



Consume Data from REST API into Azure Data Lake

Table of Contents

1. Introduction.....	3
2. Solution Components.....	3
3. Solution Diagram	3
4. Detailed Steps	4
5. References	8

1. Introduction

This document outlines how to invoke a REST API using Azure PaaS components (Azure Data Factory i.e. ADF) and then ingest the data into ADLS or Azure SQL DB into tabular structure for further use.

2. Solution Components

A **RESTful API** is an application program interface (API) that uses HTTP requests to GET, PUT, POST and DELETE data and is also referred to as a RESTful web service -- is based on representational state transfer (REST) technology, an architectural style and approach to communications often used in web services development.

Azure Data Factory is the platform for these kinds of scenarios. It is a cloud-based data integration service that allows you to create data-driven workflows in the cloud that orchestrate and automate data movement and data transformation.

Microsoft **Azure Data Lake** is a highly scalable public cloud service that allows developers, scientists, business professionals and other Microsoft customers to gain insight from large, complex data sets.

Microsoft **Azure SQL Database** is a managed cloud database provided as part of Microsoft Azure. A cloud database is a database that runs on a cloud computing platform, and access to it is provided as a service. Managed database services take care of scalability, backup, and high availability of the database.

3. Solution Diagram



Fig-1

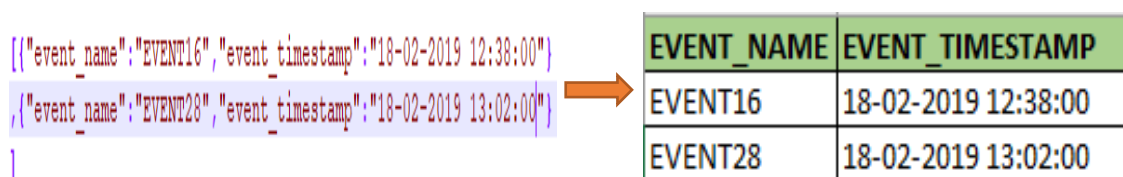


Fig-2

4. Detailed Steps

Step 1 - Identify the REST API you need to consume along with the pre-requisites (credentials, api keys or auth tokens etc.) handy.

E.g - You can register yourself in <https://developer.nytimes.com/apis> and generate the required API key/tokens to access the REST API hosted there but here I will demonstrate the simple case (no authentication needed) with anonymous user.

Step 2 – Create the ADF pipeline with HTTP as source and SQL DB as destination.

The image displays three screenshots from the Azure Data Factory (ADF) portal illustrating the initial setup steps:

- New data factory:** The 'Create new' form is shown with the following details:
 - Name: *Name*
 - Subscription: *Free Trial*
 - Resource Group: *my-demo-rg* (selected under 'Use existing')
 - Version: *V1*
 - Location: *East US*The 'Create' button is visible at the bottom.
- Data factories list:** The 'my-demo-adf' data factory is listed. In the 'Actions' section, the 'Copy data (PREVIEW)' icon is circled in red.
- Copy Data (my-demo-adf) configuration:** The 'Specify HTTP connection' screen is shown with the following settings:
 - Connection name: *Source-Http-c1p*
 - Source: *Source-Http-c1p*
 - DataSourceType: *Cloud Data Source*
 - URL (required): *https://jsonplaceholder.typicode.com/users/* (This field is circled in red with a callout: "Pass access tokens/API keys as URL parameter (if required)")
 - Server Certification Validation: *Enable*
 - Authentication type: *Anonymous Authentication*
 - Advanced settings: *Http Method: Get*The 'Next' button is highlighted in blue.

Copy Data (test-adf-sandip)

1 Properties
One-time copy

2 Source
Connection
Dataset

3 Destination
Settings
Summary
Deployment

File format settings

File format: JSON format

File pattern: Array of objects

Export as is to JSON files or Cosmos DB collection

JSONPath settings

Copy only nested JSON array: None

Batch mode

Column name	JSONPath expression	Remove
id	\$.id	Remove
name	\$.name	Remove
username	\$.username	Remove
email	\$.email	Remove
phone	\$.phone	Remove

PREVIEW SCHEMA

Previous Next

Pivot around the root element under which there are multiple nested elements

Verify the individual elements & their relative paths during extraction

Step 3 – Select SQL DB as destination of the ADF pipeline. If you just want to dump the data into storage, then you can choose either Blob or ADLS as destination instead of SQL DB.

Copy Data (test-adf-sandip)

1 Properties
One-time copy

2 Source
HTTP

3 Destination
Connection
Dataset

4 Settings
Fault tolerance, Performance

5 Summary

6 Deployment

Destination data store

Specify the destination data store for the copy task. You can use an existing data store connection or specify a new data store. Click [HERE](#) to suggest new copy destinations or give comments.

FROM EXISTING CONNECTIONS CONNECT TO A DATA STORE

Azure Blob Storage	Azure Data Lake Store	Azure Cosmos DB	Azure SQL Database	Azure SQL Data Warehouse	Azure Table Storage

Previous Next

Copy Data (test-adf-sandip)

1 Properties
One-time copy

2 Source
HTTP

3 Destination
Connection
Dataset

4 Settings
Fault tolerance, Performance

5 Summary

6 Deployment

Specify the Azure SQL database

Connection name (required)
Destination-SQLAzure-q9i

Server / database selection method (required)
From Azure subscriptions

Azure subscription (required)
Free Trial (e46bee66-4fc6-4188-8d80-06f6385a265f)

Server name (required)
my-demo-sql-db-server

Database name (required)
my-demo-sql-db

User name (required)
sandip

Password (required)

Previous Next

Copy Data (my-demo-adf)

1 Properties
One time copy

2 Source
HTTP

3 Destination
Azure SQL Database

4 Settings
Fault tolerance, Performance

5 Summary

6 Deployment

Table mapping

For each table you have selected to copy in the source data store, select a corresponding table in the destination data store or specify the stored procedure to run at the destination.

Source	Destination
Http location: /	[dbo].[users]

☐ Use stored procedure

PREVIEW	SCHEMA																																												
<table> <thead> <tr> <th>Id</th> <th>full_name</th> <th>username</th> <th>email</th> <th>phone</th> <th>website</th> <th>address_street</th> <th>address_suite</th> <th>address_city</th> <th>address_zipcode</th> <th>adc</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>Leanne Graham</td> <td>Bret</td> <td>Sincere@aprilbiz</td> <td>1-770-730-8031 x56442</td> <td>hilegard.org</td> <td>Kulas Light</td> <td>Apt. 556</td> <td>Gwenborough</td> <td>92998-3874</td> <td>-37</td> </tr> <tr> <td>2</td> <td>Ervin Howell</td> <td>Antonette</td> <td>Shanna@melissa.tv</td> <td>010-092-6593 x09125</td> <td>anastasia.net</td> <td>Victor Plains</td> <td>Suite 879</td> <td>Wiskeyburgh</td> <td>90566-7771</td> <td>-43</td> </tr> <tr> <td>3</td> <td>Clemantine Bauch</td> <td>Samantha</td> <td>Nathan@yesenia.net</td> <td>1-463-123-4447</td> <td>ramiro.info</td> <td>Douglas Extension</td> <td>Suite 847</td> <td>McKenziehaven</td> <td>59590-4157</td> <td>-68</td> </tr> </tbody> </table>	Id	full_name	username	email	phone	website	address_street	address_suite	address_city	address_zipcode	adc	1	Leanne Graham	Bret	Sincere@aprilbiz	1-770-730-8031 x56442	hilegard.org	Kulas Light	Apt. 556	Gwenborough	92998-3874	-37	2	Ervin Howell	Antonette	Shanna@melissa.tv	010-092-6593 x09125	anastasia.net	Victor Plains	Suite 879	Wiskeyburgh	90566-7771	-43	3	Clemantine Bauch	Samantha	Nathan@yesenia.net	1-463-123-4447	ramiro.info	Douglas Extension	Suite 847	McKenziehaven	59590-4157	-68	
Id	full_name	username	email	phone	website	address_street	address_suite	address_city	address_zipcode	adc																																			
1	Leanne Graham	Bret	Sincere@aprilbiz	1-770-730-8031 x56442	hilegard.org	Kulas Light	Apt. 556	Gwenborough	92998-3874	-37																																			
2	Ervin Howell	Antonette	Shanna@melissa.tv	010-092-6593 x09125	anastasia.net	Victor Plains	Suite 879	Wiskeyburgh	90566-7771	-43																																			
3	Clemantine Bauch	Samantha	Nathan@yesenia.net	1-463-123-4447	ramiro.info	Douglas Extension	Suite 847	McKenziehaven	59590-4157	-68																																			

Previous

Next

Step 4 – Specify the settings

Copy Data (my-demo-adf)

1 Properties
One time copy

2 Source
HTTP

3 Destination
Azure SQL Database

4 Settings
Fault tolerance, Performance

5 Summary

6 Deployment

Settings

More options for data movement

^ Fault tolerance settings

Error handling for incompatible rows between source and destination ⓘ

Actions

^ Performance settings

^ Advanced settings

Parallel copy ⓘ

Cloud units

Parallel copies

Previous

Next

Step 5 – Review the Summary and initiate Deployment.

Copy Data (my-demo-adf)

1 Properties
One time copy

2 Source
HTTP

3 Destination
Azure SQL Database

4 Settings
Fault tolerance, Performance

5 Summary

6 Deployment

Summary

You are running one time pipeline to copy data from HTTP to Azure SQL Database.

HTTP / Region: Unknown
 Copy Run Time Region: West Europe
→
 Azure SQL Database 1 table(s) Region: West Europe

Properties

Task name CopyPipeline-c1p

Task description <no description has been provided>

Expiration time 3:00:00:00

Source

Connection HTTP

Connection name Source-Http-c1p

Dataset name InputDataset-c1p

Http location /

Region Unknown

Minimum size in MB

Data delay 00:00:00

Maximum retry

Previous

Next

Copy Data (my-demo-adf)

- 1 Properties
One time copy
- 2 Source
HTTP
- 3 Destination
Azure SQL Database
- 4 Settings
Fault tolerance, Performance
- 5 Summary
- 6 Deployment

HTTP / Region: Unknown → Copy Run Time Region: West Europe → Azure SQL Database / 1 table(s) / Region: West Europe

Deployment complete

- Validating runtime environment ✓
- Validation passed ✓
- Registering Connections ✓
- Creating Pipelines ✓
- Provisioning Activities ✓

[Click here to monitor copy pipeline](#)

Data factory

RESOURCE EXPLORER

- Data Factories
 - my-demo-adf
 - Pipelines
 - CopyPipeline-clip
 - Datasets
 - Linked services
 - Destination-BlobStor...
 - Destination-SQLAzur...
 - Destination-SQLAzur...
 - Source-Http-clip
 - Source-Http-v54
 - Source-Http-v5k
 - Gateways

my-demo-adf / CopyPipeline-clip

No scheduled runs for this pipeline

This pipeline has no content to display.

ACTIVITY WINDOWS

1 filter applied

Pipeline	Activity	Window Start	Window End	Status	Type	Last Attempt Start	Last Attempts End	Duration	Retry Attempts
CopyPipeline-clip	Activity-0-Http L...	03/04/2019 12:00 AM UTC	03/05/2019 12:00 AM UTC	Ready	Copy	03/07/2019 12:00 AM UTC	03/07/2019 12:00 AM UTC	00:01:00	1

Activity Window Explorer

MAR 2019

S	M	T	W	T	F	S
24	25	26	27	28	1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31	1	2	3	4	5	6

03/04/2019 12:00 AM UTC - 03/05/2019 12:00 AM UTC

Activity Window

Start & End Time
03/04/2019 12:00 AM UTC - 03/05/2019 ...

Status
Ready

Activity
Activity-0-Http location...-[dbol_users]

Pipeline
CopyPipeline-clip

Input Datasets
InputDataset-clip

Step 6 – Once the deployment status is ready, go to SQL DB window and check whether records from REST API response have been captured there or not.

Dashboard > my-demo-sql-db - Query editor (preview)

my-demo-sql-db - Query editor (preview)

Query 1

```
select * from users;
```

Results

ID	FULL NAME	USERNAME	EMAIL	PHONE	WEBSITE	AD
1	Leanne Graham	Bret	Sincere@april.biz	1-770-736-8031 x56442	hildegard.org	Kid
2	Ervin Howell	Antonette	Shanna@melissa.tv	010-692-6593 x09125	anastasia.net	Vic
3	Clementine Bauch	Samantha	Nathan@yvesnia.net	1-463-123-4447	ramiro.info	Do
4	Patricia Lebsack	Karianna	Julianne.OConner@kory.org	493-170-9423 x156	kalle.biz	Ho
5	Chelsey Dietrich	Kamren	Lucio_Hettinger@annie.ca	(254)954-1289	demarco.info	Skil
6	Mrs. Dennis Schulist	Leopoldo_Corkey	Karley_Dach@jasper.info	1-477-935-8478 x6430	ola.org	No

Query succeeded | 0s

5. References

<https://docs.microsoft.com/en-us/azure/data-factory/connector-http>

<https://docs.microsoft.com/en-us/azure/data-factory/tutorial-copy-data-tool>

<https://docs.microsoft.com/en-us/azure/sql-database/>