

# **Time Series Data Engine Issues**

- Frequent Issues, on page 1
- Frequently Asked Questions, on page 10

# **Frequent Issues**

**1.** Empty HTTP API Response

Symptom	HTTP API response is OK (status 200), but there is no data in the response.
	E.g. Getting an empty response for the below location URL:
	Token needs to be added at http header.
	https:// <tenant.com>:<port>/dataeng/api/v1.0/config/location</port></tenant.com>
Environment	Can occur in both Prod and QA cluster.
Possible Causes	There is no data in impala cimdata database, or locations are not loaded from SDP server.
	Edge-server.log
	2018-06-27T07:05:25,212 INFO (pool-6-thread-1) [CimSDPServer(refresh:219)] source.SDPServer.produs-ckc access token refreshed: 0433e389-c042-34d4-9a8e-38845a0d21a4 2018-06-27T07:05:25,777 WARN (pool-6-thread-1) [HttpClientImpl(handleResponse:330)] response code is: 401 2018-06-27T07:05:25,778 ERROR (pool-6-thread-1) [CimSDPServer(run:107)] source.SDPServer.produs-ckc SDP refresh: guery SDP location api: null
Troubleshooting	Check SDP server status.
	Otherwise, perform following steps.
	• Check HBase table, if HBase table has data, it means data is missed during ETL job, get the workflow log and contact Data Engine team.
	• If HBase table is already mapped to cim database, user can query with impala SQL.

Post Verification	If Core location server is in running state, but data engine still reports empty response for about an hour, raise CDETS for Data Engine.
Verification	Rerun HTTP API to check if data is coming.
	• If there is no error in the edge log, please check device engine http and websocke work status, to see if there is relevant data output. If not, please contact device engine/extension team.
	• If Kafka does not have the data, then check edge-server.log in the Edge host. If there is any error, send to Data engine team.
	Kafka-console-consumerbootstrap-server data-01.novalocal:9092 topic cimfrom-beginning   grep "your-keyword-here"
	<b>Note</b> #data-01.novalocal is a Kafka broker hostname, 9092 is Kafka service port. This command #will output all Kafka messages, and only show messages containing your expected keyword.
	• If HBase does not have the data, check Kafka with Kafka consumer (utility). If kafka has data, but HBase does not, it means Flume has not put the data into HBase User should get all of flume agents logs and send to the Data Engine Team.

## 2. HTTP API Response Login Failure Error

Symptom	Login Failure
~JP	
	E.g. Getting an empty response for the below location URL:
	https:// <tenant.com>:<port>/dataeng/api/v1.0/config/location</port></tenant.com>
	Internal server error: Login failure for hdfs@CIM.IVSG.AUTH from keytab /usr/local/cim/conf/hdfs.keytab
Environment	Can occur in both Prod and QA cluster.
Possible Causes	Network Failure
	• Kerberos auth failure.
	• CDH not in good health condition.
Troubleshooting	• Make sure Kerberos server is running, TCP & UDP port 88 are listening on, and opened in firewall
	For other errors, perform the following steps.
	• Test API again in Edge host
	In AWS cluster, there is a web gateway between webserver in edge host and user browser. If API works in edge host, check the gateway status.
	Check SDP Server's connectivity.
	• Check edge-server.log. Refer Issue 9 <i>No route to Host</i> , if there is any similar issue, fix it.
	For any other errors, send the logs to the Data Engine team

	• Check Cloudera cluster health status, refer <i>What to do if many Services are in bad health ?</i> and <i>What to do if only one service is in bad health status?</i> in the FAQ section.
	• Check impala service with impala-shell. Check the SQL response time. SQL timeout would most likely the root cause. Also, large quantity of small files would be the root cause of the SQL timeout.
	Reach out to the Data Engine team in case of deviation from above mentioned behaviour.
Verification	Rerun HTTP API to check if data is present in the response.
Post Verification	If not a network issue, raise a CDET for Data Engine.

## **3.** HTTP API Location ID Error

Symptom	Data of the location id is not available.
	data of the location id not available:[location id]
Environment	Can occur in both Prod and QA cluster.
Possible Causes	Core location server not responding.
	• CDH cluster not in good health condition.
	• LocationID is not a valid one.
Troubleshooting	Make sure locationId is a valid city id.
	• Check SDP server status refer issue 1. Empty HTTP API Response.
Verification	Rerun the location api to check the response.
Post Verification	If location server is running and location id is valid, data engine response with no location result. Check CDH status is fine, raise CDETS against data engine

4. HTTP API Response Tenant Details Error

Symptom	Tenant details not appearing in location API, post data engine registration.
Environment	Can occur in both Prod and QA cluster.
Possible Causes	<ul> <li>Core location server not responding.</li> <li>CDH cluster not in good health condition.</li> <li>LocationID is not a valid one.</li> </ul>
Troubleshooting	<ul> <li>Make sure locationId is a valid city id.</li> <li>Check SDP server status refer issue 1. <i>Empty HTTP API Response</i>.</li> </ul>
Verification	Rerun the location api to check the response.

Post Verification	If location server is running, location id is valid and CDH status is fine,
	still data engine response has no location result then raise CDETS against
	data engine

# 5. HTTP API Response Contains unreasonable data

Symptom	Percentage value is negative, or Average, Minimum, Maximum values are unreasonable.
Environment	Can occur in both Prod and QA cluster.
Possible Causes	Ingestion of unreasonable data.
	• Invalid sensor counted in the data engine.
	• Data Engine ETL failure.
Troubleshooting	• Check data from HBase table: query from cim database with impala-shell.
	• If avg, min, max values are not reasonable, check the sensor count from impala-shell.
	• If raw data has negative values, check with Device Engine/Extension for data input issue.
	Otherwise, send the API URL and raw data DB dump for that time period to the Data Engine team.
Verification	If unreasonable data is in Raw Data Table, try TQL live to double confirm.
	If sensor count is not right, check the sensor status.
Post Verification	If Device Engine provides unreasonable data, raise CDETS for Device Engine.
	Check sensor status, if still no issue found, raise CDETS for Data Engine.

## 6. ETL Workflow is Stuck

Symptom	ETL workflow takes too long for completion, say, more than a day.
Environment	Can occur in both Prod and QA cluster.
Possible Causes	If Data Engine runs V1.1.0.3rb1, Data engine will do small file merge periodically. It may cause Impala to get stuck. Impala memory shortage causes the query to get stuck.
Troubleshooting	If Data Engine runs V1.1.0.3rb1, upgrade data engine. for other versions, refer <i>How to check if Impala is stuck</i> in the FAQ section.
Verification	Open Cloudera Manager Web console -> Impala -> Queries: Check all query duration, it is expected that all queries would finish in 30 min.

7. ETL Workflow is Killed

G (	
Symptom	ETL workflow does not finish successfully, ETL state: killed
Environment	Can occur in both Prod and QA cluster.
Possible Causes	This may be caused due to RSH issue, Port issue or HAProxy.
	RSH Issue Error Logs:
	java.lang.Exception: Error in checking user& group for city cimdata from host data-01.novalocal at com.cisco.cim.oozie.action.CheckUserAction.main
	(CheckUserAction.java:53) at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method) at
	<pre>sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:57)</pre>
	at sun.reflect.DelegatingMethodAccessorImpl.invoke (DelegatingMethodAccessorImpl.java:43) at java.lang.reflect.Method.invoke(Method.java:606)
	at org.apache.oozie.action.hadoop.JavaMain.run(JavaMain.java:55) at org.apache.oozie.action.hadoop.LauncherMain.run(LauncherMain.java:64) at org.apache.oozie.action.hadoop.JavaMain.main(JavaMain.java:35) at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method) at
	<pre>sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:57)</pre>
	at sun.reflect.DelegatingMethodAccessorImpl.invoke (DelegatingMethodAccessorImpl.java:43) at java.lang.reflect.Method.invoke(Method.java:606)
	at org.apache.oozie.action.hadoop.LauncherMapper.map(LauncherMapper.java:234)
	<pre>at org.apache.hadoop.mapred.MapRunner.run(MapRunner.java:54) at org.apache.hadoop.mapred.MapTask.runOldMapper(MapTask.java:459) at org.apache.hadoop.mapred.MapTask.run(MapTask.java:343) at org.apache.hadoop.mapred.YarnChild\$2.run(YarnChild.java:164) at java.security.AccessController.doPrivileged(Native Method) at javax.security.auth.Subject.doAs(Subject.java:415) at org.apache.hadoop.security.UserGroupInformation.doAs (UserGroupInformation.java:1920) at org.apache.hadoop.mapred.YarnChild.main(YarnChild.java:158) Failing Oozie Launcher, Main class [org.apache.oozie.action.hadoop.JavaMain], main() threw exception,</pre>
	HAProxy Error
	Caused by: java.sql.SQLException: Could not open client transport with JDBC Uri: jdbc:hive2://cm-hue-01.novalocal:10000/;principal= hive/cm-hue-01.novalocal@CIM.IVSG.AUTH:java.net.ConnectException: Connection refused
Troubleshooting	Get ETL job log and check the errors in it.
	• If an RSH issue, run "update etl conf" action of Cim Tool Service.
	• If a Port issue, refer Issue 9 No route to host.
	If HAProxy is not working, restart HAProxy.

	if hosts get restarted by some unknown reason, make sure all dependent services are restarted.
Verification	Restart ETL after one ETL cycle, check the ETL status
Post Verification	If RSH is working fine and network also looks good, raise CDETS for Data Engine.

## **8.** Failure in ETL Creation

Chvironment       Can occur in both Prod and QA cluster.         Cossible Causes       This may be caused due to RSH failure.         RSH timeout Error example:       java.lang.Exception: Process timeout for command:sudo -u oozie rsh -l root data-01.novalocal "rm -rf /usr/local/etl/;mkdir -p /usr/local/etl/;chown -R oozie:oozie /usr/local/etl/"			
Possible Causes       This may be caused due to RSH failure.         RSH timeout Error example:       java.lang.Exception: Process timeout for command:sudo -u oozie rsh -l root data-01.novalocal "rm -rf /usr/local/etl/;mkdir -p /usr/local/etl/;chown -R oozie:oozie /usr/local/etl/"         at com.cisco.cim.oozie.util.CLITool.executeLinuxCommand(CLITool.java:66) at com.cisco.cim.oozie.util.CLITool.executeRSHCommandAsRootByOozie(CLITool.java:101) at com.cisco.cim.oozie.util.InitialOperation.main(InitialOperation.java:199) at com.cisco.cim.oozie.util.InitialOperation.main(InitialOperation.java:122)         Proubleshooting       • Run "update etl conf" action in Cim Tool Service.         • If RSH timeout exception is thrown, create ETL manually with command: #run in cm-hue host cd /opt/cloudera/parcels/cim_etl_oozie/oozie/./runoozie http://oozie-install-hostname-here:11000/oozie         • If RSH is still failing, install RSH manually.         //erification       Re-run ETL creation in cim-tool to see if ETL workflow is successfully created and running.	Symptom	Error is observed when ETL creation is run from cim-tool.	
RSH timeout Error example:         java.lang.Exception: Process timeout for command:sudo -u oozie rsh -l root         data-01.novalocal "rm -rf /usr/local/etl/;mkdir -p /usr/local/etl/;chown -R         oozie:oozie /usr/local/etl/"         at com.cisco.cim.oozie.util.CLITool.executeLinuxCommand(CLITool.java:66)         at         com.cisco.cim.oozie.util.CLITool.executeRSHCommandAsRootByOozie(CLITool.java:101)         at         com.cisco.cim.oozie.util.UserManager.syncKeytabFileToHost(UserManager.java:199)         at com.cisco.cim.oozie.util.InitialOperation.main(InitialOperation.java:122)         Proubleshooting         • Run "update etl conf" action in Cim Tool Service.         • If RSH timeout exception is thrown, create ETL manually with command:         #run in cm-hue host         cd /opt/cloudera/parcels/cim_etl_oozie/oozie/         ./runoozie http://oozie-install-hostname-here:11000/oozie         • If RSH is still failing, install RSH manually.         /erification       Re-run ETL creation in cim-tool to see if ETL workflow is successfully created and running.         /rost       If it is not a RSH issue, raise CDETS for Data Engine.	Environment	Can occur in both Prod and QA cluster.	
java.lang.Exception:       Process timeout for command:sudo -u oozie rsh -l root data-01.novalocal "rm -rf /usr/local/etl/;mkdir -p /usr/local/etl/;chown -R oozie:oozie /usr/local/etl/"	Possible Causes	This may be caused due to RSH failure.	
data-01.novalocal "rm -rf /usr/local/etl/;mkdir -p /usr/local/etl/;chown -R         oozie:oozie /usr/local/etl/"         at com.cisco.cim.oozie.util.CLITool.executeLinuxCommand(CLITool.java:66)         at         com.cisco.cim.oozie.util.CLITool.executeRSHCommandAsRootByOozie(CLITool.java:101)         at         com.cisco.cim.oozie.util.UserManager.syncKeytabFileToHost(UserManager.java:199)         at com.cisco.cim.oozie.util.InitialOperation.main(InitialOperation.java:122)         Proubleshooting         • Run "update etl conf" action in Cim Tool Service.         • If RSH timeout exception is thrown, create ETL manually with command:         #run in cm-hue host         cd /opt/cloudera/parcels/cim_etl_oozie/oozie/         ./runoozie http://oozie-install-hostname-here:11000/oozie         • If RSH is still failing, install RSH manually.         Verification       Re-run ETL creation in cim-tool to see if ETL workflow is successfully created and running.         Post       If it is not a RSH issue, raise CDETS for Data Engine.		RSH timeout Error example:	
at         com.cisco.cim.oozie.util.CLITool.executeRSHCommandAsRootByOozie(CLITool.java:101)         at         com.cisco.cim.oozie.util.UserManager.syncKeytabFileToHost(UserManager.java:199)         at com.cisco.cim.oozie.util.InitialOperation.main(InitialOperation.java:122)         Proubleshooting         • Run "update etl conf" action in Cim Tool Service.         • If RSH timeout exception is thrown, create ETL manually with command:         #run in cm-hue host         cd /opt/cloudera/parcels/cim_etl_oozie/oozie/         ./runoozie http://oozie-install-hostname-here:11000/oozie         • If RSH is still failing, install RSH manually.         Verification       Re-run ETL creation in cim-tool to see if ETL workflow is successfully created and running.         Post       If it is not a RSH issue, raise CDETS for Data Engine.		<pre>data-01.novalocal "rm -rf /usr/local/etl/;mkdir -p /usr/local/etl/;chown -R oozie:oozie /usr/local/etl/"</pre>	
at         com.cisco.cim.oozie.util.UserManager.syncKeytabFileToHost (UserManager.java:199)         at com.cisco.cim.oozie.util.InitialOperation.main (InitialOperation.java:122)         Proubleshooting       • Run "update etl conf" action in Cim Tool Service.         • If RSH timeout exception is thrown, create ETL manually with command:         #run in cm-hue host         cd /opt/cloudera/parcels/cim_etl_oozie/oozie/         ./runoozie http://oozie-install-hostname-here:11000/oozie         • If RSH is still failing, install RSH manually.         /erification         Re-run ETL creation in cim-tool to see if ETL workflow is successfully created and running.         Post			
com.cisco.cim.oozie.util.UserManager.syncKeytabFileToHost (UserManager.java:199) at com.cisco.cim.oozie.util.InitialOperation.main (InitialOperation.java:122)         Proubleshooting       • Run "update etl conf" action in Cim Tool Service.         • If RSH timeout exception is thrown, create ETL manually with command: #run in cm-hue host cd /opt/cloudera/parcels/cim_etl_oozie/oozie/ ./runoozie http://oozie-install-hostname-here:11000/oozie         • If RSH is still failing, install RSH manually.         Verification         Re-run ETL creation in cim-tool to see if ETL workflow is successfully created and running.         Post         If it is not a RSH issue, raise CDETS for Data Engine.		com.cisco.cim.oozie.util.CLITool.executeRSHCommandAsRootByOozie(CLITool.java:101)	
<ul> <li>If RSH timeout exception is thrown, create ETL manually with command: #run in cm-hue host cd /opt/cloudera/parcels/cim_etl_oozie/oozie/ ./runoozie http://oozie-install-hostname-here:11000/oozie If RSH is still failing, install RSH manually.</li> <li>Werification Re-run ETL creation in cim-tool to see if ETL workflow is successfully created and running.</li> <li>Post If it is not a RSH issue, raise CDETS for Data Engine.</li> </ul>		com.cisco.cim.oozie.util.UserManager.syncKeytabFileToHost(UserManager.java:199)	
<ul> <li>#run in cm-hue host cd /opt/cloudera/parcels/cim_etl_oozie/oozie/ ./runoozie http://oozie-install-hostname-here:11000/oozie</li> <li>If RSH is still failing, install RSH manually.</li> <li>Werification Re-run ETL creation in cim-tool to see if ETL workflow is successfully created and running.</li> <li>Post If it is not a RSH issue, raise CDETS for Data Engine.</li> </ul>	Troubleshooting	Run "update etl conf" action in Cim Tool Service.	
cd /opt/cloudera/parcels/cim_etl_oozie/oozie/         ./runoozie http://oozie-install-hostname-here:11000/oozie         • If RSH is still failing, install RSH manually.         Verification       Re-run ETL creation in cim-tool to see if ETL workflow is successfully created and running.         Post       If it is not a RSH issue, raise CDETS for Data Engine.		• If RSH timeout exception is thrown, create ETL manually with command:	
Verification       Re-run ETL creation in cim-tool to see if ETL workflow is successfully created and running.         Post       If it is not a RSH issue, raise CDETS for Data Engine.		cd /opt/cloudera/parcels/cim_etl_oozie/oozie/	
Post If it is not a RSH issue, raise CDETS for Data Engine.		• If RSH is still failing, install RSH manually.	
	Verification	Re-run ETL creation in cim-tool to see if ETL workflow is successfully created and running.	
	Post Verification	If it is not a RSH issue, raise CDETS for Data Engine.	

# 9. No Route to Host

Symptom	Such error may occur while creating ETL, or from workflow log as shown below:
	2018-05-04 10:14:41,664 WARN [main] org.apache.hadoop.mapred.YarnChild: Exception running child : java.net.NoRouteToHostException: No Route to Host from data-01.novalocal/10.10.22.35 to data-03.novalocal: 39025 failed on socket timeout exception:
	java.net.NoRouteToHostException: No route to host; For more details see: http://wiki.apache.org/hadoop/NoRouteToHost
Environment	Can occur in both Prod and QA cluster.

I

Possible Causes	Network connectivity issue or port is not opened.
Troubleshooting	• Check the list of open ports and related configuration, for example
	iptables rules or /proc/sys/net/ipv4/ip_local_port_range
	<pre>#ip_local_port_range value should be same as this cat /proc/sys/net/ipv4/ip_local_port_range 62000 64000</pre>
Verification	Create or restart ETL to check if ETL creation or workflow finishes successfully.
Post Verification	If there is no network or configuration issue identified, raise CDETS for Data Engine.

## 10. LeaderNotAvailable Exception

Symptom	Kafka broker stops data ingestion and Kafka consumer throws LeaderNotAvailable Exception.
Environment	Can occur in both Prod and QA cluster.
Possible Causes	One node of Kafka Broker restarts with unclean leader election causing the issue.
	It is a known issue for Kafka 0.11.x
Troubleshooting	Restart Kafka cluster for recovery.
	• Enable unclean leader election in the Kafka configuration.
Verification	Check Kafka ingestion diagram in CM after Kafka has restarted, the data should start coming in.
Post Verification	If configuration is fine, raise CDETS for Data Engine.

## 11. DoNotRetry IOException

Symptom	• When data is queried from some external tables of cim database, such errors may be encountered.
	* The external table are tables mapped from HBase, and store too many data records.
	Error Log:
	DoNotRetryIOException: Failed after retry of OutOfOrderScannerNextException: was there a rpc timeout? CAUSED BY: OutOfOrderScannerNextException: org.apache.hadoop.HBase.exceptions.OutOfOrderScannerNextException:
	Expected nextCallSeq: 1 But the nextCallSeq got from client: 0; request=scanner_id: 572818 number_of_rows: 1024 close_scanner: false next_call_seq: 0 client_handles_partials: true client_handles_heartbeats: true track_scan_metrics: false renew: false at

	org.apache.hadoop.HBase.regionserver.RSRpcServices.scan (RSRpcServices.java:2443)
Environment	Can occur in both Prod and QA cluster.
<b>Possible Causes</b>	RPC timeout due to configuration error.
Troubleshooting	Do following configuration update:
	1. Cloudera Manager -> Impala -> Configuration -> hbase.rpc.timeout -> 3 Minutes. (The default value is 3 seconds)
	2. If all hosts memory size are 32GB. Cloudera Manager -> HBase -> Configuration -> Java Heap Size of HBase RegionServer in Bytes -> 4 GB.
	3. Cloudera Manager -> HBase -> Configuration -> HBase Client Scanner Caching -> 50
	4. Cloudera Manager -> HBase -> Configuration -> HBase Client Scanner Timeout -> 2 min
	5. Cloudera Manager-> Impala -> Configuration -> HBase RPC timeout -> 2 min
Verification	Rerun the same SQL to check if it gets completed successfully.
Post Verification	If configuration is fine, raise CDETS for Data Engine.

## **12.** Memory Limit Exceeded

Symptom	HTTP API fails.
	http-api.log file:
	Internal server error: Memory limit exceeded Codegen failed to reserve '5555712' bytes for optimization
	Internal server error: Memory limit exceeded Failed to allocate tuple buffer
	Internal server error: Memory limit exceeded
	Internal server error: Memory limit exceeded The memory limit is set too low to initialize spilling operator (id=7).
	The minimum required memory to spill this operator is 264.00 MB.
Environment	Can occur in both Prod and QA cluster.
Possible Causes	The Impala memory limit is set too low.
Troubleshooting	• Make sure Impala Daemon Memory limit is not less than 3.5 GB.
	• If less, increase to at least 3.5 GB. If in 32 G memory configuration, set it to 8G

	• Restart Impala service from CM.
	Get in touch with Data Engine team to get some checks done before restarting Impala service.
Verification	Rerun the same SQL to check if it gets completed successfully.
Post Verification	If configuration is fine, raise CDETS for Data Engine.

## **13.** Clock Offset Warn / NTP issue

Symptom	Error Message is seen in the logs:
	E0320 07:06:16.662436 4178 authentication.cc:160] SASL message (Kerberos (internal)): GSSAPI Error: Unspecified GSS failure. Minor code may provide more information (Clock skew too great)
	Status summary shows 'Warning' for Clock Offset and one of the Hosts in bad health.
Environment	Can occur in both Prod and QA cluster.
Possible Causes	NTP service is not working correctly in the host.
Troubleshooting	Restart NTPD service in the bad health host.
Verification	Check the timeskew in the host. Then check if Cluster is recovered from the bad health. If not, restart the cluster.
Post Verification	If recovered from NTP service interruption, no action required

## 14. HBase Master no active instance warn

Symptom	While upgrading, HBase hits such an error
	Bad: Master summary: cm-hue-01.novalocal(Availability:Unknown,Health:Good), name-01.novalocal(Availability:Unknown,Health:Good). This health test is bad because the Service Monitor did not find an active Master.
Environment	Can occur in both Prod and QA cluster.
Possible Causes	Configuration Error.
Troubleshooting	Perform following configuration changes:
	• CM -> zookeeper -> configuration -> search "Enable auto-creation of data directories"
	• Check on it -> click "save changes".
	• Restart zookeeper instances one after other.
Verification	HBase should recover from bad health.

Post Verification	If configuration setting is correct, raise CDETS for Data Engine.
-------------------	---

**15.** Kafka storage occupying too much space

Symptom	var/local/kafka occupying too much disk space on one specific node.
Environment	Can occur in both Prod and QA cluster.
Possible Causes	Two Kafka topic replica in the same node cause the over capacity.
Troubleshooting	Enable Kafka broker service on one more nodes and follow ' <i>How to move Kafka partion from one node to another</i> ' in FAQ section to migrate one replica to that node.
Verification	HBase should recover from bad health.
Post Verification	Check /var/local/kafka size on all the nodes running Kafka broker.

### 16. No Write Operations on Hbase Table

Symptom	Alert received from monitoring script: "no write operations on HBase Table".
Environment	Can occur in any of the Prod and QA cluster with low disk IO performance.
Possible Causes	Due to the low disk write speed, Hbase write cannot cache up.
Troubleshooting	Add flume log configuration
	log4j.appender.RFA.Threshold=INFO log4j.logger.com.cisco=DEBUG,ciscoFlm
	Add following flume configuration
	<pre>Run "sed -Ei 's/(.*).type = com.cisco.cim.flume.sink.CimHBaseSink/&amp;\n\1.batchSize = 10/g' /opt/cloudera/parcels/cim_flume_plugin/conf/flume.conf" in all hosts, and restart flume service</pre>
	Add following Kafka configuration
	<pre>run "sed -i '/kafka.batchSize/ s/.*/kafka.batchSize = 50/' /opt/cloudera/parcels/cim_flume_plugin/conf/generateConfig.properties" in all hosts, then execute action "update flume conf" in cim-tool service and restart flume agent.</pre>
Verification	No more alert should be observed after restarting the Flume.
Post Verification	If issue still persists, raise a CDET for Data Engine.

# **Frequently Asked Questions**

• 1. How to check Impala status ?

From Cloudera Health status, Impala may appear good(Green), but sometimes query from impala may get stuck. Here are 2 SQL, we can test with it:

```
Select count(*) from cim.locations;
select count(*) from cimdata.datelist half;
```

#### • 2. How to turn off false warning?

For CDH 5.11.0, Impala may throw a false warn, Impala Bad Health: Query Monitoring Status

This can be turned off as follows:

- CM -> Impala-> Configuration -> uncheck "Impala Daemon Query Collection Status Health Check",
- Save the update.

#### • 3. How to check if Impala is stuck?

• CM -> Impala -> Queries: Check Impala Query status, there should not any running SQL whose duration is over 1hour/day.

Restart Impala service from CM, should work for this issue.

• Get in touch with Data engine team to do some checks before restarting impala service.

#### • 4. What to do if many Services are in bad health ?

If there are many services in bad health status in Cloudera Manager, check Host's status.

When Host is in bad status, service instances running in it will be affected too, for example NTP issue

#### • 5. What to do if only one service is in bad health status in Cloudera Manager?

- CM -> click Name of bad health service -> click "show" in "Health History" table list.
- Send warn/error info to Data engine team.

#### • 6. How to query with Impala shell?

```
kinit -kt /usr/local/etl/etl.keytab impala/cm-hue-01.novalocal@CIM.IVSG.AUTH
```

```
impala-shell -k
```

#some messages output here .....

[cm-hue-01.novalocal:21000] > use your-own-database-name;

Query: use cim

```
[cm-hue-01.novalocal:21000] >
run your SQL query from here.
```

#### • 7. How to check Device Engine status?

Check device engine status as below:

#get device engine registered info of edge-server

```
[root@edge-01 ~]# curl http://127.0.0.1/api/v1.0/config/devengine
[{
   "schema": "default",
        "timezone": "UTC",
        "active": "true",
        "timestampUtc": 1501035136433,
        "store": "true",
        "type": "DeviceEngine",
```

```
"ttl": "false",
    "multiThread": "true",
    "review": "false",
    "name": "dmz-crdc-01",
    "id": "dmz-crdc-01",
    "category": "source",
    "config": {
        "liveQuery": "false",
        "urlPath": "",
        "version": "",
        "timeout": 180000,
        "wsUrlPath": "",
        "port": "9011",
        "wss": "false",
        "namespace": "india.com",
        "host": "10.194.185.244",
        "interval": 1800000,
        "https": "false",
        "liveUrlEndPoint": "",
        "wsUrlEndPoint": "fid-CIMDataEngineWsInterface"
    },
    "group": "default",
    "timestamp": "2017-07-26T02:12:16.433Z",
    "status": {
        "outCount": "4155216",
        "inCount": "25592458"
    }
}][root@edge-01 ~]#
```

#in the above example, wss is false, means ssl is disabled, webscoket url is:

```
#ws://10.194.185.244:9011/ fid-CIMDataEngineWsInterface
```

#if wss enabled, websocket url is:

#wss://10.194.185.244:9011/ fid-CIMDataEngineWsInterface

#Now user can debug websocket server with url.

#### • 8. How to check SDP server availability?

Edge as a SDP client will query locations data from SDP server periodically.

If you get an error, such as HTTP API Location ID Error, it generally means SDP is stuck.

We can check locations which get stored in edge with this API:

```
curl -g http://localhost//api/v1.0/config/location
#it's output example:
{
    "name": "child2150121",
   "active": "true",
    "id": "singapore.com 11504",
    "timestampUtc": 1532063266548,
    "category": "metadata",
    "type": "location",
    "config": {
        "entityType": "k112150114",
        "entityName": "child2150121",
        "hierarchy": [
            {
                "entityType": "k112150114",
                "entityName": "parent12150120",
                "tenantId": "singapore.com",
                "entityId": "11502"
```

},

• If there is no output, it means SDP is not available.

- User should check SDP server status.
- 9. How to get UTC timestamp of hour timeid from hourly table?
- Save below as one python file, for example named as str2utctimestamp.py
- run : python str2utctimestamp.py 2018062014,

you will get value 1529503200000, it is a utc timestamp of 2018062014

#### • 10. How to move Kafka partition from one node to another?

If you need to move one partition from DataName (218) to Datacmhue (217), post which it will be distributed between Data02 (219) and Datacmhue (217), following steps should be performed:

1. cd /opt/cloudera/parcels/KAFKA/lib/kafka/bin

2. /kafka-reassign-partitions.sh --zookeeper data-01.novalocal:2182/kafka --topics-to-move-json-file new-topic.json --broker-list "217,218,219,220" --generate

3. Update the new-topic.json to :-

```
{"version":1,"partitions":[{"topic":"cim","partition":0,"replicas":[220,218],
"log_dirs":["any","any"]},{"topic":"cim","partition":1,"replicas":[217,219],
"log_dirs":["any","any"]}]
```

4. kafka-reassign-partitions --zookeeper data-01.novalocal:2182/kafka --reassignment-json-file new-topic.json -execute

5. ./kafka-reassign-partitions.sh --zookeeper data-01.novalocal:2182/kafka --reassignment-json-file new-topic.json --verify

I