

Monitoring Services

Cloudera Manager's Service Monitoring feature monitors dozens of service health and performance services and role instances running on your cluster.

Service Monitoring includes the following functions:

- Presents health and performance data in a variety of formats, including interactive charts.
- Monitors metrics against configurable thresholds.
- Generates events related to system and service health and critical log entries and makes them available for searching and alerting.
- Maintains a complete record of service-related actions and configuration changes. The following topics describe how to monitor the services and role instances installed on your cluster:

Monitoring Service Status

From a service page, you can monitor the status of services, manage services and roles, and add services. From a service page, you can:

- Monitor the status of the services running on your clusters.
- Manage the services and roles in your clusters.
- Add new services.
- Access the client configuration files generated by Cloudera Manager that enable Hadoop client users to work with the HDFS, MapReduce, HBase, and YARN services you added. (These configuration files are normally deployed automatically when you install a cluster or add a service).
- View the maintenance mode status of a cluster.

You can also pull down a menu from an individual service name to go directly to one of the tabs: Status, Instances, Commands, Configuration, Audits, or Charts Library tabs.

Viewing the URLs of the Client Configuration Files

To allow Hadoop client users to work with the services you created, Cloudera Manager generates configuration files that contain the relevant configuration files with the settings from your services.

About this task

These files are deployed automatically by Cloudera Manager based on the services you have installed. You can add a service, or when you add a Gateway role on a host. You can manually download and distribute the configuration files to the users of a service, if necessary.

Procedure

1. Go to the Status page for a cluster. (Clusterscluster name).
2. Click ActionsView Client Configuration URLs. The Actions button is not enabled if you are viewing status for a point of time in the past. A pop-up window displays links to the client configuration zip files.
3. Click a link to download a zip file.

Viewing the Status of a Service Instance

You can view the status of a service instance from the Status page or the Clusters menu.

Procedure

1. Click the Cloudera Manager logo to go to the Home page.
2. Access the service.
 - In the Status tab, select ClusterName ServiceName.
 - Select Clusters ClusterName ServiceName.

Results

The Status page opens. On the Status page you can view a variety of information about a service's performance.

Viewing the Health and Status of a Role Instance

You can view the health and status of a role instance by clicking the role instance under the Role Counts link. If there is just one instance of this role, this opens the Status tab for the role instance. If there are multiple instances of a role, clicking the role link under Role Counts will open the Status page for the service, showing instances of the role type you have selected.

If you are viewing a point in time in the past, the Role Count links will be greyed out, but the behavior will depend on whether historical data is available for the role instance.

Related Information

Viewing Role Instance Status

Viewing the Maintenance Mode Status of a Cluster

For any cluster, you can view the components (service, roles, or hosts) that are in maintenance mode.

About this task

Minimum Required Role: Cluster Administrator (also provided by Full Administrator)

Procedure

- From the Cloudera Manager Home page, select the cluster that you want to view the maintenance mode status for.
- Click Actions View Maintenance Mode Status.... This pops up a dialog box that shows the components in your cluster that are in maintenance mode, and indicates which are in effective maintenance mode as well as those that have been explicitly placed into maintenance mode. From this dialog box you can select any of the components shown there and remove them from maintenance mode.

If individual services are in maintenance mode, you will see the maintenance mode icon next to the service name and a button for that service.

Note: The Actions button is not enabled if you are viewing status for a point of time in the past.

Viewing Service Status.

You can view a summary of the status for each service.

Procedure

- Click the Cloudera Manager logo to go to the Home page.
- Access the status summary.
 - In the Home Status tab, if the cluster is displayed in full form, click ServiceName in a ClusterName table.
 - In the Home Status tab, click ClusterName and then click ServiceName.
 - Select Clusters ClusterName ServiceName. For all service types there is a Status Summary that shows, for each configured role, the overall status and health of the role instance(s). Note: Not all service types provide complete monitoring and health information. Hive, Hue, Oozie, Solr, and YARN only provide the basic Status Summary.

Each service that supports monitoring provides a set of monitoring properties where you can view health tests and events, and set thresholds for tests and modify thresholds for the status of the service. The HDFS, MapReduce, HBase, ZooKeeper, and Flume services also provide additional information of service-specific metrics, health test results, health history, and a set of charts that provide metrics of interest.

Viewing Past Status

You can expand the Time Range Selector to view historical health, status, and chart data. The health and status information on the Status page represents the state of the service or role at a point in time. The charts (and the Logs and Events under Diagnostics) represent the time range selected by the Time Range Selector, which defaults to the past 30 minutes. You can view health, status, and charts by expanding the Time Range Selector. Click the mini line chart under "admin" and move the time marker to a point in the past.

When you move the time marker to a point in the past (for services and roles that support historical status), the Status page updates to the time selected. A Now button () allows you to quickly return to the current status of the service. The Actions menu is disabled while viewing a past status to ensure that you can only view outdated status information.

Related Information

Time Line

Status Summary.

The Status Summary shows the status of each service instance being managed by Cloudera Manager. Even services such as Hue, Oozie, or YARN (which are not monitored by Cloudera Manager) show a status summary. The overall status for a service is a roll-up of the health test results for the service. The Status can be:

Table 2: Status

Indicator	Status	Description
Started with outdated configuration		

For a service, this indicates the service is running, but at least one of its roles is running with a configuration that does not match the current configuration settings in Cloudera Manager. For a role, this indicates a configuration change has been made that requires a restart, and that restart has not yet occurred.

Starting The entity is starting up but is not yet running.

Stopping The entity is stopping but has not stopped yet.

Stopped The entity is stopped, as expected.

Down The entity is not running, but it is expected to be running.

History not available Cloudera Manager is in historical mode, and the entity does not have health monitoring support. This is the case for services other than HDFS, MapReduce and HBase such as ZooKeeper, Oozie, and Hue.

None The entity does not have a status. For example, it is not something that can be running and it cannot have health. Examples are the HDFS Balancer (which runs from the HDFS Rebalance action) or Gateway roles. The Start and Stop commands are not applicable to these instances.

Good health The entity is running with good health. For a specific health test, the returned normal or within the acceptable range. For a role or service, this means all health tests for that role or service are Good.

Concerning health The entity is running with concerning health. For a specific health test, indicates a potential problem. Typically this means the test result has gone above (or below) a configured Warning threshold. For a role or service, this means that at least one health test is Concerning.

Bad health The entity is running with bad health. For a specific health test, the test failed. The returned result indicates a serious problem. Typically this means the test result has gone above (or below) a configured Critical threshold. For a role or service, this means that at least one health test is Bad.

Disabled health The entity is running, but all of its health tests are disabled.

Unknown health The status of a service or role instance is unknown. This can occur for a number of reasons, such as the Service Monitor is not running, or connectivity to the Agent doing the health monitoring has been lost.

To see the status of one or more role instances, click the role type link under Status Summary. For a single instance of the role type, the link directs you to the Status page of the role instance.

If there are multiple role instances (such as for DataNodes, TaskTrackers, and RegionServers), clicking the role type link directs you to the Role Instances page for that role type. Click on each instance, under Role Instances, to view the corresponding Status page.

To display the results for each health test that applies to this role type, expand the Health Tests section. You can expand Good Health, Warnings, Bad Health, or Disabled Health. Health test results that have failed or are unavailable appear as unavailable.

Related Information

Stale Configurations

Viewing Role Instance Status

Service Summary

Some services (specifically HDFS, MapReduce, HBase, Flume, and ZooKeeper) provide additional information about their operation and performance. These are shown in a Summary panel at the left side of the service page. The contents of this panel depend on the service:

- The HDFS Summary shows disk space usage.
- The MapReduce Summary shows statistics on slot usage, jobs and so on.
- The Flume Summary provides a link to a page of Flume metric details.
- The ZooKeeper Summary provides links to the ZooKeeper role instances (nodes) as well as Zxid information if you have a ZooKeeper Quorum (multiple ZooKeeper servers). For example:

Other services such as Hue, Oozie, Impala, and Cloudera Manager itself, do not provide a Summary panel.

Health Tests and Health History

The Health Tests panel shows health test results in an expandable and collapsible list, typically showing the test name, the test status, and the metrics that the test returned. You can Expand All or Collapse All from the links at the top of the panel.

The Health Tests and Health History panels appear for HDFS, MapReduce, HBase, Flume, Impala, and Hive. They do not appear for the Cloudera Manager Service. Other services such as Hue, Oozie, and YARN do not provide a Health Tests panel.

- The color of the text (and the background color of the field) for a Health Test result indicates the status of the results. The tests are sorted by their health status – Good, Concerning, Bad, or Disabled. The entries are collapsed by default. Click the arrow to the left of an entry to expand the entry and display further information.
- Clicking the Details link for a health test displays further information about the test, such as the meaning of the test and its possible results, suggestions for actions you can take or how to make configuration changes related to the test. The help text may include a link to the relevant monitoring configuration section for the service.
- In the Health Tests panel:
 - Clicking displays the lists of health tests that contributed to the health test.
 - Clicking the Details link displays further information about the health test.
- In the Health History panel:
 - Clicking displays the lists of health tests that contributed to the health history.
 - Clicking the Show link moves the time range to the historical time period. Related Information Configuring Monitoring Settings

Viewing Service Instance Details

You can view service instance details such as the name of the role instance, the host on which it is running, its rack assignment, and more.

Procedure

1. Do one of the following:
 - In the HomeStatus tab, if the cluster is displayed in full form, click ServiceName in a ClusterName table.
 - In the HomeStatus tab, click ClusterName and then click ServiceName.
 - Select ClustersClusterNameServiceName.
2. Click the Instances tab on the service's navigation bar. This shows all instances of all role types configured for the selected service.

Results

The Instances page displays the results of the configuration validation checks it performs on this service.

Note: The information on this page is always the Current information for the selected service. The page does not support a historical view: thus, the Time Range Selector is not available.

The information on this page shows:

- The name of the role instance. Click the name to view the role status for that role.
- The host on which it is running. Click the hostname to view the host status details for the host.
- The rack assignment.
- The status. A single value summarizing the state and health of the role instance.
- Whether the role is currently in maintenance mode. If the role has been set into maintenance mode explicitly, you will see the following icon (). If it is in effective maintenance mode due to the service or its host having been

set into maintenance mode, the icon will be this ().

- Whether the role is currently decommissioned.

Role Instance Reference

The following tables contain reference information on the status, role state, and health of a role instance.

Table 3: Status

Indicator	Status	Description
Started with outdated configuration	outdated	

For a service, this indicates the service is running, but at least one of its roles is running with outdated configuration.

with a configuration that does not match the current configuration settings in Cloudera Manager. For a role, this indicates a configuration change has been made that requires a restart, and that restart has not yet occurred.

Starting The entity is starting up but is not yet running.

Stopping The entity is stopping but has not stopped yet.

Stopped The entity is stopped, as expected.

Down The entity is not running, but it is expected to be running.

History not available Cloudera Manager is in historical mode, and the entity does not have health monitoring support. This is the case for services other than HDFS, MapReduce and HBase such as ZooKeeper, Oozie, and Hue.

None The entity does not have a status. For example, it is not something that can be running and it cannot have health. Examples are the HDFS Balancer (which runs from the HDFS Rebalance action) or Gateway roles. The Start and Stop commands are not applicable to these instances.

Good health The entity is running with good health. For a specific health test, the returned result is normal or within the acceptable range. For a role or service, this means all health tests for that role or service are Good.

Concerning health The entity is running with concerning health. For a specific health test, the returned result indicates a potential problem. Typically this means the test result has gone above (or below) a configured Warning threshold. For a role or service, this means that at least one health test is Concerning.

Bad health The entity is running with bad health. For a specific health test, the test failed. The returned result indicates a serious problem. Typically this means the test result has gone above (or below) a configured Critical threshold. For a role or service, this means that at least one health test is Bad.

Disabled health The entity is running, but all of its health tests are disabled.

Unknown health The status of a service or role instance is unknown. This can occur for a number of reasons, such as the Service Monitor is not running, or connectivity to the Agent doing the health monitoring has been lost.

Viewing Role Instance Status

You can view the status for a role instance.

Procedure

1. Select a service instance to display the Status page for that service.
2. Click the Instances tab.
3. From the list of roles, select one to display that role instance's Status page.

The Actions Menu.....

The Actions menu provides a list of commands relevant to the role type you are viewing. These

include Stopping, Starting, or Restarting the role instance, accessing the Web UI for the role instance, and other commands, depending on the role you are viewing.

Minimum Required Role: Operator (also provided by Configurator, Cluster Administrator, Limited Administrator, and Full Administrator)

The Actions menu is available from the Role Status page only when you are viewing Current time. It is disabled if you are viewing a point of time in the past.

Viewing Past Status.....

You can expand the Time Range Selector to view historical health, status, and chart data. The status and health information shown on this page represents the state of the service or role at a point in time. The exceptions are the charts tabs, which show information for the time range selected in the Time Range Selector, which defaults to the past 30 minutes. By default, the information shown is for the current time. You can view status for a past point in time by moving the time marker () to a point in the past. When you move the time marker to a point in the past (for Services/Roles that support health status), the Status clearly indicates that it is referring to a past time. A Now button () enables you to view the current state of the service. In addition, the Actions menu is disabled while you are viewing past status to ensure that you cannot accidentally take an action based on outdated status information. You can also view past status by clicking the Show link in the Health Tests and Health History panels. For more information, see Related Information.

Time Line

Health Tests and Health History.....

Summary.....

The Summary panel provides basic information about the role instance, where it resides, and its configuration. All role types provide the Summary panel. Some role instances related to HDFS, MapReduce, and Tez provide a Health Tests panel and associated charts.

Health Tests and Health History

The Health Tests panel shows health test results in an expandable and collapsible list, typically showing the metrics that the test returned. You can Expand All or Collapse All from the links at the top of the panel.

The Health Tests and Health History panels appear for HDFS, MapReduce, HBase, Flume, Impala, and Tez, which are the Cloudera Manager Service. Other services such as Hue, Oozie, and YARN do not provide a Health Tests panel.

- The color of the text (and the background color of the field) for a Health Test result indicates the status of the results. The tests are sorted by their health status – Good, Concerning, Bad, or Disabled. The entries are collapsed by default. Click the arrow to the left of an entry to expand the entry and display further information.

- Clicking the Details link for a health test displays further information about the test, such as the meaning of the test and its possible results, suggestions for actions you can take or how to make configuration changes related to the test. The help text may include a link to the relevant monitoring configuration section for the service.
- In the Health Tests panel:
 - Clicking displays the lists of health tests that contributed to the health test.
 - Clicking the Details link displays further information about the health test.
- In the Health History panel:
 - Clicking displays the lists of health tests that contributed to the health history.
 - Clicking the Show link moves the time range to the historical time period.

Related Information
 Configuring Monitoring Settings

Status Summary.....

The Status Summary panel reports a roll-up of the status of all the roles.

Charts

Charts are shown for roles that are related to HDFS, MapReduce, HBase, ZooKeeper, Flume, and Management Service. Roles related to other services such as Hue, Hive, Oozie, and YARN, do not have charts. See the topic Viewing Charts for Cluster, Service, Role, and Host Instances for detailed information. Charts are presented, and the ability to search and display metrics of your choice.

Related Information
 Viewing Charts for Cluster, Service, Role, and Host Instances

The Processes Tab

The Processes page shows the processes that run as part of this service role, with a variety of filters and processes.

Procedure

1. To view the processes running for a role instance, select a service instance to display the Status page for that service.
2. Click the Instances tab.
3. From the list of roles, select one to display that role instance's Status page.
4. Click the Processes tab.

What to do next

- To see the location of a process' configuration files, and to view the Environment variable settings, click the Show link under Configuration Files/Environment.
- If the process provides a Web UI (as is the case for the NameNode, for example) click the link to open the Web UI for that process.
- To see the most recent log entries, click the Show Recent Logs link.
- To see the full log, stderr, or stdout log files, click the appropriate links.

Running Diagnostic Commands for Roles

You can run diagnostic utility tools such as "collect stack traces" and "heap dump" against

About this task

Minimum Required Role: Cluster Administrator (also provided by Full Administrator)

Cloudera Manager allows administrators to run the following diagnostic utility tools against processes:

- List Open Files (lsOf) - Lists the open files of the process.
- Collect Stack Traces (jstack) - Captures Java thread stack traces for the process.
- Heap Dump (jmap) - Captures a heap dump for the process.
- Heap Histogram (jmap -histo) - Produces a histogram of the heap for the process.

These commands are found on the Actions menu of the Cloudera Manager page for the instance of interest. For example, to run diagnostics commands for the HDFS active NameNode, perform these steps:

Procedure

1. Click the HDFS service on the Home/Status tab or select it on the Clusters menu.
2. Click Instances/NameNode (Active).
3. Click the Actions menu.
4. Choose one of the diagnostics commands listed in the lower section of the Actions menu.
5. Click the button in the confirmation dialog box to confirm your choice.
6. When the command is executed, click Download Result Data and save the file to view the command output.

Periodic Stacks Collection

Periodic stacks collection allows you to enable and configure the periodic collection of thread stacks for the Cloudera Manager.

When stacks collection is enabled for a role, call stacks are output to a log file at regular intervals. This is useful for the diagnosis of performance issues such as deadlock, slow processing, or excessive numbers of threads. Stacks collection may impact performance for the processes being collected as well as other processes on the host and is turned off by default. For troubleshooting performance issues, you may be asked by Cloudera to enable stacks collection and send the resulting logs to Cloudera for analysis.

Stacks collection is available for the majority of roles in Cloudera Manager. For the HDFS service, stacks collection is available for the active NameNode and the DataNodes.

you can enable stacks collection for the DataNode, NameNode, Failover Controller, HttpFS, Job Gateway. If the Stacks Collection category does not appear in the role's configuration settings, it is not available for that role.

Configuring Periodic Stacks Collection.....

You can enable and configure periodic stacks collection.

About this task

Minimum Required Role: Configurator (also provided by Cluster Administrator, Limited Cluster Administrator, Full Administrator)

Procedure

1. Open the Cloudera Manager page for a specific service or role.
2. Access the configuration settings in one of the following ways:
 - a. From the service page in Cloudera Manager, click the Configuration tab.
 - b. Select Scope NameNode.
 - c. Select Category Stacks Collection.
 - a. From the service page in Cloudera Manager, click the Instances tab.
 - b. Click the Configuration tab.
 - c. Select Scope role type.
 - d. Select Category Stacks Collection. The configuration settings are as follows:
 - Stacks Collection Enabled - Whether or not periodic stacks collection is enabled.
 - Stacks Collection Directory - The directory in which stack logs will be placed. If not set, stacks will be logged into a stacks subdirectory of the role's log directory.
 - Stacks Collection Frequency - The frequency with which stacks will be collected.
 - Stacks Collection Data Retention - The amount of stacks data that will be retained. When the retention limit is reached, the oldest data will be deleted.
 - Stacks Collection Method - The method that will be used to collect stacks. The jstack option involves periodically running the jstack command against the role's daemon process. The servlet method is available for those roles with an HTTP server endpoint that exposes the current stacks traces of all threads. When the servlet method is selected, that HTTP endpoint is periodically scraped. Example As an example, to configure stacks collection for an HDFS NameNode, complete the following steps:

- Go to the HDFS service page.
- Click the Configuration tab.
- Select Scope NameNode.
- Select Category Stacks Collection.
- Locate the property or search for it by typing its name in the Search box.
- Modify the configuration settings if desired.
- Click Save Changes. Stacks collection configuration settings are stored in a per-role configuration file called cloudera-stacks-monitor.properties. Cloudera Manager reads the configuration file and coordinates stack collection. Changes to the configuration settings take effect after a short delay. It is not necessary to restart the role.

Viewing and Downloading Stacks Logs

Stacks are collected and logged to a compressed, rotated log file. You can view and download

About this task

A certain amount of the log data is in an uncompressed file. When that file reaches a limit, it is bzip2 compressed. Once the total number of files exceeds the configured retention limit, the oldest files are deleted. Collected stacks data is available for download through the Cloudera Manager UI and API. To view and download stacks logs through the UI, complete the following steps:

Procedure

- On the service page, click the Instances tab.
- Click the role in the Role Type column.
- Click the Stacks Logs tab.
- Click Stacks Log File to view the most recent stacks file. Click Download Stacks Logs to download a zipped bundle of the stacks logs.

Viewing Running and Recent Commands

You can view and run recent commands for a cluster, service, or role. You can also view the running commands.

Viewing Running and Recent Commands For a Cluster

You can use the command indicator to view the running and recent commands for a cluster.

Procedure

1. Click the Recent Commands in the left navigation panel. The indicator displays the number of commands currently running for all services or roles.
2. To display all commands that have run and finished recently, do one of the following:

- Click the All Recent Commands button in the window that pops up when you click the indicator. This command displays information on all running and recent commands in the same form, as described below.
- Click the Cloudera Manager logo in the main navigation bar and click the All Recent Commands tab. The command indicator shows the number of commands running on all clusters you are managing. Likewise, All Recent Commands shows all commands that were run and finished within the search time range you specified, across all your managed clusters.

3. Select a value from the pager

to control how many commands are listed, or click the arrows to view pages.

Viewing Running and Recent Commands for a Service or Role

You can view running and recent commands for a service or role instance in the Commands tab.

About this task

For a selected service or role instance, the Commands tab shows which commands are running on that instance, and what the status, progress, and results are. For example, if you go to the Clusters tab after you have installed your cluster and look at the Commands tab, you will see recent commands such as `hadoop fs -ls /` for the NameNode, `hadoop fs -ls /` for the Secondary NameNode, and `hadoop fs -ls /` for the DataNode. You can also see the command that initially formatted HDFS on the NameNode. This information is useful if a service or role is taking a long time to start up or shut down, or if services or roles are not running or have not started correctly. You can view both the status and progress of currently running commands, and the results of commands run in the past.

Procedure

1. Click the Clusters tab on the top navigation bar.
2. Click the service name to go to the Status tab for that service.
3. For a role instance, click the Instances tab and select the role instance name to go to its Status tab.
4. Click the Commands tab.

Command Details

You can view command details. The details available for a command depend on whether the command is currently running or recently completed.

Running Commands

The Running Commands area shows commands that are in progress.

While the status of the command is In Progress, an Abort button displays so that you can abort the command if necessary.

The Commands status information is updated automatically while the command is running.

After the command has finished running (all its subcommands have finished), the status is up

disappear, and the information for Recent Commands appears as described below.

Recent Commands

The Recent Commands area shows commands that were run and finished within the search time range. If no commands were run during the selected time range, you can double the time range selection link, expanding the time range selection link. If you are in the "current time" mode, the beginning

are looking at a time range in the past, both the beginning and ending times of the range are displayed. You can change the time range using the options described in the topic Time Line. Select a value from the pager

to control how many

commands are listed, or click the arrows to view pages.

Commands are shown with the most recent ones at the top.

The icon associated with the status (which typically includes the time that the command finished) tells you whether the command succeeded or failed. If the command failed, it indicates the subcommands that actually failed. In many cases, multiple subcommands result from the top-level command. The First Run command runs during the initial startup of your cluster. Click this link to view the cluster startup.

Command Details

In the Running Commands dialog box or Recent Commands page, click a command in the Command column to display its details and any subcommands. The page title is the name of the command.

The Summary section at the top shows information about the command:

- The current status
- The context, which can be a cluster, service, host, or role
- The time the command started
- The duration of the command
- A message about the command completion
- If the context is a role, links to role instance logs The Details section shows how many steps, if any, the selected command has and lists any subcommands. Expand a command to view subcommands. In the Running Commands dialog box, each subcommand also has an Abort button that is present as long as the subcommand is in progress. You can perform the following actions:
 - Select the option to display all the subcommands or only failed or running commands.
 - Click the link in the Context column to go to the Status page for the component (host, service, or role instance) to which this command is related.
 - Click a Role Log tab to display the log for that role, and stdout and stderr if available for the role. Related Information Time Line