Experiment

I did the following Spark experiment(s) on k8s and yarn.

- Launch Spark pi with 4 to 5 executors
- Wait till the job launches and driver starts processing data from executors
- Kill executors one at a time till there are no executors (at least temporarily in Yarn).
- See if Spark scheduler and driver recover from executor failures

Kubernetes

Summary

On killing executors, Driver was making progress but seems like it doesn't recognise that executors have failed. At least I don't see anything in the logs which confirms that. Driver manages with reduced number of executors until there are none and then it is stuck. Driver pod is still running but its logs are not moving and the following are last few log lines.

```
2017-02-15 13:42:02 INFO  TaskSetManager:54 - Finished task 105399.0 in stage 0.0 (TID 105402) in 5 ms on
spark-pi-1487165634566-exec-1 (executor 1) (105400/1000000)
2017-02-15 13:42:02 INFO  TaskSetManager:54 - Starting task 105401.0 in stage 0.0 (TID 105404,
spark-pi-1487165634566-exec-1, executor 1, partition 105401, PROCESS_LOCAL, 7544 bytes)
2017-02-15 13:44:36 WARN  HeartbeatReceiver:66 - Removing executor 1 with no recent heartbeats: 148604 ms
exceeds timeout 120000 ms
2017-02-15 13:44:36 ERROR TaskSchedulerImpl:70 - Lost an executor 1 (already removed): Executor heartbeat
timed out after 148604 ms
2017-02-15 13:44:36 INFO  KubernetesClusterSchedulerBackend:54 - Requesting to kill executor(s) 1
```

```
2017-02-15 13:44:36 WARN  KubernetesClusterSchedulerBackend:66 - Executor to kill 1 does not exist!
2017-02-15 13:44:36 INFO  KubernetesClusterSchedulerBackend:54 - Actual list of executor(s) to be killed
```

No indication in driver logs that driver was able to identify that it lost executors.

```
└─[1] <> cat /tmp/driver-log | grep -i lost
2017-02-15 13:44:36 ERROR TaskSchedulerImpl:70 - Lost an executor 1 (already removed): Executor heartbeat
timed out after 148604 ms
```

Spark scheduler is not actively trying to create more executor pods for driver to continue. Job is stuck. You can also see that the driver pod is running for an hour and will potentially forever.

```
└─[0] <> kubectl get pods | grep spark-pi
spark-pi-1487165634566                1/1       Running            0          1h
```

Indication that driver was able to talk to multiple executors at some point.

```
─[vkatta@Varuns-MacBook-Pro] - [~] - [2017-02-15 06:38:35]
└─[0] <> cat /tmp/1 | grep -i "executor 2" | tail -3
2017-02-15 13:38:26 INFO  TaskSetManager:54 - Starting task 51695.0 in stage 0.0 (TID 51697,
spark-pi-1487165634566-exec-2, executor 2, partition 51695, PROCESS_LOCAL, 7544 bytes)
2017-02-15 13:38:26 INFO  TaskSetManager:54 - Finished task 51693.0 in stage 0.0 (TID 51695) in 16 ms on
spark-pi-1487165634566-exec-2 (executor 2) (51693/1000000)
2017-02-15 13:38:26 INFO  TaskSetManager:54 - Starting task 51696.0 in stage 0.0 (TID 51698,
spark-pi-1487165634566-exec-2, executor 2, partition 51696, PROCESS_LOCAL, 7544 bytes)
┌─[vkatta@Varuns-MacBook-Pro] - [~] - [2017-02-15 06:40:24]
└─[0] <> cat /tmp/1 | grep -i "executor 4" | tail -3
2017-02-15 13:35:57 INFO  TaskSetManager:54 - Starting task 2255.0 in stage 0.0 (TID 2255,
spark-pi-1487165634566-exec-4, executor 4, partition 2255, PROCESS_LOCAL, 7544 bytes)
2017-02-15 13:35:57 INFO  TaskSetManager:54 - Finished task 2254.0 in stage 0.0 (TID 2254) in 17 ms on
spark-pi-1487165634566-exec-4 (executor 4) (2254/1000000)
2017-02-15 13:35:57 INFO  TaskSetManager:54 - Starting task 2257.0 in stage 0.0 (TID 2257,
spark-pi-1487165634566-exec-4, executor 4, partition 2257, PROCESS_LOCAL, 7544 bytes)
┌─[vkatta@Varuns-MacBook-Pro] - [~] - [2017-02-15 06:40:28]
└─[0] <> cat /tmp/1 | grep -i "executor 1" | tail -3
2017-02-15 13:42:02 INFO  TaskSetManager:54 - Starting task 105401.0 in stage 0.0 (TID 105404,
spark-pi-1487165634566-exec-1, executor 1, partition 105401, PROCESS_LOCAL, 7544 bytes)
2017-02-15 13:44:36 WARN  HeartbeatReceiver:66 - Removing executor 1 with no recent heartbeats: 148604 ms
exceeds timeout 120000 ms
2017-02-15 13:44:36 ERROR TaskSchedulerImpl:70 - Lost an executor 1 (already removed): Executor heartbeat
timed out after 148604 m
```

Launch command

```
bin/spark-submit \
  --deploy-mode cluster \
  --class org.apache.spark.examples.SparkPi \
  --master k8s://https://192.168.6.154:6443 \
  --kubernetes-namespace default \
  --conf spark.executor.instances=5 \
  --conf spark.app.name=spark-pi \
  --conf spark.kubernetes.driver.docker.image=docker:5000/spark-driver:varun_2_14 \
  --conf spark.kubernetes.executor.docker.image=docker:5000/spark-executor:varun_2_14 \
  examples/target/original-spark-examples_2.11-2.2.0-SNAPSHOT.jar 1000000
```

Killing executors

Executors were killed by killing the pods hosting them. Example

```
kubectl delete pod spark-pi-1487165634566-exec-5
```

Yarn

Summary

Yarn Spark job was able to recover from executor failures by identifying them as they happened and launching new executors as needed.

Job ultimately succeeded

```
17/02/15 06:08:06 DEBUG Client:
         client token: N/A
         diagnostics: N/A
         ApplicationMaster host: 192.168.6.169
         ApplicationMaster RPC port: 0
         queue: root.root
         start time: 1487166762933
         final status: SUCCEEDED
         tracking URL:
http://cloudera-mgr-testing-2-n1.pepperdata.com:8088/proxy/application_1486773628179_0010/history/applicat
ion_1486773628179_0010/2
         user: root
17/02/15 06:08:06 DEBUG Client: The ping interval is 60000 ms.
17/02/15 06:08:06 DEBUG Client: Connecting to cloudera-mgr-testing-2-n1.pepperdata.com/192.168.6.169:8020
17/02/15 06:08:06 DEBUG Client: IPC Client (922180374) connection to
cloudera-mgr-testing-2-n1.pepperdata.com/192.168.6.169:8020 from root: starting, having connections 2
17/02/15 06:08:06 DEBUG Client: IPC Client (922180374) connection to
cloudera-mgr-testing-2-n1.pepperdata.com/192.168.6.169:8020 from root sending #956
17/02/15 06:08:06 DEBUG Client: IPC Client (922180374) connection to
cloudera-mgr-testing-2-n1.pepperdata.com/192.168.6.169:8020 from root got value #956
17/02/15 06:08:06 DEBUG ProtobufRpcEngine: Call: getFileInfo took 2ms
17/02/15 06:08:06 INFO ShutdownHookManager: Shutdown hook called
17/02/15 06:08:06 INFO ShutdownHookManager: Deleting directory
/tmp/spark-bcb3afb9-8372-438b-a59e-0d23c67d1049
17/02/15 06:08:06 DEBUG Client: stopping client from cache: org.apache.hadoop.ipc.Client@2b0d0ce5
```

Yarn driver totally registered 10 different executors in its lifetime

```
cat application_1486773628179_0010/container_1486773628179_0010_01_000001/stderr | grep -i info | grep -i
'registered executor'
17/02/15 05:52:51 INFO YarnClusterSchedulerBackend: Registered executor:
AkkaRpcEndpointRef(Actor[akka.tcp://sparkExecutor@cloudera-mgr-testing-2-n3.pepperdata.com:38613/user/Exec
utor#-75079307]) with ID 2
```

```
17/02/15 05:52:51 INFO YarnClusterSchedulerBackend: Registered executor:
AkkaRpcEndpointRef(Actor[akka.tcp://sparkExecutor@cloudera-mgr-testing-2-n2.pepperdata.com:43753/user/Exec
utor#-716937683]) with ID 1
17/02/15 05:54:59 INFO YarnClusterSchedulerBackend: Registered executor:
AkkaRpcEndpointRef(Actor[akka.tcp://sparkExecutor@cloudera-mgr-testing-2-n2.pepperdata.com:40566/user/Exec
utor#1248002028]) with ID 3
17/02/15 05:55:51 INFO YarnClusterSchedulerBackend: Registered executor:
AkkaRpcEndpointRef(Actor[akka.tcp://sparkExecutor@cloudera-mgr-testing-2-n2.pepperdata.com:38402/user/Exec
utor#466822047]) with ID 4
17/02/15 05:56:06 INFO YarnClusterSchedulerBackend: Registered executor:
AkkaRpcEndpointRef(Actor[akka.tcp://sparkExecutor@cloudera-mgr-testing-2-n2.pepperdata.com:32828/user/Exec
utor#-1665316825]) with ID 5
17/02/15 05:57:09 INFO YarnClusterSchedulerBackend: Registered executor:
AkkaRpcEndpointRef(Actor[akka.tcp://sparkExecutor@cloudera-mgr-testing-2-n2.pepperdata.com:42201/user/Exec
utor#1194943407]) with ID 6
17/02/15 05:57:21 INFO YarnClusterSchedulerBackend: Registered executor:
AkkaRpcEndpointRef(Actor[akka.tcp://sparkExecutor@cloudera-mgr-testing-2-n2.pepperdata.com:44862/user/Exec
utor#1269185194]) with ID 7
17/02/15 05:57:39 INFO YarnClusterSchedulerBackend: Registered executor:
AkkaRpcEndpointRef(Actor[akka.tcp://sparkExecutor@cloudera-mgr-testing-2-n2.pepperdata.com:35885/user/Exec
utor#896046475]) with ID 8
17/02/15 05:57:51 INFO YarnClusterSchedulerBackend: Registered executor:
AkkaRpcEndpointRef(Actor[akka.tcp://sparkExecutor@cloudera-mgr-testing-2-n3.pepperdata.com:39422/user/Exec
utor#-1224382930]) with ID 9
17/02/15 05:57:57 INFO YarnClusterSchedulerBackend: Registered executor:
AkkaRpcEndpointRef(Actor[akka.tcp://sparkExecutor@cloudera-mgr-testing-2-n3.pepperdata.com:39968/user/Exec
utor#1502651287]) with ID 10
```

Yarn driver seems to recognise that it lost executors

```
# root @ cloudera-mgr-testing-2-n1 in /var/log/hadoop-yarn/container [6:49:55] C:1
$ cat application_1486773628179_0010/container_1486773628179_0010_01_000001/stderr | grep -i info  | grep
lost
17/02/15 05:54:52 INFO DAGScheduler: Executor lost: 1 (epoch 0)
17/02/15 05:55:46 INFO DAGScheduler: Executor lost: 3 (epoch 0)
17/02/15 05:56:01 INFO DAGScheduler: Executor lost: 4 (epoch 0)
17/02/15 05:57:05 INFO DAGScheduler: Executor lost: 5 (epoch 0)
17/02/15 05:57:17 INFO DAGScheduler: Executor lost: 6 (epoch 0)
17/02/15 05:57:34 INFO DAGScheduler: Executor lost: 7 (epoch 0)
17/02/15 05:57:47 INFO DAGScheduler: Executor lost: 2 (epoch 0)
17/02/15 05:57:52 INFO DAGScheduler: Executor lost: 9 (epoch 0)
```

Spark scheduler was able to launch new executors through Yarn RM as it kept losing executors as observed in the logs.

Launch Command

spark-submit --files /etc/spark/conf/log4j.properties --class org.apache.spark.examples.SparkPi --deploy-mode cluster --num-executors 4 --master yarn /opt/cloudera/parcels/CDH-5.5.6-1.cdh5.5.6.p0.2/lib/spark/lib/spark-examples.jar 100000

Killing Executors

Executors were killed by killing the yarn containers hosting them. Example.
Sudo kill <container-pid>