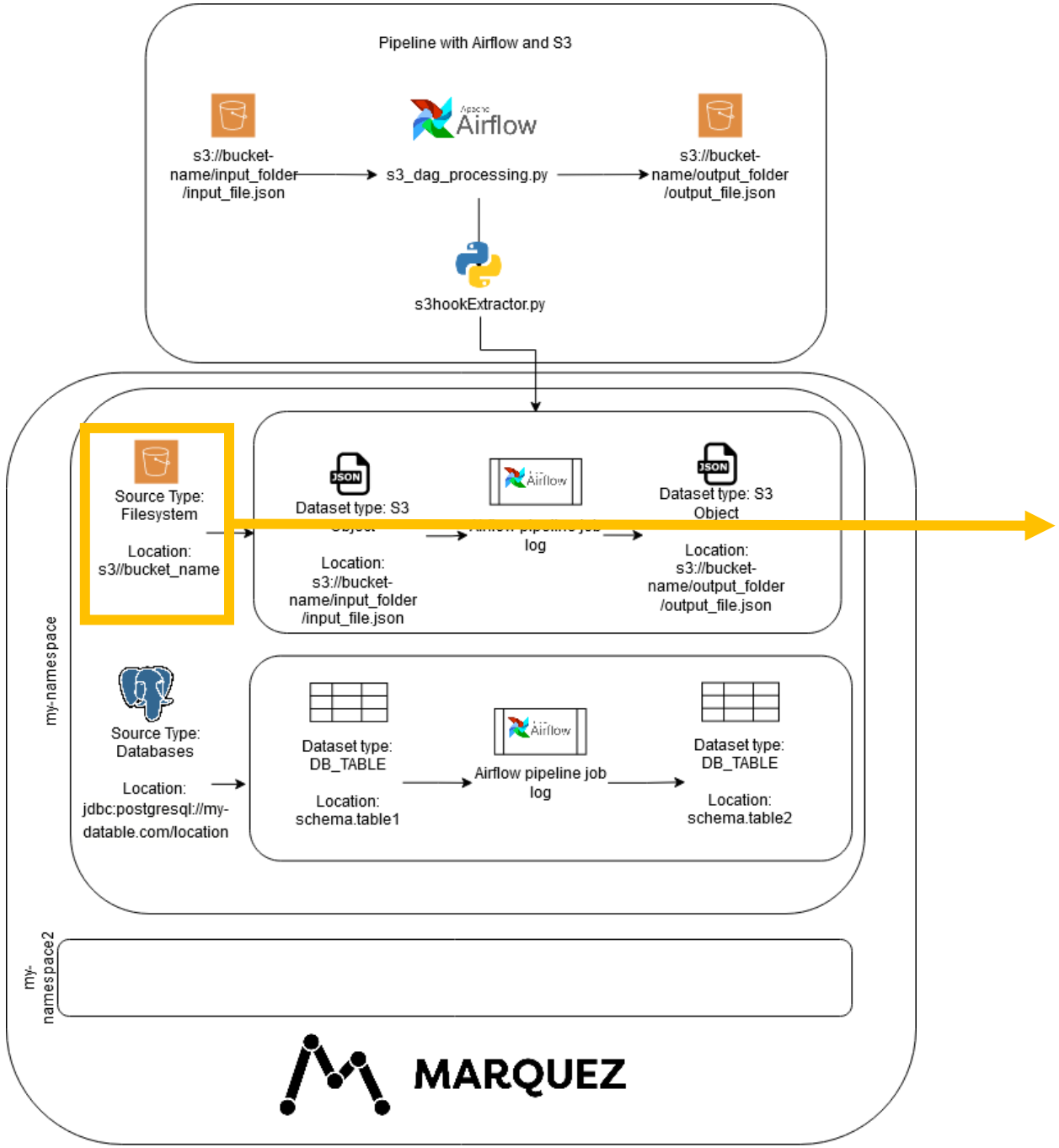


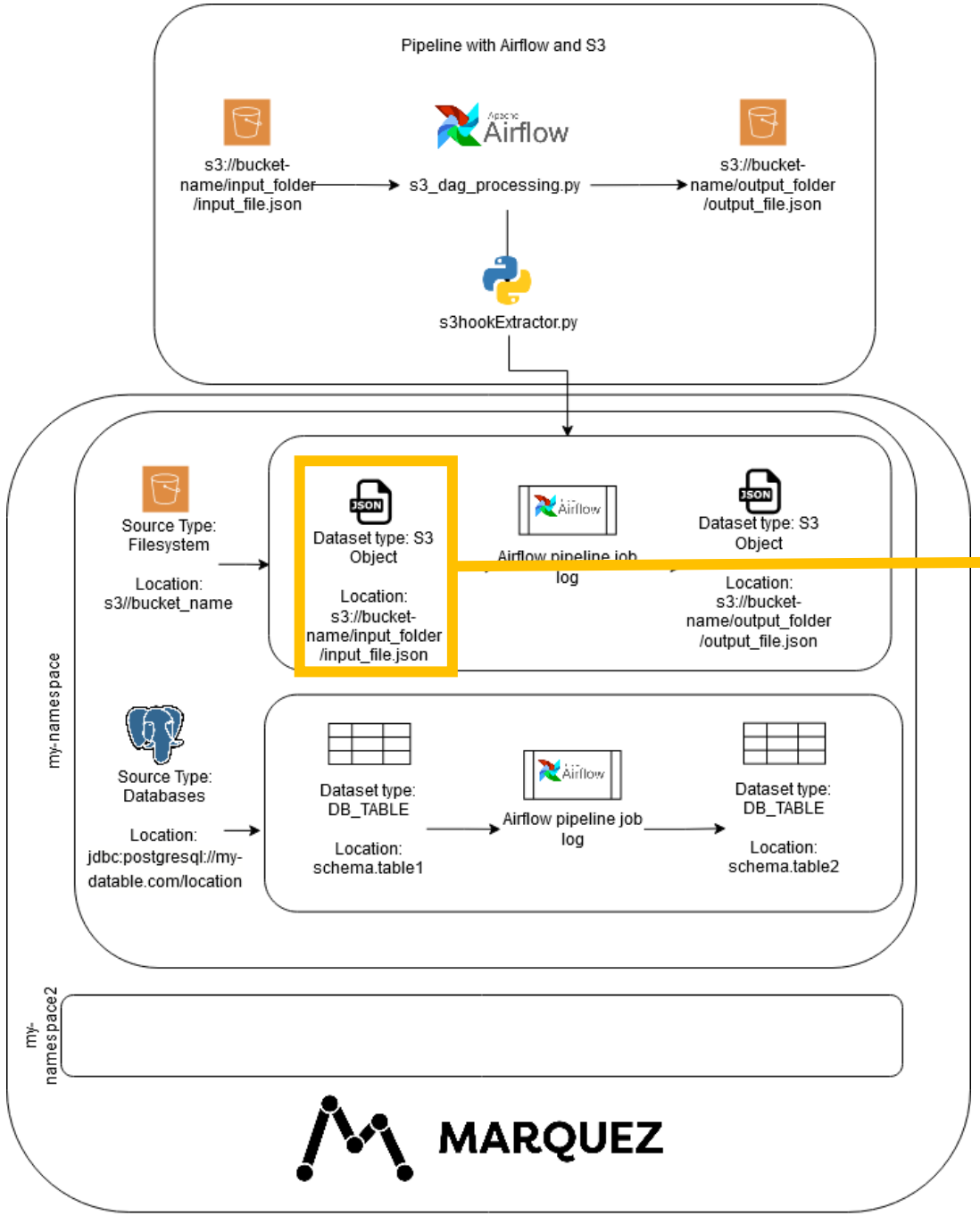
Marquez API:
PUT: http://localhost:5000/api/v1/sources/s3-source

Payload:

```
{
    "type": "S3",
    "qualifier":"FILESYSTEM",
    "connectionUrl": "s3://my-bucket-name/",
    "description": “Added S3 Bucket as source"
}
```

Suggested changes:

Possibility to add S3 URI or keep the option to add http urls.



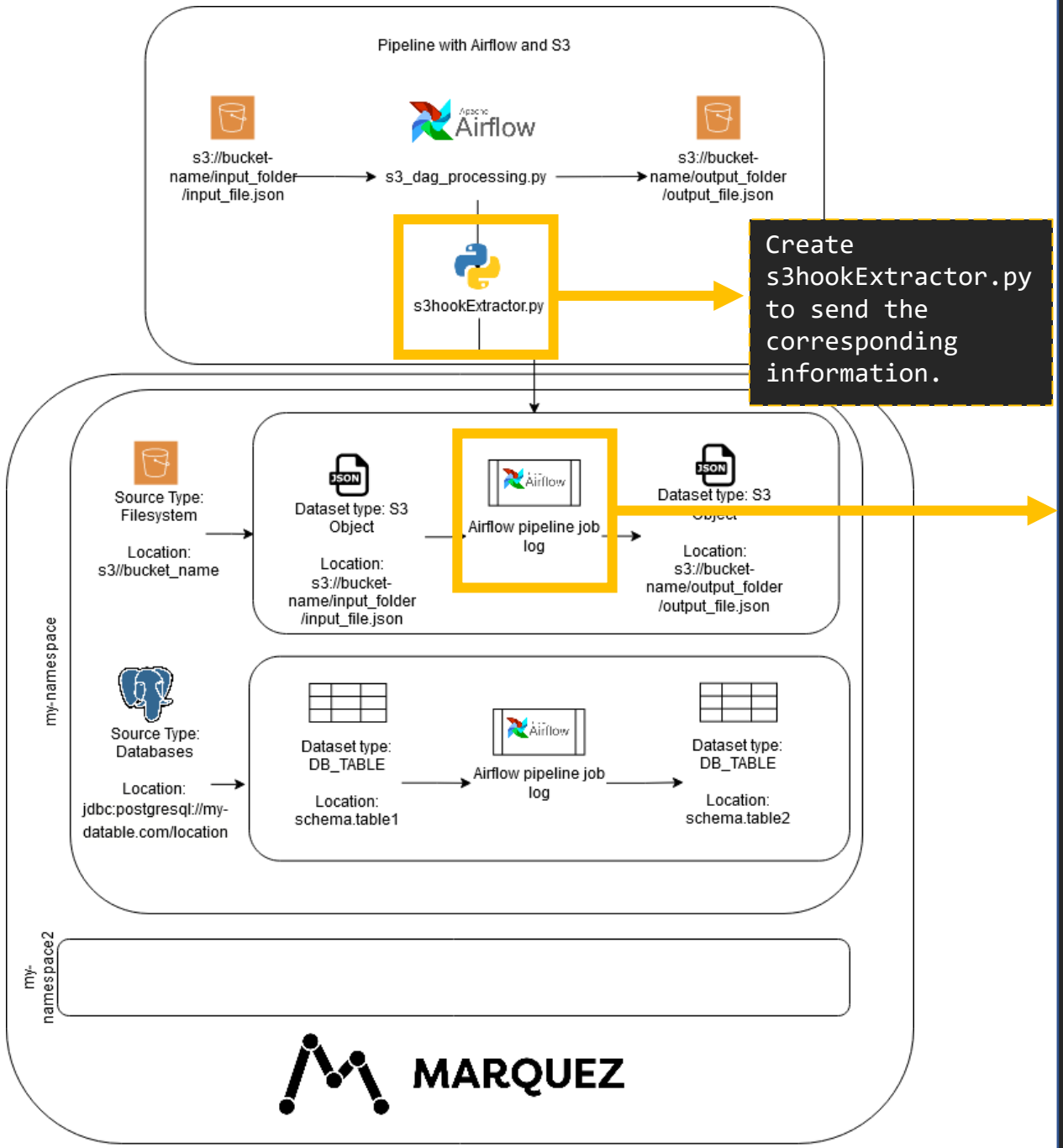
Marquez API:
PUT: http://localhost:5000/api/v1/namespaces/my-namespace/datasets/s3-dataset-input

Payload:

```
{
    "type": “OBJECT",
    "physicalName": " input_file.json",
    "sourceName": "s3-source",
    "schemaLocation": "",
    "objectLocation" :"s3://bucket-name/input_folder/input_file.json",
    "fields": [],
    "description": "s3 object for input dataset"
  }
```

Suggested changes:

- Add dataset type “OBJECT” for S3, GCP, etc.
- Add the field: “objectLocation”.



Marquez API:
PUT: http://localhost:5000/api/v1/namespaces/my-namespace/jobs/s3-job-test-1

Payload (based on improvements using Open Lineage api structure https://github.com/MarquezProject/marquez/blob/main/api/src/test/resources/open_lineage/event_full.json):


```
{
    "type": "BATCH",
    "inputs": [
      {
        "namespace": "my-namespace",
        "name": "s3-dataset-input",
        "facets":{
            "documentation": {
                "_producer": "https://github.com/Airflow/dag.py",
                "_schemaURL": "",
                "_objectURL": "https://bucket-name.s3.region.amazonaws.com/input_file.json",
                "description": "Registering input file"
              },
              "dataSource": {
                "_producer": "https://github.com/OpenLineage/OpenLineage/blob/v1-0-0/client",
                "_schemaURL": "",
                "name": "input_file.json",
                "uri": "https://bucket-name.s3.region-name.amazonaws.com/input_file.json"
              }
        }

      }
    ],
    "outputs": [
        {
          "namespace": "my-namespace",
          "name": "s3-dataset-output",
          "facets":{
            "documentation": {
                "_producer": "https://github.com/Airflow/dag.py",
                "_schemaURL": "",
                "_objectURL": "https://bucket-name.s3.region-name.amazonaws.com/output_file.json",
                "description": "Registering output file"
              },
              "dataSource": {
                "_producer": "https://github.com/Airflow/dag.py",
                "_schemaURL": "",
                "name": "output_file.json",
                "uri": "https://bucket-name.s3.region.amazonaws.com/output_file.json"
              }
        }
        }
    ],
    "location": "https://github.com/my-repo/airflow/dags/my-dag/my-dag-code.py",
    "description": "Test S3 job task!"
}
```


Suggested changes: Add field “_objectURL” and create s3hookExtractor.py