

Connecting ADF to ADLS Gen1

Okay, let's walk through the steps to connect ADF to ADLS.

Connecting ADF to ADLS is completed through the use of the built-in Linked Services connector.

A linked service can be thought of as a data connector and defines the specific information required to connect to that data source i.e. ADLS, Azure Blob Storage, Azure SQL etc.


Create the Linked Service

Within the Data Factory portal select **Connections -> Linked Services** and then **Data Lake Storage Gen1**:

New Linked Service



Data Store Compute

 data lake

All Azure Database File Generic Protocol NoSQL Services and apps



Azure Data Lake Storage
Gen1



Azure Data Lake Storage
Gen2 (Preview)

Cancel

Continue

Click **Continue** and we're prompted to provide the Data Lake store's details. Assuming you already have a data lake store created go ahead and select your store:

Edit Linked Service (Azure Data Lake Storage Gen1) ×

Name *

AzureDataLakeStore1

Description

Connect via integration runtime *



AutoResolveIntegrationRuntime

Data Lake Store selection method



From Azure subscription

Azure subscription



Select all

Data Lake Store account name *

ak1

Tenant *

17782221-0000-4488-8888-888888888888

Authentication type *

Service Principal

Service principal ID *



Cancel

Test connection

Finish

Most of these settings are self-explanatory but the complexities are around the authentication type.

In this example we are going to use a ***Service Principal*** (SPN) in Azure Active Directory (Azure AD). Creating an SPN allows us to grant access to the Data Lake Store by Data Factory.

Selecting the SPN option prompts us for the following information:

Authentication type *

Service Principal

Service principal ID *

Service principal key *

Service principal key

Azure Key Vault

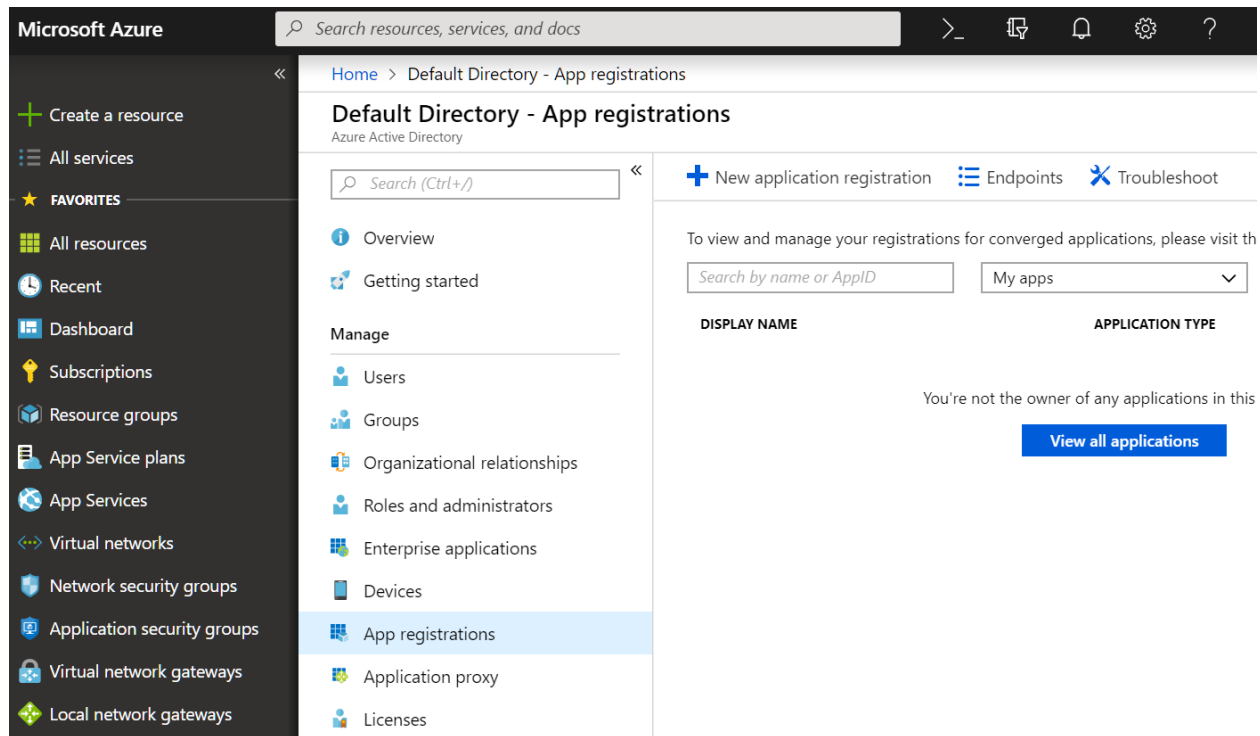
Best practice is to also store the SPN key in Azure Key Vault but we'll keep it simple in this example.

Create the Service Principal

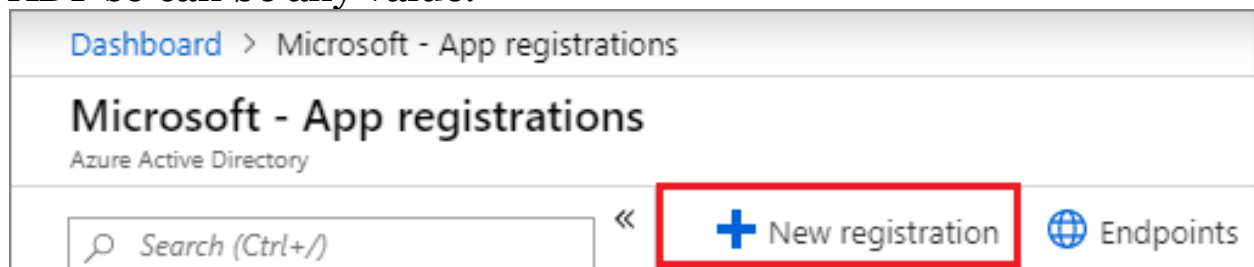
The next step is to create the SPN in Azure AD (you'll need the appropriate Azure AD permissions to do this).

At this point it's easiest to open a new browser tab and open another copy of the Azure Portal (we want to come back to the Data Factory config in moment).

Within the Azure Portal select **Azure Active Directory** -> **App registrations** and then **New application registration**:




This brings up the **Create** blade. Provide a meaningful name and ensure the type is set to **Web app / API**. The URL isn't used by ADF so can be any value:



Dashboard > Microsoft - App registrations > Register an application

Register an application

 If you are building an application for external users that will be distributed by Microsoft, you must register as a first party application to meet all security, privacy, and compliance policies. [Read our decision guide](#)

*** Name**

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

example-app ✓

Supported account types

Who can use this application or access this API?

☒ Accounts in this organizational directory only (Microsoft)

☐ Accounts in any organizational directory

☐ Accounts in any organizational directory and personal Microsoft accounts (e.g. Skype, Xbox, Outlook.com)

[Help me choose...](#)

Redirect URI (optional)

We'll return the authentication response to this URI after successfully authenticating the user. Providing this now is optional and it can be changed later, but a value is required for most authentication scenarios.

Web ▼ https://contoso.org/exampleapp ✓

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

Register

Click the **Register** button to complete the registration of the new application. Once created the applications details will be displayed:

Dashboard > Microsoft - App registrations > example-app

example-app

Overview

Quickstart

Manage

Branding

Authentication

Certificates & secrets

Delete

Endpoints

Display name
example-app

Application (client) ID
36fa6513-66cd-4c90-a7b2-3fdbe89d1630

Directory (tenant) ID
72f988bf-86f1-41af-91ab-2d7cd011db47

Object ID
4a48ed60-a01c-466f-a3bb-29a530707347

Copy to clipboard

Dashboard > Microsoft - App registrations > example-app

example-app

Overview

Quickstart

Manage

Branding

Authentication

Certificates & secrets

Delete

Endpoints

Display name
example-app

Application (client) ID
a8d0a934-4bf4-4a53-bfca-292751f2bd53

Directory (tenant) ID

Object ID

Copy to clipboard

The **Application ID** is actually what ADF refers to as the **Service principal ID**. **Copy** the Application ID, switch back to the other browser tab with Data Factory and **paste** the ID into the field:

Authentication type *

Service Principal

Service principal ID *



4934d3b0-b820-44ef-9d8d-65b30b6bd678

Service principal key

Azure Key Vault

Service principal key *

|



Generate a Service Principal Key

The next step is to generate the SPN key. Back in the **Registered app** blade for our SPN select **Settings** and then **Certificates & Secrets**:

[Dashboard](#) > [Microsoft - App registrations](#) > example-app

example-app

«

Overview

Quickstart

Manage

Branding

Authentication

Certificates & secrets

Delete

Endpoints

Display name

example-app

Application (client) ID

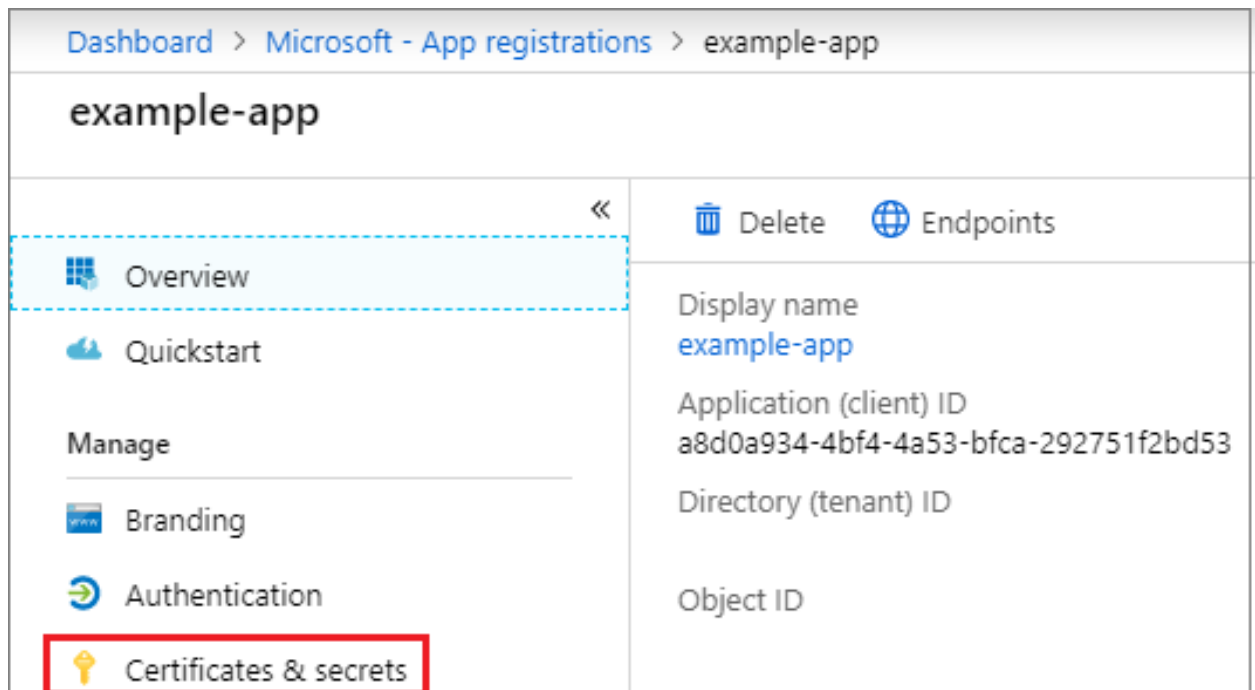
a8d0a934-4bf4-4a53-bfca-292751f2bd53

Directory (tenant) ID

Object ID

Enter a password **description** and **expiry**:

1. Select **Certificates & secrets**.



Dashboard > Microsoft - App registrations > example-app

example-app

Overview

Quickstart

Manage

Branding

Authentication

Certificates & secrets

«

Delete

Endpoints

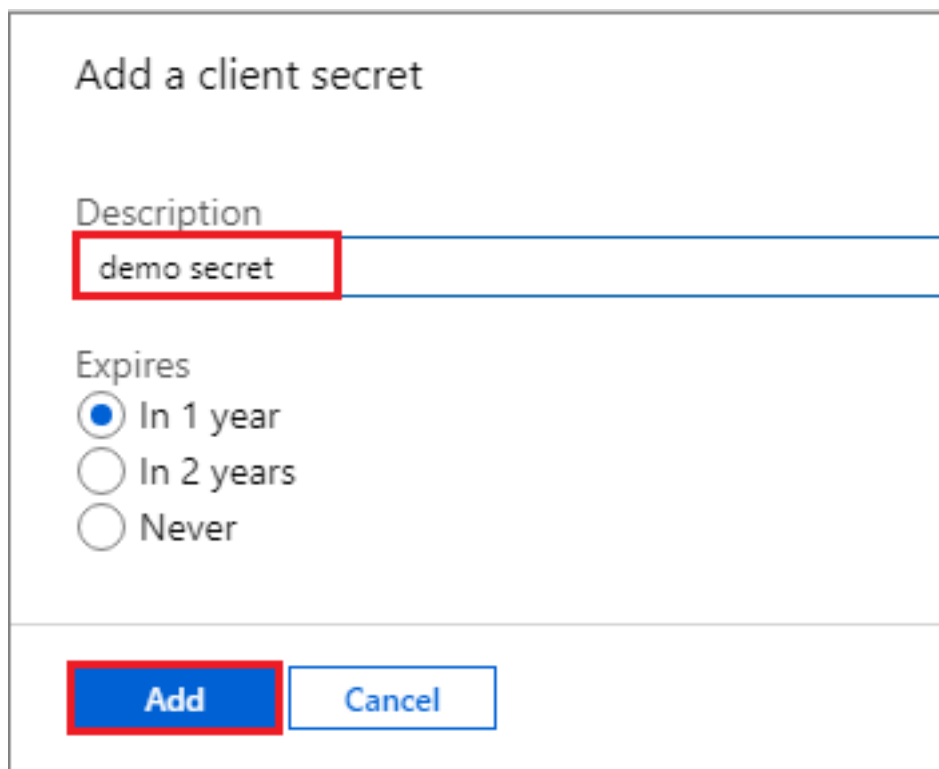
Display name
example-app

Application (client) ID
a8d0a934-4bf4-4a53-bfca-292751f2bd53

Directory (tenant) ID

Object ID

2. Select **Client secrets** -> **New client secret**.
3. Provide a description of the secret, and a duration. When done, select **Add**.



Add a client secret

Description

demo secret

Expires


☒ In 1 year

☐ In 2 years

☐ Never

Add Cancel

After saving the client secret, the value of the client secret is displayed. Copy this value because you aren't able to retrieve the key later. You provide the key value with the application ID to sign in as the application. Store the key value where your application can retrieve it.

Client secrets		
A secret string that the application uses to prove its identity when requesting a token. Also can be referred to as application password.		
+ New client secret		
DESCRIPTION	EXPIRES	VALUE
demo secret	5/14/2020	nWu9HVZ7Rnj.2y7XSkVyUngZ][x9Z:e 

Switch back to Data Factory and **paste** in the key:

Authentication type *

Service Principal


Service principal ID *

4934d3b0-b820-44ef-9d8d-65b30b6bd678

Service principal key

Azure Key Vault

Service principal key *

.....| 

Click **Finish** to close the linked service. We'll come back to it later to retest once we've assigned the required permissions.

We now need to explicitly grant permissions to Data Factory so that it can manipulate data in our Data Lake.

As you'll see below assigning the permissions is quite convoluted so I've separated these into steps 1–3 below.

Permissions—Step 1

Return to the registered app in **Azure AD**, select **Settings - > Required Permissions**:

*note—if you've navigated away from the Azure AD blade you'll find your app under **App registrations** in Azure AD*

Home > Default Directory - App registrations > adf-adlsgen1-app > Settings

adf-adlsgen1-app

Registered app

Settings Manifest Delete

Display name	Application ID
adf-adlsgen1-app	4934d3b0-b820-44ef-9d8d-65b30b6bd678
Application type	Object ID
Web app / API	a74a34e6-8dea-4ac7-8f43-8f54a03ac2a9
Home page	Managed application in local directory
http://a.dummy.url	adf-adlsgen1-app

Settings

Filter settings

GENERAL

- Properties
- Reply URLs
- Owners

API ACCESS

- Required permissions
- Keys

TROUBLESHOOTING + SUPPORT

- Troubleshoot
- New support request

Click on **Add**:

Required permissions

+ Add Grant permissions

API	APPLICATION PERMI...	DELEGATED PERMIS...
Windows Azure Active Directory (Microsoft.Azure.Act...	0	1

Under **Select an API** highlight **Azure Data Lake** and click **Select**:

Add API access

1 Select an API
Azure Data Lake

2 Select permissions
0 role, 1 scope

Select an API

Search for other applications with Service Principal name

Windows Azure Active Directory (Microsoft.Azure.ActiveDirectory)

Microsoft Graph

Azure Key Vault

Windows Azure Service Management API

Azure Data Lake

Azure DevOps (Microsoft.VisualStudio.Online)

Office 365 Management APIs

Under **Select permissions** ensure **Delegated Permissions** are highlighted. Click **Select** and then **Done**:

Add API access

1 Select an API
Azure Data Lake

2 Select permissions
0 role, 1 scope

Enable Access

☐ APPLICATION PERMISSIONS

No application permissions available.

☒ DELEGATED PERMISSIONS

☒ Have full access to the Azure Data Lake service

The first set of permissions have now been added:

Required permissions

+ Add

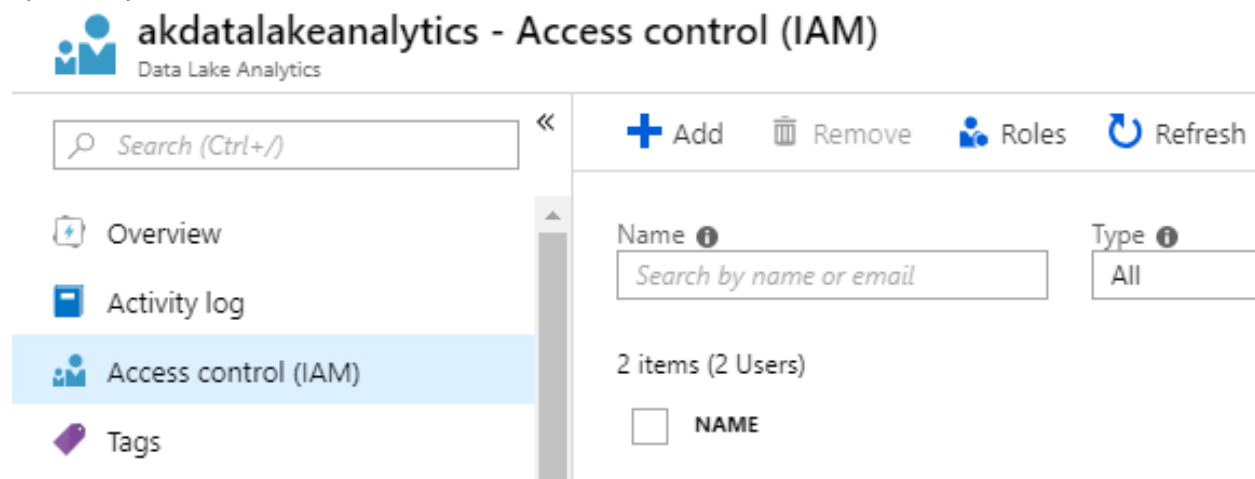
Grant permissions

API	APPLICATION PERMI...	DELEGATED PERMIS...
Windows Azure Active Directory (Microsoft.Azure.Act...	0	1
Azure Data Lake	0	1

Permissions—Step 2

We now need to grant the Registered App (SPN) permissions to our Data Lake. This is completed via the Access control (IAM) blade for the Data Lake Analytics account.

Locate your Data Lake Analytics account, select **Access control (IAM)** and click **Add**:



Select the **Data Lake Analytics Developer** role, enter the Registered App name and click **Save**:

note: you may need type the App name manually if it doesn't appear in the list of objects


Add permissions

Role ⓘ
Data Lake Analytics Developer

Assign access to ⓘ
Azure AD user, group, or application

Select ⓘ
adf-adlsgen1-app

Selected members:



adf-adlsgen1-app

Remove

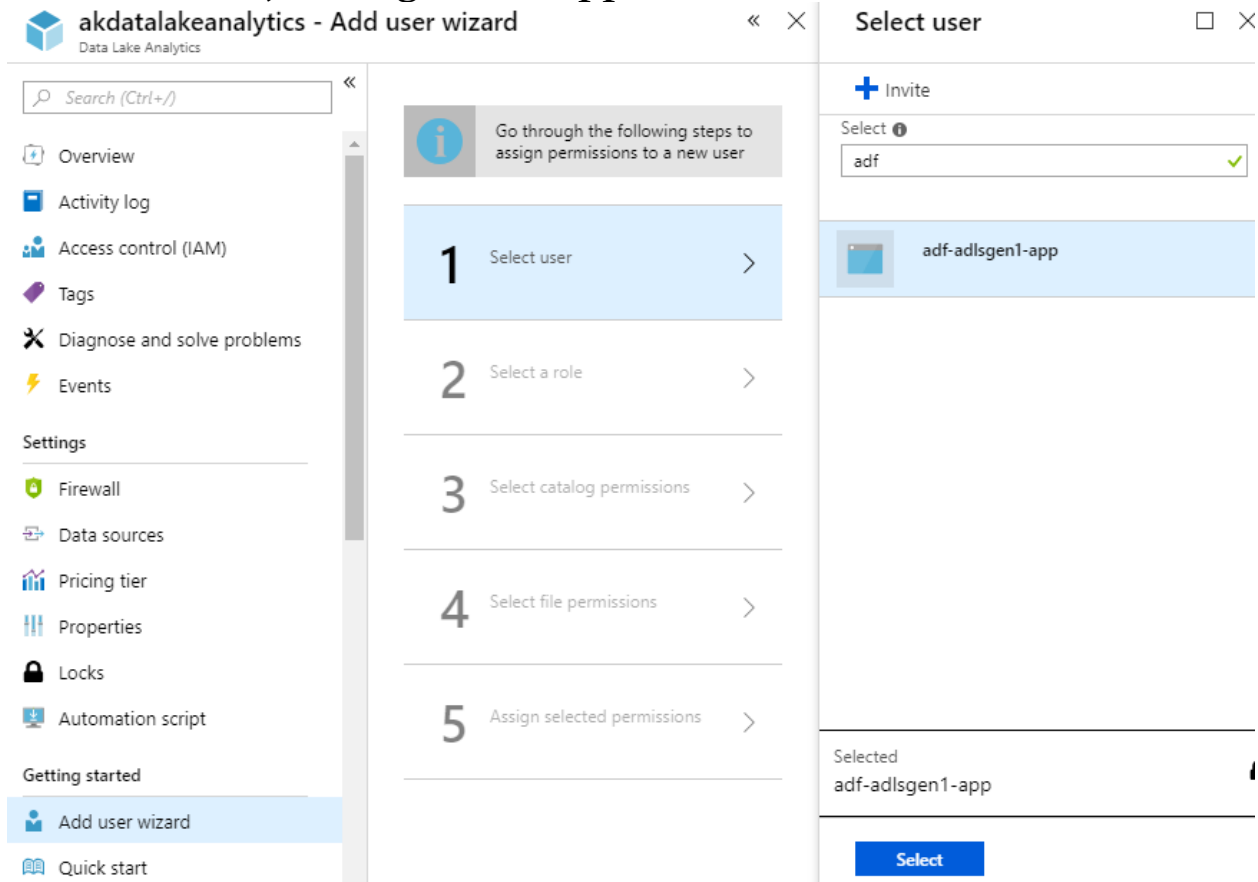
Save

Discard

Permissions—Step 3

In the previous step we added Azure resource permissions. We now need to add the registered user to the Data Lake instance itself.

Still from within the **Data Lake Analytics** blade select **Add user wizard**, the registered app and click **Select**:



Select the **Data Lake Analytics Developer** role:

user wizard
« ×

Select a role
□ ×

Go through the following steps to assign permissions to a new user

1
Select user
adf-adlsgen1-app
✓

2
Select a role
Data Lake Analytics Developer
>

Owner ⓘ

Contributor ⓘ

Reader ⓘ

Data Lake Analytics Developer ⓘ

Ensure the catalog permissions for the **database** are set to **Read and Write** and click **Select**:

user wizard
« ×

Select catalog permissions
□ ×

Go through the following steps to assign permissions to a new user

1
Select user
adf-adlsgen1-app
✓

2
Select a role
Data Lake Analytics Developer
✓

3
Select catalog permissions
>

Select U-SQL Catalog and Database permissions needed by adf-adlsgen1-app. Permissions for the Catalog and the master database are required for job submission.

SCOPE	PERMISSIONS
akdatalakeanalytics (Catalog)	Read and write ▼
master (Database)	Read and write ▼

On the **Select file permissions** step ensure **apply to** is set to **This folder and all children**, click **Select**:

Accounts ▶

First, browse and select files or folders to assign permissions. To select a row, hover over the row then click the check box to the left.


NAME
<input checked="" type="checkbox"/> ak1

Next, determine the permissions to be assigned on the selected files and folders:






ACCOUNT	PATH	READ	WRITE	EXECUTE	APPLY TO	
ak1	/	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	This folder and all children	<input type="checkbox"/>
ak1	/system	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	This folder and all children	<input type="checkbox"/>

We now need to apply the previously selected permissions.
Click **Run** and wait for the job to finish:

Assign selected permissions □ ×



Assign permissions to adf-adlsgen1-app
Click run to grant this user the selected permissions below.

TASK	STATUS	
Assign Data Lake Analytics Developer role to account akdatalakeanalytics	<div><div></div>Pending</div>	
Assign Read and write permissions to akdatalakeanalytics (Catalog)	<div><div></div>Pending</div>	
Assign Read and write permissions to master (Database)	<div><div></div>Pending</div>	
Assign adf-adlsgen1-app rwx permissions to '/' and all its children on ak1.	<div><div></div>Pending</div>	
Assign adf-adlsgen1-app rwx permissions to '/system' and all its children on ak1.	<div><div></div>Pending</div>	

5 tasks pending

Run

Result, we are now finished with permissions! Click **Done** to close the blade:



adf-adlsgen1-app now has the selected permissions.
Click here to open Data Explorer to give this user permissions to additional data.

TASK	STATUS	
Assign Data Lake Analytics Developer role to account akdatalakeanalytics	✓ Completed	🗑️
Assign Read and write permissions to akdatalakeanalytics (Catalog)	✓ Completed	🗑️
Assign Read and write permissions to master (Database)	✓ Completed	🗑️
Assign adf-adlsgen1-app rwx permissions to '/' and all its children on ak1.	✓ Completed. 36 succeeded, 0 failed.	🗑️
Assign adf-adlsgen1-app rwx permissions to '/system' and all its children on ak1.	✓ Completed. 2 succeeded, 0 failed.	🗑️




5 tasks completed

Run

Done

Let's retest our Linked Service!

Within Data Factory reopen the Linked Service we created previously (the pencil icon):

Connections		
Linked Services		
Integration Runtimes		
+ New		
Name	Actions	Type
AzureDataLakeStore1	  	Azure Data Lake Storage Gen1

Click **Test Connection** and confirm everything is working:

Service principal key

Azure Key Vault

Service principal key *

.....

Annotations

+ New | Delete

☐ NAME

Click new to start adding new annotations

✓ Connection successful

Cancel

Test connection

Finish

Summary

I hope you found this useful. The various Azure documentation pages hint at what's required but don't provide an end-to-end view of the steps.

In summary these are -

1. Create Linked Service within Data Factory
2. Create the Service Principal within Azure AD
3. Add the Service Principal details to the ADF Linked Service
4. Within Azure AD, assign Data Lake permissions
5. Within Data Lake Analytics, grant developer permissions to the SPN

6. Within Data Lake Analytics, use the Add user wizard to grant catalog permissions
7. Test!