Tables and their source types in Azure HDInsight Log Analytics

AIM: This documentation contains the names of tables and their log types present in each HDInsight cluster and its different node types(head node, worker node and zookeeper node).

Any user can use this to find out the names of all the main tables present in the log analytics workspace and their sources/log types.

CONTENTS

1. Tables in Spark Cluster
2. Source types(Log Types) of main tables in spark cluster
3. Tables in Interactive Query Cluster
4. Source types(Log Types) of main tables in Interactive Query cluster
5. Tables in HBase Cluster
6. Source types(Log Types) of main tables in HBase cluster
7. Tables in Hadoop Cluster
8. Source types(Log Types) of main tables in Hadoop cluster

SPARK CLUSTER

TABLES IN **SPARK CLUSTER**

There are 7 tables inside spark cluster which are:

1. HDInsight Ambari Cluster Alerts
2. HDInsight Ambari System Metrics
3. HDInsight Hadoop And Yarn Logs
4. HDInsight Hadoop And Yarn Metrics
5. HDInsight Security Logs
6. HDInsight Spark Logs
7. HDInsight Oozie Logs

Configuration File mapping for Spark Cluster Tables

|  |  |  |
| --- | --- | --- |
| S.no | Table Name | Configuration File Name |
| 1. | HDInsightAmbariClusterAlerts | AmbariClusterAlertsOutput.conf |
| 2. | HDInsightAmbariSystem Metrics | AmbariSystemMetricsOutput.conf |
| 3. | HDInsightHadoopAndYarnLogs | YarnLogOutput.conf |
| 4. | HDInsight Security Logs | SecurityLogOutput.conf |
| 5. | HDInsight Spark Logs | SparkLogOutput.conf |
| 6. | HDInsight Hadoop And Yarn Metrics | YarnMetricsOutput.conf |
| 7. | HDInsight Oozie Logs | OozieLogOutput.conf |

Note: These configuration files are fluentd config files which collect data from various sources and tag them with a common tag, for adding them into one table.

These configuration files are responsible for populating data inside the source table.

DIFFERENT LOG TYPES(SOURCES) INSIDE **SPARK** TABLES

|  |  |  |
| --- | --- | --- |
| S.no | Table Name | Log Types |
| 1. | HDInsightAmbariCluster Alerts | No log types |
| 2. | HDInsightAmbariSystem Metrics | No log types |
| 3. | HDInsightHadoopAnd YarnLogs | Head Node: MRJobSummary, ResourceManager, TimelineServer  Worker Node: NodeManager |
| 4. | HDInsightSecurityLogs | AmbariAuditLog, AuthLog |
| 5. | HDInsightSparkLogs | Head Node: JupyterLog, LivyLog, SparkThriftDriverLog  Worker Node: SparkExecutorLog, SparkDriverLog |
| 6. | HDInsightHadoopAnd YarnMetrics | No log types |
| 7. | HDInsightOozieLogs | Oozie |

NOTE:

The data from different log types add up to make one table. User can disable log types which they do not want.

The tables which do not have any log types means, they have 0 log types.

Example of disabling a particular log type

HDInsightSparkLogs: JupyterLog SparkThriftDriverLog, HDInsightHadoopAndYarnMetrics –disable

The log types are separated using space and Tables are separated using comma. The action is specified using -- like this, --disable or --enable.

INTERACTIVE QUERY CLUSTER

TABLES IN **INTERACTIVE QUERY** CLUSTER

There are 7 tables inside interactive hive cluster which are:

1. HDInsight Ambari Cluster Alerts
2. HDInsight Ambari System Metrics
3. HDInsight Hadoop And Yarn Logs
4. HDInsight Hadoop And Yarn Metrics
5. HDInsight Hive And LLAP logs
6. HDInsight Hive And LLAP metrics
7. HDInsight Hive Tez App Stats
8. HDInsight Security Logs

Configuration File mapping for Interactive Query Cluster Tables

|  |  |  |
| --- | --- | --- |
| S.no | Table Name | Configuration File Name |
| 1. | HDInsightAmbariClusterAlerts | AmbariClusterAlertsOutput.conf |
| 2. | HDInsightAmbariSystem Metrics | AmbariSystemMetricsOutput.conf |
| 3. | HDInsightHadoopAndYarnLogs | YarnLogOutput.conf |
| 4. | HDInsightHadoopAndYarnMetrics | YarnMetricsOutput.conf |
| 5. | HDInsight Hive And LLAP logs | InteractiveHiveLogOutput.conf |
| 6. | HDInsight Hive And LLAP metrics | InteractiveHiveMetricsOutput.conf |
| 7. | HDInsight Hive Tez App Stats | AppStatsHiveTezOutput.conf |
| 8. | HDInsight Security Logs | SecurityLogOutput.conf |

Note: These configuration files are fluentd config files which collect data from various sources and tag them with a common tag, for adding them into one table.

These configuration files are responsible for populating data inside the source table.

DIFFERENT LOG TYPES(SOURCES) INSIDE **INTERACTIVE QUERY** TABLES

|  |  |  |
| --- | --- | --- |
| S.no | Table Name | Log Types |
| 1. | HDInsightAmbariClusterAlerts | No other log types |
| 2. | HDInsightAmbariSystem Metrics | No other log types |
| 3. | HDInsightHadoopAndYarnLogs | **Head Node**: MRJobSummary, ResourceManager, TimelineServer  **WorkerNode:** NodeManager |
| 4. | HDInsightHadoopAndYarnMetrics | No log types |
| 5. | HDInsight Hive And LLAP logs | Head Node: InteractiveHiveHSILog,  InteractiveHiveMetastoreLog,  ZeppelinLog |
| 6. | HDInsight Hive And LLAP metrics | No log types |
| 7. | HDInsight Hive Tez App Stats | No log types |
| 8. | HDInsight Security Logs | **Head Node:** AmbariAuditLog, AuthLog  **Zookeeper Node, Worker Node:** AuthLog |

NOTE:

The data from different log types add up to make one table. User can disable log types which he/she does not want.

**Example of disabling a particular log type**

HDInsightHBaseLogs: PhoenixServerLog HDFSGarbageCollectorLog, HDInsightHadoopAndYarnMetrics –disable

The log types are separated using space and Tables are separated using comma. The action is specified using -- like this, --disable or --enable.

HBASE CLUSTER

TABLES IN **HBASE CLUSTER**

There are 7 tables inside spark cluster which are:

1. HDInsight Ambari Cluster Alerts
2. HDInsight Ambari System Metrics
3. HDInsight Hadoop And Yarn Logs
4. HDInsight Hadoop And Yarn Metrics
5. HDInsight HBase Logs
6. HDInsight HBase Metrics
7. HDInsight Security Logs

Configuration File mapping for Spark Cluster Tables

|  |  |  |
| --- | --- | --- |
| S.no | Table Name | Configuration File Name |
| 1. | HDInsightAmbariClusterAlerts | **Head Node, Worker Node, Zookeeper Node**: AmbariClusterAlertsOutput.conf |
| 2. | HDInsightAmbariSystem Metrics | **Head Node, Worker Node, Zookeeper Node:** AmbariSystemMetricsOutput.conf |
| 3. | HDInsightHadoopAndYarnLogs | **Head Node, Worker Node:** YarnLogOutput.conf |
| 4. | HDInsightSecurityLogs | **Head Node, Worker Node, Zookeeper Node**: SecurityLogOutput.conf |
| 5. | HDInsightHBaseLogs | **Head Node, Worker Node, Zookeeper Node**: HBaseLogOutput.conf |
| 6. | HDInsightHBaseMetrics | **Worker Node, Zookeeper Node**: HBaseMetricsOutput.conf |
| 7. | HDInsightHadoopAndYarn Metrics | **Head Node, Worker Node:** YarnMetricsOutput.conf |

Note: These configuration files are fluentd config files which collect data from various sources and tag them with a common tag, for adding them into one table.

These configuration files are responsible for populating data inside the source table.

DIFFERENT LOG TYPES(SOURCES) INSIDE **HBASE** TABLES

|  |  |  |
| --- | --- | --- |
| S.no | Table Name | Log Types |
| 1. | HDInsightAmbariClusterAlerts | No other log types |
| 2. | HDInsightAmbariSystem Metrics | No other log types |
| 3. | HDInsightHadoopAndYarnLogs | **Head Node**: MRJobSummary, ResourceManager, TimelineServer  **WorkerNode:** NodeManager |
| 4. | HDInsightSecurityLogs | **Head Node:** AmbariAuditLog, AuthLog  **Worker Node:** AuthLog  **Zookeper Node:** AuthLog |
| 5. | HDInsightHBaseLogs | **Head Node**: HDFSGarbageCollectorLog, HDFSNameNodeLog  **WorkerNode:** PhoenixServerLog, HBaseRegionServerLog, HBaseRestServerLog  **Zookeeper Node:** HBaseMasterLog |
| 6. | HDInsightHBaseMetrics | No log types |
| 7. | HDInsightHadoopAndYarn Metrics | No log types |

NOTE:

The data from different log types add up to make one table. User can disable log types which he/she does not want.

**Example of disabling a particular log type**

HDInsightHBaseLogs: PhoenixServerLog HDFSGarbageCollectorLog, HDInsightHadoopAndYarnMetrics –disable

The log types are separated using space and Tables are separated using comma. The action is specified using -- like this, --disable or --enable.

HADOOP CLUSTER

TABLES IN **HADOOP CLUSTER**

There are 7 tables inside spark cluster which are:

1. HDInsight Ambari Cluster Alerts
2. HDInsight Ambari System Metrics
3. HDInsight Hadoop And Yarn Logs
4. HDInsight Hadoop And Yarn Metrics
5. HDInsight Hive And LLAP Logs
6. HDInsight Hive And LLAP Metrics
7. HDInsight Security Logs

Configuration File mapping for Hadoop Cluster Tables

|  |  |  |
| --- | --- | --- |
| S.no | Table Name | Configuration File Name |
| 1. | HDInsightAmbariClusterAlerts | AmbariClusterAlertsOutput.conf |
| 2. | HDInsightAmbariSystem Metrics | AmbariSystemMetricsOutput.conf |
| 3. | HDInsightHadoopAndYarnLogs | YarnLogOutput.conf |
| 4. | HDInsightHadoopAndYarnMetrics | YarnMetricsOutput.conf |
| 5. | HDInsight Hive And LLAP Logs | HiveLogOutput.conf |
| 6. | HDInsight Hive And LLAP Metrics | HiveMetricsOutput.conf |
| 7. | HDInsight Security Logs | SecurityLogOutput.conf |

Note: These configuration files are fluentd config files which collect data from various sources and tag them with a common tag, for adding them into one table.

These configuration files are responsible for populating data inside the source table.

DIFFERENT LOG TYPES(SOURCES) INSIDE **HADOOP** TABLES

|  |  |  |
| --- | --- | --- |
| S.no | Table Name | Log Types |
| 1. | HDInsightAmbariClusterAlerts | No log types |
| 2. | HDInsightAmbariSystem Metrics | No log types |
| 3. | HDInsightHadoopAndYarnLogs | **Head Node:** MRJobSummary, ResourceManager, TimelineServer  Worker Node: NodeManager |
| 4. | HDInsightHadoopAndYarnMetrics | No Log Types |
| 5. | HDInsight Hive And LLAP Logs | **Head Node:** HiveMetastoreLog, HiveServer2Log, WebHcatLog |
| 6. | HDInsight Hive And LLAP Metrics | No log types |
| 7. | HDInsight Security Logs | Head Node: AmbariAuditLog, AuthLog  Zookeeper Node: AuthLog |

NOTE:

The data from different log types add up to make one table. User can disable log types which he/she does not want.

The tables which do not have any log types means, they have 0 log types.

Example of disabling a particular log type

HDInsightSparkLogs: JupyterLog SparkThriftDriverLog, HDInsightHadoopAndYarnMetrics –disable

The log types are separated using space and Tables are separated using comma. The action is specified using -- like this, --disable or --enable.