	Run start ↑	Run end	Duration	Triggered by	Status	Run
<input type="checkbox"/> pl_rerunning	12/6/20, 1:53:02 PM	12/6/20, 1:53:18 PM	00:00:15	Manual trigger	✖ Failed	Original
<input type="checkbox"/> pl_rerunning	<input type="button" value="Edit"/>	<input type="button" value="Rerun"/>	12/6/20, 1:52:43 PM	12/6/20, 1:53:03 PM	00:00:19	Manual trigger ✖ Failed Original





Azure subscription 1 | Resource providers

Subscription

Search (Cmd+/)
Register
Unregister
Refresh

insights

Provider	Status
Microsoft.OperationalInsights	Registered
Microsoft.D365CustomerInsights	NotRegistered
microsoft.insights	NotRegistered
Microsoft.PolicyInsights	NotRegistered
Microsoft.SecurityInsights	NotRegistered
Microsoft.TimeSeriesInsights	NotRegistered

Cost Management

- Cost analysis
- Cost alerts
- Budgets
- Advisor recommendations

Billing

- Invoices

Settings

- Programmatic deployment
- Billing properties
- Resource groups
- Resources
- Preview features
- Usage + quotas
- Policies
- Management certificates
- My permissions

Resource providers

- Deployments
- Resource locks

Support + troubleshooting

- New support request



Microsoft Azure | ADF-Cookbook-DataFactory

» Dashboards

Runs

Pipeline runs

Trigger runs

Runtimes & sessions

Integration runtimes

Data flow debug

Notifications

Alerts & metrics

Alerts & metrics

Refresh Metrics New alert rule

Configure alert logic

Failed activity runs metrics

Show history

Over the last 6 hours

Total

Selecting the dimension values will help you filter to the right time series.

Dimension	Values
ActivityType	Select a value
ActivityName	Select a value
PipelineName	Select a value
FailureType	Select a value

Alert logic

Condition *

Greater than

Time aggregation *

Total

Update criteria Cancel

Add notification

Learn more about [Pricing](#) and [Privacy statement](#).

Action name *

Failed Activities Alert

Select which notifications you'd like to receive

Email

oncall-engineers@my-company.com

SMS

Country code

Phone number *

1

▼

8135551212

Carrier charges may apply.

Azure app push notifications

Enter your email used to log into your Azure account. [Learn about connecting to your Azure resources using the Azure app.](#)

email@example.com

Voice

Country code

Phone number *

Add notification

Cancel

New alert rule

Alert rule name *

Pipeline Failures Detected

Description

Alert on-call engineers that one or more pipelines failed

Severity *

Sev2

Target criteria

Actions

Whenever Activity Failed Runs metric is Greater Than to 0



+ Add criteria

There will be a monthly rate for the configured criteria. Learn more about [Pricing](#).

Notifications

Action group type

Actions

On-Call Engineers



+ Configure notification

Enable rule upon creation



On

[Create alert rule](#)

[Cancel](#)

Alerts & metrics



Refresh



Metrics



New alert rule

ALERT

ENABLED

RESOURCE TYPE

RESOURCES

ACTIONS

Pipeline Failures Detected



On

Activity





You've been added to an Azure Monitor action group

You are now in the on-call action group and will receive notifications sent to the group.

[View details on Azure Monitor action groups >](#)

Account information

Subscription ID: [REDACTED]

Resource group name: AD [REDACTED] RG

Action group name: On-Call Engineers

The screenshot shows the Azure Data Factory pipeline editor. At the top, there are two 'Copy data' components connected by a green arrow. The left component is titled 'Backup SQL Data' and the right one is 'Copy From Blob To Azure SQL'. Below the pipeline, a navigation bar includes tabs for General, Source, Sink, Mapping, Settings, and User properties. The 'Source' tab is currently selected. Under the 'Source' tab, there is a 'Source dataset *' dropdown set to 'AzureSQLTables', along with 'Open', 'New', and 'Preview data' buttons. Below this, a table titled 'Dataset properties' lists a single entry: 'tableName' with a value of 'Country123'. A red box highlights the 'NAME' column for this row. At the bottom of the table, there is a link 'Add dynamic content [Alt+P]'. The entire interface has a light gray background with blue and green accents for the UI elements.



⚠ Your Azure Monitor alert was triggered

Azure monitor alert rule Pipeline Failure Detected was triggered for ADF-Cookbook-DataFactory at September 28, 2020 21:17 UTC.

Alert rule description	Alert on-call engineers that one or more pipeline failures detected
Rule ID	/subscriptions/ a6/resourcegroups/ADF-Orchestration-RG/providers/Microsoft.Insights/metricalerts/Pipeline%20Failure%20Detected View Rule >
Resource ID	/subscriptions/e' a6/resourcegroups/ADF-Orchestration-RG/providers/Microsoft.DataFactory/factories/ADF-Cookbook-DataFactory View Resource >

Alert Activated Because:

Metric name	ActivityFailedRuns
Metric namespace	factories/ADF-Cookbook-DataFactory
Dimensions	ResourceId = /SUBSCRIPTIONS/ 'RESOURCEGROUPS/ADF-ORCHESTRATION-RG/PROVIDERS/MICROSOFT.DATAFACTORY/FACTORIES/ADF-COOKBOOK-DATAFACTORY'
Time Aggregation	Total
Period	Over the last 1 mins
Value	1
Operator	GreaterThan
Threshold	0
Criterion Type	StaticThresholdCriterion

Pipeline runs

Triggered Debug Rerun Cancel Refresh List Gantt

Search by run ID or name Pacific Time (US & C... : Last 7 days Pipeline name : All Status : All Runs : Latest runs Add filter Copy filters

Showing 1 - 4 items

Pipeline name	Run start	Run end	Duration	Triggered by	Status	Run	Parameters	Annotations	Errors
pl_copy_from_mult_sources	12/9/20, 7:47:49 PM	--	00:00:00	Manual trigger	Queued	Original			
pl_troubleshooting	12/9/20, 7:47:42 PM	--	00:00:08	Manual trigger	In progress	Original			
pl_orchestration_recipe_5_par...	12/9/20, 7:47:36 PM	--	00:00:15	Manual trigger	In progress	Original			
pl_troubleshooting	12/9/20, 7:45:16 PM	12/9/20, 7:45:34 PM	00:00:18	Manual trigger	Failed	Original			

Pipeline runs

Triggered Debug Rerun

Search by run ID or name

Showing 1 - 5 items

- Pipeline name
- pl_orchestration_recipe_5_chil...
- pl_copy_from_mult_sources
- pl_troubleshooting
- pl_orchestration_recipe_5_par...
- pl_troubleshooting

Pipeline runs

Triggered Debug Rerun Cancel Refresh Edit columns List Gantt

Search by run ID or name Local time : Last 24 hours Status : All Runs : Latest runs Add filter Copy filters

Showing 1 - 100 items

Pipeline name	Duration	Status	Error	Run start	Run end	Triggered by	Parameters	Annotations
Copy from SSIS to Cosmos DB	00:00:07	Failed	☒	10/3/20, 6:28:11 PM	10/3/20, 6:28:19 PM	Manual trigger		
Copy from SSIS to Cosmos DB	00:00:02	Succeeded	✓	10/3/20, 6:27:09 PM	10/3/20, 6:27:11 PM	Manual trigger		
Copy from SSIS to Cosmos DB	00:00:01	Succeeded	✓	10/3/20, 6:26:44 PM	10/3/20, 6:26:46 PM	Manual trigger		
Copy from SSIS to Cosmos DB	00:00:07	Succeeded	✓	10/3/20, 6:00:01 PM	10/3/20, 6:00:08 PM	Hourly Data Movement		
Data Movement CosmosDB to...	00:00:37	Succeeded	✓	10/3/20, 6:00:01 PM	10/3/20, 6:00:38 PM	Hourly Data Movement		
Copy from SSIS to Cosmos DB	00:00:12	Succeeded	✓	10/3/20, 5:00:01 PM	10/3/20, 5:00:13 PM	Hourly Data Movement		
Data Movement CosmosDB to...	00:01:32	Succeeded	✓	10/3/20, 5:00:01 PM	10/3/20, 5:01:34 PM	Hourly Data Movement		
Data Movement CosmosDB to...	00:01:30	Succeeded	✓	10/3/20, 4:00:01 PM	10/3/20, 4:01:31 PM	Hourly Data Movement		
Copy from SSIS to Cosmos DB	9/30/20, 10:59:59 AM	9/30/20, 11:00:02 AM	00:00:02	Hourly Data Movement	Succeeded			
Data Movement C...	9/30/20, 10:00:00 AM	9/30/20, 10:00:21 AM	00:00:20	Hourly Data Movement	Succeeded			
Copy from SSIS to Cosmos DB	9/30/20, 10:00:00 AM	9/30/20, 10:00:06 AM	00:00:05	Hourly Data Movement	Succeeded			

Pipeline run consumption

X

	Quantity	Unit
Pipeline orchestration		
Activity Runs	4	Activity runs
Pipeline execution		
Azure integration runtime		
Data movement activities	0.0000	DIU-hour
Pipeline activities	0.0167	Execution hours
External activities	0.0333	Execution hours
Learn more		
Pricing calculator		

