**AWS , STREAMSETS & VAULT:**

* **give a brief explanation for below scenario**
  + **Lets say some process pushes the yaml files to s3 bucket, as soon as its pushed need to convert it json and push it back same bucket.**

Solution:

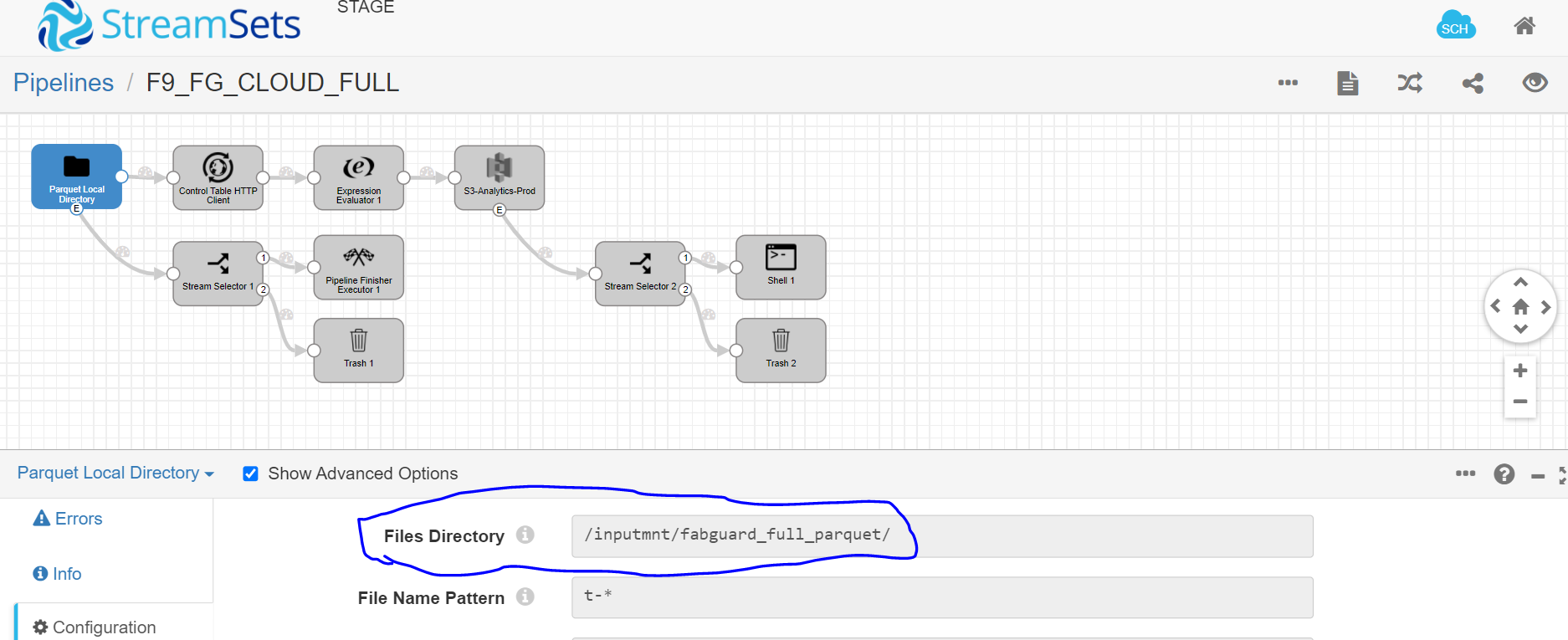
* Any files irrespective of extension can be pushed to S3
* If you need to convert yaml to json then we can use simple python convertor & tag the folder.

|  |
| --- |
| Link : <https://stackoverflow.com/questions/50846431/converting-a-yaml-file-to-python-json-object/50846530>  pip3 install pyyaml  import yaml, json  with open('./file.yaml') as f:  print(json.dumps(yaml.load(f))) |

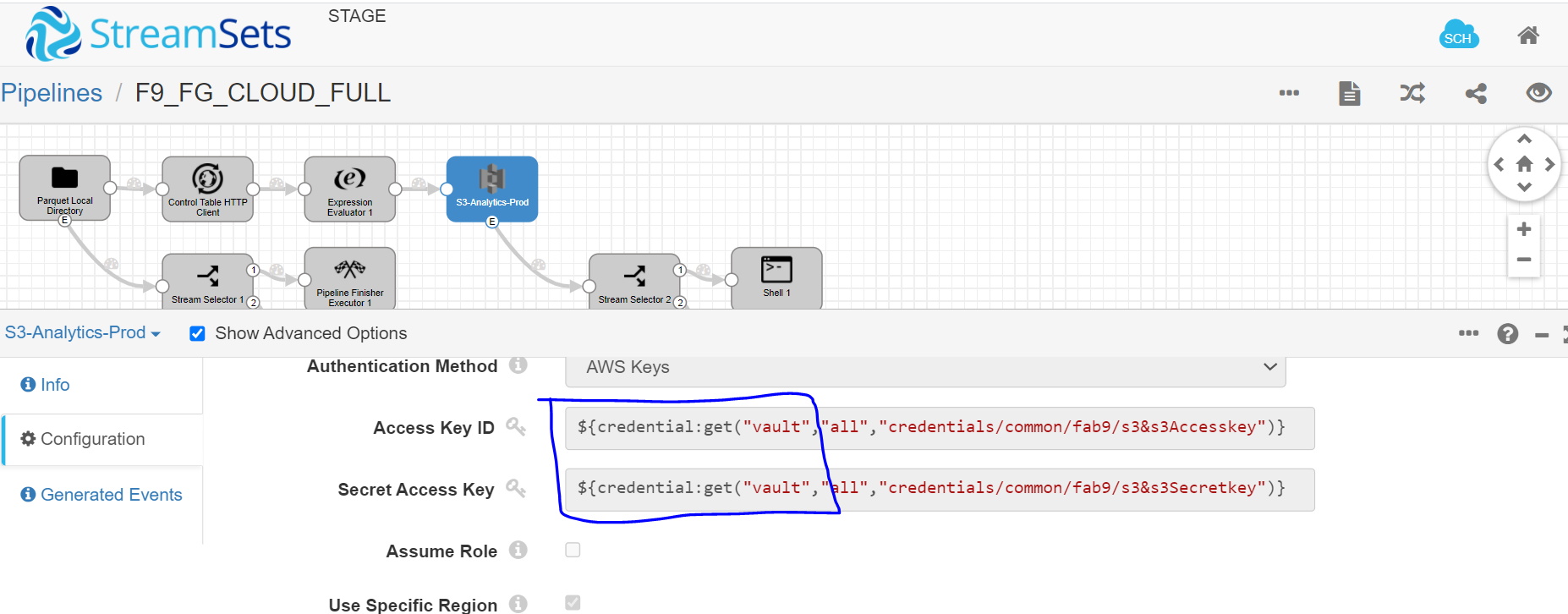
* Use streamsets which is a open source application which is used extensively in BIGDATA projects
* Below are the screenshots of the application & process flow which I have created for one of projects
* Streamsets is packaged as docker Image. So, its easy to deploy on Kubernetes & make use of extensive capabilities in BIGDATA platform.
* For Security we need to use “HASHICORP VAULT” which is widely recognized as best security for applications.
* HASHICORP VAULT can be deployed on Kubernetes & can be accessed apps deployed on Kubernetes.[Demonstrated in the below images]

Note: I have deployed & configured both Vault & streamsets. Below are the diagrams which I have worked in my project.

Input Files:



Security with VAULT



S3 Bucket Configuration:

