

### Task:

构建 Flink Demo

验证环境:

Windows: WSL

windows10-(Linux (ubuntu 18.04))

openjdk version "11.0.8" 2020-07-14

OpenJDK Runtime Environment (build 11.0.8+10-post-Ubuntu-0ubuntu120.04)

OpenJDK 64-Bit Server VM (build 11.0.8+10-post-Ubuntu-0ubuntu120.04, mixed mode, sharing)

### Demo:

要想运行 Flink，我们需要下载并解压 Flink 的二进制包，下载地址如下：

<https://flink.apache.org/downloads.html>

我们可以选择 Flink 与 Scala 结合版本，这里我们选择最新的 1.11.1 版本 [Apache Flink 1.11.1 for Scala 2.12](#) 进行下载。

Windows 用户可以在 Cygwin 或 WSL 中运行 Flink。为了能够运行 Flink，唯一的要求是安装有效的 **Java 8 或 11**。

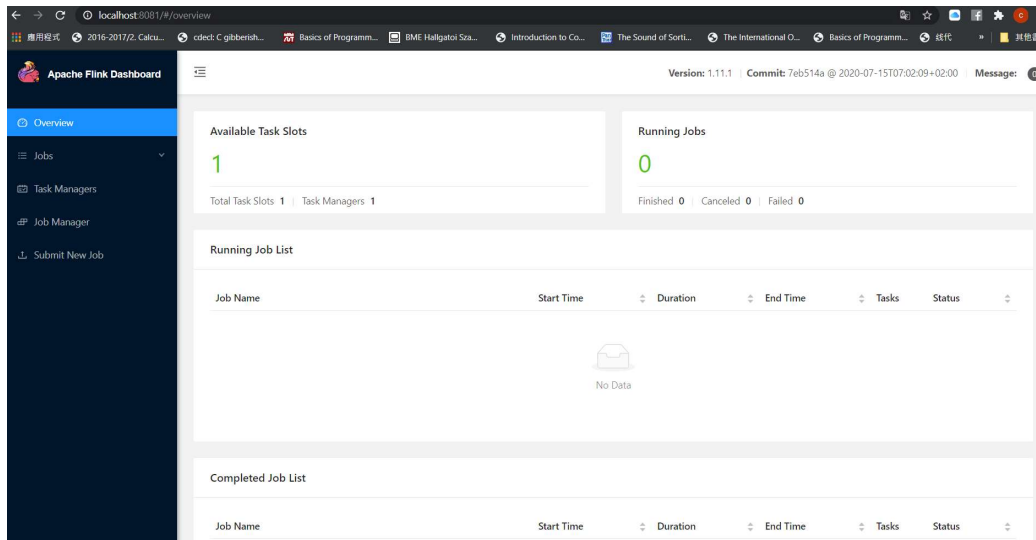
网上有大量的通过 .bat 运行，但最新版本的 Flink 并没有提供 .bat 文件，官网上显示可以通过 WSL 执行。

```
/flink-1.11.1/bin$ ./start-cluster.sh
Starting cluster.
Starting standalone session daemon on host LAPTOP-2HMAC39H.
Starting taskexecutor daemon on host LAPTOP-2HMAC39H.
```

运行 flink 示例程序

在 <http://localhost:8081> 上检查 Dispatcher 的 Web 前端，并确保一切正常并正在运行。Web 前端应报告一个可用的 TaskManager 实例。

**调度程序：概述**



还可以通过检查 logs 目录中的日志文件来验证系统是否正在运行：

```
$ tail log/flink-*-standalonesession-*.log
INFO ... - Rest endpoint listening at localhost:8081
INFO ... - http://localhost:8081 was granted leadership ...
INFO ... - Web frontend listening at http://localhost:8081.
INFO ... - Starting RPC endpoint for StandaloneResourceManager at
akka://flink/user/resourcemanager .
INFO ... - Starting RPC endpoint for StandaloneDispatcher at
akka://flink/user/dispatcher .
INFO ... - ResourceManager akka.tcp://flink@localhost:6123/user/resourcemanager
was granted leadership ...
INFO ... - Starting the SlotManager.
INFO ... - Dispatcher akka.tcp://flink@localhost:6123/user/dispatcher was granted
leadership ...
INFO ... - Recovering all persisted jobs.
INFO ... - Registering TaskManager ... at ResourceManager
```

批处理示例：

提交 flink 的批处理 examples 程序：

```
./bin/flink run examples/batch/WordCount.jar
```

这是 flink 提供的 examples 下的批处理例子程序，统计单词个数。

```
./bin/flink run examples/batch/WordCount.jar
Executing WordCount example with default input data set.
Use --input to specify file input.
Printing result to stdout. Use --output to specify output path.
Job has been submitted with JobID 13e0472d51f00d7f87fb29e962e064a4
Program execution finished
Job with JobID 13e0472d51f00d7f87fb29e962e064a4 has finished.
Job Runtime: 381 ms
```

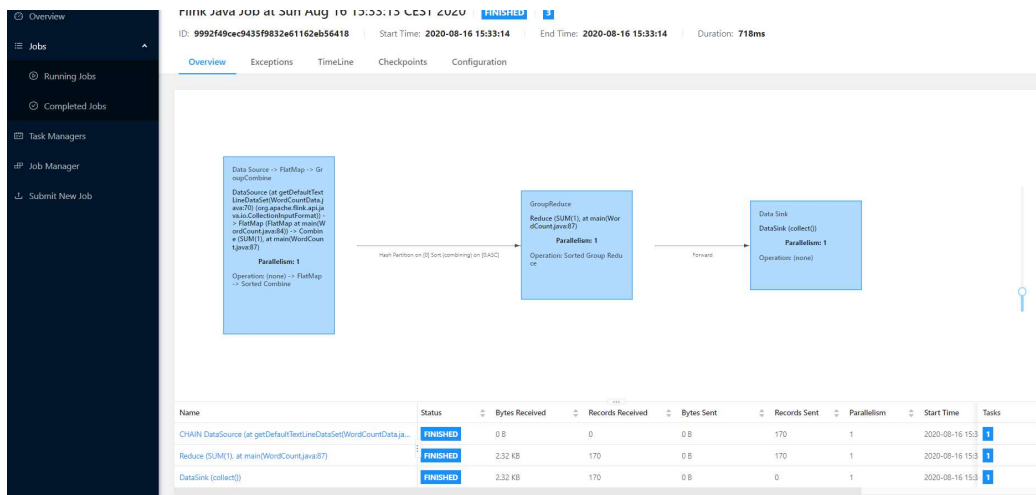
### Accumulator Results:

- 58a01edc3dec184b852bd1f96e0c9f2f (java.util.ArrayList) [170 elements]

(a,5)  
(action,1)  
(after,1)  
(against,1)  
(all,2)  
(and,12)  
(arms,1)  
(arrows,1)  
(awry,1)  
(ay,1)  
(bare,1)  
(be,4)  
(bear,3)  
(bodkin,1)  
(bourn,1)  
(but,1)  
(by,2)

得到结果，这里统计的是默认的数据集，可以通过--input --output 指定输入输出。

我们可以在页面中查看运行的情况：



流处理示例：

启动 nc 服务器：

nc -l 9000

提交 flink 的批处理 examples 程序：

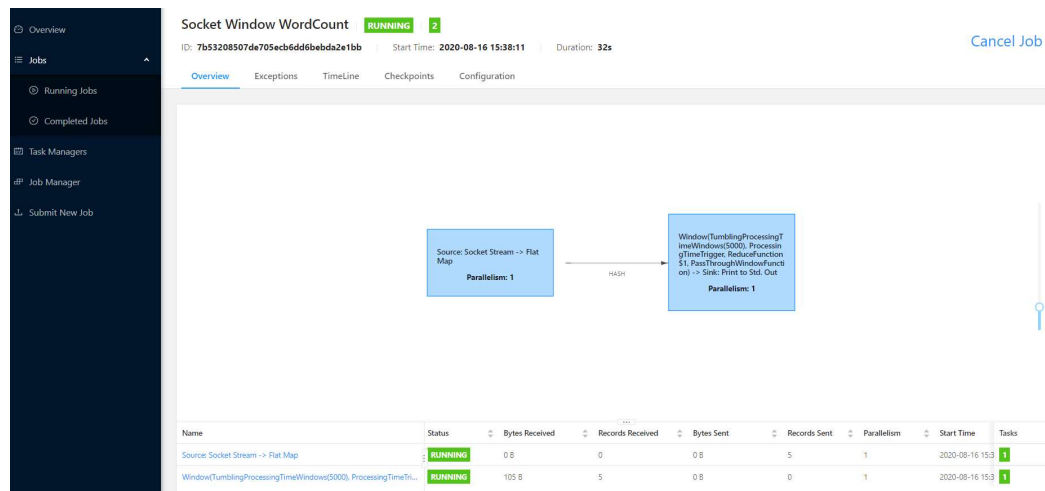
```
bin/flink run examples/streaming/SocketWindowWordCount.jar --port 9000
```

这是 flink 提供的 examples 下的流处理例子程序，接收 socket 数据传入，统计单词个数。

在 nc 端写入单词

```
$ nc -l 9000
lorem ipsum
impsum hello haha
bye BYE
```

可以在 Web 查看运行状态



输出在日志中

```
$ tail -f log/flink-*-taskexecutor-*.out
lorem : 1
ipsum : 1
impsum : 1
haha : 1
hello : 1
bye : 2
```

停止 flink

```
$ ./bin/stop-cluster.sh
```

参考 Link:

Flink doc: <https://ci.apache.org/projects/flink/flink-docs-master/ops/deployment/local.html>

Flink 入门（三）——环境与部署: <https://cloud.tencent.com/developer/article/1554210>

Flink release v1.11 doc: <https://ci.apache.org/projects/flink/flink-docs-release-1.11/>

[https://ci.apache.org/projects/flink/flink-docs-release-1.11/try-flink/local\\_installation.html](https://ci.apache.org/projects/flink/flink-docs-release-1.11/try-flink/local_installation.html)

Windows 环境下 Flink 消费 Kafka 实现热词统计: <https://developer.aliyun.com/article/698298>