

Introduction to Apache Spark APIs for Data Processing Monitoring and Measuring Spark Jobs Execution

Luca Canali

CERN IT, Data Analytics and Spark Service



Web UI

- Main entry point to Spark instrumentation
- Spark Web UI provides information on
 - Jobs, stages, tasks
 - Executors and used resources
 - DataFrame and SQL operations, Streaming
- Connect to the Web UI URL
 - Default: <http://driver-node-host:4040>

Web UI

Active Jobs (1)

Page: 1 1 Pages. Jump to 1 Show 100 items in a page. Go

Job Id	Description	Submitted	Duration	Stages: Succeeded/Total	Tasks (for all stages): Succeeded/Total
7	count at <console>:26 count at <console>:26 (kill)	2019/08/10 17:50:13	17 s	0/2	0/5 (4 running)

Page: 1 1 Pages. Jump to 1 Show 100 items in a page. Go

Completed Jobs (7)

Page: 1 1 Pages. Jump to 1 Show 100 items in a page. Go

Job Id	Description	Submitted	Duration	Stages: Succeeded/Total	Tasks (for all stages): Succeeded/Total
6	show at <console>:26 show at <console>:26	2019/08/10 17:49:30	0.4 s	1/1	1/1
5	show at <console>:28 show at <console>:28	2019/08/10 17:48:32	0.8 s	3/3	9/9
4	show at <console>:28 show at <console>:28	2019/08/10 17:47:40	2 s	3/3	9/9

Executors

Show Additional Metrics

- Select All
- On Heap Memory
- Off Heap Memory

Summary

	RDD Blocks	Storage Memory	Disk Used	Cores	Active Tasks	Failed Tasks	Complete Tasks	Total Tasks	Task Time (GC Time)	Input	Shuffle Read	Shuffle Write	Blacklisted
Active(3)	0	5.9 KiB / 1.1 GiB	0.0 B	2	0	0	5	5	4 s (0.2 s)	0.0 B	0.0 B	0.0 B	0
Total(3)	0	5.9 KiB / 1.1 GiB	0.0 B	2	0	0	5	5	4 s (0.2 s)	0.0 B	0.0 B	0.0 B	0
Dead(0)	0	0.0 B / 0.0 B	0.0 B	0	0	0	0	0	0.0 ms (0.0 ms)	0.0 B	0.0 B	0.0 B	0

Executors

Show 20 entries

Search:

Executor ID	Address	Status	RDD Blocks	Storage Memory	Disk Used	Cores	Active Tasks	Failed Tasks	Complete Tasks	Total Tasks	Task Time (GC Time)	Input	Shuffle Read	Shuffle Write	Logs	Thread Dump
1	10.12.221.27:55834	Active	0	2 KiB / 366.3 MiB	0.0 B	1	0	0	3	3	2 s (0.1 s)	0.0 B	0.0 B	0.0 B	stdout stderr	Thread Dump
0	10.12.221.27:55835	Active	0	2 KiB / 366.3 MiB	0.0 B	1	0	0	2	2	2 s (94.0 ms)	0.0 B	0.0 B	0.0 B	stdout stderr	Thread Dump
driver	10.12.221.27:55827	Active	0	2 KiB / 366.3 MiB	0.0 B	0	0	0	0	0	0.0 ms (0.0 ms)	0.0 B	0.0 B	0.0 B		Thread Dump

Showing 1 to 3 of 3 entries

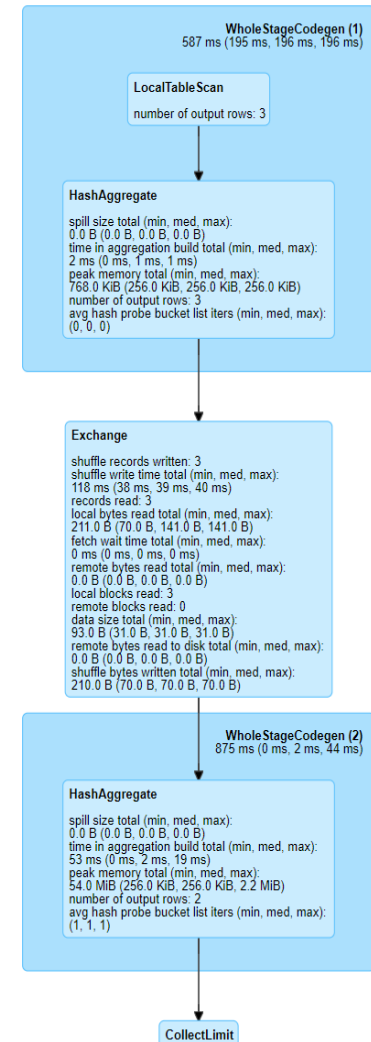
Previous 1 Next

Details for Query 2

Submitted Time: 2019/11/20 09:31:38

Duration: 1 s

Succeeded Jobs: 1 2 3 4 5

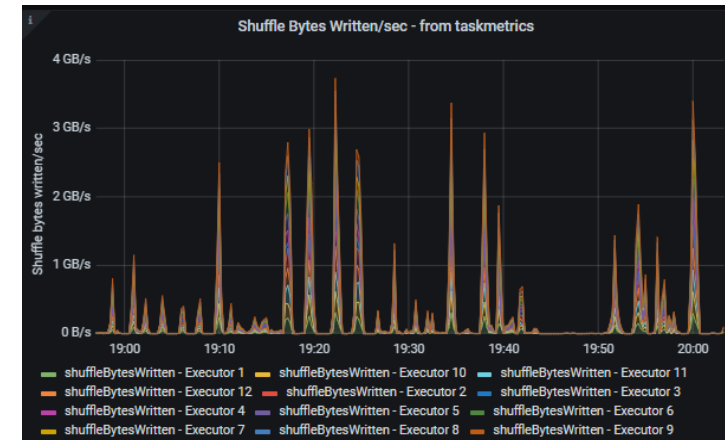
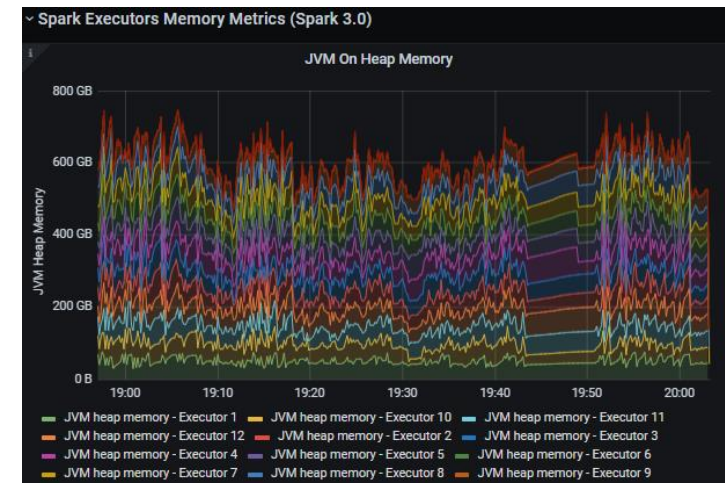
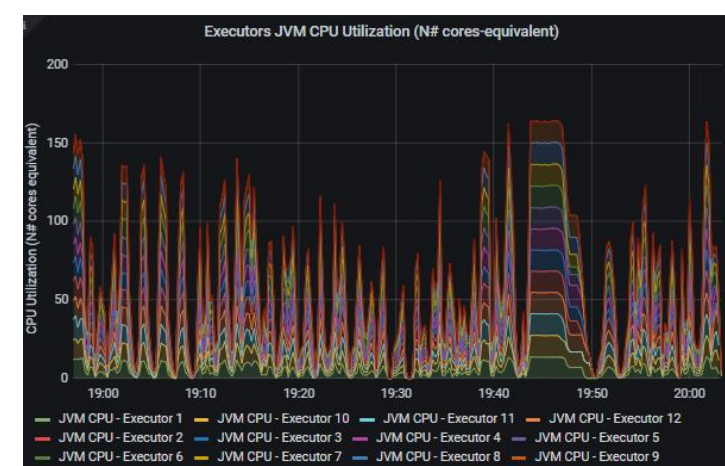


Details



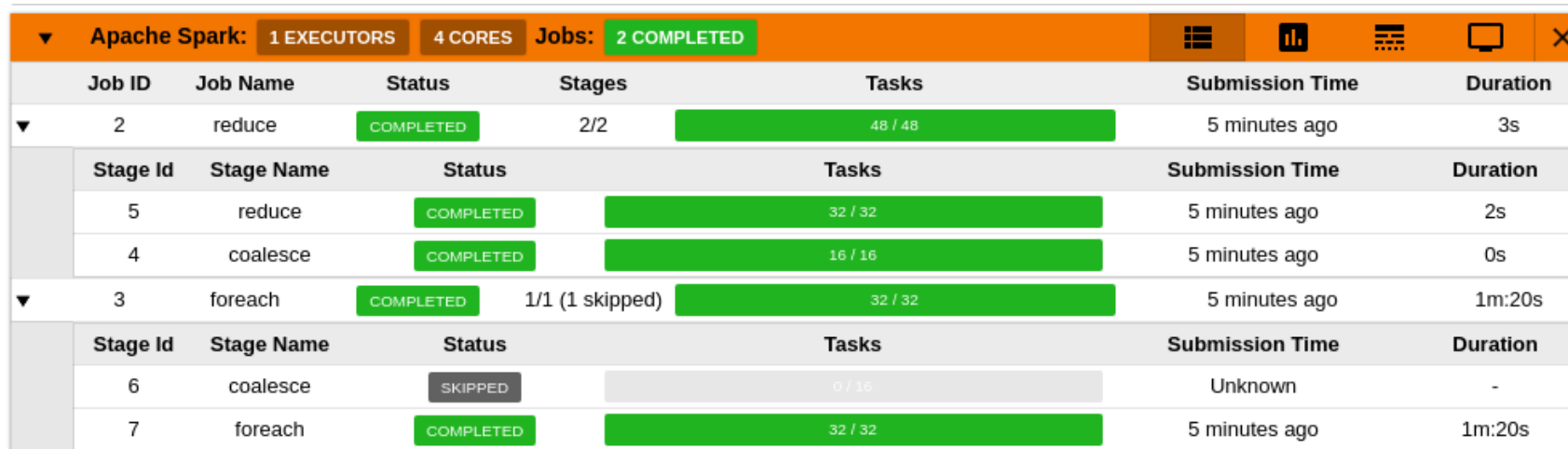
Spark Performance Dashboard

- Visualize Spark metrics
 - **Real-time** + historical data
 - Summaries and time series of key metrics
 - Data for root-cause **analysis**
 - See <https://github.com/cerndb/spark-dashboard>



Spark Monitor

- Automatically displays a live monitoring tool below cells that run Spark jobs in a Jupyter notebook (and Jupyter lab)
 - A table of **jobs and stages with progress bars**
 - A graph showing number of **active tasks & executor cores vs time**
 - <https://github.com/swan-cern/sparkmonitor>
 - Integrated with SWAN notebooks



The screenshot displays the Spark Monitor interface. At the top, a summary bar shows 'Apache Spark: 1 EXECUTORS 4 CORES Jobs: 2 COMPLETED'. Below this is a table with columns: Job ID, Job Name, Status, Stages, Tasks, Submission Time, and Duration. Job 2 (reduce) is completed, showing 48/48 tasks. Job 3 (foreach) is also completed, showing 32/32 tasks, but includes a skipped stage (6) with 0/16 tasks. Each stage and task row includes a progress bar and a status indicator (COMPLETED or SKIPPED).

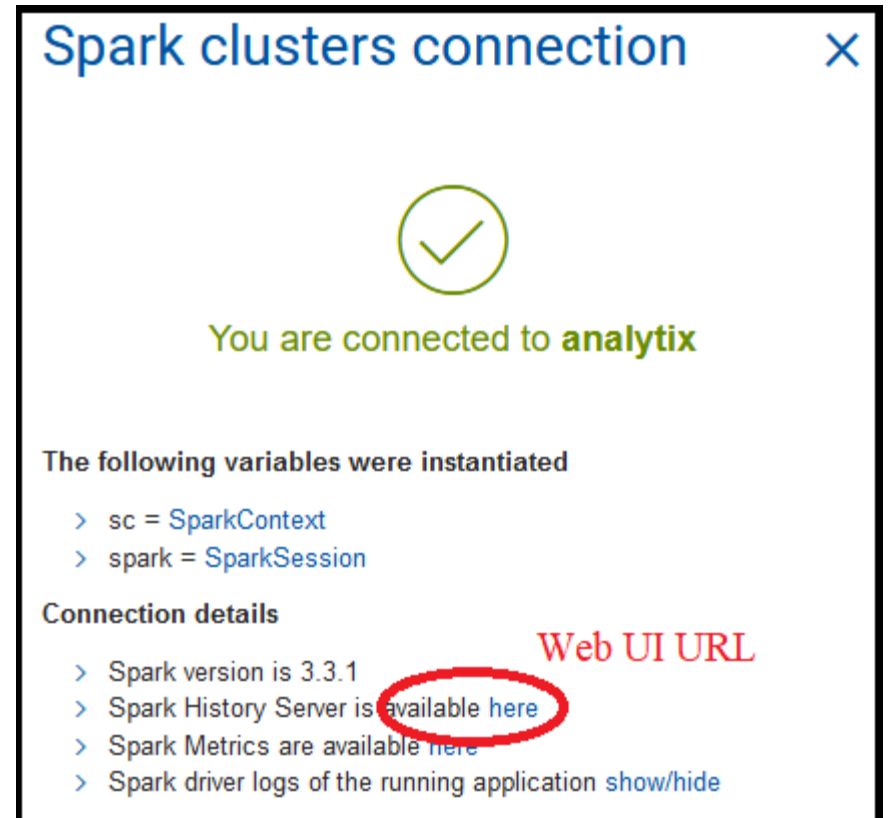
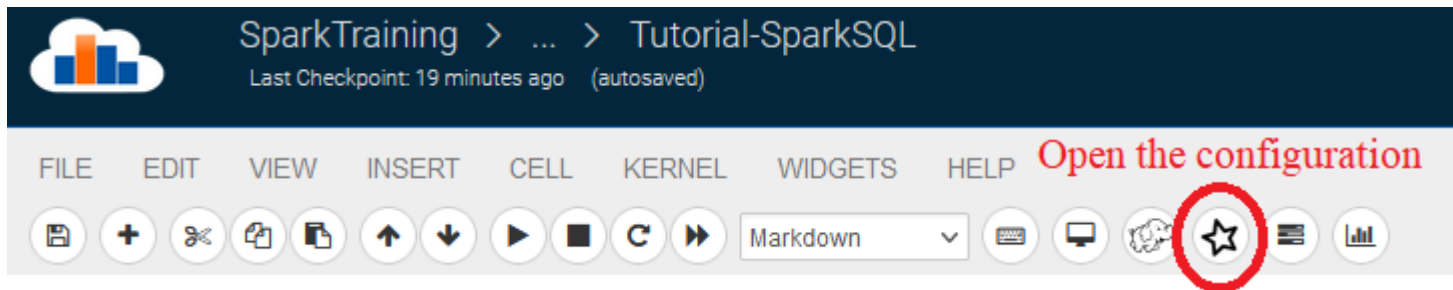
Job ID	Job Name	Status	Stages	Tasks	Submission Time	Duration
2	reduce	COMPLETED	2/2	48 / 48	5 minutes ago	3s
	Stage Id	Stage Name	Status	Tasks	Submission Time	Duration
	5	reduce	COMPLETED	32 / 32	5 minutes ago	2s
	4	coalesce	COMPLETED	16 / 16	5 minutes ago	0s
3	foreach	COMPLETED	1/1 (1 skipped)	32 / 32	5 minutes ago	1m:20s
	Stage Id	Stage Name	Status	Tasks	Submission Time	Duration
	6	coalesce	SKIPPED	0 / 16	Unknown	-
	7	foreach	COMPLETED	32 / 32	5 minutes ago	1m:20s

SWAN and Spark Monitor

- Spark monitor is active by default
 - Job view, Tasks view and Event timeline
- ADD PICTURE

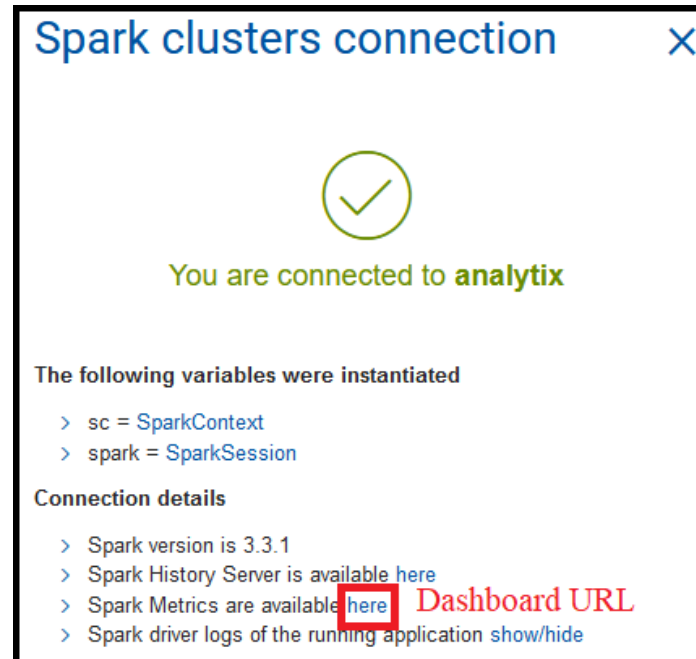
The Spark Web UI from SWAN

- Open the configuration window (“star” button)
- Follow the link to the Web UI



Spark Metrics Dashboard from SWAN

- Configuration
 - Configure the extra metrics logging at Spark session configuration
- Open the dashboard
- Follow the URL



Spark clusters connection

You are connected to **analytix**

The following variables were instantiated

- > sc = SparkContext
- > spark = SparkSession

Connection details

- > Spark version is 3.3.1
- > Spark History Server is available [here](#)
- > Spark Metrics are available [here](#) **Dashboard URL**
- > Spark driver logs of the running application [show/hide](#)

Spark clusters connection

You are going to connect to:
analytix

You can configure the following options.
Environment variables can be used via {ENV_VAR_NAME}.

Add a new option

Write the option name...

Bundled configurations

These options will be overwritten by non-bundled options if specified

- Include CMSSpark options
- Include SparkMetrics options **Activate this to use the metrics dashboard**
- Include PropagateUserPythonModules options
- Include ShipKerberosToExecutors options

Selected configuration

- ⚙️ SparkMetrics
 - ⚙️ spark.cern.grafana.url
https://hadoop-grafana.web.cern.ch/d/1/sparkmetrics
 - ⚙️ spark.metrics.conf.driver.sink.graphite.class
org.apache.spark.metrics.sink.GraphiteSink
 - ⚙️ spark.metrics.conf.executor.sink.graphite.class
org.apache.spark.metrics.sink.GraphiteSink
 - ⚙️ spark.metrics.conf.*.sink.graphite.host
dbod-sparkm.cern.ch
 - ⚙️ spark.metrics.conf.*.sink.graphite.port
8292

Monitoring Spark on Hadoop/YARN

- URL of the YARN Web UI
 - Find the URL in the doc: https://hadoop-user-guide.web.cern.ch/infra/list_of_clusters/
 - URL is of the form:
 - `https://<YARN-RM-HOST>.cern.ch:8088/cluster/apps/RUNNING`
 - List running applications for your cluster
 - Find there your application

Spark Measure

- Custom package, measure execution metrics
 - For advanced troubleshooting and performance studies
 - <https://github.com/LucaCanali/sparkMeasure>

```
Spark Context default degree of parallelism = 8
Aggregated Spark stage metrics:
numStages => 3
numTasks => 17
elapsedTime => 13520 (14 s)
stageDuration => 13411 (13 s)
executorRunTime => 100020 (1.7 min)
executorCpuTime => 98899 (1.6 min)
executorDeserializeTime => 4358 (4 s)
...
```

Logs, and logging levels settings

- Executors logs
 - Spark Web UI under the “executors” tab
 - YARN UI -> look for container logs
- Change Spark logging verbosity:
 - `spark.sparkContext.setLogLevel("INFO")`
- Fine grained logging config
 - Edit `$SPARK_CONF_DIR/log4j2.properties`

Key Learning Points

- Spark job execution can be complex
- Monitoring, instrumentation and logging are key
- Tools from Apache Spark and ecosystem
 - Spark **Web UI** is the main entry point
 - SWAN integrates the “**spark monitor**” widget
 - Advanced troubleshooting: Spark **dashboard**, sparkMeasure, configuration and verbose logging

Demo

- See the video: monitoring Spark on SWAN