

Sharepoint to S3 (SFTP/SSH)

Overview: The goal of this development work is to **automate** the processing of files from a client's Sharepoint directory (part of our Teams implementation) into their appropriate AWS-S3 bucket.

High-Level Processing Steps:

1. Access client's Sharepoint directory and see if there are any files for processing.
2. If files (CSV's) are present, check header record to verify that files are for Production run AND are set to be processed for today's date (not future dated). Each CSV file can have differing header information so must be validated individually).
3. If files are being processed today, 1) strip header information before transfer, 2) initiate bridge process and 3) transfer files into appropriate S3 bucket.
4. Ensure that all processing steps are logged and saved either in the Sharepoint or S3 space.

Sharepoint File Location: <https://arizonastateu.sharepoint.com/:f:/r/sites/O365CLASSchoolFuel-VentureDevilsFacultyInitiative-LittletonElementarySchoolDistrict/Shared%20Documents/Littleton%20Elementary%20School%20District?csf=1&web=1&e=Uypbpk>

Walk-Me Video: <https://www.youtube.com/watch?v=VW6ggVsvOeQ>

The AWS Transfer Family setup documentation is here:

<https://docs.aws.amazon.com/transfer/latest/userguide/create-server-sftp.html>

Other Details:

In the GitHub directory is transferUserPolicy.json which is an example AWS Policy document. You have to setup an AWS Role and Attach the Policy to the role. Note that this one references a specific bucket. You will have to change it to your bucket for your specific needs.

In the GitHub directory is a jpeg which is an example of how I setup the SSH-**SFTP** connector.

Make certain you **encrypt the S3 bucket** that you setup for this. That way you have encryption at rest.

Part I: Transferring files before filtering with a lambda function

Preparation

1 Create a bucket and a subfolder in the bucket

2 Create an IAM

2.1 Create a policy

2.2 Create an IAM with the above policy. Select **Transfer** for use case

The screenshot shows the AWS IAM console 'Create role' page. At the top, a blue banner reads: 'Introducing the new IAM roles experience. We've redesigned the IAM roles experience to make it easier to use. [Let us know what you think.](#)' Below this, the breadcrumb 'IAM > Roles > Create role' is visible. On the left, a sidebar shows the steps: 'Step 1: Select trusted entity' (active), 'Step 2: Add permissions', and 'Step 3: Name, review, and create'. The main content area is titled 'Select trusted entity'. Under 'Trusted entity type', there are five options: 'AWS service' (selected), 'AWS account', 'Web identity', 'SAML 2.0 federation', and 'Custom trust policy'. The 'AWS service' option is highlighted with a blue border. Below this, the 'Use case' section is titled 'Allow an AWS service like EC2, Lambda, or others to perform actions in this account.' It lists 'Common use cases' with 'EC2' and 'Lambda' options. Under 'Use cases for other AWS services', the 'Transfer' option is selected and highlighted with a blue border. A dark, abstract graphic is visible in the bottom right corner of the console interface.

3 Create SSH keys with WinSCP

Go to terminal, enter '**ssh-keygen**' and press enter, press enter until a key's randomart image is shown

```
Command Prompt
C:\Users\Shuwei>ssh-keygen
Generating public/private rsa key pair.
Enter file in which to save the key (C:\Users\Shuwei\.ssh/id_rsa):
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in C:\Users\Shuwei\.ssh/id_rsa.
Your public key has been saved in C:\Users\Shuwei\.ssh/id_rsa.pub.
The key fingerprint is:
SHA256:Ruj6vdiQq80PSuudMJfxQsQmtAdWGs8QaKI4Z3+KsX8 shuwei@DESKTOP-P117I3V
The key's randomart image is:
+---[RSA 3072]-----+
| .Bo. |
|o+ X . |
|= + B . . |
|+ o= . . |
|+ .o . S |
| ..=.o |
| +=o+ |
|.O*+=E* |
|.o*==+. |
+----[SHA256]-----+
C:\Users\Shuwei>
```

4 Create an SFTP server in AWS Transfer Family. Go to AWS Transfer Family, click **Create Server**.

4.1 Choose **SFTP** as protocol, click **Next**.

4.2 Choose an identity provider.

Currently, I am choosing **Service managed** for testing, which “create and manage users within the service”.

4.3 In **Endpoint configuration**, choose **Publicly accessible**, click **Next**.

4.4 In **Choose a domain**, select **AWS S3**, click **Next**.

4.5 In **CloudWatch logging**, select **Choose an existing role**. Click **Next**.

4.6 Create users. Enter public key while creating users

Creating a flow

This part documents the process of creating two types of flows, automated cloud flows and scheduled flows.


1 Create an automated cloud flow

1.1 Go to **Power Automate**

1.2 On the left panel, click **Create**. Click **Automated cloud flow**

1.3 In **Build and automated cloud flow**, Under **Choose your flow's trigger**, search “sharepoint”, select **when a file is created in a folder**, click **Create**.

Build an automated cloud flow



Free yourself from repetitive work just by connecting the apps you already use—automate alerts, reports, and other tasks.

Examples:

- Automatically collect and store data in business solutions
- Generate reports via custom queries on your SQL database

Flow name
 Add a name or we'll generate one

Choose your flow's trigger *

sharepoint

- ☒ When a file is created in a folder SharePoint
- ☐ When an item or a file is modified SharePoint
- ☐ When an item is created SharePoint
- ☐ When an item is deleted SharePoint
- ☐ When a file is classified by a content... SharePoint
- ☐ When a file is created (properties on... SharePoint

[Skip](#) [Create](#) [Cancel](#)

Select a site address and a folder ID

When a file is created in a folder

* Site Address
 O365 CLAS School Fuel - Venture Devils Faculty Initiative - Client-Littleton Elementary School District -
<https://arizonastateu.sharepoint.com/sites/O365CLASSchoolFuel-VentureDevilsFacultyInitiative-LittletonElementarySchoolDistrict>

* Folder ID
 /Shared Documents/Littleton Elementary School District

[Show advanced options](#)

1.4 Click **+ New Step**. In **Choose an operation**, choose **get file content** for SharePoint. Fill the following form and Click **+ New step**.

Get file content

* Site Address
 O365 CLAS School Fuel - Venture Devils Faculty Initiative - Client-Littleton Elementary School District -
<https://arizonastateu.sharepoint.com/sites/O365CLASSchoolFuel-VentureDevilsFacultyInitiative-LittletonElementarySchoolDistrict>

* File Identifier
 x-ms-file-id x

[Show advanced options](#)

1.5 Select **Create file** connector for SFTP-SSH

SFTP - SSH

Test connection failed. Details: SSH host key finger print should be provided when the SSH host key validation is enabled. clientRequestId: 0ac4f3b8-604d-55cb-9772-862e918ec326

* Connection name:

* Host server address:

* User name:

Password:

SSH private key:

SSH private key passphrase:

Port number:

Disable SSH host key validation: ☒ Disable SSH host key validation? (True/False)

SSH host key finger print:

Root folder path:

Create **Cancel**

After connection succeeds, fill the following form. Click **Save**

Create file

* Folder path:

* File name:

* File content:

Show advanced options

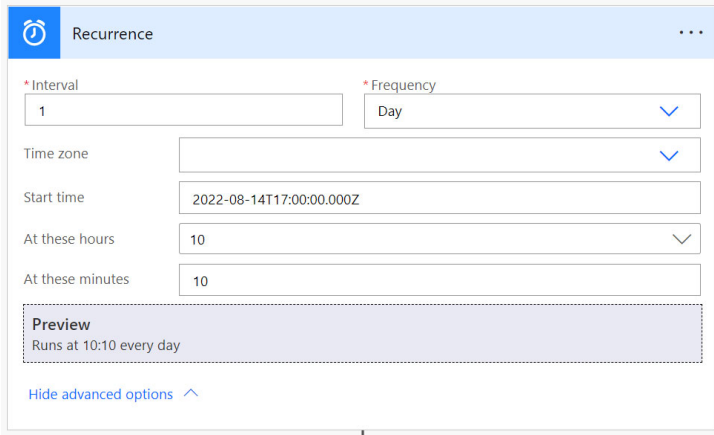
1.6 Select **Delete file** as an action

2 Create a scheduled cloud flow

2.1 Go to **Power Automate**

2.2 On the left panel, click **Create**. Click **Scheduled cloud flow**

2.3 In **Recurrence**, fill the form select **when a file is created in a folder**, Click **+ New step**.



Recurrence

* Interval: 1

* Frequency: Day

Time zone: [Dropdown]

Start time: 2022-08-14T17:00:00.000Z

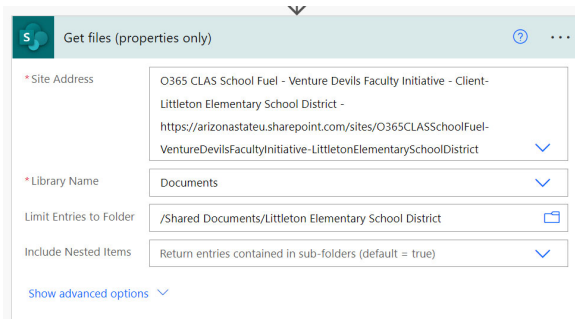
At these hours: 10

At these minutes: 10

Preview
Runs at 10:10 every day

[Hide advanced options](#)

2.4 Select **Get files(properties only)** for Sharepoint. Fill the following form and Click **+ New Step**.



Get files (properties only)

* Site Address: O365 CLAS School Fuel - Venture Devils Faculty Initiative - Client-Littleton Elementary School District - https://arizonastateu.sharepoint.com/sites/O365CLASchoolFuel-VentureDevilsFacultyInitiative-LittletonElementarySchoolDistrict

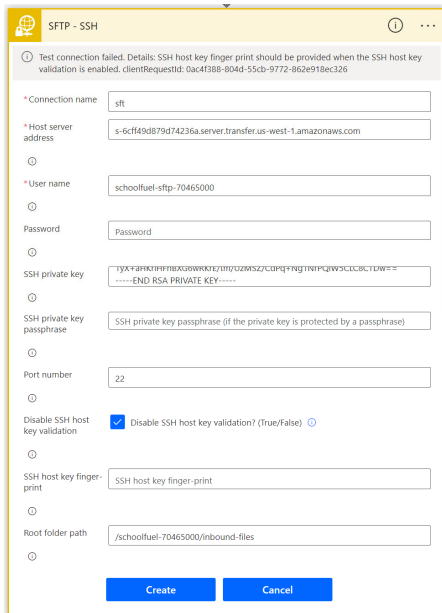
* Library Name: Documents

Limit Entries to Folder: /Shared Documents/Littleton Elementary School District

Include Nested Items: Return entries contained in sub-folders (default = true)

[Show advanced options](#)

2.5 Select **Apply to each**, fill **Select an output from previous steps** with a dynamic parameter, **Value**. Click **add an action** and select **Get file content** for SharePoint. Fill that form and Click **add an action**. Select **Create file** for SFTP-SSH. Configure the SFTP connector. After creating files, add **Delete file** as an action



SFTP - SSH

Test connection failed. Details: SSH host key fingerprint should be provided when the SSH host key validation is enabled. clientRequestId: 0ac4f388-804d-55cb-9772-862e918ec326

* Connection name: sft

* Host server address: s-6c7f49d879d74236a.server.transfer.us-west-1.amazonaws.com

* User name: schoolfuel-sftp-70465000

Password: [Field]

SSH private key: [Field]

SSH private key passphrase: [Field]

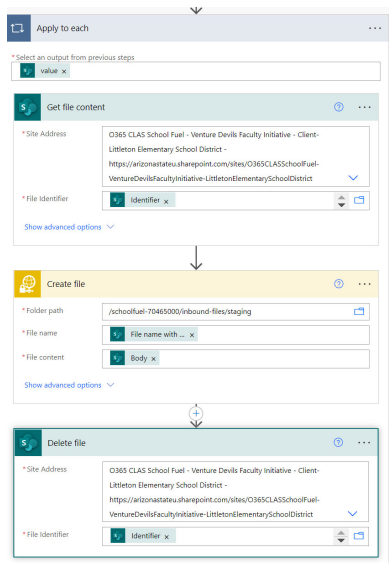
Port number: 22

Disable SSH host key validation: ☒ Disable SSH host key validation? (True/False)

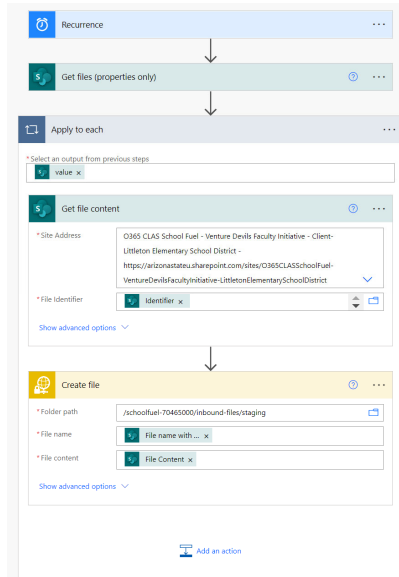
SSH host key fingerprint: [Field]

Root folder path: /schoolfuel-70465000/inbound-files

Create **Cancel**



The whole flow:



Filter files after file has been transferred using Lambda Function

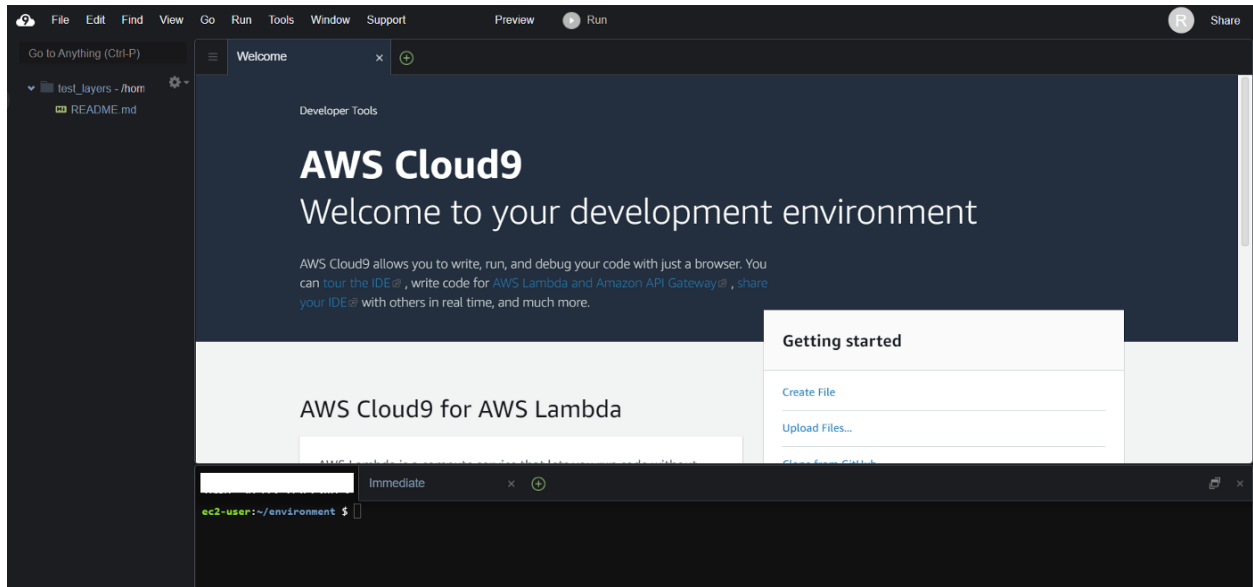
1 Go to **Cloud9**, create an environment to install **pandas** and **openpyxl** package.

1.1 Search for **Cloud9** in the AWS Services, Click **Create Environment**

1.2 Name your environment and click **next step**

1.3 Keep the environment default settings (Create a new EC2 instance for environment; t2.micro ; Amazon Linux etc..) and click **next step**

1.4 Click **Create Environment** and you're ready to go



1.5 Type the following code line by line into the terminal at the bottom. The pip install command can be replaced with a package of your choosing. You can also install more than 1 package.

```
mkdir folder
cd folder
virtualenv v-env
source ./v-env/bin/activate
pip install pandas
pip install openpyxl
deactivate
```

Then type the following code line by line to create your layer

```
mkdir python
cd python
cp -r ../v-env/lib64/python3.7/site-packages/* .
cd ..
zip -r panda_layer.zip python
aws lambda publish-layer-version --layer-name pandas_openpyxl --
zip-file fileb://pandas_openpyxl_layer.zip --compatible-runtimes
python3.7
```

1.6 Download this layer and add this layer in Lambda Function. Delete the above created cloud9 enviroment.

2 Creating a Lambda function and Adding the Pandas Openpyxl layer to the Lambda function

2.1 Go to the AWS Lambda service and click **Create Function**

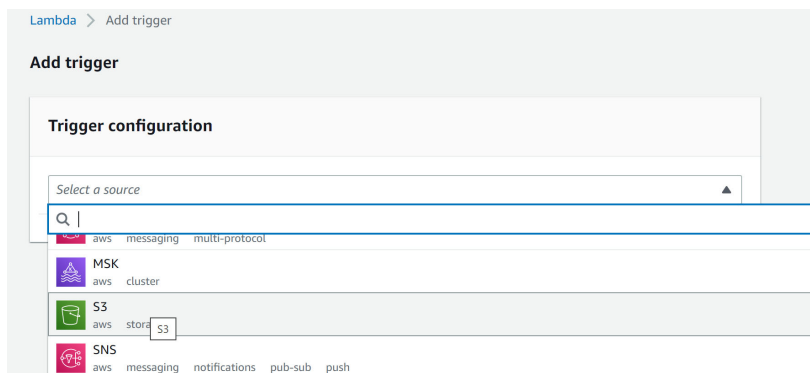
2.2 Name the function, set the runtime to Python 3.7. Select a role with two permission policies, one to execute Lambda function and another to access S3 bucket. Click **Create Function**

2.3 Click on **Layers** in the function designer, then click **add a layer**

2.4 On the name dropdown, you should see your Pandas and Openpyxl layer. Click **Add**.

2.5 Add a trigger for the function.

In Function overview, click **Add trigger**. In Trigger configuration, select S3



In the bucket column, select a trigger bucket.

2.6 Edit code in the Lambda function.

Part II: Transferring files with filtering within a flow

Notes

1 The file to be transferred must be in table format.

FileHomeInsertDrawPage LayoutFormulasDataReviewViewHelp

Paste

Clipboard

Calibri

11

B

I

U

A^

A^

Font

Alignment

General

\$

%

,

0.00

00

Number

M7

X

✓

fx

	A	B	C	D	E	F	G	H	I	J	K	L
1	ID	Boolean	Date	Student	SchoolID	SchoolY	DateofA	Grade	Absence	Absence	Absence	
2	70465000	Yes	8/11/2022	1216	70465101	2018-2019	7/25/2018	8	Yearly Abs	IL	12	
3	70465000	Yes	8/11/2022	20011	70465105	2018-2019	7/25/2018	1	Yearly Abs	IL	68.5	
4	70465000	Yes	8/11/2022	1263	70465101	2018-2019	7/25/2018	7	Yearly Abs	IL	26	
5	70465000	Yes	8/11/2022	19306	70465107	2018-2019	7/25/2018	1	Yearly Abs	IL	53	
6												
7												

FileHomeInsertDrawPage LayoutFormulasDataReviewViewAutomateHelp

Calibri

11

B

Font

Alignment

General

0.00

00

D20

X

✓

fx

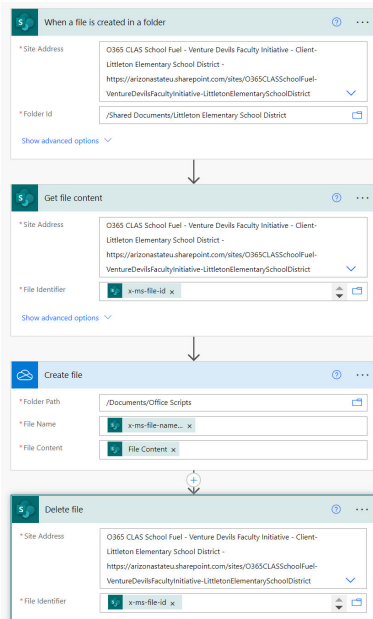
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	ID	Boolean	Date	Student	SchoolID	SchoolY	Incident	Incident	Incident	StartDa	Disposi	EndDat	Violatic	Violatic	Disposi	Disposi	Incident	Referra	Locatio
2	70465000	Yes	#####	8632	70465108	2021-2022	3/1/2022	#####	6925	3/1/2022	3/1/2022	3/5/2022	N/A	Fighting	N/A	OSS-Long	3	PachMs	CL
3				8632	70465108	2021-2022	#####	#####	6925	#####	#####	#####	N/A	Defiance c N/A	ISS-Short	1	0.5	PachMs	CL
4				8632	70465108	2021-2022	#####	#####	6925	#####	#####	#####	N/A	Minor Agg	N/A	ISS-Short	1	PachMs	CL
5				11290	70465104	2021-2022	#####	#####	6974	#####	#####	#####	N/A	Defiance c N/A	ISS-Short	1	1	RaseRa	CL
6				11290	70465104	2021-2022	#####	#####	6974	#####	#####	#####	N/A	Defiance c N/A	ISS-Short	1	0.75	RaseRa	CL
7				11290	70465104	2021-2022	#####	#####	6974	#####	#####	#####	N/A	Defiance c N/A	ISS-Short	1	0.5	RaseRa	CL
8				10862	70465109	2021-2022	#####	#####	6974	#####	#####	#####	N/A	Defiance c N/A	ISS-Short	1	0.5	RaseRa	CL
9				10862	70465109	2021-2022	#####	#####	6974	#####	#####	#####	N/A	Defiance c N/A	ISS-Short	1	0.5	RaseRa	CL
10				10859	70465107	2021-2022	#####	#####	6974	#####	#####	#####	N/A	Defiance c N/A	ISS-Short	1	0.25	RaseRa	CL
11				1216	70465101	2021-2022	#####	#####	6974	#####	#####	#####	N/A	Defiance c N/A	ISS-Short	1	0.25	RaseRa	CL
12				1263	70465101	2021-2022	#####	#####	6974	#####	#####	#####	N/A	Defiance c N/A	ISS-Short	1	0.25	RaseRa	CL
13																			
14																			

2 Power Automate has built in connectors for table in OneDrive only. There are no built-in connectors to work with excel files in SharePoint. Not sure if it is possible to create a customized connector for tables in SharePoint.

Process

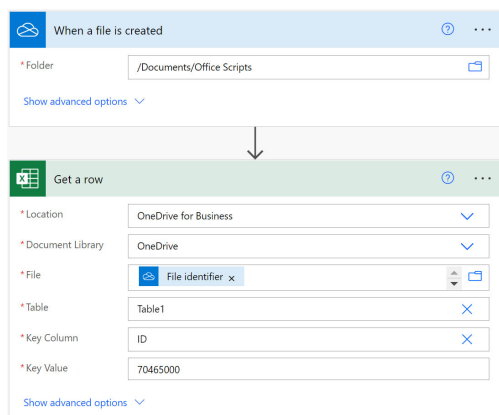
Since Power Automate does not have built-in connectors to work with excel files in SharePoint, the process consists of two parts, a first flow to transfer files from SharePoint to OneDrive, and a second flow to transfer files from OneDrive to S3 buckets.

1 Create a flow that transfers files from SharePoint to OneDrive



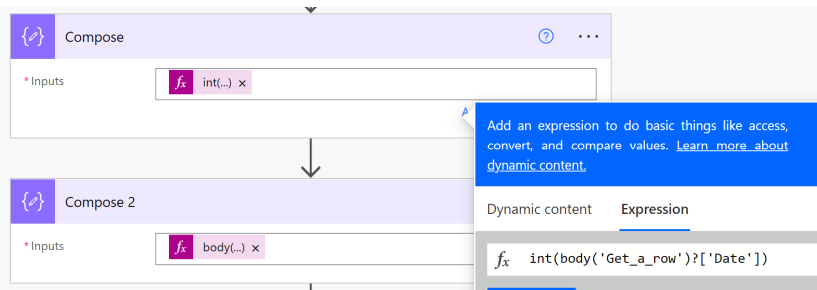
2 Create a flow that transfers files from OneDrive to S3 buckets with filtering

2.1 Select **When a file is created** as a trigger. Add **get a row** as an action. For **Location** and **Document Library**, select one for each from the drop down menu. For **File**, add an dynamic content, **File Identifier**. For **Table**, hardcode Table1. For Key column and key value, hardcode a column name and value.

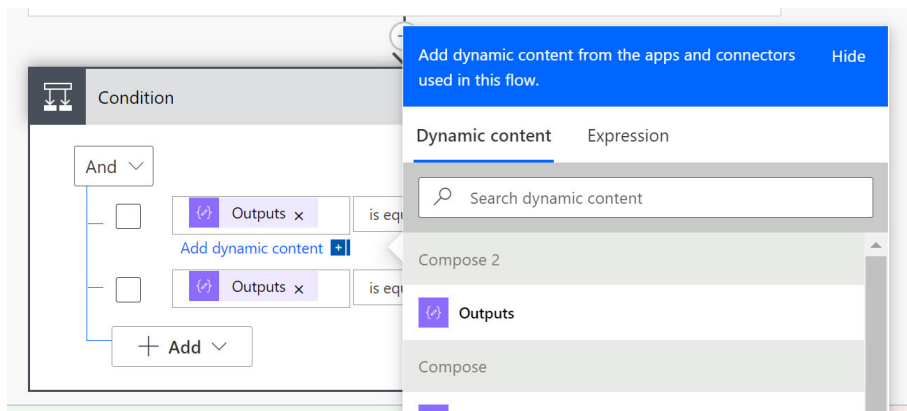


2.2 Compose two value from the outputs from **Get a row** action. Extract “Date” and “Boolean Value”. Add **Compose** as an action. Click **Add dynamic content**, click **expression**. Type `int(body('Get_a_row')?['Date'])` in the **fx** column, click **update**.

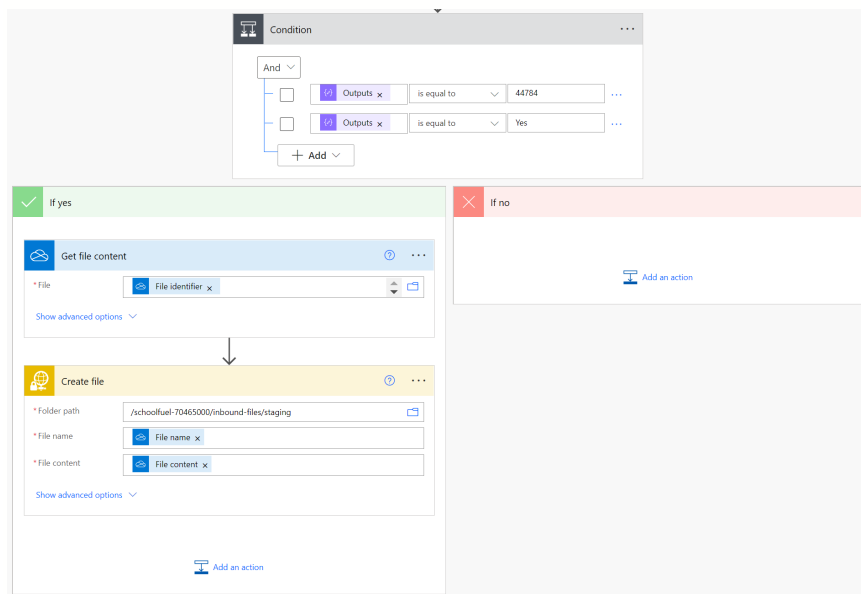
Add a second **Compose**, type `body('Get_a_row')?['Boolean Value']` as input.



2.3 select **Condition** as an action. Add **Outputs** from **Compose** and **Compose 2**, fill the form with condition.



2.4 If yes, transfer the file to S3 bucket.



Notes: This flow works well except one problem, deleting the newly created file in OneDrive. I do not have access to deleting the file. The delete function works well with SharePoint.

