

HP OMi Management Pack for Hadoop

For the Linux and Windows® operating systems

Software Version: 1.00

[Online Help](#)

Document Release Date: June 2013

Software Release Date: June 2013



Legal Notices

Warranty

The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

The information contained herein is subject to change without notice.

Restricted Rights Legend

Confidential computer software. Valid license from HP required for possession, use or copying. Consistent with FAR 12.211 and 12.212, Commercial Computer Software, Computer Software Documentation, and Technical Data for Commercial Items are licensed to the U.S. Government under vendor's standard commercial license.

Copyright Notice

© Copyright 2013 Hewlett-Packard Development Company, L.P.

Trademark Notices

Adobe® is a trademark of Adobe Systems Incorporated.

Microsoft® and Windows® are U.S. registered trademarks of Microsoft Corporation.

UNIX® is a registered trademark of The Open Group.

Documentation Updates

The title page of this document contains the following identifying information:

- Software Version number, which indicates the software version.
- Document Release Date, which changes each time the document is updated.
- Software Release Date, which indicates the release date of this version of the software.

To check for recent updates or to verify that you are using the most recent edition of a document, go to:

<http://h20230.www2.hp.com/selfsolve/manuals>

This site requires that you register for an HP Passport and sign in. To register for an HP Passport ID, go to:

<http://h20229.www2.hp.com/passport-registration.html>

Or click the **New users - please register** link on the HP Passport login page.

You will also receive updated or new editions if you subscribe to the appropriate product support service. Contact your HP sales representative for details.

Support

Visit the HP Software Support Online web site at:

<http://www.hp.com/go/hpsoftwaresupport>

This web site provides contact information and details about the products, services, and support that HP Software offers.

HP Software online support provides customer self-solve capabilities. It provides a fast and efficient way to access interactive technical support tools needed to manage your business. As a valued support customer, you can benefit by using the support web site to:

- Search for knowledge documents of interest
- Submit and track support cases and enhancement requests
- Download software patches
- Manage support contracts
- Look up HP support contacts
- Review information about available services
- Enter into discussions with other software customers
- Research and register for software training

Most of the support areas require that you register as an HP Passport user and sign in. Many also require a support contract. To register for an HP Passport ID, go to:

<http://h20229.www2.hp.com/passport-registration.html>

To find more information about access levels, go to:

http://h20230.www2.hp.com/new_access_levels.jsp

Disclaimer for PDF Version of Online Help

This document is a PDF version of the online help. This PDF file is provided so you can easily print multiple topics from the help information or read the online help in PDF format.

Note: Some topics do not convert properly to PDF, causing format problems. Some elements of online help are completely removed from the PDF version. Those problem topics can be successfully printed from within the online help.

Contents

Online Help	1
Contents	6
Introduction	8
Hadoop Management Templates	9
Deploying Hadoop Management Templates	9
Essential Hadoop Management Template	9
User Interface (UI) Reference	10
Hadoop Aspects	14
Grouping of Hadoop Aspects	16
User Interface Reference	16
DataNode DFS Statistics	17
DataNode Operations	17
Hadoop Base	18
Hadoop Discovery	18
Hadoop Heart Beats	18
Hadoop JVM Statistics	19
Hadoop Network Statistics	20
Hadoop Service Availability	20
Job Tracker Black Listed Task Trackers	21
Job Tracker Jobs Status	21
Job Tracker Map Reduce Statistics	22
NameNode Blocks Statistics	22
NameNode Capacity Trend	23
NameNode File Operations	23
Task Tracker Operations	24
Parameters	25
Types of Parameters	25
Hadoop Parameters	25

Tuning Parameters	26
Tools	27
Customizing OMi MP for Hadoop	28
Customizing OMi MP for Hadoop before Deployment	28
Editing Hadoop Management Templates	28
Editing Parameters	28
Creating Hadoop Management Templates	29
Deployment Scenario	31
Task 1: Adding Nodes to the BSM console	31
Task 2: Deploying the Hadoop Discovery Aspect	31
Task 3: Deploying Essential Hadoop Management Template	32
Appendix	32
Data Source and Metrics	32
Eventable Metrics in Hadoop	35

Chapter 1

Introduction

The OMi Management Pack for Hadoop (OMi MP for Hadoop) works with OMi and enables you to monitor Hadoop environment using the Business Service Management (BSM). The OMi MP for Hadoop provides out of the box Management Template for monitoring Hadoop in a Big Data environment. The Management Template comprises of a wide range of Aspects which enable monitoring the Hadoop core components - Hadoop Distributed File System (HDFS) and MapReduce. The administrators can seamlessly deploy the out of the box Management Templates for monitoring Hadoop in an environment.

The Subject Matter Experts (SMEs) and developers can easily customize the Management Templates.

The OMi MP for Hadoop also provides the following additional functionalities to support a unified monitoring solution:

- Configuration Item (CI) based deployment and configuration
- Supports agent monitoring by integrating with the product- HP Operations Agent
- Supports Run Time Service Model (RTSM) based deployment

The OMi MP for Hadoop includes the following artifacts for monitoring Hadoop in an environment:

- Hadoop Management Templates
- Hadoop Aspects
- Policy Templates

These artifacts are parameterized for easy deployment and can be modified before or after deployment.

Chapter 2

Hadoop Management Templates

The Hadoop Management Templates provide a complete management solution for monitoring Hadoop. The Management Templates comprises of several Aspects which enables you to monitor Hadoop clusters in an environment. You can deploy the out of the box Management Templates with the default settings or you can customize the Management Templates based on your requirements. In addition, you can also create Management Templates based on the monitoring requirements.

The OMi MP for Hadoop comprises the [Hadoop Management Template](#).

To access


Open the Management Templates and Aspects manager:

Select **Admin > Operations Management > Monitoring > Management Templates and Aspects**

In the Management Templates and Aspects pane:

Select **Configuration Folders > Hadoop Management > Hadoop Management Templates**

Deploying Hadoop Management Templates

1. Open the Management Templates and Aspects manager:
Admin > Operations Management > Monitoring > Management Templates & Aspects
2. In the Management Templates and Aspects pane:
Configuration Folders > Hadoop Management > Hadoop Management Templates
3. In the **Hadoop Management Templates** folder, select the Management Template that you want to deploy, and then click . The Assign and Deploy wizard opens.
4. In the Configuration Item page, click the CI to which you want to assign the Management Template, and then click **Next**.
5. In the Parameter page, click **Next**.
6. *Optional.* If you do not want to enable the assignment immediately, clear the **Enable Assigned Objects** check box. You can then enable the assignment later using the Assignments and Parameters manager.
7. Click **Finish**.

Essential Hadoop Management Template

The Hadoop Essential Management Template can be used to monitor the Hadoop components in an environment. It comprises of essential Hadoop Aspects and Infrastructure Aspects for monitoring the availability, health, and performance of Hadoop environments.

To use and deploy Infrastructure Aspects, you must install OMi Management Pack for Infrastructure software.

To access

Open the Management Templates and Aspects manager:

Select **Admin > Operations Management > Monitoring > Management Templates and Aspects**

Select **Configuration Folders > Hadoop Management > Hadoop Management Templates > Essential Hadoop Management Template**

User Interface (UI) Reference

Management Template - General

Provides an overview of the attributes of the Management Template.

UI Element	Description
Name	Essential Hadoop Management Template
Description	Monitors the Hadoop components - HDFS and MapReduce
ID	3cd36519-065f-55ec-2ca1-6679d88e1376
Version ID	A unique identifier for this version of the Management Template. In this case, a5a89e78-6cf2-9e39-8b6e-a7d95be51809
Version	The current version of the Management Template. In this instance, the version of the Management Template is 1.0.
Change Log	Text that describes what is new or modified in this version of the Management Template.

Management Template - Topology View

UI Element	Description
Topology View	Hadoop Topology is the Topology View for Essential Hadoop Management Templates. It contains the Hadoop related CI types that you want to manage using the Management Template.
CI Type	The type of CIs that the Essential Hadoop Management Template enables you to manage. This is the type of CI to which the Management Template can be assigned. The Essential Hadoop Management Templates contains Hadoop Master, Hadoop Slave, Computer, Node, and Cluster CI Types.

Management Template - Aspects

The Essential Hadoop Management Template contains the following Aspects:

[DataNode DFS Statistics](#)[DataNode Operations](#)[Hadoop Base](#)[Hadoop Discovery](#)[Hadoop HeartBeats](#)[Hadoop JVM Statistics](#)[Hadoop Network Statistics](#)[Hadoop Service Availability](#)[JobTracker BlackListed TaskTrackers](#)[JobTracker Jobs Status](#)[JobTracker MapReduce Statistics](#)[NameNode Blocks Statistics](#)[NameNode Capacity Trend](#)[NameNode File Operations](#)[TaskTracker Operations](#)

The Essential Hadoop Management Template comprises of the following Infrastructure Aspects:

Bandwidth Utilization and Network IOPS

Monitors I/O operations, and performance of the systems in the network. It monitors the network I/O operations and performance based on the bandwidth used, outbound queue length and average bytes transferred per second.

CI Type	Policy Template	Description	Policy Type
Computer	SI-NetworkUsageAndPerformance	This policy monitors the network usage of the system and shows error rates and collisions to identify potential network bottlenecks. This policy template monitors the physical NICs of only the vMA machines. It does not monitor performance data for package collision on the Windows operating system, as the BYNETIF_COLLISION metric is not available on it.	Measurement Threshold Template

CPU Performance

Monitors I/O operations, and performance of the systems in the network. It monitors the network I/O operations and performance based on the bandwidth used, outbound queue length and average bytes transferred per second.

CI Type	Policy Template	Description	Policy Type
	SI-CPUSpikeCheck	This policy template monitors the variation in processor performance. A system experiences CPU spike when there is a sharp rise in the CPU usage immediately followed by a decrease in usage. SI-CPUSpikeCheck policy template monitors CPU time spent in user mode and system mode and total CPU time when the CPU is busy.	Measurement Threshold Template

Memory and Swap Utilization

Monitors memory performance of the system. Memory performance monitoring is based on Memory utilization (in percentage), Swap space utilization (in percentage), Free memory available (in MBs) and Free swap space available (in MBs).

CI Type	Policy Template	Description	Policy Type
Computer	SI-MemoryUsageAndPerformance	This policy template monitors the memory usage of the system and shows error rates and collisions to identify potential memory bottlenecks.	Measurement Threshold Template
	SI-SwapcapacityMonitor	This policy template monitors the swap space utilization of the system.	Measurement Threshold Template

Remote Disk Space Utilization

Monitors the space utilization of remote disk.

CI Type	Policy Template	Description	Policy Type
Computer	SI-LinuxNFSUtilizationMonitor	This policy template monitors space utilization level for NFS remote filesystems on Linux platforms.	Measurement Threshold Template
	SI-LinuxCIFSUtilizationMonitor	This policy template monitors space utilization level for CIFS remote filesystems on Linux platforms.	

Space Availability and Disk IOPS

Monitors the disk I/O operations and space utilization of the system.

CI Type	Policy Template	Description	Policy Type
Computer	SI-File-SystemUtilizationMonitor	Monitors the FileSystem Utilization	Measurement Threshold Template

System Infrastructure Discovery

Discovers and gathers information regarding the system resources, operating system, and applications on a managed node.

CI Type	Policy Template	Description	Policy Type
Computer	SI-SystemDiscovery	This policy template gathers service information from the managed nodes such as hardware resources, operating system attributes, and applications.	Measurement Threshold Template

Chapter 3

Hadoop Aspects

Hadoop Aspects can be used to monitor the building blocks or units of Hadoop. A Hadoop aspect comprises of policy templates, instrumentation, and parameters for monitoring the health and performance of Hadoop. Each Hadoop aspect provides the ability to monitor an Hadoop configuration item (CI).








To access




Open the Management Templates and Aspects manager:



Select **Admin > Operations Management > Monitoring > Management Templates and Aspects**

Select **Configuration Folders > Hadoop Management > Hadoop Aspects**

Creating Hadoop Aspects



1. Open the Management Templates and Aspects manager:
Admin > Operations Management > Monitoring > Management Templates & Aspects > Hadoop Management > Hadoop Aspects
2. In the Configuration Folders pane, click the configuration folder in which you want to create the new aspect. If you need to create a new configuration folder, click .
3. In the Management Templates & Aspects pane, click , and then click  **Aspect**. The Create Aspect wizard opens.
4. In the General page, type a unique **Name** for the new aspect.
Click **Next**.
5. In the CI Types page, select one or more **Available CI Type(s)** to which this aspect can be assigned, and then click  to add them to the list of assigned CI types. (Press **CTRL** to select several CI types.). Click **Next**.
6. In the Instrumentation page, click  to add instrumentation to the aspect. The Add Instrumentation dialog box opens, which enables you to select the instrumentation that you want to add. Click **Next**.
7. *Optional.* In the Aspects page, click , and then click  **Add Existing Aspect**. The Add Existing Aspect dialog box opens, which enables you to select an existing aspect that you want to nest within this aspect. Click an aspect, and then click **OK**.

If suitable Aspects do not exist, click , and then click  **Add New Aspect** to create them from here. Click **Next**.
8. In the Policy Templates page, click . The Add Policy Template to Aspect dialog box opens. Select the policy templates that you want to add, and then click **OK**. (Press **CTRL** to select several policy templates.)

If suitable policy templates do not exist, click , and then click  **Add New Policy Template** to create them from here.

9. In the Policy Templates page, select the **Version** of the policy templates that you want to add.


Each modification to a policy template is stored in the database as a separate version. Aspects contain specific versions of policy templates. If a new version of a policy template becomes available later, you have to update the aspect to include the latest version, if that is what you want.

10. *Optional.* In the Policy Templates page, click the policy template to which you want to add a deployment condition, click the , and click the  **Edit Deployment Condition**. The Edit Deployment Condition dialog box opens, which enables you to specify deployment conditions for the selected policy template. Set the condition and then click **OK**.

In the Policy Templates page, click **Next**.

11. In the Parameters page, you see a list of all the parameters from the policy templates that you added to this aspect.


To combine parameters:

- a. Press **CTRL** and click the parameters that you want to combine.
- b. Click . The Edit/Combine Parameters dialog box opens.
- c. Type a **Name** for the combined parameters.
- d. *Optional.* Specify a **Description**, **Default Value**, and whether the combined parameter is **Read Only**, an **Expert Setting**, or **Hidden**.

You can set either a specific default value, or you can click **From CI Attribute** and then browse for a CI attribute. When you specify a CI attribute, Operations Management sets the parameter value automatically during deployment of the policy templates, using the actual value of this attribute from the CI. You can also set conditional parameter values here.

Read Only prevents changes to the parameter value when the aspect is assigned to a configuration item. Hidden also prevents changes, but additionally makes the parameter invisible. Users can choose whether to show expert settings when they make an assignment.





- e. Click **OK**.

You can also edit the parameters without combining them, to override the defaults in the policy template. Click one parameter, and then click . The Edit/Combine Parameters dialog box opens.

12. In the Create Aspect wizard, click **Finish** to save the aspect and close the wizard. The new aspect appears in the Management Templates & Aspects pane.

Deploying Hadoop Aspects

Note: Developers can deploy Aspects only for testing. If you want to monitor an application or service, deploy a Management Template.

1. Open the Management Templates and Aspects manager:
Admin > Operations Management > Monitoring > Management Templates & Aspects > Hadoop Management > Hadoop Aspects
2. In the **Management Templates & Aspects** pane, click the aspect that you want to deploy, and then click . The Assign and Deploy wizard opens.
3. In the Configuration Item page, click the configuration item to which you want to assign the aspect, and then click **Next**.
4. In the Parameter page, specify a value for each parameter:
 - a. *Optional*. By default, the list shows only mandatory parameters. To see optional parameters, click . You can also click  to see expert parameters.
 - b. Select a parameter in the list, and then click . The Edit Parameter dialog box opens.
 - c. Click **Value**, specify the value, and then click **OK**.
 Click **Next**.
5. *Optional*. If you do not want to enable the assignment immediately, clear the **Enable Assigned Objects** check box. You can then enable the assignment later using the Assignments and Parameters manager.
6. Click **Finish**.

Grouping of Hadoop Aspects

Hadoop Aspects can be used to monitor the core components of Hadoop - HDFS and Map Reduce. The Aspects are categorized based on the nodes in the Hadoop cluster - Hadoop Master and Hadoop Slave.

User Interface Reference

General	Provides an overview of the general attributes of the Hadoop Aspects.
CI Type	The type of configuration items that the Aspect can be assigned to. This is the type of CI to which the Management Template can be assigned. The Hadoop Aspects contain the Computer, Node, Cluster, Hadoop Master, and Hadoop Slave CI types.
Instrumentation	Provides a single package which contains the binaries for discovery, collection, and data logging.
Aspects	Provides an overview of any Aspects that the Hadoop Aspect contains. You can expand each item in the list to see more details about the nested aspect. The Hadoop Base aspect is part of all the other Aspects.
Policy Templates	Provides an overview of the policy templates that the Hadoop Aspect contain. You can expand each item in the list to see more details about the policy template.

The OMi MP for Hadoop comprises of the following Aspects:

DataNode DFS Statistics

Monitors the Hadoop Slave DataNode DFS Statistics.

CI Type	Policy Template	Policy Description	Policy Type
Hadoop Slave	Hadoop_DataNode_Capacity	Monitors the DataNode DFS capacity in bytes.	Measurement Threshold Template
	Hadoop_DataNode_Dfsused	Monitors the DataNode DFS used in bytes.	
	Hadoop_DataNode_Remaining	Monitors the DataNode DFS remaining in bytes.	

DataNode Operations

Monitors the datanode operations.

CI Type	Policy Template	Policy Description	Policy Type
Hadoop Slave	Hadoop_DataNode_ReadBlockAverageTime	Monitors the average time for read block operations.	Measurement Threshold Template
	Hadoop_DataNode_ReadBlockOperations	Monitors the number of read block operations.	
	Hadoop_DataNode_ReadsLocalClient	Monitors the number of reads from local client.	
	Hadoop_DataNode_ReadsRemoteClient	Monitors the number of reads from the remote client.	
	Hadoop_DataNode_WriteBlockAverageTime	Monitors the average time for write block operations.	
	Hadoop_DataNode_WriteBlockOperations	Monitors the average time for write block operations.	
	Hadoop_DataNode_WritesLocalClient	Monitors the number of writes from local client.	
	Hadoop_DataNode_WritesRemoteClient	Monitors the number of writes from remote client.	
	Hadoop_DataNode_ReadThroughputOperation	Monitors the throughput of read operations.	
	Hadoop_DataNode_WriteThroughputOperation	Monitors the throughput of write operations.	

Hadoop Base

This is the base aspect for Hadoop monitoring.

CI Type	Policy Template	Policy Description	Policy Type
Hadoop Master, Hadoop Slave	Hadoop_Configuration	This policy template contains the Hadoop configuration.	ConfigFile Template
	Hadoop_High	Runs JMX collector/analyzer every high schedule.	Schedule Task Template
	Hadoop_Low	Runs JMX collector/analyzer low schedule.	Schedule Task Template
	Hadoop_Medium	Runs JMX collector/analyzer every medium schedule.	Schedule Task Template
	Hadoop_Messages	Interception of messages submitted by Hadoop programs.	Open Message Interface Template
	Hadoop_VeryHigh	Runs JMX collector/analyzer every very high schedules.	Schedule Task Template

Hadoop Discovery

Discovers the Hadoop components - NameNode, SecondaryNameNode, JobTracker, DataNode, TaskTracker, and Clusters.

CI Type	Policy Template	Policy Description	Policy Type
Computer, Node	Hadoop Discovery	Discovers the Hadoop components.	Service Auto-Discovery Template
	Hadoop_Messages	Interception of messages submitted by Hadoop programs.	Open Message Interface Template

Hadoop Heart Beats

Monitors Hadoop Master and Slave components for HeartBeats.

CI Type	Policy Template	Policy Description	Policy Type
Hadoop Master	Hadoop_DataNode_HeartBeatAverageTime	Monitors the average time for heartbeat for the Hadoop DataNode.	Measurement Threshold Template
	Hadoop_DataNode_HeartBeatOperations	Monitors the number of operations for heartbeat for the Hadoop DataNode.	
	Hadoop_JobTracker_HeartBeatAverageTime	Monitors the average time for heartbeat for the Hadoop Job Tracker.	
	Hadoop_JobTracker_HeartBeatOperations	Monitors the number of operations for heartbeat for the Hadoop JobTracker.	
	Hadoop_NameNode_HeartBeatAverageTime	Monitors the average time for heartbeat for the Hadoop NameNode.	
	Hadoop_NameNode_HeartBeatOperations	Monitors the number of operations for heartbeat for the Hadoop NameNode.	

Hadoop JVM Statistics

Monitors Hadoop components for Java Heap usage and Garbage Collection.

CI Type	Policy Template	Policy Description	Policy Type
Hadoop Master	Hadoop_DataNode_GC_TimeSpent	Monitors the total GC time in seconds for the Hadoop DataNode.	Measurement Threshold Template
	Hadoop_DataNode_HeapMemoryUsage	Monitors the heap memory used in MB for the Hadoop DataNode.	
	Hadoop_JobTracker_GC_TimeSpent	Monitors the Java heap usage and garbage collection for the Job Tracker.	
	Hadoop_JobTracker_HeapMemoryUsage	Monitors the Heap memory used in Mb for the Job Tracker.	
	Hadoop_NameNode_GC_TimeSpent	Monitors the total GC time in milliseconds for the Name Node.	
	Hadoop_NameNode_HeapMemoryUsage	Monitors the heap memory used in MB for the NameNode.	
	Hadoop_TaskTracker_GC_TimeSpent	Monitors the total GC time in seconds for the TaskTracker.	
	Hadoop_TaskTracker_HeapMemoryUsage	Monitors the heap memory used in Mb for the TaskTracker.	

Hadoop Network Statistics

Monitors Hadoop Master NameNode, JobTracker Bytes Transfer, Hadoop Slave DataNode, and TaskTracker Bytes Transfer.

Hadoop Master	Hadoop_JobTracker_ReceivedBytes	Monitors the RPC received bytes count for the JobTracker.	Measurement Threshold Template
	Hadoop_JobTracker_SentBytes	Monitors the RPC sent bytes count for the JobTracker.	
	Hadoop_TaskTracker_ReceivedBytes	Monitors the RPC received bytes count for the TaskTracker.	
	Hadoop_TaskTracker_SentBytes	Monitors the RPC sent bytes count for the TaskTracker.	
	Hadoop_NameNode_ReceivedBytes	Monitors the RPC received bytes count for the NameNode.	
	Hadoop_NameNode_SentBytes	Monitors the RPC sent bytes count for the NameNode.	
	Hadoop_DataNode_ReceivedBytes	Monitors the RPC received bytes count for the DataNode.	
	Hadoop_DataNode_SentBytes	Monitors the RPC sent bytes count for the DataNode.	

Hadoop Service Availability

Monitors the availability of Hadoop Component - NameNode, JobTracker, SecondaryNameNode, DataNode, and TaskTracker.

CI Type	Policy Template	Policy Description	Policy Type
Hadoop Master, Hadoop Slave	Hadoop_DataNode_Availability	Monitors the Hadoop DataNode availability metric	Measurement Threshold Template
	Hadoop_DataNode_Availability_Check	Checks the Hadoop DataNode service every 5 minutes.	Scheduled Task Template
	Hadoop_JobTracker_Availability	Monitors Hadoop JobTracker availability metric.	Measurement Threshold Template
	Hadoop_JobTracker_Availability_Check	Monitors Hadoop JobTracker service availability every 5 minutes.	
	Hadoop_NameNode_Availability	Monitors the Hadoop NameNode availability metric.	
	Hadoop_NameNode_Availability_Check	Monitors the Hadoop NameNode service availability every 5 minutes.	Scheduled Task Template
	Hadoop_SecondaryNameNode_Availability	Monitors the Hadoop Secondary Name Node Availability Metric.	Measurement Threshold Template
	Hadoop_SecondaryNameNode_AvailabilityCheck	Checks Hadoop Secondary Name Node Service Availability every 5 minutes.	Scheduled Task Template
	Hadoop_TaskTracker_Availability	Monitors the Hadoop Task Tracker Availability Metric.	Measurement Threshold Template
	Hadoop_TaskTracker_Availability_Check	Checks the availability of the Task Tracker every five minutes.	Scheduled Task Template

Job Tracker Black Listed Task Trackers

Monitors the availability statistics of Task Trackers in a cluster.

CI Type	Policy Template	Policy Description	Policy Type
Hadoop Master	Hadoop_JobTracker_BlackListedTrackers	Monitors the number of blacklisted trackers.	Measurement Threshold Template

Job Tracker Jobs Status

Monitors the jobs of the job tracker.

CI Type	Policy Template	Policy Description	Policy Type
Hadoop Master	Hadoop_JobTracker_JobsCompleted	Monitors the number of jobs completed	Measurement Threshold Template
	Hadoop_JobTracker_JobsFailed	Monitors the number of jobs failed	Measurement Threshold Template
	Hadoop_JobTracker_JobsSubmitted	Monitors the number of jobs submitted	Measurement Threshold Template
	Hadoop_ThroughputRunningJobs	Monitors the throughput of running jobs	Measurement Threshold Template

Job Tracker Map Reduce Statistics

Monitors the job tracker map reduce statistics.

CI Type	Policy Template	Policy Description	Policy Type
Hadoop Master	Hadoop_JobTracker_MapsLaunched	Monitors the number of maps launched	Measurement Threshold Template
	Hadoop_JobTracker_MapsSlots	Monitors the number of map slots	Measurement Threshold Template
	Hadoop_JobTracker_ReducedLaunched	Monitors the number of reduces launched	Measurement Threshold Template
	Hadoop_JobTracker_ReducesSlots	Monitors the number of reduces slots	Measurement Threshold Template
	Hadoop_JobTracker_WaitingMaps	Monitors the number of waiting maps	Measurement Threshold Template
	Hadoop_JobTracker_WaitingReduces	Monitors the number of waiting reduces	Measurement Threshold Template
	Hadoop_JobTracker_MapSlotUsagePct	Monitors the percentage of map slots used	Measurement Threshold Template
	Hadoop_JobTracker_ReduceSlotUsagePct	Monitors the percentage of reduced slots used	Measurement Threshold Template

NameNode Blocks Statistics

Monitors the NameNode Blocks Statistics.

CI Type	Policy Template	Policy Description	Policy Type
Hadoop Master	Hadoop_NameNode_BlockCapacity	Monitors the block capacity of the Namenode	Measurement Threshold Template
	Hadoop_NameNode_BlocksTotal	Monitors the total blocks of the Namenode	Measurement Threshold Template
	Hadoop_NameNode_CorruptBlocks	Monitors the corrupt blocks of the Namenode	Measurement Threshold Template
	Hadoop_NameNode_ExcessBlocks	Monitors the excess blocks of the Namenode	Measurement Threshold Template
	Hadoop_NameNode_MissingBlocks	Monitors the missing blocks of the Namenode	Measurement Threshold Template
	Hadoop_NameNode_PendingReplicationBlocks	Monitors the Pending replication blocks of the Namenode	Measurement Threshold Template
	Hadoop_NameNode_UnderReplicatedBlocks	Monitors the Underreplicated blocks of the Namenode	Measurement Threshold Template

NameNode Capacity Trend

This aspect monitors the capacity of the name node.

CI Type	Policy Template	Policy Description	Policy Type
Hadoop Master	Hadoop_CapacityLoadHDFS	Monitors the capacity load on HDFS	Measurement Threshold Template
	Hadoop_NameNode_CapacityTotalGB	Monitors the total capacity of the namenode in GB	Measurement Threshold Template
	Hadoop_NameNode_CapacityUsedGB	Monitors the capacity used in GB	Measurement Threshold Template

NameNode File Operations

Monitors the Name Node File Operations.

CI Type	Policy Template	Policy Description	Policy Type
Hadoop Master	Hadoop_NameNode_FilesCreated	Monitors the number of files created	Measurement Threshold Template
	Hadoop_NameNode_FilesDeleted	Monitors the number of files deleted	Measurement Threshold Template
	Hadoop_NameNode_FilesRenamed	Monitors the number of files renamed	Measurement Threshold Template
	Hadoop_NameNode_FilesTotal	Monitors the number of total files	Measurement Threshold Template

Task Tracker Operations

Monitors the Task Tracker activities.

CI Type	Policy Template	Policy Description	Policy Type
Hadoop Slave	Hadoop_TaskTrackers_TasksFailedTimeout	Monitors the number of tasks failed due to timeout	Measurement Threshold template

Chapter 4

Parameters

Parameters are variables that are integral components of Hadoop Management Templates, Hadoop Aspects, and Policy Templates. Each parameter corresponds to a variable. Parameters contain default values that are used for monitoring the different components of Hadoop. You can modify the values of the variables to suit your monitoring requirements.

Types of Parameters

The parameters are grouped as follows:

Mandatory Parameters - These parameters contain the essential information required by policy templates. For example, Hadoop instance name is a mandatory parameter.

Expert Parameters - These parameters can be used by SMEs and Administrators.

Hadoop Parameters

The following table provides information about the Hadoop Parameters:

Parameter	Parameter Type	Description	Default Value
Application Instance	Mandatory	Hadoop Instance Name where Hadoop instance is running.	
Frequency	Mandatory	Frequency of monitoring Hadoop metrics by a policy template. For example, the frequency of monitoring Hadoop DataNode Service availability.	
Threshold	Mandatory	Threshold of a policy template. For example, the threshold of monitoring Hadoop DataNode Service availability.	

Parameter	Parameter Type	Description	Default Value
Severity	Mandatory	Severity level of a policy template. For example, the severity levels of monitoring Hadoop DataNode Service availability.	
Frequency of High Scheduler	Expert	Frequency of the scheduler, which is expected to run for long intervals (in minutes).	15
Frequency of Low Scheduler	Expert	Frequency for the scheduler which is expected to run for short intervals (in hours).	24
Frequency of Medium Scheduler	Expert	Frequency for the scheduler which is expected to run for medium intervals (in hour).	1
Frequency of Very High Scheduler	Expert	Frequency for the scheduler which is expected to run for very high intervals (in minutes).	5



Note: The frequency, threshold, and severity parameters are defined for each policy template. For example, the Hadoop_AvIDbNode policy template contains the following parameters:

- Frequency of monitoring Hadoop DataNode Service availability
- Threshold of monitoring Hadoop DataNode Service availability
- Severity of monitoring Hadoop DataNode Service availability

Tuning Parameters

You can edit the parameters of the Hadoop Management Templates that are already deployed to the CIs.

1. Open the Assignments and Tuning manager:
Select **Admin > Operations Management > Monitoring > Assignments and Tuning**.
2. In the Browse Views tab, select the Hadoop Topology that contains the CI for which you want to tune parameters. Alternatively, you can use the Search tab to find a CI.

3. In the list of Hadoop CIs, select a CI. The Assignments pane shows details of any existing assignments for the Hadoop CI.
4. Select the assignment for which you want to tune parameters. The Details of Assignment pane shows the current parameter values.
5. In the Details of Assignment pane, change the parameters by following these steps:
 - a. *Optional.* By default, the list shows only mandatory parameters. To see all parameters, click  .
 - b. Select a parameter in the list, and then click  .
 - For standard parameters, the Edit Parameter dialog box opens.
Click **Value**, specify the value, and then click **OK**.
 - For instance parameters, the Edit Instance Parameter dialog box opens.
Change the instance values if necessary, and then for each instance value, change dependent parameter values. After you change the instances and dependent parameter values, click **OK**.
6. In the Details of Assignment pane, click **Save Changes**. Operations Management deploys the new parameter values to the relevant HP Operation Agents.

Tools

The OMi MP for Hadoop is packaged with tools which enable administering and monitoring the Hadoop CIs. It comprises the following tools:

Tool	Description
Start Hadoop Monitoring	Starts Hadoop monitoring on the managed node.
Stop Hadoop Monitoring	Stops Hadoop monitoring on the managed node.
Restart Hadoop Monitoring	Restarts Hadoop monitoring on the managed node.

Chapter 5

Customizing OMi MP for Hadoop

OMi MP for Hadoop can be customized to suit your monitoring requirements. You can edit the existing Hadoop Management Templates or create new Hadoop Management Templates to monitor Hadoop in your environment.

Customizing OMi MP for Hadoop before Deployment

You can customize OMi MP for Hadoop to optimally and seamlessly monitor Hadoop in your environment. OMi MP for Hadoop provides the following customization scenarios:

- Editing Hadoop Management Templates
- Creating Hadoop Management Templates

Editing Hadoop Management Templates

You can edit the Hadoop Management Templates to modify the following artifacts:

- Parameters
- Aspects

Editing Parameters

Use Case: You are using Essential Hadoop Management Template to monitor Hadoop in a Big Data environment. You are monitoring the service availability of DataNode in the environment and want to modify the corresponding parameters.


To closely monitor the service availability of DataNode in your environment, you must modify the parameter - Frequency of DataNode Availability Scheduler.

1. Open the Management Templates and Aspects manager:

Admin > Operations Management > Monitoring > Management Templates & Aspects

2. In the Management Templates & Aspects pane:






Configuration Folders > Hadoop Management > Hadoop Management Templates



3. Select the **Hadoop Management Template** from the list, and then click . The Edit Management Template dialog box opens.
4. Click the **Parameters** tab. The list of parameters appear.
5. Double-click the **Frequency of DataNode Availability Scheduler** parameter. The Edit/Combine Parameters window appears.
6. Select the appropriate value from the drop down.

7. Click **OK**. The Edit Management Template dialog box opens.
8. Click **OK**. The version of the Hadoop Management Template is incremented.

Note: The version number of the Hadoop Management Template is incremented when any customizations are made to the Hadoop Management Template.

Creating Hadoop Management Templates


1. Open the Management Templates and Aspects manager:
Admin > Operations Management > Monitoring > Management Templates & Aspects
2. In the Management Templates & Aspects pane:
Configuration Folders > Hadoop Management > Hadoop
3. Select the Hadoop configuration folder and if you need to create a new configuration folder, click . The Create Configuration Folder opens.
4. Type the name of the new configuration folder and the description. For example, you can type the new configuration folder name as **Test**.
5. Click **OK**. The new configuration folder is created.
Configuration Folders > Hadoop Management > Hadoop Management Templates > Test
6. In the Management Templates & Aspects pane, select the new configuration folder and click  and then click  **Management Template**. The Create Management Template wizard opens.
7. In the General page, type a **Name** for the new Hadoop Management Template. Click **Next**.
8. A Hadoop Management Template enables you to manage Hadoop configuration items and all the related dependent CIs. Select **Hadoop Topology** from the list as the Topology View. The Hadoop_Topology shows the Hadoop CIs and all the related CI types.
9. Select an item in the topology map to select the **CI Type** of the CIs that this Management Template enables you to manage. This is the type of CI to which the Management Template can be assigned. For example, you can select Hadoop to monitor Hadoop. Click **Next**.
10. In the Aspects page, click , and then click  **Add Existing Aspect** to add existing Aspects to the new Hadoop Management Template. The Add Existing Aspect dialog box opens. Select the Aspects that you want to add, and then click **OK**.

If suitable Aspects do not exist, click , and then click  **Add New Aspect** to create them from here.
11. For each aspect that you add, you must specify at least one **Target CI**.

Click an aspect in the list, and then in the topology map click the CI types you want the aspect to monitor when this Management Template is assigned. (Press **CTRL** to select several CI types.) Each CI type that you select here must correspond to one of the CI types assigned within the aspect itself (or a child of one of those CI types). For example, you can select Hadoop CI from the topology map.

12. In the Parameters page, you see a list of all the parameters from the Aspects that you added to this Management Template.


To combine parameters:

- a. Press **CTRL** and click the parameters that you want to combine.
- b. Click the . The Edit/Combine Parameters dialog box opens.
- c. Type a **Name** for the combined parameters.
- d. *Optional.* Specify a **Description**, **Default Value**, and whether the combined parameter is **Read Only**, an **Expert Setting**, or **Hidden**.

You can specify either a specific default value, or you can click **From CI Attribute** and then browse for a CI attribute. When you specify a CI attribute, Operations Management sets the parameter value automatically during the deployment of the underlying policy templates, using the actual value of this attribute from the CI. You can also change values of conditional parameters. (The conditions are read-only and cannot be changed at Management Template level.)

Read Only prevents changes to the parameter value when the Management Template is assigned to a configuration item. Hidden also prevents changes, but additionally makes the parameter invisible when the Management Template is assigned, and during parameter tuning. Users can choose whether to show expert settings when they make an assignment.

- e. Click **OK**.

You can also edit the parameters without combining them, to override the defaults in the Aspects or policy templates. Click one parameter, and then click . The Edit/Combine Parameters dialog box opens.

13. In the Create Management Template wizard, click **Finish** to save the Management Template and close the wizard. The new Management Template appears in the Management Templates & Aspects pane.


Chapter 6

Deployment Scenario

Use Case: You want to monitor a Big Data environment, which comprises several Hadoop clusters. In such an environment, you want to monitor the Hadoop components - HDFS and MapReduce.

Task 1: Adding Nodes to the BSM console

Before you begin monitoring, you need to add the nodes to the BSM console.

1. Open the Monitored Nodes manager from the Operations Management Administration:
Admin > Operations Management > Setup > Monitored Nodes
2. In the Node Views pane, click  and then click **Generic node** to open the Create New Monitored Node dialog box.
3. Specify the Primary DNS Name, IP address, Operating System, and Processor Architecture of the node and click **OK**.

Task 2: Deploying the Hadoop Discovery Aspect


To discover the Hadoop CIs on the managed nodes, you must deploy the Hadoop Discovery Aspect.

1. Open the Management Templates and Aspects manager:
Admin > Operations Management > Monitoring > Management Templates and Aspects
2. In the Management Templates and Aspects pane:
Configuration Folders > Hadoop Management > Hadoop Aspects
3. In the **Hadoop Aspects** folder, right-click the Hadoop discovery aspect, and then click **Assign and Deploy** item to open the Assign and Deploy Wizard.
4. In the Configuration Item page, click the **configuration item** to which you want to deploy the Discovery Aspect and then click **Next**.
5. Click **Next**.
6. *Optional.* If you do not want to enable the assignment immediately, clear the **Enable Assigned Objects** check box. You can then enable the assignment later using the Assignments and Parameters manager.
7. Click **Finish**.

Task 3: Deploying Essential Hadoop Management Template

Before deploying the Essential Hadoop Management Template, you must discover the Hadoop CIs on all the managed nodes using the Discovery Aspect.

You can deploy the Essential Hadoop Management Template using the following steps:

1. Open the Management Templates and Aspects manager:
Admin > Operations Management > Monitoring > Management Templates and Aspects
2. In the Management Templates and Aspects pane:
Configuration Folders > Hadoop Management > Hadoop Management Templates
3. In the **Hadoop Management Templates** folder, click the Management Template that you want to deploy, and then click . The Assign and Deploy wizard opens.
4. In the Configuration Item page, click the CI to which you want to assign the Management Template, and then click **Next**.
5. In the Parameter page, click **Next**.
6. *Optional.* If you do not want to enable the assignment immediately, clear the **Enable Assigned Objects** check box. You can then enable the assignment later using the Assignments and Parameters manager.
7. Click **Finish**.

Appendix

Data Source and Metrics

The following table provides information about the metrics that are logged into the data source.

Data Source: HADOOP_AVAIL

Class Name	Data Type	Category Type	Metric
HADOOP_NN_AVA	KEY	UTF8	NN_INSTANCE_NAME
HADOOP_NN_AVA	GGE	R64	NN_AVAILABILITY
HADOOP_JT_AVA	KEY	UTF8	JT_INSTANCE_NAME

Class Name	Data Type	Category Type	Metric
HADOOP_JT_AVA	GGE	R64	JT_AVAILABILITY
HADOOP_DN_AVA	KEY	UTF8	DN_INSTANCE_NAME
HADOOP_DN_AVA	GGE	R64	DD_AVAILABILITY
HADOOP_TT_AVA	KEY	UTF8	TT_INSTANCE_NAME
HADOOP_TT_AVA	GGE	R64	TT_AVAILABILITY
HADOOP_SN_AVA	KEY	UTF8	SN_INSTANCE_NAME
HADOOP_SN_AVA	GGE	R64	SN_AVAILABILITY

Data Source: HADOOP_DATA

Class Name	Data Type	Category Type	Metric
HADOOP_NAMENODE	KEY	UTF8	NN_INSTANCE_NAME
HADOOP_NAMENODE	GGE	R64	NN_HEAP_USAGE
HADOOP_NAMENODE	GGE	R64	NN_GC_TIMESPENT
HADOOP_NAMENODE	GGE	R64	NN_FILES_CREATED
HADOOP_NAMENODE	GGE	R64	NN_FILES_DELETED
HADOOP_NAMENODE	GGE	R64	NN_FILES_RENAMED
HADOOP_NAMENODE	GGE	R64	NN_FILES_TOTAL
HADOOP_NAMENODE	GGE	R64	NN_CAPACITY_USED
HADOOP_NAMENODE	GGE	R64	NN_CAPACITY_TOTAL
HADOOP_NAMENODE	GGE	R64	NN_CAPACITY_LOAD
HADOOP_NAMENODE	GGE	R64	NN_PENDREP_BLOCK
HADOOP_NAMENODE	GGE	R64	NN_UNDERREP_BLOCK
HADOOP_NAMENODE	GGE	R64	NN_BLOCK_CAPACITY
HADOOP_NAMENODE	GGE	R64	NN_BLOCK_CORRUPT

Class Name	Data Type	Category Type	Metric
HADOOP_NAMENODE	GGE	R64	NN_BLOCK_MISSING
HADOOP_NAMENODE	GGE	R64	NN_BLOCK_TOTAL
HADOOP_NAMENODE	GGE	R64	NN_BLOCK_EXCESS
HADOOP_NAMENODE	GGE	R64	NN_HEARTBEAT_OPT
HADOOP_NAMENODE	GGE	R64	NN_HRTBEAT_AVGTIME
HADOOP_NAMENODE	GGE	R64	NN_SENT_BYTES
HADOOP_NAMENODE	GGE	R64	NN_RECEIVED_BYTES
HADOOP_JOBTRACKER	KEY	UTF8	JT_INSTANCE_NAME
HADOOP_JOBTRACKER	GGE	R64	JT_HEAP_USAGE
HADOOP_JOBTRACKER	GGE	R64	JT_GC_TIMESPENT
HADOOP_JOBTRACKER	GGE	R64	JT_WAITING_MAPS
HADOOP_JOBTRACKER	GGE	R64	JT_WAITING_REDUCE
HADOOP_JOBTRACKER	GGE	R64	JT_MAPS_LAUNCHED
HADOOP_JOBTRACKER	GGE	R64	JT_MAPS_SLOTS
HADOOP_JOBTRACKER	GGE	R64	JT_MAPSLOT_USE_PER
HADOOP_JOBTRACKER	GGE	R64	JT_REDUCE_LAUNCHED
HADOOP_JOBTRACKER	GGE	R64	JT_REDUCE_SLOTS
HADOOP_JOBTRACKER	GGE	R64	JT_REDSLOT_USE_PER
HADOOP_JOBTRACKER	GGE	R64	JT_BLACKLTD_TRACKERS
HADOOP_JOBTRACKER	GGE	R64	JT_JOBS_SUBMITTED
HADOOP_JOBTRACKER	GGE	R64	JT_JOBS_COMPLETED
HADOOP_JOBTRACKER	GGE	R64	JT_RUNJOB_THGPUT
HADOOP_JOBTRACKER	GGE	R64	JT_JOBS_FAILED
HADOOP_JOBTRACKER	GGE	R64	JT_HEARTBEAT_OPT
HADOOP_JOBTRACKER	GGE	R64	JT_HRTBEAT_AVGTIME
HADOOP_JOBTRACKER	GGE	R64	JT_SENT_BYTES
HADOOP_JOBTRACKER	GGE	R64	JT_RECEIVED_BYTES
HADOOP_DATANODE	KEY	UTF8	DN_INSTANCE_NAME

Class Name	Data Type	Category Type	Metric
HADOOP_DATANODE	GGE	R64	DN_HEAP_USAGE
HADOOP_DATANODE	GGE	R64	DN_GC_TIMESPENT
HADOOP_DATANODE	GGE	R64	DN_READS_LCLIENT
HADOOP_DATANODE	GGE	R64	DN_READS_RCLIENT
HADOOP_DATANODE	GGE	R64	DN_READ_BLK_OPTN
HADOOP_DATANODE	GGE	R64	DN_READ_OPR_THGPUT
HADOOP_DATANODE	GGE	R64	DN_READ_BLK_AVGTIME
HADOOP_DATANODE	GGE	R64	DN_WRITES_LCLIENT
HADOOP_DATANODE	GGE	R64	DN_WRITES_RCLIENT
HADOOP_DATANODE	GGE	R64	DN_WRITE_BLK_OPTN
HADOOP_DATANODE	GGE	R64	DN_WRITE_OPR_THGPUT
HADOOP_DATANODE	GGE	R64	DN_WRITE_BLK_AVGTIME
HADOOP_DATANODE	GGE	R64	DN_HEARTBEAT_OPT
HADOOP_DATANODE	GGE	R64	DN_HRTBEAT_AVGTIME
HADOOP_DATANODE	GGE	R64	DN_SENT_BYTES
HADOOP_DATANODE	GGE	R64	DN_RECEIVED_BYTES
HADOOP_DATANODE	GGE	R64	DN_DFS_USED
HADOOP_DATANODE	GGE	R64	DN_CAPACITY
HADOOP_DATANODE	GGE	R64	DN_REMAINING
HADOOP_TASKTRACKER	KEY	UTF8	TT_INSTANCE_NAME
HADOOP_TASKTRACKER	GGE	R64	TT_HEAP_USAGE
HADOOP_TASKTRACKER	GGE	R64	TT_GC_TIMESPENT
HADOOP_TASKTRACKER	GGE	R64	TT_TASKFAIL_TIMEOUT
HADOOP_TASKTRACKER	GGE	R64	TT_SENT_BYTES
HADOOP_TASKTRACKER	GGE	R64	TT_RECEIVED_BYTES

Eventable Metrics in Hadoop

Metric_ID	Metric_Name
-----------	-------------

10001	NameNode_HeapMemoryUsage
10002	NameNode_GC_TimeSpent
10007	NameNode_CapacityUsedGB
10009	NameNode_PendingReplicationBlocks
10010	NameNode_UnderReplicatedBlocks
10012	NameNode_CorruptBlocks
10013	NameNode_MissingBlocks
10016	NameNode_HeartBeatOperations
10020	JobTracker_HeapMemoryUsage
10021	JobTracker_GC_TimeSpent
10028	JobTracker_BlackListedTrackers
10032	JobTracker_HeartBeatOperations
10036	DataNode_HeapMemoryUsage
10037	DataNode_GC_TimeSpent
10046	DataNode_HeartBeatOperations
10050	DataNode_DfsUsed
10053	TaskTracker_HeapMemoryUsage
10054	TaskTracker_GC_TimeSpent

We appreciate your feedback!

If an email client is configured on this system, by default an email window opens when you click on the bookmark “Comments”.

In case you do not have the email client configured, copy the information below to a web mail client, and send this email to **docfeedback@hp.com**

Product name:

Document title:

Version number:

Feedback:

