

Consuming Data from s3 using boto3

Create policy to read from s3 Bucket: To read the data from specific s3 bucket we need to create the policy and that policy is attached to the group through the user.

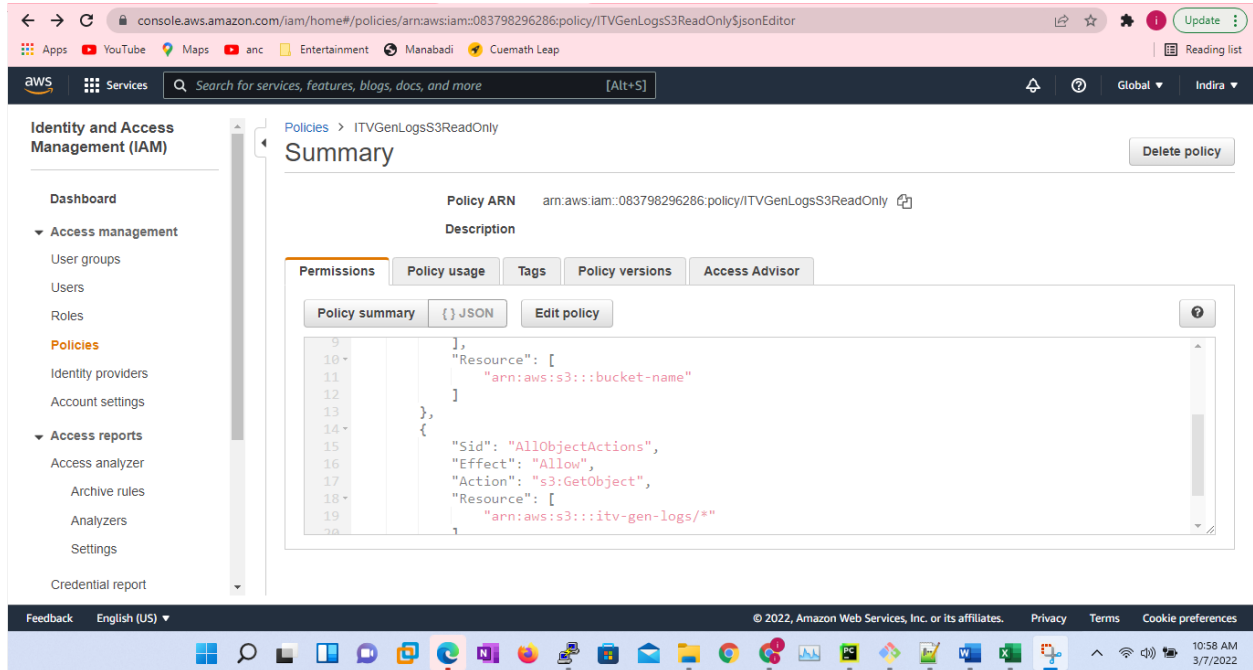


Figure 1 Created the policy with read only permissions

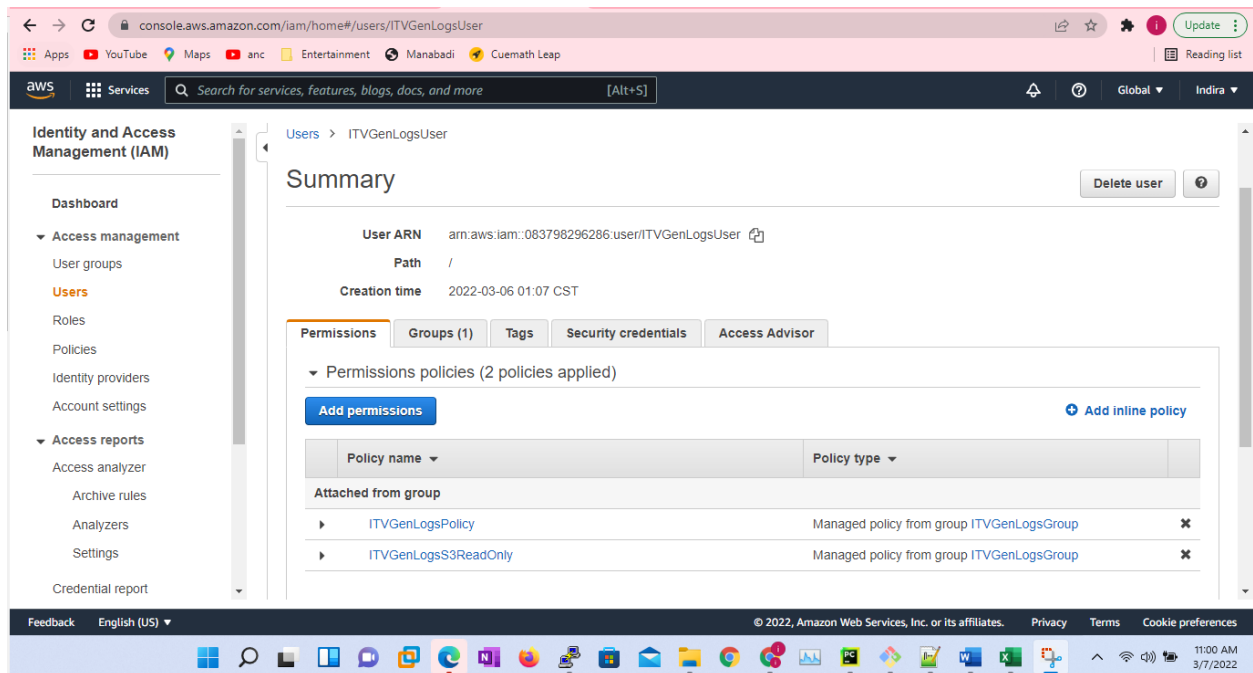


Figure 2 ITVGenLogsReadOnly policy attached to the user

Validating whether the genlogs user have the permissions to access the objects with in the s3 bucket ITVGenLogs using AWS CLI

```
pcbra@DESKTOP-06V8C9M MINGW64 /c/development/practice
$ aws configure --profile ITVGenLogsUser
AWS Access Key ID [None]: AKIARHAWEQ3PP473QY76
AWS Secret Access Key [None]: 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
Default region name [None]:
Default output format [None]:

pcbra@DESKTOP-06V8C9M MINGW64 /c/development/practice
$ aws s3 ls itv-gen-logs --profile ITVGenLogsUser

An error occurred (AccessDenied) when calling the ListObjectsV2 operation: Access Denied

pcbra@DESKTOP-06V8C9M MINGW64 /c/development/practice
$ aws s3 ls itv-gen-logs --profile ITVGenLogsUser
PRE logs /

pcbra@DESKTOP-06V8C9M MINGW64 /c/development/practice
$ aws s3 ls itv-gen-logs/logs --profile ITVGenLogsUser
PRE logs /

pcbra@DESKTOP-06V8C9M MINGW64 /c/development/practice
$
```

Setup Python Virtual Environment to Explore boto3

Create the virtual environment using the below commands:

```
python venv genlogss3-venv
genlogss3 -venv/scripts/activate
pip install boto3
pip install jupyter lab
```

```
MINGW64/c/development/practice/genlogss3
pcbra@DESKTOP-06V8C9M MINGW64 /c/development/practice/genlogss3
$ python -m venv genlogss3-venv

pcbra@DESKTOP-06V8C9M MINGW64 /c/development/practice/genlogss3
$ genlogss3-venv/scripts/activate

pcbra@DESKTOP-06V8C9M MINGW64 /c/development/practice/genlogss3
$ pip install boto3
Collecting boto3
  Downloading boto3-1.21.13-py3-none-any.whl (132 kB)
Collecting jmespath<1.0.0,>=0.7.1
  Using cached jmespath-0.10.0-py2.py3-none-any.whl (24 kB)
Collecting botocore<1.25.0,>=1.24.13
  Downloading botocore-1.24.13-py3-none-any.whl (8.6 MB)
Collecting s3transfer<0.6.0,>=0.5.0
  Downloading s3transfer-0.5.2-py3-none-any.whl (79 kB)
Collecting urllib3<1.27,>=1.25.4
  Using cached urllib3-1.26.8-py2.py3-none-any.whl (138 kB)
Collecting python-dateutil<3.0.0,>=2.1
  Using cached python-dateutil-2.8.2-py2.py3-none-any.whl (247 kB)
Collecting six<=1.5
  Using cached six-1.16.0-py2.py3-none-any.whl (11 kB)
Installing collected packages: jmespath, urllib3, six, python-dateutil, botocore, s3transfer, boto3
Successfully installed boto3-1.21.13 botocore-1.24.13 jmespath-0.10.0 python-dateutil-2.8.2 s3transfer-0.5.2 six-1.16.0 urllib3-1.26.8
WARNING: You are using pip version 20.1.1; however, version 22.0.3 is available.
You should consider upgrading via the 'c:\users\pcbra\appdata\local\programs\python\python37\python.exe -m pip install --upgrade pip' command.

pcbra@DESKTOP-06V8C9M MINGW64 /c/development/practice/genlogss3
$ pip install jupyterlab
Collecting jupyterlab
  Downloading jupyterlab-3.3.0-py3-none-any.whl (8.7 MB)
Collecting Jinja2<=2.1
  Downloading Jinja2-3.0.3-py3-none-any.whl (133 kB)
Collecting jupyterlab-server<=2.10
  Downloading jupyterlab_server-2.10.3-py3-none-any.whl (61 kB)
Collecting jupyter-server<=1.4
  Downloading jupyter_server-1.13.5-py3-none-any.whl (397 kB)
Collecting jupyter-core
  Downloading jupyter_core-4.9.2-py3-none-any.whl (86 kB)
Collecting nbclassic<=0.2
  Downloading nbclassic-0.3.6-py3-none-any.whl (13 kB)
Collecting packaging
  Downloading packaging-21.3-py3-none-any.whl (40 kB)
Collecting ipython
  Downloading ipython-7.32.0-py3-none-any.whl (793 kB)
Collecting tornado<=6.1.0
  Downloading tornado-6.1-cp37-cp37m-win_amd64.whl (422 kB)
Collecting MarkupSafe<=2.0
```

Untitled.ipynb - JupyterLab | ProfileNotFound Traceback (most recent call last): | Amazon S3 buckets — Boto3 | +

localhost:8888/lab/tree/Untitled.ipynb

Apps | YouTube | Maps | anc | Entertainment | Manabadi | Cuemath Leap | Reading list

File | Edit | View | Run | Kernel | Tabs | Settings | Help

Filter files by name

Name	Last Modified
genlogss3-venv	an hour ago
Untitled.ipynb	seconds ago

Untitled.ipynb

```
[6]: import boto3

[7]: import os
    os.environ.setdefault('AWS_PROFILE', 'ITVGenLogsUser')

[7]: 'ITVGenLogsUser'

[8]: s3_client = boto3.client('s3')

[9]: s3_client

[9]: <botocore.client.S3 at 0x199fcff4c48>

[ ]: s3_client.list_buckets()

[12]: s3_client.list_objects(
    Bucket='itv-gen-logs'
)

[12]: {'ResponseMetadata': {'RequestId': 'VWTAEDC9D80PEECG',
    'HostId': 'muitHSm1mK9F3b5213CdxbiuCO8zGUjmtZfre24nBQ+v043OT95IvWVNIaUwXg9e+xlNq+2E=',
    'HTTPStatusCode': 200,
    'HTTPHeaders': {'x-amz-id-2': 'muitHSm1mK9F3b5213CdxbiuCO8zGUjmtZfre24nBQ+v043OT95IvWVNIaUwXg9e+xlNq+2E='}}
```

Simple | 0 | 1 | Python 3 (ipykernel) | Idle | Mode: Command | Ln 1, Col 1 | Untitled.ipynb

12:52 PM 3/7/2022