Known Issues and Workarounds in Cloudera Manager 5 (#known_issues)

The following sections describe the current known issues in Cloudera Manager 5.

Hive Replication fails when Impala is SSL enabled but Hadoop services are not (#rn_OPSAPS-34354-584)

Support for Impala Replication with SSL enabled was added in Cloudera Manager 5.7.4, 5.8.2, 5.9.0 and higher.

Workaround:

1. In Cloudera Manager, go to the HDFS service.
2. Click the Configuration tab.
3. Search for and enter the appropriate values for the following parameters:
   - Cluster-Wide Default TLS/SSL Client Truststore Location
   - Cluster-Wide Default TLS/SSL Client Truststore Password

Impala metatdata replication is not supported, but will be supported in a later maintenance release.

This support also ensures that connection to Impala is successful when SSL is enabled even if Kerberos is not enabled.

Fixed in Cloudera Manager , , , 5.8.4, .

python-psycopg2 Dependency (#sles_dependency)

Cloudera Manager 5.8 and higher has a new dependency on the package `python-psycopg2`. This package is not available in standard SLES 11 and SLES 12 repositories. You need to add this repository or install it manually before you install or upgrade Cloudera Manager.

Workaround

To add the `python-psycopg2` repository, either add the repository from one of the following URLs:

SLES 11 SP4: http://download.opensuse.org/repositories/devel:/languages:/python/SLE_11_SP4/devel:languages:python.repo (https://www.apache.org/dyn/closer.cgi?path=/kafka/0.10.0.0/RELEASE_NOTES.html)

SLES 12 SP2: http://download.opensuse.org/repositories/devel:/languages:/python/SLE_12_SP2/devel:languages:python.repo (https://www.apache.org/dyn/closer.cgi?path=/kafka/0.10.0.0/RELEASE_NOTES.html)

Or manually install the `python-psycopg2` package from the following URLs:

SLES 11 SP4: http://download.opensuse.org/repositories/devel:/languages:/python/SLE_11_SP4/x86_64/python-psycopg2-2.6.2-1.3.x86_64.rpm (https://www.apache.org/dyn/closer.cgi?path=/kafka/0.10.0.0/RELEASE_NOTES.html)

SLES 12 SP2: http://download.opensuse.org/repositories/devel:/languages:/python/SLE_12_SP2/x86_64/python-psycopg2-2.6.2-1.4.x86_64.rpm (https://www.apache.org/dyn/closer.cgi?path=/kafka/0.10.0.0/RELEASE_NOTES.html)

Pauses in Cloudera Manager after adding peer (#concept_wjv_bnz_ry)

Pauses and slow performance of the Cloudera Manager Admin Console occur after creating a peer.

Workaround:

Increase the memory allocated to Cloudera Manager:

1. On the Cloudera Manager server host, edit the `/etc/default/cloudera-scm-server` file.
2. In the `Java Options` section, change the value of the `-Xmx` argument of the `CMF_JAVA_OPTS` property from `-Xmx2G` to `-Xmx4G`. For example:

   ```
   export CMF_JAVA_OPTS="-Xmx4G -XX:MaxPermSize=256m -XX:+HeapDumpOnOutOfMemoryError -XX:HeapDumpPath=/tmp"
   ```

3. Restart the Cloudera Manager server:

   ```
   sudo service cloudera-scm-server restart
   ```

Cannot select time period in custom charts in C5.9 (#rn_OPSAPS-37190)

The quick time selection (30m, 1h, and so on) on custom dashboards does not work in 5.9.x.

Fixed in Cloudera Manager 5.9.1 and higher.

Cloudera Manager 5.8.2 allows you to select nonexistent CDH 5.8.1 package installation (#rn_OPSAPS-36860)

The Cloudera manager 5.8.2 install/upgrade wizard allows you to select CDH 5.8.1 as a package installation, even though CDH 5.8.1 does not exist. The installation fails with an error message similar to the following:

```
[Errno 14] HTTPS Error 404 - Not Found
```

Workaround

Return to the package selection page, and select Latest Release of CDH5 component compatible with this version of Cloudera Manager or CDH 5.8.2.

Error when distributing parcels: No such torrent (#opeaps-37183)

Parcel distribution might fail with an error message conforming to:

```
Error when distributing to <host>: No such torrent: <parcel_name>.torrent
```

Workaround

Remove the file /opt/cloudera/parcel-cache/<parcel_name>.torrent from the host.

Hive Replication Metadata Transfer Step fails with Temporary AWS Credential Provider (#concept_kws_ymb_tx)

Hive Replication Schedules that use Amazon S3 as the Source or Destination fail when using temporary AWS credentials. The following error displays:

```
Message: Hive Replication Metadata Transfer Step Failed -
com.cloudera.com.amazonaws.services.s3.model.AmazonS3Exception:
Status Code: 403, AWS Service: Amazon S3, AWS Request ID: 76D1F6A02792908A,
AWS Error Code: null, AWS Error Message: Forbidden,
S3 Extended Request ID: Xy3nAS4HSPKLA6hHKvpqReBud7M1Fhk7On0HttYGE0eKPHKwiFkTPQxEVU82OZq5d8omSrdbhcI=.
```

Hive table Views do not get restored from S3 (#concept_tst_ymb_tx)

When creating a Hive Replication schedule that copies Hive data from S3 and you select the Reference Data From Cloud option, Hive table Views are not restored correctly and result in a Null Pointer Exception when querying data from the view.

Fixed in Cloudera Manager 5.9.1 and higher.

ACLs are not replicated when restoring Hive data from S3 (#concept_cxv_ymb_tx)

ACLs are never replicated when the Enable Access Control Lists option in the configuration of the HDFS service is not selected the first time a Replication Schedule that replicates from S3 to Hive runs. Enabling it and re-running the restore operation does not restore the ACLs.

Snapshot diff is not working for Hive to S3 replication when data is deleted on source (#concept_fll_gqb_tx)

If you have enabled snapshots on an HDFS folder and a Hive table uses an external file in that folder, and then you replicate that data to S3 and delete the file on the source cluster, the file is not deleted in subsequent replications to S3, even if the Delete Permanently option is selected.

Block agents from heartbeating to a Cloudera Manager with different UUID until agent restart (#rn_OPSAPS-34847)

If for some reason Cloudera Manager server's identity (Cloudera Manager guid) is changed, agents drop heartbeat requests and will not follow any requested commands from Cloudera Manager server. As a result, the agents report bad health. This situation can be fixed by taking either of the following approaches:

• Restore the previous Cloudera Manager server guid

OR

• Remove the cm_guid file from each of the agents and then restart the agent.

Cloudera Manager set catalogd default jvm memory to 4G can cause out of memory error on upgrade to Cloudera Manager 5.7 or higher (#rn_OPSAPS-34039)

After upgrading to 5.7 or higher, you might see a reduced Java heap maximum on Impala Catalog Server due to a change in its default value. Upgrading from Cloudera Manager lower than 5.7 to Cloudera Manager 5.8.2 no longer causes any effective change in the Impala Catalog Server Java Heap size.

When upgrading from Cloudera Manager 5.7 or later to Cloudera Manager 5.8.2, if the Impala Catalog Server Java Heap Size is set at the default (4GB), it is automatically changed to either 1/4 of the physical RAM on that host, or 32GB, whichever is lower. This can result in a higher or a lower heap, which could cause additional resource contention or out of memory errors, respectively.

Cloudera Manager 5.7.4 installer does not show Key Trustee KMS (#OPSAPS-37132)

A fresh install of Cloudera Manager tries to install Key Trustee KMS 5.8.2 when trying to install the latest version. You must either choose 5.7.0 as the Key Trustee KMS version, or manually provide a link to the 5.7.4 bits.

Class Not Found Error when upgrading to Cloudera Manager 5.7.2 (#cnfe_572)

When you upgrade to version 5.7.2 of Cloudera Manager, the client configuration for all services is marked *stale*.

Workaround:

From the Cluster menu, select Deploy Client Configuration to redeploy the client configuration.

Kerberos setup fails on Debian 8.2 (#concept_dgg_3kx_fw)

This issue is due to the following Debian bug: **https://bugs.debian.org/cgi-bin/bugreport.cgi?bug=777579;msg=5;att=0 (https://bugs.debian.org/cgi-bin/bugreport.cgi?bug=777579;msg=5;att=0)** .

Workaround:

1. Log in to the host where the Cloudera Manager server is running.
2. Edit the `systemd/system/krb5-admin-server.service` file and add `/etc/krb5kdc` to the `ReadWriteDirectories` section.
3. Run the following commands:

   ```
   systemctl daemon-reload
   sudo service krb5-admin-server restart
   ```

4. Generate the credentials.

Password in Cloudera Manager's `db.properties` file is not redacted (#concept_ic2_sq2_lw)

The `db.properties` file is managed by customers and is populated manually when the Cloudera Manager Server database is being set up for the first time. Since this occurs before the Cloudera Manager Server has even started, encrypting the contents of this file is a completely different challenge as compared to that of redacting configuration files.

Releases affected: 5.3 and higher

Cluster provisioning fails [(#concept_k2m_2gm_xv)](#concept_k2m_2gm_xv)

In some cases, provisioning of a cluster may fail at the start of the process. This does not happen in all cases and is mainly noticed on RHEL 6 and especially when some hosts are reporting bad health.

Releases affected: 5.5.0-5.5.3, 5.6.0-5.6.1, 5.7.0

Releases containing the fix: 5.5.4, 5.7.1

For releases containing the fix, parcel activation and first run command now completes as expected, even when some hosts report bad health.

This issue is fixed in Cloudera Manager 5.5.4 and 5.7.1 and higher.

Cloudera Manager can run out of memory if a remote repository URL is unreachable [(#concept_ovh_t3d_3v)](#concept_ovh_t3d_3v)

If one of the URLs specified in on the Parcel Settings page (Hosts > Parcels > Configuration) becomes unreachable, Cloudera Manager may run out of memory.

Workaround:

Do one of the following:

- If the URL is incorrect, enter the correct URL.
- Deselect the Automatically Download New Parcels setting on the Parcel Settings page.
- Set the value of the Parcel Update Frequency on the Parcel Settings to a large interval such as several days.

Clients can run Hive on Spark jobs even if Hive dependency on Spark is not configured [(#concept_o1k_cjb_fv)](#concept_o1k_cjb_fv)

In CDH 5.7 and higher, when Hive and Spark on YARN both are configured, but Hive is not configured to depend on Spark on YARN, clients can set the execution engine to `spark` and Hive on Spark jobs will still be executed but will run in an unsupported mode. These jobs may not appear in the Spark History Server.

Workaround: Configure Hive to depend on Spark on YARN.

Known Issues for the DSSD D5 Hadoop Plugin [(#concept_o33_gv2_bv)](#concept_o33_gv2_bv)

Default Directory for Data Directory for a DSSD DataNode Should be Reconfigured (#concept_ilp_hj5_nw)

Note: This issue applies only to clusters that use the EMC DSSD D5 storage appliance as the storage for HDFS DataNodes.

The DSSD DataNodes operating in the Hadoop cluster store version and ID information on a local disk directory. The default directory is `/tmp/hadoop-hdfs`. This default directory may be deleted by the operating system over time or when the operating system is restarted. If the DSSD DataNode process is restarted after the default directory is deleted, it will not be able to locate replicas stored on the DSSD D5 appliance. This causes the HDFS service to report under-replicated or even missing HDFS blocks.

Therefore it is best practice to change the default directory to a different value such as `/var/lib/hadoop-hdfs/dssddn`. The directory should be created and configured prior to initiating the installation of the Hadoop cluster using the following procedure:

1. Log into each host that will run the DataNode process.
2. Create the directory by executing the command:

   ```
   mkdir -p /var/lib/hadoop-hdfs/dssddn
   ```

3. Change the directory ownership by executing the command:l

   ```
   chown -R hdfs:hadoop /var/lib/hadoop-hdfs
   ```

4. Change the directory permissions by executing the command:

   ```
   chmod -R 700 /var/lib/hadoop-hdfs
   ```

5. In Cloudera Manager, configure the `dfs.datanode.data.dir` property in the HDFS Service Advanced Configuration Snippet (Safety Valve) for hdfs-site.xml configuration property. This property must to be set after the initial setup is completed but prior to writing any data to HDFS.

Note: Changing the value of the `dfs.datanode.data.dir` property after data has been written to HDFS will result is under-replicated or lost HDFS blocks.

Cloudera Manager reports that DSSD DataNode version is different from NameNode (#concept_vgw_wkc_rw)

The following warning log entry for the NameNode displays when using the DSSD Hadoop Plugin version 1.2.0 . This is expected and does not affect normal HDFS operation:

```
Jun 24, 8:12:41.846 AM WARN BlockStateChange
BLOCK* processReport: Report from the DataNode (8fa7fed1-1044-433b-a9fc-682bc08d1e25) is unsorted.
```

```
This will cause overhead on the NameNode which needs to sort the Full BR.
Please update the DataNode to the same version of Hadoop HDFS as the NameNode (2.6.0-cdh5.8.0).
```

HDFS caching (#concept_mty_px2_bv)

HDFS caching is not supported when using the DSSD D5 plugin.

Short Circuit Reads and Access Control with Impala or HBase (#concept_vtk_qx2_bv)

Enabling short-circuit reads for HBase or Impala on an HDFS cluster that uses DSSD D5 DataNodes requires that the processes associated with these applications be granted `hdfs` group membership. When short-circuit reads are enabled for Impala (for example), Impala process that act as short-circuit read clients (like `impalad`) are able to read and write all data stored in the DSSD D5. Cloudera Manager applies the `hdfs` group membership on a per-service basis, and applications that do not require short-circuit reads or for which short-circuit reads have not been enabled will have the same granularity of access control as present on a traditional HDFS cluster. Whether short-circuit reads are enabled or not, access control that is enforced by the application rather than at the file system level is identical for DSSD D5 DataNode HDFS clusters and traditional HDFS clusters.

See **DSSD D5 and Short-Circuit Reads (dssd_scr.html#dssd_scr)** .

The YARN NodeManager connectivity health test does not work for CDH 5 (#concept_xfq_fdj_r5)

The NodeManager connectivity is always GOOD (green) even if the ResourceManager considers the NodeManager to be LOST or DECOMMISSIONED.

Workaround: None.

HDFS HA clusters see NameNode failures when KDC connectivity is bad (#concept_brv_1q4_15)

When KDC connectivity is bad, the JVM takes 30 seconds before retrying or declaring failure to connect. Meanwhile, the JournalNode write timeout (which needs KDC authentication for the first write, or under troubled connectivity), is only 20 seconds.

Workaround: In `krb5.conf`, set the `kdc_timeout` parameter value to 3 seconds. In Cloudera Manager, perform the following steps:

1. Go to Administration > Settings > Kerberos.

2. Add the `kdc_timeout` parameter to the **Advanced Configuration Snippet (Safety Valve) for [libdefaults] section of krb5.conf** property. This should give the JVM enough time to try connecting to a KDC before the JournalNode timeout.

The HDFS File browser in Cloudera Manager fails when HDFS federation is enabled (#concept_s3d_1l5_xt)

Workaround: Use the command-line `hdfs dfs` commands to directly manipulate HDFS files when federation is enabled. CDH supports HDFS federation.

Hive Metastore canary fails to drop database (#concept_hwf_pfg_xs)

The Hive Metastore canary fails to drop the database due to **HIVE-11418 (cdh_rn_hive_ki.html#concept_vzc_mtf_xs)** .

Workaround: Turn off the Hive Metastore canary by disabling the Hive Metastore Canary Health Test:

1. Go to the Hive service.
2. Click the Configuration tab.
3. Select SCOPE > Hive Metastore Server.
4. Select CATEGORY > Monitoring.
5. Deselect the Hive Metastore Canary Health Test checkbox for the Hive Metastore Server Default Group.
6. Click Save Changes to commit the changes.

Cloudera Manager upgrade fails due to incorrect Sqoop 2 path (#concept_u1n_rzb_ws)

Sqoop 2 will not start or loses data when upgrading from Cloudera Manager 5.4.0 or 5.4.1 to Cloudera Manager 5.4.3, or when upgrading Cloudera Manager 3 or 4 to Cloudera Manager 5.4.0 or 5.4.1. This is due to Cloudera Manager erroneously configuring the Derby path with "`repositoy`" instead of "`repository`". To upgrade Cloudera Manager, use one of the following workarounds:

- **Workaround for Upgrading from Cloudera Manager 3 or 4 to Cloudera Manager 5.4.0 or 5.4.1**

   1. Log in to your Sqoop 2 server host using SSH and move the Derby database files to the new location, usually from `/var/lib/sqoop2/repository` to `/var/lib/sqoop2/repositoy`.
   2. Start Sqoop2. If you found this problem while upgrading CDH, run the Sqoop 2 database upgrade command using the Actions drop-down menu for Sqoop 2.

- **Workaround for Upgrading from Cloudera Manager 5.4.0 or 5.4.1 to Cloudera Manager 5.4.3**

1. Log in to your Sqoop 2 server host using SSH and move the Derby database files to the new location, usually from `/var/lib/sqoop2/repositoy` to `/var/lib/sqoop2/repository`.

2. Start Sqoop2, or if you found this problem while upgrading CDH, run the Sqoop 2 database upgrade command using the Actions drop-down menu for Sqoop 2.

NameNode incorrectly reports missing blocks during rolling upgrade [(#concept_us1_yyt_js)](#concept_us1_yyt_js)

During a rolling upgrade to any of the CDH releases listed below, the NameNode may report missing blocks after rolling back multiple DataNodes. This is caused by a race condition with block reporting between the DataNode and the NameNode. No permanent data loss occurs, but data can be unavailable for up to six hours before the problem corrects itself.

Releases affected: CDH 5.0.6, 5.1.5, 5.2.5, 5.3.3, 5.4.1, 5.4.2.

Releases containing the fix:: CDH 5.2.6, 5.3.4, 5.4.3

Workaround:

- **To avoid the problem** - Cloudera advises skipping the affected releases and installing a release containing the fix. For example, do not upgrade to CDH 5.4.2; upgrade to CDH 5.4.3 instead.
- **If you have already completed an upgrade to an affected release, or are installing a new cluster** - You can continue to run the release, or upgrade to a release that is not affected.

Using ext3 for server dirs easily hit inode limit [(#concept_oyp_xtt_2s)](#concept_oyp_xtt_2s)

Using the ext3 filesystem for the Cloudera Manager command storage directory may exceed the maximum subdirectory size of 32000.

Workaround: Either decrease the value of the Command Eviction Age property so that the directories are more aggressively cleaned up, or migrate to the ext4 filesystem.

Backup and disaster recovery replication does not set MapReduce Java options [(#concept_usc_qht_zr)](#concept_usc_qht_zr)

Replication used for backup and disaster recovery relies on system-wide MapReduce memory options, and you cannot configure the options using the Advanced Configuration Snippet.

Kafka 1.2 CSD conflicts with CSD included in Cloudera Manager 5.4 [(#concept_tvn_2hg_qr)](#concept_tvn_2hg_qr)

If the Kafka CSD was installed in Cloudera Manager to 5.3 or lower, the old version must be uninstalled, otherwise it will conflict with the version of the Kafka CSD bundled with Cloudera Manager 5.4.

Workaround: Remove the Kafka 1.2 CSD before upgrading Cloudera Manager to 5.4:

1. Determine the location of the CSD directory:
   a. Select Administration > Settings.
   b. Click the **Custom Service Descriptors** category.
   c. Retrieve the directory from the **Local Descriptor Repository Path** property.
2. Delete the Kafka CSD from the directory.

Recommission host does not deploy client configurations [(#concept_ehv_kyt_4r)](#concept_ehv_kyt_4r)

The failure to deploy client configurations can result in client configuration pointing to the wrong locations, which can cause errors such as the NodeManager failing to start with "Failed to initialize container executor".

Workaround: Deploy client configurations first and then restart roles on the recommissioned host.

Hive on Spark is not supported in Cloudera Manager and CDH 5.4 and CDH 5.5 [(#concept_sks_cyt_4r)](#concept_sks_cyt_4r)

You can configure Hive on Spark, but it is not recommended for production clusters.

CDH 5 requires JDK 1.7 [(#concept_olg_dsn_4r)](#concept_olg_dsn_4r)

JDK 1.6 is not supported on any CDH 5 release, but before CDH 5.4.0, CDH libraries have been compatible with JDK 1.6. As of CDH 5.4.0, CDH libraries are no longer compatible with JDK 1.6 and **applications using CDH libraries must use JDK 1.7.**

In addition, you must upgrade your cluster to a supported version of JDK 1.7 before upgrading to CDH 5. See **[Upgrading to Oracle JDK 1.7 before Upgrading to CDH 5 (http://www.cloudera.com/documentation/enterprise/latest/topics/cdh_cm_upgrading_to_jdk7.html)](http://www.cloudera.com/documentation/enterprise/latest/topics/cdh_cm_upgrading_to_jdk7.html)** for instructions.

Upgrade wizard incorrectly upgrades the Sentry DB [(#concept_icg_pf2_nr)](#concept_icg_pf2_nr)

There's no Sentry DB upgrade in 5.4, but the upgrade wizard says there is. Performing the upgrade command is not harmful, and taking the backup is also not harmful, but the steps are unnecessary.

Cloudera Manager does not correctly generate client configurations for services deployed using CSDs [(#concept_bbf_1d2_nr)](#concept_bbf_1d2_nr)

HiveServer2 requires a Spark on YARN gateway on the same host in order for Hive on Spark to work. You must deploy Spark client configurations whenever there's a change in order for HiveServer2 to pick up the change.

CSDs that depend on Spark will get incomplete Spark client configuration. Note that Cloudera Manager does not ship with any such CSDs by default.

Workaround: Use `/etc/spark/conf` for Spark configuration, and ensure there is a Spark on YARN gateway on that host.

Solr, Oozie and HttpFS fail when KMS and TLS/SSL are enabled using self-signed certificates [(#concept_avt_n25_xt)](#concept_avt_n25_xt)

When the KMS service is added and TLS/SSL is enabled, Solr, Oozie and HttpFS are not automatically configured to trust the KMS's self-signed certificate and you might see the following error.

```
org.apache.oozie.service.AuthorizationException: E0501: Could not perform authorization operation,
sun.security.validator.ValidatorException: PKIX path building failed: sun.security.provider.certpath.SunCertPathBuilderException:
unable to find valid certification path to requested target
```

Workaround: You must explicitly load the relevant truststores with the KMS certificate to allow these services to communicate with the KMS. To do so, edit the truststore location and password for Solr, Oozie and HttpFS (found under the HDFS service) as follows.

1. Go to the Cloudera Manager Admin Console.
2. Go to the Solr/Oozie/HDFS service.
3. Click the Configuration tab.
4. Search for "<service> TLS/SSL Certificate Trust Store File" and set this property to the location of truststore file.
5. Search for "<service> TLS/SSL Certificate Trust Store Password" and set this property to the password of the truststore.
6. Click Save Changes to commit the changes.

Cloudera Manager 5.3.1 upgrade fails if Spark standalone and Kerberos are configured [(#concept_yby_f5c_dr)](#concept_yby_f5c_dr)

CDH upgrade fails if Kerberos is enabled and Spark standalone is installed. Spark standalone does not work in a kerberized cluster.

Workaround: To upgrade, remove the Spark standalone service first and then proceed with upgrade.

Adding Key Trustee KMS 5.4 to Cloudera Manager 5.5 displays warning (#concept_dmp_2gz_45)

Adding the Key Trustee KMS service to a CDH 5.4 cluster managed by Cloudera Manager 5.5 displays the following message, even if Key Trustee KMS is installed:

"The following selected services cannot be used due to missing components: keytrustee-keyprovider. Are you sure you wish to continue with them?"

Workaround: Verify that the Key Trustee KMS parcel or package is installed and click OK to continue adding the service.

KMS and Key Trustee ACLs do not work in Cloudera Manager 5.3 (#concept_k1p_v2s_5q)

ACLs configured for the KMS (File) and KMS (Navigator Key Trustee) services do not work since these services do not receive the values for `hadoop.security.group.mapping` and related group mapping configuration properties.

Workaround:

KMS (File): Add all configuration properties starting with `hadoop.security.group.mapping` from the NameNode `core-site.xml` to the KMS (File) property, **Key Management Server Advanced Configuration Snippet (Safety Valve) for core-site.xml**

KMS (Navigator Key Trustee): Add all configuration properties starting with `hadoop.security.group.mapping` from the NameNode `core-site.xml` to the KMS (Navigator Key Trustee) property, **Key Management Server Proxy Advanced Configuration Snippet (Safety Valve) for core-site.xml**.

Exporting and importing Hue database sometimes times out after 90 seconds (#concept_ts3_wns_5q)

Executing 'dump database' or 'load database' of Hue from Cloudera Manager returns "command aborted because of exception: Command timed-out after 90 seconds". The Hue database can be exported to JSON from within Cloudera Manager. Unfortunately, sometimes the Hue database is quite large and the export times out after 90 seconds.

Workaround: Ignore the timeout. The command should eventually succeed even though Cloudera Manager reports that it timed out.

Changing the Key Trustee Server hostname requires editing keytrustee.conf [(#concept_jcj_2wn_mq)](#concept_jcj_2wn_mq)

If you change the hostname of your active or passive Key Trustee Server, you must edit the `keytrustee.conf` file. This issue typically arises if you replace an active or passive server with a server having a different hostname. If the same hostname is used on the replacement server, there are no issues.

Workaround: Use the same hostname on the replacement server.

Hosts with Impala Llama roles must also have at least one YARN role [(#concept_ocl_1xk_bq)](#concept_ocl_1xk_bq)

When integrated resource management is enabled for Impala, host(s) where the Impala Llama role(s) are running must have at least one YARN role. This is because Llama requires the `topology.py` script from the YARN configuration. If this requirement is not met, you may see errors such as:

```
"Exception running /etc/hadoop/conf.cloudera.yarn/topology.py
java.io.IOException: Cannot run program "/etc/hadoop/conf.cloudera.yarn/topology.py"
```

in the Llama role logs, and Impala queries may fail.

Workaround: Add a YARN gateway role to each Llama host that does not already have at least one YARN role (of any type).

The high availability wizard does not verify that there is a running ZooKeeper service [(#concept_lkl_s1v_1q)](#concept_lkl_s1v_1q)

If one of the following is true:

- 1. ZooKeeper present and not running and the HDFS dependency on ZooKeeper dependency is not set
- 2. ZooKeeper absent

the enable high-availability wizard fails.

Workaround: Before enabling high availability, do the following:

1. Create and start a ZooKeeper service if one does not exist.

2. Go to the HDFS service.

3. Click the Configuration tab.

4. Select Scope > Service-Wide

5. Set the ZooKeeper Service property to the ZooKeeper service.

6. Click Save Changes to commit the changes.

Cloudera Manager Installation Path A fails on RHEL 5.7 due to PostgreSQL conflict (#concept_irk_t5f_tp)

On RHEL 5.7, `cloudera-manager-installer.bin` fails due to a PostgreSQL conflict if PostgreSQL 8.1 is already installed on your host.

Workaround: Remove PostgreSQL from host and rerun `cloudera-manager-installer.bin`.

Spurious warning on Accumulo 1.6 gateway hosts (#concept_odr_1d5_t4)

When using the Accumulo shell on a host with only an Accumulo 1.6 Service gateway role, users will receive a warning about failing to create the directory **/var/log/accumulo**. The shell works normally otherwise.

Workaround: The warning is safe to ignore.

Accumulo 1.6 service log aggregation and search does not work (#concept_dms_1d5_t4)

Cloudera Manager log aggregation and search features are incompatible with the log formatting needed by the Accumulo Monitor. Attempting to use either the "Log Search" diagnostics feature or the log file link off of an individual service role's summary page will result in empty search results.

Severity: High

Workaround: Operators can use the Accumulo Monitor to see recent severe log messages. They can see recent log messages below the WARNING level via a given role's process page and can inspect full logs on individual hosts by looking in `/var/log/accumulo`.

Cloudera Manager incorrectly sizes Accumulo Tablet Server max heap size after 1.4.4-cdh4.5.0 to 1.6.0-cdh4.6.0 upgrade (#concept_wfm_xpn_t4)

Because the upgrade path from Accumulo 1.4.4-cdh4.5.0 to 1.6.0-cdh4.6.0 involves having both services installed simultaneously, Cloudera Manager will be under the impression that worker hosts in the cluster are oversubscribed on memory and attempt to downsize the max heap size allowed for 1.6.0-cdh4.6.0 Tablet Servers.

Severity: High

Workaround: Manually verify that the Accumulo 1.6.0-cdh4.6.0 Tablet Server max heap size is large enough for your needs. Cloudera recommends you set this value to the sum of 1.4.4-cdh4.5.0 Tablet Server and Logger heap sizes.

Accumulo installations using LZO do not indicate dependence on the GPL Extras parcel (#concept_q51_fg4_r4)

Accumulo 1.6 installations that use LZO compression functionality do not indicate that LZO depends on the GPL Extras parcel. When Accumulo is configured to use LZO, Cloudera Manager has no way to track that the Accumulo service now relies on the GPL Extras parcel. This prevents Cloudera Manager from warning administrators before they remove the parcel while Accumulo still requires it for proper operation.

Workaround: Check your Accumulo 1.6 service for the configuration changes mentioned in the Cloudera documentation for using Accumulo with CDH prior to removing the GPL Extras parcel. If the parcel is mistakenly removed, reinstall it and restart the Accumulo 1.6 service.

Created pools are not preserved when Dynamic Resource Pools page is used to configure YARN or Impala (#concept_fmm_d3x_wn)

Pools created on demand are not preserved when changes are made using the Dynamic Resource Pools page. If the Dynamic Resource Pools page is used to configure YARN or Impala services in a cluster, it is possible to specify pool placement rules that create a pool if one does not already exist. If changes are made to the configuration using this page, pools created as a result of such rules are not preserved across the configuration change.

Workaround: Submit the YARN application or Impala query as before, and the pool will be created on demand once again.

User should be prompted to add the AMON role when adding MapReduce to a CDH 5 cluster (#concept_akd_tt2_qn)

When the MapReduce service is added to a CDH 5 cluster, the user is not asked to add the AMON role. Then, an error displays when the user tries to view MapReduce activities.

Workaround: Manually add the AMON role after adding the MapReduce service.

Enterprise license expiration alert not displayed until Cloudera Manager Server is restarted (#concept_gyk_jgk_qn)

When an enterprise license expires, the expiration notification banner is not displayed until the Cloudera Manager Server has been restarted. The enterprise features of Cloudera Manager are not affected by an expired license.

Workaround: None.

Configurations for decommissioned roles not migrated from MapReduce to YARN (#concept_rgm_zkm_mn)

When the Import MapReduce Configuration wizard is used to import MapReduce configurations to YARN, decommissioned roles in the MapReduce service do not cause the corresponding imported roles to be marked as decommissioned in YARN.

Workaround: Delete or decommission the roles in YARN after running the import.

The HDFS command Roll Edits does not work in the UI when HDFS is federated (#concept_p2g_q3m_mn)

The HDFS command Roll Edits does not work in the Cloudera Manager UI when HDFS is federated because the command does not know which nameservice to use.

Workaround: Use the API, not the Cloudera Manager UI, to execute the Roll Edits command.

Cloudera Manager reports a confusing version number if you have oozie-client, but not oozie installed on a CDH 4.4 node (#concept_o1z_qvg_fn)

In CDH versions before 4.4, the metadata identifying Oozie was placed in the client, rather than the server package. Consequently, if the client package is not installed, but the server is, Cloudera Manager will report Oozie has been present but as coming from CDH 3 instead of CDH 4.

Workaround: Either install the oozie-client package, or upgrade to at least CDH 4.4. Parcel based installations are unaffected.

Cloudera Manager does not work with CDH 5.0.0 Beta 1 (#concept_ryv_cdj_bn)

When you upgrade from Cloudera Manager 5.0.0 Beta 1 with CDH 5.0.0 Beta 1 to Cloudera Manager 5.0.0 Beta 2, Cloudera Manager won't work with CDH 5.0.0 Beta 1 and there's no notification of that fact.

Workaround: None. Do a new installation of CDH 5.0.0 Beta 2.

On CDH 4.1 secure clusters managed by Cloudera Manager 4.8.1 and higher, the Impala Catalog server needs advanced configuration snippet update (#concept_r5c_5h4_wm)

Impala queries fail on CDH 4.1 when Hive "Bypass Hive Metastore Server" option is selected.

Workaround: Add the following to Impala catalog server advanced configuration snippet for `hive-site.xml`, replacing *Hive_Metastore_Server_Host* with the host name of your Hive Metastore Server:

```
<property>
<name>hive.metastore.local</name>
<value>false</value>
</property>
<property>
<name>hive.metastore.uris</name>
<value>thrift://Hive_Metastore_Server_Host:9083</value>
</property>
```

Rolling Upgrade to CDH 5 is not supported. (#concept_gzl_jh3_bm)

Rolling upgrade between CDH 4 and CDH 5 is not supported. Incompatibilities between major versions means rolling restarts are not possible. In addition, rolling upgrade will *not* be supported from CDH 5.0.0 Beta 1 to any later releases, and may not be supported between *any future beta versions* of CDH 5 and the General Availability release of CDH 5.

Workaround: None.

Error reading .zip file created with the Collect Diagnostic Data command. (#concept_xrb_w4j_zl)

After collecting Diagnostic Data and using the Download Diagnostic Data button to download the created zip file to the local system, the zip file cannot be opened using the FireFox browser on a Macintosh. This is because the zip file is created as a Zip64 file, and the unzip utility included with Macs does not support Zip64. The zip utility must be version 6.0 or later. You can determine the zip version with `unzip -v`.

Workaround: Update the unzip utility to a version that supports Zip64.

After JobTracker failover, complete jobs from the previous active JobTracker are not visible. (#concept_cvp_qvd_vl)

When a JobTracker failover occurs and a new JobTracker becomes active, the new JobTracker UI does not show the completed jobs from the previously active JobTracker (that is now the standby JobTracker). For these jobs the "Job Details" link does not work.

Severity: Med

Workaround: None.

After JobTracker failover, information about rerun jobs is not updated in Activity Monitor. [(#concept_hvp_qvd_vl)](#concept_hvp_qvd_vl)

When a JobTracker failover occurs while there are running jobs, jobs are restarted by the new active JobTracker by default. For the restarted jobs the Activity Monitor will not update the following: 1) The start time of the restarted job will remain the start time of the original job. 2) Any Map or Reduce task that had finished before the failure happened will not be updated with information about the corresponding task that was rerun by the new active JobTracker.

Severity: Med

Workaround: None.

Installing on AWS, you must use private EC2 hostnames. [(#concept_zyp_qvd_vl)](#concept_zyp_qvd_vl)

When installing on an AWS instance, and adding hosts using their public names, the installation will fail when the hosts fail to heartbeat.

Severity: Med

Workaround:

Use the Back button in the wizard to return to the original screen, where it prompts for a license.

Rerun the wizard, but choose "Use existing hosts" instead of searching for hosts. Now those hosts show up with their internal EC2 names.

Continue through the wizard and the installation should succeed.

If HDFS uses Quorum-based Storage without HA enabled, the SecondaryNameNode cannot checkpoint. [(#concept_dbq_qvd_vl)](#concept_dbq_qvd_vl)

If HDFS is set up in non-HA mode, but with Quorum-based storage configured, the `dfs.namenode.edits.dir` is automatically configured to the Quorum-based Storage URI. However, the SecondaryNameNode cannot currently read the edits from a Quorum-based Storage URI, and will be unable to do a checkpoint.

Severity: Medium

Workaround: Add to the NameNode's advanced configuration snippet the `dfs.namenode.edits.dir` property with both the value of the Quorum-based Storage URI as well as a local directory, and restart the NameNode. For example,

```
<property> <name>dfs.namenode.edits.dir</name>
<value>qjournal://jn1HostName:8485;jn2HostName:8485;jn3HostName:8485/journalhdfs1,file:///dfs/edits</value>
</property>
```

Changing the rack configuration may temporarily cause mis-replicated blocks to be reported. [(#concept_lbq_qvd_vl)](#concept_lbq_qvd_vl)

A rack re-configuration will cause HDFS to report mis-replicated blocks until HDFS rebalances the system, which may take some time. This is a normal side-effect of changing the configuration.

Severity: Low

Workaround: None

Cannot use '/' as a mount point with a Federated HDFS Nameservice. [(#concept_vbq_qvd_vl)](#concept_vbq_qvd_vl)

A Federated HDFS Service does not support nested mount points, so it is impossible to mount anything at '/'. Because of this issue, the root directory will always be read-only, and any client application that requires a writeable root directory will fail.

Severity: Low

Workaround:

1. In the CDH 4 HDFS Service > Configuration tab of the Cloudera Manager Admin Console, search for "nameservice".
2. In the Mountpoints field, change the mount point from "/" to a list of mount points that are in the namespace that the Nameservice will manage. (You can enter this as a comma-separated list - for example, "/hbase, /tmp, /user" or by clicking the plus icon to add each mount point in its own field.) You can determine the list of mount points by running the command hadoop fs -ls / from the CLI on the NameNode host.

Historical disk usage reports do not work with federated HDFS. [(#concept_pcq_qvd_vl)](#concept_pcq_qvd_vl)

Severity: Low

Workaround: None.

(CDH 4 only) Activity monitoring does not work on YARN activities. [(#concept_ucq_qvd_vl)](#concept_ucq_qvd_vl)

Activity monitoring is not supported for YARN in CDH 4.

Severity: Low

Workaround: None

HDFS monitoring configuration applies to all Nameservices (#concept_ddq_qvd_vl)

The monitoring configurations at the HDFS level apply to all Nameservices. So, if there are two federated Nameservices, it's not possible to disable a check on one but not the other. Likewise, it's not possible to have different thresholds for the two Nameservices.

Severity: Low

Workaround: None

Supported and Unsupported Replication Scenarios and Limitations (#concept_p2g_c43_bn)

See **Data Replication (http://www.cloudera.com/documentation/enterprise/latest/topics/cm_bdr_replication_intro.html)** .

Restoring snapshot of a file to an empty directory does not overwrite the directory (#concept_wgy_21f_jn)

Restoring the snapshot of an HDFS file to an HDFS path that is an empty HDFS directory (using the Restore As action) will result in the restored file present inside the HDFS directory instead of overwriting the empty HDFS directory.

Workaround: None.

HDFS Snapshot appears to fail if policy specifies duplicate directories. (#concept_ypc_z2w_yl)

In an HDFS snapshot policy, if a directory is specified more than once, the snapshot appears to fail with an error message on the Snapshot page. However, in the HDFS Browser, the snapshot is shown as having been created successfully.

Severity: Low

Workaround: Remove the duplicate directory specification from the policy.

Hive replication fails if "Force Overwrite" is not set. (#OPSAPS13548)

The Force Overwrite option, if checked, forces overwriting data in the target metastore if there are incompatible changes detected. For example, if the target metastore was modified and a new partition was added to a table, this option would force deletion of that partition, overwriting the table with the version found on the source. If the Force Overwrite option is not set, recurring replications may fail.

Severity: Med

Workaround: Set the Force Overwrite option.

Cloudera Manager set cataloged default JVM memory to 4G can cause an out of memory error during upgrade to Cloudera Manager 5.7 and higher (#rn_OPSAPS-34039a)

When upgrading from Cloudera Manager 5.7 or higher to Cloudera Manager 5.8.2, if the Impala Catalog Server Java Heap Size is set at the default (4GB), it is automatically changed to either 1/4 of the physical RAM on that host, or 32GB, whichever is lower. This can result in a higher or a lower heap, which could cause additional resource contention or out of memory errors, respectively.

Categories: Amazon (../categories/hub_amazon.html) | Cloudera Manager (../categories/hub_cloudera_manager.html) | EC2 (../categories/hub_ec2.html) | Known Issues (../categories/hub_known_issues.html) | Preview Features (../categories/hub_preview_features.html) | Release Notes (../categories/hub_release_notes.html) | Troubleshooting (../categories/hub_troubleshooting.html) | Upgrading (../categories/hub_upgrading.html) | All Categories (../categories/hub.html)

- **About Cloudera (http://www.cloudera.com/about-cloudera.html)**
- **Resources (http://www.cloudera.com/resources.html)**
- **Contact (http://www.cloudera.com/contact-us.html)**
- **Careers (http://www.cloudera.com/about-cloudera/careers.html)**
- **Press (http://www.cloudera.com/about-cloudera/press-center.html)**
- **Documentation (http://www.cloudera.com/documentation.html)**

United States: +1 888 789 1488
Outside the US: +1 650 362 0488

- [(https://www.linkedin.com/company/cloudera)](https://www.linkedin.com/company/cloudera)

- [(https://www.facebook.com/cloudera)](https://www.facebook.com/cloudera)

- [(https://twitter.com/cloudera)](https://twitter.com/cloudera)

- [(http://www.cloudera.com/contact-us.html)](http://www.cloudera.com/contact-us.html)

[Terms & Conditions (http://www.cloudera.com/legal/terms-and-conditions.html)](http://www.cloudera.com/legal/terms-and-conditions.html) | [Privacy Policy (http://www.cloudera.com/legal/policies.html)](http://www.cloudera.com/legal/policies.html)

Page generated February 8, 2017.