

Cloud Computing-5573

Date: 10/20/2024

Group8:

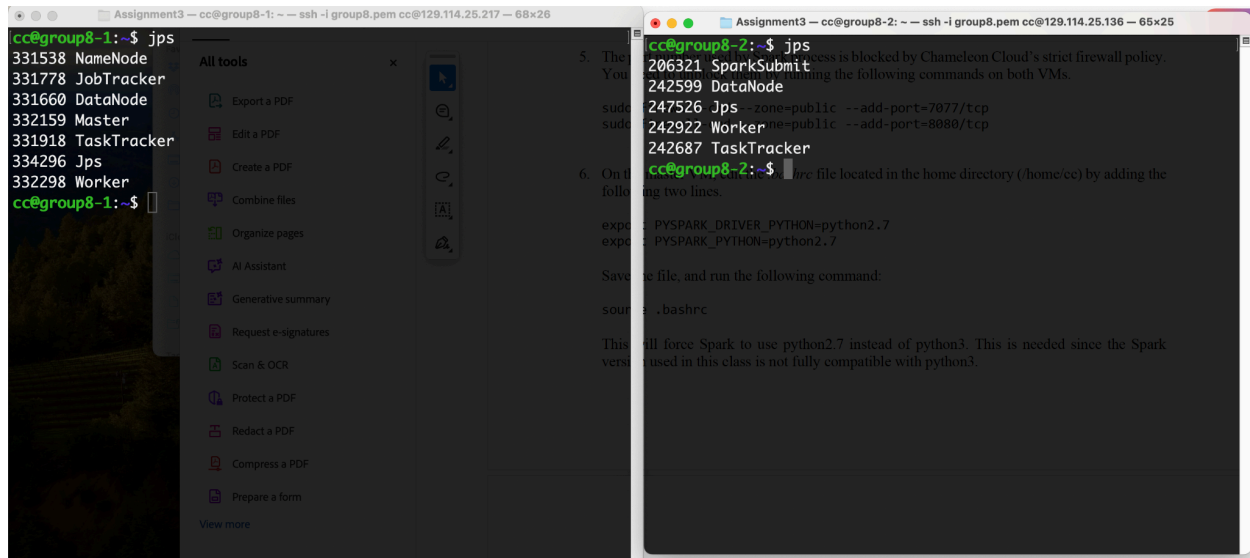
Amjad Alqahtani

Caleb Alva

Nicholas Winkelmann

Assignment 2: Spark Setup and Programming


Part 1: We have followed the instructions provided in the assignment handout from step 1 through step 6. Finally, we executed the `jps` command to display all daemon processes.



Part 2: We have completed step 7, and the screenshot of the output is provided below.

```
SUCCESS: 0.0
The location with the most crimes: BROOKLYN. number of crimes: 106202
The top 3 crimes reported in July:
PETIT LARCENY: 7086 crimes
HARRASSMENT 2: 5738 crimes
ASSAULT 3 & RELATED OFFENSES: 4804 crimes
Number of DANGEROUS WEAPONS crimes reported in July: 701
```

Part 3: Screenshot of Web UI



Spark Master at spark://10.56.0.205:7077

URL: spark://10.56.0.205:7077

REST URL: spark://10.56.0.205:6066 (cluster mode)

Alive Workers: 2

Cores in use: 4 Total, 0 Used

Memory in use: 5.7 GB Total, 0.0 B Used

Applications: 0 Running, 6 Completed

Drivers: 0 Running, 0 Completed

Status: ALIVE

Workers

Worker id	Address	State	Cores	Memory
worker-20241019163248-10.56.2.193-42255	10.56.2.193:42255	ALIVE	2 (0 Used)	2.8 GB (0.0 B Used)
worker-20241019163249-10.56.0.205-41669	10.56.0.205:41669	ALIVE	2 (0 Used)	2.8 GB (0.0 B Used)

Running Applications

Application ID	Name	Cores	Memory per Node	Submitted Time	User	State	Duration
----------------	------	-------	-----------------	----------------	------	-------	----------

Completed Applications

Application ID	Name	Cores	Memory per Node	Submitted Time	User	State	Duration
app-20241019170655-0005	NYC_Crime_Analysis	4	1024.0 MB	2024/10/19 17:06:55	cc	FINISHED	16 s
app-20241019165524-0004	NYC_Crime_Analysis	4	1024.0 MB	2024/10/19 16:59:24	cc	FINISHED	17 s
app-20241019165536-0003	PySparkShell	4	1024.0 MB	2024/10/19 16:55:36	cc	FINISHED	3.0 min
app-20241019164639-0002	NYC_Crime_Analysis	4	1024.0 MB	2024/10/19 16:46:39	cc	FINISHED	17 s
app-20241019163938-0001	NYC_Crime_Analysis	4	1024.0 MB	2024/10/19 16:39:38	cc	FINISHED	17 s
app-20241019163649-0000	NYC_Crime_Analysis	4	1024.0 MB	2024/10/19 16:36:49	cc	FINISHED	17 s

Part 4: Team contributions

Amjad Alqahtani	Caleb Alva	Nicholas Winkelmann
<div>1. Setup the cluster</div> <div>2. Review code</div> <div>3. Prepare Report</div> <div>4. Keep the timeline of the team</div>	<div>1. Debug the code</div> <div>2. Fixed an error</div> <div>3. Test and evaluate the code</div> <div>4. Participate in the report</div>	<div>1. Write the code</div> <div>2. Testing the cluster setup</div> <div>3. Participate in the report</div>