

Stream Processing Engine (SPE) support in stream2gym

1. Versions:
 - Apache Kafka: 2.13-2.8.0
 - Apache Flink: 1.17.1
 - Kafka is Pre-Installed and Apache Flink must be installed into pyflink dir
2. Kafka-Flink connector is placed in the 'dependency' directory in the project repository. Apache Flink-1.17.1 is installed to the pyflink directory, pyflink directory is added to PYTHONPATH at spawnSPEClients in emuLoad.py

Running a sample application with SPE support

Check the following resources to know how to specify stream processing application related attributes:

1. SPE parameters: [stream2gym/documentation/config-parameters.pdf](#)
2. Sample application with SPE support: [stream2gym/use-cases/app-testing/word-count/flink_input.graphml](#)

```
<node id="h3">
  <data key="streamProcType">Flink</data>
  <data key="streamProcCfg">use-cases/app-testing/word-count/yamlConfig/spe-flink.yaml</data>
</node>
```

streamProcType attribute is used to specify SPE type (e.g. Spark/Flink/KStream)

streamProcCfg attribute is used to specify configuration details using a YAML file.

flink-spe.yaml contains the stream processing application path and any other parameters required for the application.

```
# Stream processing engine configuration
---
app : use-cases/app-testing/word-count/flink-word-count/flink-word-count.py
...
```

P.S. currently stream2gym only supports passing the application path. To support other parameters, please update readStreamProcConfig method inside configParser script.