Hive Related Trobleshoots.

1@.

Unable to see new database under Hive with Sentry enabled

[Options](https://community.cloudera.com/t5/CDH-Manual-Installation/Unable-to-see-new-database-under-Hive-with-Sentry-enabled/m-p/48122)

‎11-29-2016 06:13 AM

I have enabled Sentry and Kerberos on CDH 5.9. I have created an admin role and grant it to the admin group.My admin user is able to manage the Hive server.

However the normal user account can't see the new database I created.

I have run steps as below:

1. Create a new database call marketing

2. Create a new role call marketing-senior

3. Grant all priviliges on the database marketing to the role marketing-senior

4. Grant role marketing-senior to one AD group call marketing-senior as well.

5. Use one normal user account log into Hue. My normal user account is in the group marketing-senior.

6. The normal user account can't see database marketing.

6. Ran statement 'use marketing', it gave an error message as below

"Error while compiling statement: FAILED: SemanticException No valid privileges User hduser1 does not have privileges for SWITCHDATABASE The required privileges: Server=server1->Db=\*->Table=+->Column=\*->action=select;Server=server1->Db=\*->Table=+->Column=\*->action=insert;"

## **Re: Unable to see new database under Hive with Sentry enabled**

[Options](https://community.cloudera.com/t5/CDH-Manual-Installation/Unable-to-see-new-database-under-Hive-with-Sentry-enabled/m-p/48122)

‎11-29-2016 06:32 AM

Sorry guys. The user saw the database after sometime. Maybe the Sentry needs sometime to fresh the permission settings.

# 2@

# OutOfMemoryError is very frequent on Hive Metastore server

[Options](https://community.cloudera.com/t5/CDH-Manual-Installation/OutOfMemoryError-is-very-frequent-on-Hive-Metastore-server/m-p/40444/highlight/true)

‎05-04-2016 02:50 AM

 Hi All,

We started getting very frequently OutofMemory error for **Hive** Metastore database. Can you please let us know what could be the cause of this?

Exception in thread "pool-1-thread-145" java.lang.OutOfMemoryError: Java heap space  
at java.nio.ByteBuffer.wrap(ByteBuffer.java:350)  
at java.lang.StringCoding$StringDecoder.decode(StringCoding.java:137)  
at java.lang.StringCoding.decode(StringCoding.java:173)  
at java.lang.String.<init>(String.java:443)  
at java.lang.String.<init>(String.java:515)  
at org.apache.thrift.protocol.TBinaryProtocol.readStringBody(TBinaryProtocol.java:355)  
at org.apache.thrift.protocol.TBinaryProtocol.readString(TBinaryProtocol.java:347)  
at org.apache.hadoop.hive.metastore.api.FieldSchema$FieldSchemaStandardScheme.read(FieldSchema.java:490)  
at org.apache.hadoop.hive.metastore.api.FieldSchema$FieldSchemaStandardScheme.read(FieldSchema.java:476)  
at org.apache.hadoop.hive.metastore.api.FieldSchema.read(FieldSchema.java:410)  
at org.apache.hadoop.hive.metastore.api.StorageDescriptor$StorageDescriptorStandardScheme.read(StorageDescriptor.java:1309)  
at org.apache.hadoop.hive.metastore.api.StorageDescriptor$StorageDescriptorStandardScheme.read(StorageDescriptor.java:1288)  
at org.apache.hadoop.hive.metastore.api.StorageDescriptor.read(StorageDescriptor.java:1150)  
at org.apache.hadoop.hive.metastore.api.Table$TableStandardScheme.read(Table.java:1393)  
at org.apache.hadoop.hive.metastore.api.Table$TableStandardScheme.read(Table.java:1330)  
at org.apache.hadoop.hive.metastore.api.Table.read(Table.java:1186)  
at org.apache.hadoop.hive.metastore.api.ThriftHiveMetastore$create\_table\_args$create\_table\_argsStandardScheme.read(ThriftHiveMetastore.java:19529)  
at org.apache.hadoop.hive.metastore.api.ThriftHiveMetastore$create\_table\_args$create\_table\_argsStandardScheme.read(ThriftHiveMetastore.java:19514)  
at org.apache.hadoop.hive.metastore.api.ThriftHiveMetastore$create\_table\_args.read(ThriftHiveMetastore.java:19461)  
at org.apache.thrift.ProcessFunction.process(ProcessFunction.java:25)  
at org.apache.hadoop.hive.metastore.TUGIBasedProcessor.process(TUGIBasedProcessor.java:109)  
at org.apache.thrift.server.TThreadPoolServer$WorkerProcess.run(TThreadPoolServer.java:244)  
at java.util.concurrent.ThreadPoolExecutor$Worker.runTask(ThreadPoolExecutor.java:886)  
at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:908)  
at java.lang.Thread.run(Thread.java:662)  
Exception in thread "main" java.lang.OutOfMemoryError: Java heap space  
at org.apache.thrift.transport.TServerSocket.acceptImpl(TServerSocket.java:114)  
at org.apache.hadoop.hive.metastore.TServerSocketKeepAlive.acceptImpl(TServerSocketKeepAlive.java:39)  
at org.apache.hadoop.hive.metastore.TServerSocketKeepAlive.acceptImpl(TServerSocketKeepAlive.java:32)  
at org.apache.thrift.transport.TServerTransport.accept(TServerTransport.java:31)  
at org.apache.thrift.server.TThreadPoolServer.serve(TThreadPoolServer.java:131)  
at org.apache.hadoop.hive.metastore.HiveMetaStore.startMetaStore(HiveMetaStore.java:4245)  
at org.apache.hadoop.hive.metastore.HiveMetaStore.main(HiveMetaStore.java:4147)  
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)  
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:39)  
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:25)  
at java.lang.reflect.Method.invoke(Method.java:597)  
at org.apache.hadoop.util.RunJar.main(RunJar.java:208)

Regards,

Ajay

## **Re: OutOfMemoryError is very frequent on Hive Metastore server**

[Options](https://community.cloudera.com/t5/CDH-Manual-Installation/OutOfMemoryError-is-very-frequent-on-Hive-Metastore-server/m-p/40444/highlight/true)

‎05-13-2016 12:34 AM

Hi All,

Thanks for your help. Heap memory was not sized as per recommendation given in below link and we had increased the memory and restarted the Hive metastore server which also did not help.

<http://www.cloudera.com/documentation/enterprise/latest/topics/cdh_ig_hiveserver2_configure.html>

Looks like there were some process which was holding/blocking memory and we had to restart the Complete cluster to resolve this problem.

Thank you once again for input.

Regards,

Ajay chaudhary

**Java Heap Size of Hive Metastore Server in Bytes**

Hive Metastore Server Default Group

 

**Java Heap Size of HiveServer2 in Bytes**

HiveServer2 Default Group

 

@3. Invalid Url problems?

# Invalid URL: jdbc:hive2://guna:10000/default - Hive beeline Error

[Options](https://community.cloudera.com/t5/Cloudera-Manager-Installation/Invalid-URL-jdbc-hive2-guna-10000-default-Hive-beeline-Error/m-p/36947/highlight/true)

‎02-01-2016 11:13 PM

Below are the steps are been performed

1. /etc/**hive**/conf/**hive**-site.xml

<configuration>

<property>

<name>javax.jdo.option.ConnectionURL</name>

<value>jdbc:mysql://guna:3306/metastore</value>

</property>

<property>

<name>javax.jdo.option.ConnectionDriverName</name>

<value>com.mysql.jdbc.Driver</value>

</property>

<property>

<name>javax.jdo.option.ConnectionUserName</name>

<value>hiveuser</value>

</property>

<property>

<name>javax.jdo.option.ConnectionPassword</name>

<value>password</value>

</property>

<property>

<name>datanucleus.autoCreateSchema</name>

<value>false</value>

</property>

<property>

<name>datanucleus.fixedDatastore</name>

<value>true</value>

</property>

<property>

<name>datanucleus.autoStartMechanism</name>

<value>SchemaTable</value>

</property>

<property>

<name>hive.metastore.uris</name>

<value>thrift://192.168.123.1:9083</value>

</property>

<property>

<name>hive.support.concurrency</name>

<value>true</value>

</property>

<property>

<name>hive.zookeeper.quorum</name>

<value>guna,tigerZ,hory</value>

</property>

</configuration>

2. Created  the **Hive** warehouse directory in HDFS.

3. Made the MySQL connector jar file accessible to **Hive**

4. Initialized the **Hive** Metastore in MySQL database

5. sudo service **hive**-metastore start

checked with sudo jps - able to see Runjar

6. sudo service hive-server2 start

When I execute the below beeline in my guna teriminal I am getting the below error  please help me

**beeline -u jdbc:hive2://elephant:10000/default \**  
**-n username**

beeline -u jdbc:hive2://elephant:10000/default -n username

scan complete in 3ms

Connecting to jdbc:hive2://guna:10000/default

Error: Invalid URL: jdbc:hive2://guna:10000/default (state=08S01,code=0)

Beeline version 0.12.0-cdh5.0.0 by Apache **Hive**

0: jdbc:hive2://guna:10000/default>

Cloudera manager Agent Errors:

@1.

Cloudera Agent shut down

Once what happen 1 of scm-agent of our host went down. We were unable to identify the problem since particular service on host was working fine. But It was showing critical status in our cloudera manager.

Soln>

In that case I went to cli of that host and started the service of scm-agent.

Sudo service cloudera-scm-agent start

# Hue logon is slow or hung

[Options](https://community.cloudera.com/t5/Web-UI-Hue-Beeswax/Hue-logon-is-slow-or-hung/m-p/4159/highlight/true)

‎12-19-2013 09:43 PM

Hello Team,

        Logons to Hue  is **hung**. We are using CDH4.8.

I had to restart it to fix this. But that is not long term solution as we will have many users accessing it going forward.

Thanks

Bhupal

## **Re: Hue logon is slow or hung**

[Options](https://community.cloudera.com/t5/Web-UI-Hue-Beeswax/Hue-logon-is-slow-or-hung/m-p/4159/highlight/true)

‎12-19-2013 10:36 PM

So this could be related to <https://issues.cloudera.org/browse/HUE-1367> (and in particular Beeswax or Impala been down).

Restarting Hue will restart Beeswax so it is why it might have fixed it. Since Hue 3, the UX has been improved and Beeswax has been replaced by HiveServer2 so it will be better.

This is in C5beta1.

Yarn Application:

Accepted Solution

# Yarn applications hang foreever if run in parallel

[Options](https://community.cloudera.com/t5/Batch-Processing-and-Workflow/Yarn-applications-hang-foreever-if-run-in-parallel/m-p/15184/highlight/true)

‎07-16-2014 08:47 AM

Hi,

we have a cluster with 8 nodes on CDH5 (5.0.2) with Yarn MRv2 in use and a big problem which is probably due to the Config.

In addition to Hadoop, we also use Imapala so we can not use give all ressoures to yarn.

Each of our nodes have 128GB of RAM and 12 cores.

Currently sees the Memory config for Yarn as follows:

mapreduce.map.memory.mb = 8Gib

mapreduce.reduce.memory.mb = 8Gib

yarn.app.mapreduce.am.resource.mb = 8Gib

mapreduce.map.java.opts.max.heap = 6960MiB

mapreduce.reduce.java.opts.max.heap = 6960MiB

"Java Heap Size in bytes of NodeManager" = 8Gib

yarn.nodemanager.resource.memory-mb = 80Gib

Now we get the problem that if we run multiple applications in parallel all stop and no one finished.

it looks as if they hang forever. I see no exception or errors in "/var/log/hadoop-yarn" (Debug Log Level).

I would be glad if someone can help? :)

BG

## **Re: Yarn applications hang foreever if run in parallel**

[Options](https://community.cloudera.com/t5/Batch-Processing-and-Workflow/Yarn-applications-hang-foreever-if-run-in-parallel/m-p/15184/highlight/true)

‎08-12-2014 10:19 AM

I am experiencing the same problem stated earlier.  We have a 4-node cluster using YARN on v5.1.0.  I have an Oozie workflow that uses Sqoop to import from MySQL, which is sharded with 10 tables.  Therefore, I have a coordinator that executes the same workflow with 10 simultaneous (parallel) sessions, to pull from each sharded table.

However, sometime after the workflows reach the Sqoop action step, they stop running.  The jobs are not failing, rather they stop processing, even though their status shows "Running" in the Hue workflow dashboard.  None of the jobs have had any updated status in the SysLog for more than 12 hours.

Further, if other, unrelated jobs are submitted, they also appear to hang.  I have had a job running successfully for several days, which is executing a DISTCP command to import S3 data.  This job has also hung after submitting the 10 parallel workflows.

Is there a configuration that must be set to allows the same workflow to be processed in parallel?

Thank you!

Michael Reynolds

## **Re: Yarn applications hang foreever if run in parallel**

[Options](https://community.cloudera.com/t5/Batch-Processing-and-Workflow/Yarn-applications-hang-foreever-if-run-in-parallel/m-p/15184/highlight/true)

‎08-12-2014 11:38 AM

On a small cluster, sometimes all the resources are occupied by AMs, and no real work get done. See <https://issues.apache.org/jira/browse/YARN-1913.> One workaround is to configure the `maxRunningApps' to a smaller number. See <http://hadoop.apache.org/docs/r2.4.1/hadoop-yarn/hadoop-yarn-site/FairScheduler.html.>

# YARN apps stuck, won't allocate resources

[Options](https://community.cloudera.com/t5/Batch-Processing-and-Workflow/YARN-apps-stuck-won-t-allocate-resources/m-p/23224/highlight/true)

‎12-31-2014 06:03 PM

CDH 5.2.0-1.cdh5.2.0.p0.36

We had an issue with HDFS filling up causing a number of services to fail and after we cleared space and restarted the cluster we aren't able to run any hive workflows through oozie.  It seems to get stuck **allocating**resources.

No changes were made to YARN resource configurations which seems to be the goto for troubleshooting steps.  We have plenty of resources **allocated** to YARN **containers** and there is currently no app limits set in dynamic pool resources.

When I start an oozie workflow the oozie:launcher application starts normally but the hive query that is executed is always stuck in ACCEPTED state and never transitions to RUNNING.

The oozie:launcher application is accepted and scheduled.

2015-01-01 00:47:48,472 INFO org.apache.hadoop.yarn.server.resourcemanager.scheduler.fair.FairScheduler: Accepted application application\_1420073214126\_0001 from user: admin, in queue: default, currently num of applications: 1  
2015-01-01 00:47:48,475 INFO org.apache.hadoop.yarn.server.resourcemanager.rmapp.RMAppImpl: application\_1420073214126\_0001 State change from SUBMITTED to ACCEPTED  
2015-01-01 00:47:48,475 INFO org.apache.hadoop.yarn.server.resourcemanager.ApplicationMasterService: Registering app attempt : appattempt\_1420073214126\_0001\_000001  
2015-01-01 00:47:48,476 INFO org.apache.hadoop.yarn.server.resourcemanager.rmapp.attempt.RMAppAttemptImpl: appattempt\_1420073214126\_0001\_000001 State change from NEW to SUBMITTED  
2015-01-01 00:47:48,490 INFO org.apache.hadoop.yarn.server.resourcemanager.scheduler.fair.FairScheduler: Added Application Attempt appattempt\_1420073214126\_0001\_000001 to scheduler from user: admin  
2015-01-01 00:47:48,492 INFO org.apache.hadoop.yarn.server.resourcemanager.rmapp.attempt.RMAppAttemptImpl: appattempt\_1420073214126\_0001\_000001 State change from SUBMITTED to SCHEDULED

oozie:launcher **container** is **allocated** and acquired

2015-01-01 00:47:54,514 INFO org.apache.hadoop.yarn.server.resourcemanager.rmcontainer.RMContainerImpl: container\_1420073214126\_0001\_01\_000001 **Container** Transitioned from NEW to **ALLOCATED**  
2015-01-01 00:47:54,514 INFO org.apache.hadoop.yarn.server.resourcemanager.RMAuditLogger: USER=admin OPERATION=AM **Allocated** **Container** TARGET=SchedulerApp RESULT=SUCCESS APPID=application\_1420073214126\_0001 CONTAINERID=container\_1420073214126\_0001\_01\_000001  
2015-01-01 00:47:54,514 INFO org.apache.hadoop.yarn.server.resourcemanager.scheduler.SchedulerNode: Assigned **container** container\_1420073214126\_0001\_01\_000001 of capacity <memory:1024, vCores:1> on host node:8041, which has 1 **containers**, <memory:1024, vCores:1> used and <memory:23552, vCores:11> available after **allocation**  
2015-01-01 00:47:54,516 INFO org.apache.hadoop.yarn.server.resourcemanager.security.NMTokenSecretManagerInRM: Sending NMToken for nodeId : ascn07.idc1.level3.com:8041 for **container** : container\_1420073214126\_0001\_01\_000001  
2015-01-01 00:47:54,520 INFO org.apache.hadoop.yarn.server.resourcemanager.rmcontainer.RMContainerImpl: container\_1420073214126\_0001\_01\_000001 **Container** Transitioned from **ALLOCATED** to ACQUIRED

oozie:launcher application is **allocated**, launched, and starts running

2015-01-01 00:47:54,559 INFO org.apache.hadoop.yarn.server.resourcemanager.rmapp.attempt.RMAppAttemptImpl: appattempt\_1420073214126\_0001\_000001 State change from SCHEDULED to **ALLOCATED**\_SAVING  
2015-01-01 00:47:54,568 INFO org.apache.hadoop.yarn.server.resourcemanager.rmapp.attempt.RMAppAttemptImpl: appattempt\_1420073214126\_0001\_000001 State change from **ALLOCATED**\_SAVING to **ALLOCATED**  
2015-01-01 00:47:54,575 INFO org.apache.hadoop.yarn.server.resourcemanager.amlauncher.AMLauncher: Launching masterappattempt\_1420073214126\_0001\_000001

<snip>

2015-01-01 00:47:54,834 INFO org.apache.hadoop.yarn.server.resourcemanager.rmapp.attempt.RMAppAttemptImpl: appattempt\_1420073214126\_0001\_000001 State change from **ALLOCATED** to LAUNCHED

2015-01-01 00:47:55,094 INFO org.apache.hadoop.yarn.server.resourcemanager.rmcontainer.RMContainerImpl: container\_1420073214126\_0001\_01\_000001 **Container** Transitioned from ACQUIRED to RUNNING  
2015-01-01 00:47:59,724 INFO org.apache.hadoop.yarn.server.resourcemanager.ApplicationMasterService: AM registration appattempt\_1420073214126\_0001\_000001  
2015-01-01 00:47:59,725 INFO org.apache.hadoop.yarn.server.resourcemanager.RMAuditLogger: USER=admin IP=1.1.1.1 OPERATION=Register App Master TARGET=ApplicationMasterService RESULT=SUCCESS APPID=application\_1420073214126\_0001 APPATTEMPTID=appattempt\_1420073214126\_0001\_000001  
2015-01-01 00:47:59,725 INFO org.apache.hadoop.yarn.server.resourcemanager.rmapp.attempt.RMAppAttemptImpl: appattempt\_1420073214126\_0001\_000001 State change from LAUNCHED to RUNNING

Then the next job begins, which is a hive job.  It transitions from new -> scheduled but a new **container** is never created/**allocated**.

2015-01-01 00:48:14,119 INFO org.apache.hadoop.yarn.server.resourcemanager.ClientRMService: Application with id 2 submitted by user admin  
2015-01-01 00:48:14,119 INFO org.apache.hadoop.yarn.server.resourcemanager.rmapp.RMAppImpl: Storing application with id application\_1420073214126\_0002  
2015-01-01 00:48:14,119 INFO org.apache.hadoop.yarn.server.resourcemanager.RMAuditLogger: USER=admin IP=1.1.1.1 OPERATION=Submit Application Request TARGET=ClientRMService RESULT=SUCCESS APPID=application\_1420073214126\_0002  
2015-01-01 00:48:14,120 INFO org.apache.hadoop.yarn.server.resourcemanager.rmapp.RMAppImpl: application\_1420073214126\_0002 State change from NEW to NEW\_SAVING  
2015-01-01 00:48:14,120 INFO org.apache.hadoop.yarn.server.resourcemanager.recovery.RMStateStore: Storing info for app: application\_1420073214126\_0002  
2015-01-01 00:48:14,120 INFO org.apache.hadoop.yarn.server.resourcemanager.rmapp.RMAppImpl: application\_1420073214126\_0002 State change from NEW\_SAVING to SUBMITTED  
2015-01-01 00:48:14,120 WARN org.apache.hadoop.security.UserGroupInformation: No groups available for user admin  
2015-01-01 00:48:14,120 INFO org.apache.hadoop.yarn.server.resourcemanager.scheduler.fair.FairScheduler: Accepted application application\_1420073214126\_0002 from user: admin, in queue: default, currently num of applications: 2  
2015-01-01 00:48:14,121 INFO org.apache.hadoop.yarn.server.resourcemanager.rmapp.RMAppImpl: application\_1420073214126\_0002 State change from SUBMITTED to ACCEPTED  
2015-01-01 00:48:14,121 INFO org.apache.hadoop.yarn.server.resourcemanager.ApplicationMasterService: Registering app attempt : appattempt\_1420073214126\_0002\_000001  
2015-01-01 00:48:14,121 INFO org.apache.hadoop.yarn.server.resourcemanager.rmapp.attempt.RMAppAttemptImpl: appattempt\_1420073214126\_0002\_000001 State change from NEW to SUBMITTED  
2015-01-01 00:48:14,121 INFO org.apache.hadoop.yarn.server.resourcemanager.scheduler.fair.FairScheduler: Added Application Attempt appattempt\_1420073214126\_0002\_000001 to scheduler from user: admin  
2015-01-01 00:48:14,121 INFO org.apache.hadoop.yarn.server.resourcemanager.rmapp.attempt.RMAppAttemptImpl: appattempt\_1420073214126\_0002\_000001 State change from SUBMITTED to SCHEDULED

At this point the job never progresses.  In cm->yarn applications it has a status of "Pending", on the resource manager UI it has a state of "ACCEPTED" but never transitions into "RUNNING".

This issue is mentioned in a blog post from april (#5) <http://blog.cloudera.com/blog/2014/04/apache-hadoop-yarn-avoiding-6-time-consuming-gotchas/>

The suggested fix of adding a value to "max running apps" has no effect

## **Re: YARN apps stuck, won't allocate resources**

[Options](https://community.cloudera.com/t5/Batch-Processing-and-Workflow/YARN-apps-stuck-won-t-allocate-resources/m-p/23224/highlight/true)

‎01-09-2015 02:05 PM

User error.

Everything was fine with the resource pools, but there was a default user limit set.

## **Re: YARN apps stuck, won't allocate resources**

[Options](https://community.cloudera.com/t5/Batch-Processing-and-Workflow/YARN-apps-stuck-won-t-allocate-resources/m-p/23224/highlight/true)

‎07-07-2015 03:41 PM

In our case I had accidently set a default "user limits" to 1 for "max running apps per user".  All of our jobs required more than one application to run at a time per user.

This is configured in Clusters -> Dynamic resource pools -> Configuration -> User limits -> Default settings

It could also be that your jobs are attempting to wait for resources to become available before starting.  Perhaps you have too few resources available for what is being requested?