****

**Evaluation Review**

UNIVERSITY OF MARYLAND CLUSTER| (Prod) | Nov 2017

# 

# T A B L E O F C O N T E N T S

Section 1.0 Overview 2

Section 2.0 Configuration check 4

Section 3.0 Host Inspector check 7

Section 4.0 Security Inspector check 8

Section 5.0 Health Issues check 9

Section 6.0 Configuration Issue check 11

Section 7.0 Database Settings check 12

Section 8.0 Non-Default Values check 13

Section 9.0 Resource management check 14

Section 10.0 YARN parameters check

Section 10.0 Evaluation Script Analysis 15

# SECTION 1.0 Overview

Clairvoyant team performs initial review of Hadoop cluster to verify various configurations are following best practices and scan for existence of known issues. The aim for initial review process is as follows:

1. Verifying the configuration of operating system, network, and hardware configuration.
2. Review the Hadoop deployment strategy.
3. Review cluster configuration and provide any recommendations for any changes.
4. Review and provide recommendations for security on the cluster. (if applicable)

This cluster review document describes our initial system architecture findings and related recommendations.

# Summary:

# Review from evaluation script:

1. vm.swappiness is inconsistently set across machines in the cluster. Recommended value is 1.
2. Available Entropy is below 500 and it is recommended to be above 1000. Installation of the RNGd can achieve this.
3. JDBC packages are inconsistent across cluster. JDBC package is not installed and PATH not set for JDBC connector.
4. DNS entries for host A and PTR records are not configured.
5. NSCD service is inactive and not installed on one node. Recommended to install NSCD daemon and cache hosts lookups.
6. Transparent Huge Pages (THP) is not completely disabled. Recommendation is to disable THP across all nodes.
7. $JAVA\_HOME is not set at the OS layer across all machines in the cluster. Recommended to set $JAVA\_HOME.
8. Network Maximum Transmission Unit (MTU) is set to 1500. It can be configured to use Jumbo frames (9000) which can give better performance. This is highly dependent upon network configuration.

# Review from Cloudera manager:

1. Swap should be disabled
2. YARN tuning required for cluster
3. Impala service can be removed as it is not used.
4. New Groups can be added for three different data nodes with different data disks size.
5. Hue service - Database migration from SQLite to Mysql
6. THP should be disabled.
7. Activity monitor can be removed.
8. Role groups- Data node Group1 can be removed.
9. Dynamic resource pools need to be configured for cluster.
10. Upgrade cluster to latest CM and CDH includes reviewing TSB's.

# 

# 

# SECTION 2.0 CONFIGURATION CHECK

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| # | CHECK | Overview of the Check | Existing Value | Recommendation | Other Comments |
| 1 | Cloudera Manager Version | Check Cloudera manager version  CM -> Support -> About | 5.7.1 |  |  |
| 2 | CDH Version | Check CDH version  CM -> Clusters -> select your Cluster name | CDH 5.7.1 |  |  |
| 3 | Java Version: | CM -> Search box and  Enter **Java Home Directory**  This will tell you whether JAVA HOME is overridden from Cloudera Manager or not. | /usr/java/jdk1.8.0\_92 |  |  |
| 4 | Security | Check Kerberos is enabled or disabled.  CM -> Administration -> Security | Kerberos Disabled |  |  |
| 5 | License | Check License,  CM-> Administration -> License | Not Available |  |  |

# Section 3.0 HOST INSPECTOR CHECK

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sr. | Check | Overview of the check | Existing Value | Recommendation | Other Comments |
| 1 | Host Inspector | CM-> Hosts -> All Hosts -> Inspect All Hosts | Click on Show Inspector results and check for any issues with the hosts of cluster  THP is enabled on hosts. | THP should be disabled. |  |

# Section 4.0 SECURITY INSPECTOR CHECK

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sr. | Check | Overview of the check | Existing Value | Recommendation | Other Comments |
| 1 | Host Inspector | CM-> Hosts -> Administration -> Security Inspector | Click on Show Inspector results and check for any issues with security of the cluster. | Not Applicable |  |

# Section 5.0 HEALTH ISSUES CHECK

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 13 | Check Health issues | CM -> All health issues  This check will let you know if there are any Health issues on the cluster.  e.g. Service or Hosts can be in concerning state. | No |  |  |

# 

# Section 6.0 CONFIGURATION ISSUES CHECK

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 14 | Check Configuration issues | CM -> Configuration -> Configuration Issues.  This check will let you know if there are any configuration issues on the cluster.  e.g. memory overcommit validation | No |  |  |

# 

# 

# SECTION 7.0 DATABASE SETTINGS CHECK

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sr. | Check | Overview of the check | Existing Value | Recommendation | Other Comments |
| 1 | Check  Database settings for cluster. | CM -> Configuration ->Database settings.  Select SCOPE from Filters.  HUE  Check for property, Database type | Database type = SQLite3 | MySQL |  |
| CM -> Configuration ->Database settings.  Select SCOPE from Filters.  Oozie  Check for property, Database type | Database type = MySQL |  |  |

# SECTION 8.0 NON-DEFAULT VALUES CHECK

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sr. | Check | Overview of the check | Existing Value | Recommendation | Other Comments |
| 1 | Check for Non-default Values | CM -> Configuration ->Non-default Values  Select SCOPE from Filters.  All Hosts | Enable Configuration Change Alerts  -> TRUE |  |  |
|  |  |  | JAVA\_HOME Directory ->  /usr/java/jdk1.8.0\_92 |  |  |
|  |  |  | Host memory Swapping thresholds ->  Warning:10240  Critical:40960 |  |  |
| 2 |  | CM -> Configuration ->Non-default Values  Select SCOPE from Filters.  Host | Suppress Health Test: Agent Status -> True |  |  |
| 3 |  | CM -> Configuration ->Non-default Values  Select SCOPE from Filters.  Settings  This will cover all the configuration related to Parcels,Kerberos,Active directory, TLS and custom settings if configured already. | This section has correct Parcels configuration. |  |  |
| 4 |  | CM -> Configuration ->Non-default Values  Select SCOPE from Filters.  Alert Publisher |  |  |  |
| 5 |  | CM -> Configuration ->Non-default Values  Select SCOPE from Filters.  Activity monitor | Check whether Activity monitor is available or not. | Activity Monitor service is not required as MapReduce version 1 service is not installed. | Activity monitor can be removed and it’s database deleted. |
| 6 |  | CM -> Configuration ->Non-default Values  Select SCOPE from Filters.  Host monitor | Maximum Non-Java memory for Host monitor -> 4GiB |  |  |
| 7 |  | CM -> Configuration ->Non-default Values  Select SCOPE from Filters.  Service monitor | Java Heap Size of Service monitor in Bytes -> 6 GiB  Maximum Non-Java memory for Service monitor -> 4GiB |  |  |

# 

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 8 | Check for Non-default Values for Service | CM -> Configuration ->Non-default Values  Select SCOPE from Filters.  HBASE | HBASE service Advanced configuration snippet  Name:hbase.root.dir  Value:umdcluster |  |  |
|  |  |  | Java Heap size of region server in Bytes ->10636 MiB |  |  |
|  |  |  | Host memory Swapping thresholds ->  Warning:10240  Critical:40960 |  |  |
| 9 |  | CM -> Configuration ->Non-default Values  Select SCOPE from Filters.  HDFS | dfs.ha.fencing .methods=shell(true)  Dfs.datanode.failed.volumes.tolerated = 12 and 11 for two datanode groups respectively.  Datanode.block.count.thresholds = 2000000  Java Heap size of Data nodes in Bytes = 4 GiB  Datanode Data directory permissions = 755  Suppress Health Test : Block Count -> Data node 13,14 and 15  Use Trash = True |  |  |
| 10 |  | The replication level for submitted job files. | mapreduce .client.submit.file.replication = 3 |  |  |
| 11 |  | Fraction of the number of the map tasks in the job which should be completed before reduce tasks are scheduled for the job. | Mapreduce.job.reduce.slowstart.completedmaps =1 |  |  |
|  |  |  | Mapreduce.job.reduces =288 |  |  |
|  |  |  | mapreduce.map.memory.mb=12GiB |  |  |
|  |  |  | Mapreduce.reduce.memory.mb = 12 GiB |  |  |
|  |  |  | Mapreduce.map.java.opts.max.heap = 8GiB |  |  |
|  |  |  | Mapreduce.reduce.java.opts.max.heap = 8 GiB |  |  |
|  |  |  | Java Heap size of Journalnode in Bytes = 1 GiB |  |  |
|  |  |  | dfs.namenode.handler.count=54 |  |  |
|  |  |  | Java Heap size of Namenode in Bytes = 8 GiB |  |  |
|  |  |  | Java Heap size of Secondary Namenode in Bytes = 8 GiB |  |  |
|  |  |  | Java Heap size of Balancer in Bytes = 4 GiB |  |  |
|  |  |  |  |  |  |
| 12 |  | CM -> Configuration ->Non-default Values  Select SCOPE from Filters.  HIVE | Webhdfs\_url =HttpFS |  |  |
|  |  |  | Hive metastore database type = MySql |  |  |
| 13 |  | CM -> Configuration ->Non-default Values  Select SCOPE from Filters.  Oozie | Oozie server database type = Mysql |  |  |
| 14 |  | CM -> Configuration ->Non-default Values  Select SCOPE from Filters.  YARN | YARN service advanced configuration snippet (safety valve) for Yarn-site.xml  Property:  Yarn.resourcemanager.hostname= mgrnode4.umd.edu  Yarn.nodemanager.aux-services  =mapreduce\_shuffle |  |  |
|  |  |  | Heartbeat interval =120 milliseconds |  |  |
|  |  |  | Container memory = 75000 MiB |  | yarn.nodemanager.resource.memory-mb |
|  |  |  | Container Virtual CPU cores= 44 |  | yarn.nodemanager.resource.cpu-vcores |
|  |  |  | Yarn-scheduler.maximum allocation.mb = 82342 MiB  Yarn-scheduler-maximum-allocation-vcores= 40 | YARN TUNING required. |  |
| 15 |  | CM -> Configuration ->Non-default Values  Select SCOPE from Filters.  Zookeeper | maxSessionTimeout=60000 |  |  |
| 16 | Check Non- Uniform values | CM -> Configuration -> Non-uniform values.  This check will let you know if there are any non-uniform values / configurations on the cluster.  e.g. multiple datanode role groups. | There are two role groups are present  1. Data node default group  2. Data node Group 1 | Extra role group can be deleted. |  |

# 

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 17 | Check Advanced Configuration Snippets | CM -> Configuration -> Advanced Configuration Snippets | NA |  |  |

# 

# Section 9.0 RESOURCE MANAGEMENT CHECK

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sr. | Check | Overview of the check | Existing Value | Recommendation | Other Comments |
| 1 | Static service pools | (CM->Clusters->Static Service pools) | Not configured | Should be Configured |  |
| 2 | Dynamic Service pools | (CM->Clusters->Dynamic resource pool configuration) | Not configured | Should be Configured |  |

# Section 10.0 YARN PARAMETERS CHECK

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sr. | Check | Overview of the check | Existing Value | Recommendation | Other Comments |
| 1 | yarn.nodemanager.resource.cpu-vcores | (CM->Clusters->YARN->Configurations) | 44 |  |  |
| 2 | yarn.nodemanager.resource.memory-mb | (CM->Clusters->YARN->Configurations) | 75000 |  |  |
| 4 | Job History server Heap size | (CM>YARN>instance>Jobhistoryserver>configuration) |  |  |  |
| 6 | YARN application Master Heap size | (CM>YARN>configuration) |  |  |  |
| 7 | Container Virtual CPU cores | (CM>YARN>configuration)yarn.nodemanager.resource.cpu-vcores |  |  |  |
| 8 | Application Master Memory | (CM>YARN>configuration)yarn.app.mapreduce.am.resource.mb |  |  |  |
| 9 | ApplicationMaster Java Maximum Heap Size | (CM>YARN>configuration) |  |  |  |
| 10 | Container Memory Increment | (CM>YARN>configuration)yarn.scheduler.increment-allocation-mb |  |  |  |
| 11 | Container Memory | (CM>YARN>configuration)yarn.nodemanager.resource.memory-mb |  |  |  |
| 12 | Container Memory Minimum | (CM>YARN>configuration)yarn.scheduler.minimum-allocation-mb |  |  |  |
| 13 | Map Task Memory | (CM>YARN>configuration)mapreduce.map.memory.mb |  |  |  |
| 14 | Reduce Task Memory | (CM>YARN>configuration)mapreduce.reduce.memory.mb |  |  |  |
| 15 | Map Task CPU virtual core | (CM>YARN>configuration)mapreduce.map.cpu.vcores |  |  |  |
| 16 | Reduce Task CPU Virtual Cores | (CM>YARN>configuration)mapreduce.reduce.cpu.vcores |  |  |  |
| 17 | Map Task Maximum Heap Size | (CM>YARN>configuration)mapreduce.map.java.opts.max.heap |  |  |  |
| 18 | Reduce Task Maximum Heap Size | (CM>YARN>configuration)mapreduce.reduce.java.opts.max.heap |  |  |  |
| 19 | Container usage metric | Enabled |  |  |  |
| 20 |  | (CM>YARN>configuration)yarn.app.mapreduce.am.resource.cpu-vcores |  |  |  |
| 23 |  | (CM>YARN>configuration)mapreduce.task.io.sort.mb |  |  |  |
| 24 |  | (CM>YARN>configuration)yarn.scheduler.minimum-allocation-vcores |  |  |  |
| 25 |  | (CM>YARN>configuration)yarn-scheduler.maximum-allocation-vcores |  |  |  |
| 26 |  | (CM>YARN>configuration)yarn.scheduler.increment-allocation-vcores |  |  |  |
| 27 |  | (CM>YARN>configuration)yarn.scheduler.minimum-allocation-mb |  |  |  |
| 28 |  | (CM>YARN>configuration)yarn.scheduler.maximum-allocation-mb |  |  |  |
| 29 |  | (CM>YARN>configuration)yarn.scheduler.increment-allocation-mb |  |  |  |
|  |  |  |  |  |  |

# 

# Section 10.0 EVALUATION SCRIPT ANALYSIS:

**Summery**:

1.vm.swappiness inconsistent across machines in the cluster.

2.Entropy - Available Entropy is below 500 and as per standard it should be in range of 1000 to 3000.

3.JDBC Package are inconsistent across cluster. JDBC package is not installed and PATH not set for JDBC connector.

4.DNS Lookup not configured properly.

5.NSCD service is inactive and not installed on one node. Recommended to install NSCD daemon.

6.THP is not completely disabled. Recommendation is to disable THP across all nodes.

7.JAVA\_HOME is not set across all machines in the cluster. Recommended to set JAVA\_HOME.

8.Maximum transmission unit is set to 1500, It can be configured to use (Jumbo frames) 9000 which can give better performance

CHECK 1: Swappiness is inconsistent across machines in the cluster

admin:evaluate-pre jay$ grep ^vm.swappiness \*.\*  
itadmin@192.168.5.201.out:vm.swappiness = 1  
itadmin@192.168.5.202.out:vm.swappiness = 1  
itadmin@192.168.5.203.out:vm.swappiness = 10  
itadmin@192.168.5.204.out:vm.swappiness = 1  
itadmin@192.168.5.234.out:vm.swappiness = 10  
itadmin@192.168.5.235.out:vm.swappiness = 10

CHECK 2: Available Entropy is below 500 and as per standard it should be in range of 1000 to 3000.

admin:evaluate-pre jay$ grep "\*\* available entropy:" -A1 \*.\*  
itadmin@192.168.5.201.out:\*\* available entropy:  
itadmin@192.168.5.201.out-130  
--  
itadmin@192.168.5.202.out:\*\* available entropy:  
itadmin@192.168.5.202.out-171  
--  
itadmin@192.168.5.203.out:\*\* available entropy:  
itadmin@192.168.5.203.out-150  
--

CHECK 3: JDBC Package are inconsistent across cluster. JDBC package is not installed and PATH not set for JDBC connector.

admin:evaluate-pre jay$ grep "\*\*JDBC" -A3 \*.\*  
itadmin@192.168.5.201.out:\*\*\* JDBC  
itadmin@192.168.5.201.out-package mysql-connector-java is not installed  
itadmin@192.168.5.201.out-package postgresql-jdbc is not installed  
itadmin@192.168.5.201.out-\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  
--  
itadmin@192.168.5.202.out:\*\*\* JDBC  
itadmin@192.168.5.202.out-package mysql-connector-java is not installed  
itadmin@192.168.5.202.out-package postgresql-jdbc is not installed  
itadmin@192.168.5.202.out-\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  
--

CHECK4:DNS Lookup not configured properly.

admin:evaluate-pre jay$ grep 'forward' -A4 \*.\*  
itadmin@192.168.5.201.out:\*\* forward:  
itadmin@192.168.5.201.out-Host datanode1.umd.edu not found: 3(NXDOMAIN)  
itadmin@192.168.5.201.out-\*\* reverse:  
itadmin@192.168.5.201.out-Host 3\(NXDOMAIN\) not found: 2(SERVFAIL)  
itadmin@192.168.5.201.out-\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  
--

CHECK5: NSCD service is inactive and not installed on one node. Recommended to install NSCD daemon.

admin:evaluate-pre jay$ egrep "nscd.service" \*.\* -A4 |grep active  
itadmin@192.168.5.201.out- Active: inactive (dead)  
itadmin@192.168.5.202.out- Active: inactive (dead)  
itadmin@192.168.5.203.out- Active: inactive (dead)  
itadmin@192.168.5.204.out- Active: inactive (dead)  
itadmin@192.168.5.205.out- Active: inactive (dead)  
itadmin@192.168.5.206.out- Active: inactive (dead)  
itadmin@192.168.5.207.out- Active: inactive (dead)  
itadmin@192.168.5.208.out- Active: inactive (dead)

CHECK6: THP is not completely disabled. Recommendation is to disable THP across all nodes.

itadmin@192.168.5.201.out:[always] madvise never  
itadmin@192.168.5.201.out:always madvise [never]  
itadmin@192.168.5.201.out:echo never > /sys/kernel/mm/transparent\_hugepage/enabled  
itadmin@192.168.5.201.out:echo never > /sys/kernel/mm/transparent\_hugepage/defrag  
  
itadmin@192.168.5.202.out:[always] madvise never  
itadmin@192.168.5.202.out:always madvise [never]  
itadmin@192.168.5.202.out:echo never > /sys/kernel/mm/transparent\_hugepage/enabled  
itadmin@192.168.5.202.out:echo never > /sys/kernel/mm/transparent\_hugepage/defrag

CHECK7: JAVA\_HOME is not set across all machines in the cluster. Recommended to set JAVA\_HOME.

admin:evaluate-pre jay$ grep '\*\* JAVA\_HOME' \*.\* -A1  
itadmin@192.168.5.201.out:\*\*\* JAVA\_HOME  
itadmin@192.168.5.201.out-JAVA\_HOME=  
--  
itadmin@192.168.5.202.out:\*\*\* JAVA\_HOME  
itadmin@192.168.5.202.out-JAVA\_HOME=  
--

CHECK 8: Maximum transmission unit is set to 1500, It can be configured to use (Jumbo frames) 9000 which can give better performance.

admin:evaluate-pre jay$ grep mtu \*.\*  
itadmin@192.168.5.201.out:1: lo: <LOOPBACK,UP,LOWER\_UP> mtu 65536 qdisc noqueue state UNKNOWN  
itadmin@192.168.5.201.out:2: em3: <BROADCAST,MULTICAST,UP,LOWER\_UP> mtu 1500 qdisc mq state UP qlen 1000  
itadmin@192.168.5.201.out:3: em4: <NO-CARRIER,BROADCAST,MULTICAST,UP> mtu 1500 qdisc mq state DOWN qlen 1000  
itadmin@192.168.5.201.out:4: em1: <NO-CARRIER,BROADCAST,MULTICAST,UP> mtu 1500 qdisc mq state DOWN qlen 1000  
itadmin@192.168.5.201.out:5: em2: <NO-CARRIER,BROADCAST,MULTICAST,UP> mtu 1500 qdisc mq state DOWN qlen 1000  
itadmin@192.168.5.201.out:6: p4p1: <BROADCAST,MULTICAST,UP,LOWER\_UP> mtu 1500 qdisc mq state UP qlen 1000  
itadmin@192.168.5.201.out:7: p4p2: <NO-CARRIER,BROADCAST,MULTICAST,UP> mtu 1500 qdisc mq state DOWN qlen 1000  
itadmin@192.168.5.201.out:8: idrac: <BROADCAST,MULTICAST,UP,LOWER\_UP> mtu 1500 qdisc pfifo\_fast state UNKNOWN qlen 1000