

## Time Travel in Databricks (Lakehouse)

We cannot set permissions in databricks hive\_metastore, we have to create a unity catalog for that.

ETL JSONPython ☆

FileEditViewRunHelpLast edit was 3 minutes agoProvide feedback

Run allShellunext's ClusterScheduleShare

1dbutils.fs.mount(  
2source = "wasbs://raw@saunextadls.blob.core.windows.net",  
3mount\_point = "/mnt/saunextadls/raw",  
4extra\_configs = {"fs.azure.account.key.saunextadls.blob.core.windows.net":"DsZWJs7JVHZZz1I7GKycLV8ejCdj8V2UkqMLgAp6QyV0w5rVrHvmVTgwcThdHUmWg7MXon65/8z+Astj4Yiug=="}  
)  
  
True  
  
Command took 12.23 seconds -- by shellunext\_1693422651722@npunext.onmicrosoft.com at 9/27/2023, 10:10:31 AM on Shellunext's Cluster

Cmd 3

1%fs ls dbfs:/mnt/  
  
Table +  

	path	name	size	modificationTime
1	dbfs:/mnt/saunextadls/	saunextadls/	0	1695789632000

  
1 row | 0.36 seconds runtime  
Refreshed 23 minutes ago  
  
Command took 0.36 seconds -- by shellunext\_1693422651722@npunext.onmicrosoft.com at 9/27/2023, 10:11:29 AM on Shellunext's Cluster

Cmd 4

1%fs ls dbfs:/mnt/saunextadls/raw  
  
Table +  

	path	name	size	modificationTime
1	dbfs:/mnt/saunextadls/raw/json/	json/	0	0

  
1 row | 0.37 seconds runtime  
Refreshed 22 minutes ago

ETL JSONPython ☆

FileEditViewRunHelpLast edit was 4 minutes agoProvide feedback

Run allShellunext's ClusterScheduleShare

Cmd 5

1df=spark.read.json("dbfs:/mnt/saunextadls/raw/json/")  
  
(1) Spark Jobs  
df: pyspark.sql.dataframe.DataFrame = [humidity: double, pressure: double ... 3 more fields]  
Command took 3.37 seconds -- by shellunext\_1693422651722@npunext.onmicrosoft.com at 9/27/2023, 10:13:01 AM on Shellunext's Cluster

Cmd 6

1df.display()  
  
(1) Spark Jobs  
Table +  

	humidity	pressure	sensor_id	temperature	timestamp
1	49.99	1003.39	1895b922-5c66-4d7b-bcfa-ec6bb355b0ec	26.08	1993-04-13T22:04:18
2	56.99	1010.56	c7111856-eaf2-43da-827e-ae70232842b1	20.3	1988-10-31T14:38:45
3	41.29	1012.52	5a61a7b2-724d-4d84-b4f6-f51f7039ae97	26.56	1987-04-24T22:18:32
4	47.78	1014.82	e308eb54-ed3f-4c28-8552-44c5c88e7fa1	29.73	1993-11-23T04:18:17
5	58.92	1000.05	b113908a-6963-470c-9788-510521eb8a74	21.08	2007-03-02T14:42:08
6	46.51	1008.96	80d978e1-e602-42a4-8521-e19a0130e9e0	25.6	1988-09-06T19:31:24
7	55.16	1005.49	af7d4b05-3493-44d2-af62-361fd0815467	29.36	1971-10-05T16:21:51

  
113 rows | 0.81 seconds runtime  
Refreshed 21 minutes ago  
  
Command took 0.81 seconds -- by shellunext\_1693422651722@npunext.onmicrosoft.com at 9/27/2023, 10:13:15 AM on Shellunext's Cluster

Cmd 7

1df1=df.withColumn("ingestiondate",current\_timestamp()).withColumn("path",input\_file\_name())  
  
df1: pyspark.sql.dataframe.DataFrame = [humidity: double, pressure: double ... 5 more fields]  
Command took 0.16 seconds -- by shellunext\_1693422651722@npunext.onmicrosoft.com at 9/27/2023, 10:17:35 AM on Shellunext's Cluster

## ETL JSON

Python



File Edit View Run Help Last edit was 4 minutes ago Provide feedback

Run all

Shellunext's Cluster

Schedule

Share



```
1 %sql
2 create schema if not exists json
```

OK

Command took 0.72 seconds -- by shellunext\_1693422651722@npunext.onmicrosoft.com at 9/27/2023, 10:18:27 AM on Shellunext's Cluster

Cmd 9

```
1 df1.write.mode("overwrite").option("path","dbfs:/mnt/saunextadls/raw/output/juhi/json").saveAsTable("json.bronze")
```

(6) Spark Jobs

Command took 5.51 seconds -- by shellunext\_1693422651722@npunext.onmicrosoft.com at 9/27/2023, 10:22:38 AM on Shellunext's Cluster

Cmd 10

```
1 %sql
2 select count(*) from json.bronze
```

(3) Spark Jobs

\_sqlidf: pyspark.sql.dataframe.DataFrame = [count(1): long]

Table

count(1)
113

1 row | 1.70 seconds runtime

Refreshed 11 minutes ago

This result is stored as PySpark data frame \_sqlidf and in the IPython output cache as Out[10]. Learn more

Command took 1.70 seconds -- by shellunext\_1693422651722@npunext.onmicrosoft.com at 9/27/2023, 10:23:37 AM on Shellunext's Cluster

## Delta Lake

SQL



File Edit View Run Help Last edit was 16 minutes ago Provide feedback

Run all

Shellunext's Cluster

Schedule

Share



```
1 describe history emp
```

(1) Spark Jobs

Table

version	timestamp	userid	userName	operation	operationParameters
1	0	2023-09-27T08:34:33.000+0000	4887540299005056	shellunext_1693422651722@npunext.onmicrosoft.com	CREATE TABLE

1 row | 1.57 seconds runtime

Refreshed 20 minutes ago

Command took 1.57 seconds -- by shellunext\_1693422651722@npunext.onmicrosoft.com at 9/27/2023, 2:04:59 PM on Shellunext's Cluster

Cmd 5

```
1 select * from emp
```

(1) Spark Jobs

Query returned no results

Command took 1.31 seconds -- by shellunext\_1693422651722@npunext.onmicrosoft.com at 9/27/2023, 2:05:12 PM on Shellunext's Cluster

Cmd 6

```
1 insert into table emp values(1,'a',23,'DE')
```

(6) Spark Jobs

Table

num_affected_rows	num_inserted_rows
1	1

1 row | 7.92 seconds runtime

Refreshed 19 minutes ago

Delta Lake

SQL

☆

File Edit View Run Help

Last edit was 16 minutes ago

Provide feedback

▶ Run all

● Shellunext's Cluster

📅 Schedule

Share

1

select \* from emp

▶ (2) Spark Jobs

Table

	id	name	age	dept
1	1	a	23	DE

1 row | 2.00 seconds runtime

Refreshed 19 minutes ago

Command took 2.00 seconds -- by shellunext\_1693422651722@npunext.onmicrosoft.com at 9/27/2023, 2:06:19 PM on Shellunext's Cluster

Cmd 8

1

describe history emp

▶ (1) Spark Jobs

Table

	version	timestamp	userid	userName	operation	operationParameters
1	1	2023-09-27T08:36:09.000+0000	4887540299005056	shellunext_1693422651722@npunext.onmicrosoft.com	WRITE	["mode": "Append", "statsOr
2	0	2023-09-27T08:34:33.000+0000	4887540299005056	shellunext_1693422651722@npunext.onmicrosoft.com	CREATE TABLE	["partitionBy": "[]", "descript

2 rows | 0.92 seconds runtime

Refreshed 19 minutes ago

Command took 0.92 seconds -- by shellunext\_1693422651722@npunext.onmicrosoft.com at 9/27/2023, 2:06:32 PM on Shellunext's Cluster

Cmd 9

1

insert into table emp values(2,'b',23,'DE')

▶ (6) Spark Jobs

Delta Lake

SQL

☆

File Edit View Run Help

Last edit was 17 minutes ago

Provide feedback

▶ Run all

● Shellunext's Cluster

📅 Schedule

Share

1

insert into table emp values(3,'c',23,'DE'),

2

(4,'d',23,'DE')

▶ (6) Spark Jobs

Table

	num_affected_rows	num_inserted_rows
1	2	2

1 row | 3.57 seconds runtime

Refreshed 19 minutes ago

Command took 3.57 seconds -- by shellunext\_1693422651722@npunext.onmicrosoft.com at 9/27/2023, 2:06:54 PM on Shellunext's Cluster

Cmd 11

1

describe history emp

▶ (1) Spark Jobs

Table

	version	timestamp	userid	userName	operation	operationParameters
1	3	2023-09-27T08:36:55.000+0000	4887540299005056	shellunext_1693422651722@npunext.onmicrosoft.com	WRITE	["mode": "Append", "statsOr
2	2	2023-09-27T08:36:44.000+0000	4887540299005056	shellunext_1693422651722@npunext.onmicrosoft.com	WRITE	["mode": "Append", "statsOr
3	1	2023-09-27T08:36:09.000+0000	4887540299005056	shellunext_1693422651722@npunext.onmicrosoft.com	WRITE	["mode": "Append", "statsOr
4	0	2023-09-27T08:34:33.000+0000	4887540299005056	shellunext_1693422651722@npunext.onmicrosoft.com	CREATE TABLE	["partitionBy": "[]", "descript

4 rows | 0.97 seconds runtime

Refreshed 19 minutes ago

Command took 0.97 seconds -- by shellunext\_1693422651722@npunext.onmicrosoft.com at 9/27/2023, 2:07:06 PM on Shellunext's Cluster



1 select \* from emp

(3) Spark Jobs

Table +

	id	name	age	dept
1	3	c	23	DE
2	4	d	23	DE
3	1	a	23	DE
4	2	b	23	DE

4 rows | 0.93 seconds runtime

Refreshed 19 minutes ago

Command took 0.93 seconds -- by shellunext\_1693422651722@npunext.onmicrosoft.com at 9/27/2023, 2:07:16 PM on Shellunext's Cluster

Cmd 13

1 delete from emp where id= 1

(8) Spark Jobs

Table +

	num_affected_rows
1	1

1 row | 4.64 seconds runtime

Refreshed 18 minutes ago

Command took 4.64 seconds -- by shellunext\_1693422651722@npunext.onmicrosoft.com at 9/27/2023, 2:07:40 PM on Shellunext's Cluster



Cmd 15

1 Update emp  
2 set dept='DS'  
3 where id= 4

(8) Spark Jobs

Table +

	num_affected_rows
1	1

1 row | 4.28 seconds runtime

Refreshed 18 minutes ago

Command took 4.28 seconds -- by shellunext\_1693422651722@npunext.onmicrosoft.com at 9/27/2023, 2:08:09 PM on Shellunext's Cluster

Cmd 16

1 describe history emp

(1) Spark Jobs

Table +

	version	timestamp	userid	userName	operation	operationParameters
1	5	2023-09-27T08:38:11.000+0000	4887540299005056	shellunext_1693422651722@npunext.onmicrosoft.com	UPDATE	{ "predicate": "[\"(id#4168 =
2	4	2023-09-27T08:37:42.000+0000	4887540299005056	shellunext_1693422651722@npunext.onmicrosoft.com	DELETE	{ "predicate": "[\"(id#3433 =
3	3	2023-09-27T08:36:55.000+0000	4887540299005056	shellunext_1693422651722@npunext.onmicrosoft.com	WRITE	{ "mode": "Append", "statsOr
4	2	2023-09-27T08:36:44.000+0000	4887540299005056	shellunext_1693422651722@npunext.onmicrosoft.com	WRITE	{ "mode": "Append", "statsOr
5	1	2023-09-27T08:36:09.000+0000	4887540299005056	shellunext_1693422651722@npunext.onmicrosoft.com	WRITE	{ "mode": "Append", "statsOr



```
1 select * from emp version as of 3
```

(3) Spark Jobs

Table ▼ +

	id	name	age	dept
1	3	c	23	DE
2	4	d	23	DE
3	1	a	23	DE
4	2	b	23	DE

4 rows | 0.93 seconds runtime

Refreshed 19 minutes ago

Command took 0.93 seconds -- by shellunext\_1693422651722@npunext.onmicrosoft.com at 9/27/2023, 2:08:43 PM on Shellunext's Cluster

Cmd 19

```
1 create table oldemp as
2 select * from emp version as of 3
```

(7) Spark Jobs

Query returned no results

Command took 4.44 seconds -- by shellunext\_1693422651722@npunext.onmicrosoft.com at 9/27/2023, 2:08:58 PM on Shellunext's Cluster

Cmd 20

```
1 select * from emp timestamp as of '2023-09-27T08:35:01.000+0000'
```

(1) Spark Jobs

Query returned no results

Command took 0.62 seconds -- by shellunext\_1693422651722@npunext.onmicrosoft.com at 9/27/2023, 2:09:51 PM on Shellunext's Cluster

Microsoft Azure databricks

Search data, notebooks, recent, and more... CTRL + P

juli-databricks shellunext\_1693422651722@npun...

New

Workspace

Recents

Catalog

Workflows

Compute

SQL

SQL Editor

Queries

Dashboards

Alerts

Query History

SQL Warehouses

Data Engineering

Job Runs

Data Ingestion

Delta Live Tables

Machine Learning

Experiments

Features

Models

Serving

Marketplace

Partner Connect

Catalog

Type to filter

For you All

hive\_metastore

default

json

bronze

sample

samples

query1

Run (1000)

hive\_metastore.default

Getting Started Pro 2XS Save\* Schedule Share

1 select \* from emp

2 select \* from json.bronze

Results

Bar 1

Bar 2

+

SUM(temperature)

timestamp

223 rows

Refreshed 2 minutes ago