

mXpress

Scrum 3

By Team MX

Ranil Fernando, Scott Thomas, Jonathon Florek

2020-01-14

mXpress Emergency Vehicle Routing System



Solving slow emergency response times in large cities



Fills the need for high quality, specialized routing systems for emergency dispatchers



Innovates by utilizing both batch and stream big data coupled with AI to do predictive and reactive trip planning



Potentially allows for consumer vehicle routing using batch data



Adaptable framework allows for multiple qualities of service provided



Spark Master at spark://master1:7077

URL: spark://master1:7077

Alive Workers: 3

Cores in use: 22 Total, 0 Used

Memory in use: 43.5 GB Total, 0.0 B Used

Applications: 0 Running, 2 Completed

Drivers: 0 Running, 0 Completed

Status: ALIVE

Workers (3)

Worker Id	Address	State	Cores	Memory
worker-20200105154514-192.168.1.130-42983	192.168.1.130:42983	ALIVE	5 (0 Used)	14.5 GB (0.0 B Used)
worker-20200105154602-192.168.1.136-41081	192.168.1.136:41081	ALIVE	12 (0 Used)	14.5 GB (0.0 B Used)
worker-20200105154830-192.168.1.180-42971	192.168.1.180:42971	ALIVE	5 (0 Used)	14.5 GB (0.0 B Used)

Running Applications (0)

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

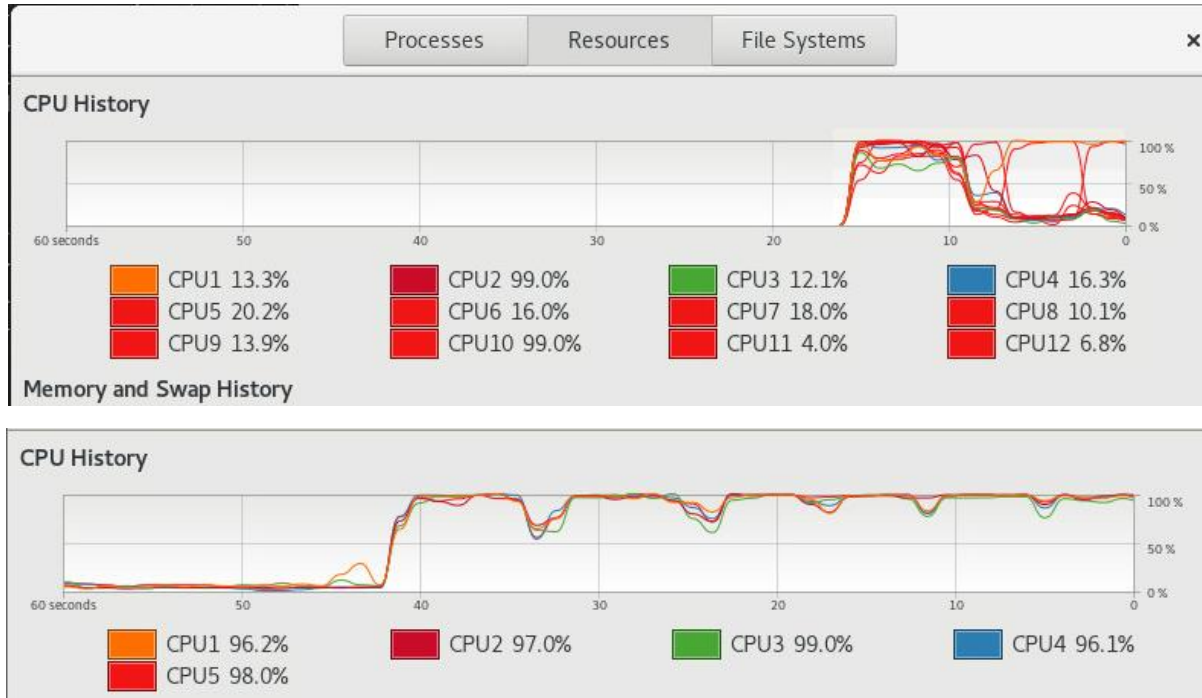
Completed Applications (2)

Application ID	Name	Cores	Memory per Executor	Submitted Time	User	State	Duration
app-20200105155318-0001	Spark Pi	22	1024.0 MB	2020/01/05 15:53:18	spark_1	FINISHED	1.8 min
app-20200105154450-0000	Spark shell	0	1024.0 MB	2020/01/05 15:44:50	spark_2	FINISHED	14 s

- Fully functional Spark & Hadoop cluster
 - Workers
 - Applications
 - Jobs
- Batch processing in Scala
- Mapping sensors to geographic position
- Functional OSRM Server
- Routetagger utility application

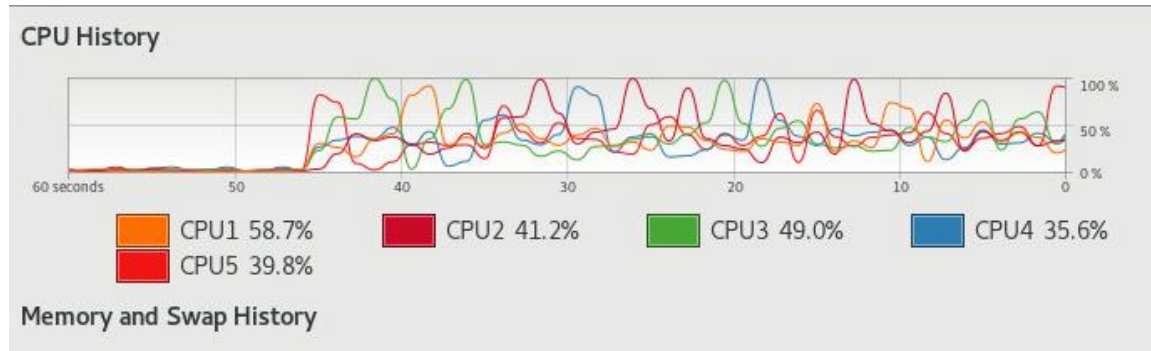
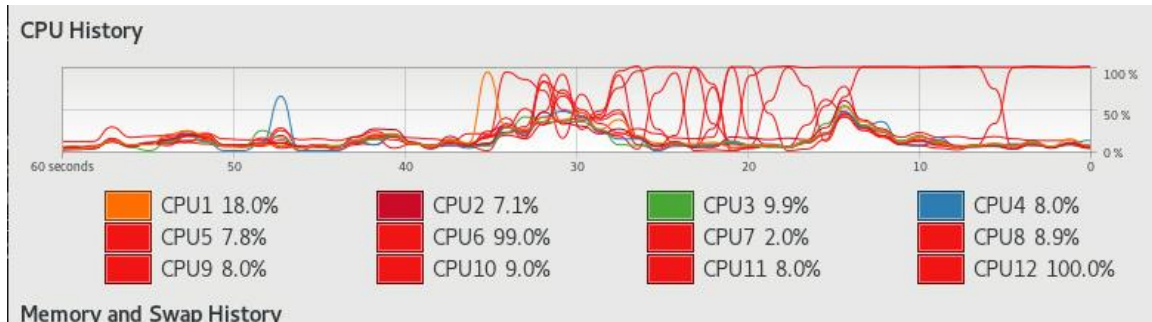
Project Progress

Single Core Execution Units



- 22 Execution Units Total
 - 12x Master Node
 - 5x + 5x Slave
- Slave Utilization 96%+
 - Near perfect scaling
- Completion time 102s
 - 1 Billion Iterations Pi Approximation

Multi Core Execution Units



- 3x Execution Units
 - 12 Core
 - 5 Core
 - 5 Core
- Poor Scaling
 - Below 30% Total Utilization
- Completion Time 430s
 - Same 1 Billion Operations

Mapping Demo

- OSRM
- Electron

Next Steps

- Pruning CSV data for batch
- Sensor simulation
- Kafka pipeline
- More advanced programming in Scala
- Output of Spark CSV -> MongoDB -> OSRM



Backbone & Testing

- 3 Node Spark Cluster
 - Distributed Compute Load
 - Multiple Execution Units
- Running Scala code on cluster
- Performance Testing
 - Ethernet Bandwidth
 - Connect/Disconnect time
 - Divide/Join time

```
spark_1@localhost:/usr/spark/spark-2.4.4-bin-hadoop2.7
File Edit View Search Terminal Tabs Help

spark_1@localhost:/usr/spark/spark-2.4.4-bin-hadoop2.7 x
spark_1@localhost:/usr/spark/spark-2.4.4-bin-hadoop2.7 x [icon] v

19/10/20 13:44:00 INFO TaskSetManager: Finished task 3.0 in stage 0.0 (TID 3) in 423 ms on localhost (executor driver) (1/10)
19/10/20 13:44:00 INFO Executor: Running task 9.0 in stage 0.0 (TID 9)
19/10/20 13:44:00 INFO Executor: Running task 7.0 in stage 0.0 (TID 7)
19/10/20 13:44:00 INFO Executor: Running task 8.0 in stage 0.0 (TID 8)
19/10/20 13:44:00 INFO TaskSetManager: Finished task 1.0 in stage 0.0 (TID 1) in 442 ms on localhost (executor driver) (2/10)
19/10/20 13:44:00 INFO TaskSetManager: Finished task 4.0 in stage 0.0 (TID 4) in 444 ms on localhost (executor driver) (3/10)
19/10/20 13:44:00 INFO TaskSetManager: Finished task 0.0 in stage 0.0 (TID 0) in 472 ms on localhost (executor driver) (4/10)
19/10/20 13:44:00 INFO TaskSetManager: Finished task 2.0 in stage 0.0 (TID 2) in 447 ms on localhost (executor driver) (5/10)
19/10/20 13:44:00 INFO Executor: Finished task 6.0 in stage 0.0 (TID 6). 824 bytes result sent to driver
19/10/20 13:44:00 INFO Executor: Finished task 5.0 in stage 0.0 (TID 5). 824 bytes result sent to driver
19/10/20 13:44:00 INFO TaskSetManager: Finished task 5.0 in stage 0.0 (TID 5) in 77 ms on localhost (executor driver) (6/10)
19/10/20 13:44:00 INFO TaskSetManager: Finished task 6.0 in stage 0.0 (TID 6) in 75 ms on localhost (executor driver) (7/10)
19/10/20 13:44:00 INFO Executor: Finished task 8.0 in stage 0.0 (TID 8). 824 bytes result sent to driver
19/10/20 13:44:00 INFO TaskSetManager: Finished task 8.0 in stage 0.0 (TID 8) in 73 ms on localhost (executor driver) (8/10)
19/10/20 13:44:00 INFO Executor: Finished task 9.0 in stage 0.0 (TID 9). 824 bytes result sent to driver
19/10/20 13:44:00 INFO TaskSetManager: Finished task 9.0 in stage 0.0 (TID 9) in 76 ms on localhost (executor driver) (9/10)
19/10/20 13:44:00 INFO Executor: Finished task 7.0 in stage 0.0 (TID 7). 824 bytes result sent to driver
19/10/20 13:44:00 INFO TaskSetManager: Finished task 7.0 in stage 0.0 (TID 7) in 84 ms on localhost (executor driver) (10/10)
19/10/20 13:44:00 INFO TaskSchedulerImpl: Removed TaskSet 0.0, whose tasks have all completed, from pool
19/10/20 13:44:00 INFO DAGScheduler: ResultStage 0 (reduce at SparkPi.scala:38) finished in 0.758 s
19/10/20 13:44:00 INFO DAGScheduler: Job 0 finished: reduce at SparkPi.scala:38, took 0.834029 s
Pi is roughly 3.141779141779142
19/10/20 13:44:00 INFO SparkUI: Stopped Spark web UI at http://192.168.1.136:4040
19/10/20 13:44:00 INFO MapOutputTrackerMasterEndpoint: MapOutputTrackerMasterEndpoint stopped!
19/10/20 13:44:00 INFO MemoryStore: MemoryStore cleared
19/10/20 13:44:00 INFO BlockManager: BlockManager stopped
19/10/20 13:44:00 INFO BlockManagerMaster: BlockManagerMaster stopped
19/10/20 13:44:00 INFO OutputCommitCoordinator$OutputCommitCoordinatorEndpoint: OutputCommitCoordinator stopped!
19/10/20 13:44:00 INFO SparkContext: Successfully stopped SparkContext
19/10/20 13:44:00 INFO ShutdownHookManager: Shutdown hook called
19/10/20 13:44:00 INFO ShutdownHookManager: Deleting directory /tmp/spark-b01ef10d-2a57-4f6b-bb6f-c314444e40fd
19/10/20 13:44:00 INFO ShutdownHookManager: Deleting directory /tmp/spark-1e2bb37f-b14f-410d-9402-98b3c06edcc3
[root@localhost spark-2.4.4-bin-hadoop2.7]#
```




Group Reflection

- Do you feel you are on track?
 - Yes, lots of progress has been made over the break. We feel we are on track to have the main application completed and ready for more testing/documentation before the start of the reading week.
- Do you feel there are barriers to your success (if any)?
 - Architecture Convolutions (DB's)
 - Enterprise grade hardware requirements – Execution units/RAM
- Do you need any help going forward?
 - Nope
- Any other questions or concerns?
 - Nope