Overview

SPM supports monitoring of both **MRv1** (0.22 and earlier, 1.0, 1.1) and **YARN** (0.23, 2.*) based Hadoop versions. Since the architecture is different, SPM uses different application types for them and different reports are available.

Common reports for all Hadoop types:

- Overview
- NameNode
- DataNode
- CPU & Mem
- Disk
- Network
- JVM
- GC

In addition to that, MRv1 versions also get the following reports:

- JobTracker
- JobTracker Queues
- TaskTracker

While reports specific for YARN versions are:

- ResourceManager
- ResourceManager Queues
- NodeManager

In some cases, some reports will be empty because particular Hadoop version doesn't expose some metrics over JMX. For instance, 0.20, 0.21, 0.22 **MRv1** versions of Hadoop will not have data in **JobTracker**, **JobTracker Queues** and **TaskTracker** reports (while 1.0 and 1.1 will have all reports populated). **NOTE**: regardless of this, you can monitor JVM stats of **JobTracker** and **TaskTracker** processes under JVM report for all **MRv1** versions (0.20, 0.21, 0.22 included). Also, since **SecondaryNameNode** doesn't expose specific metrics, it doesn't have a special report, but it can also be monitored under JVM report (for instance, you can create an alert to notify you when its heap size reaches some limit or drops to 0, meaning the process likely died).

All **YARN** versions (0.23, 2.*) will display all available reports and we expect all new Hadoop versions to continue behaving like that.

YARN versions don't have separate reports for the following components (since they don't expose specific metrics):

- **HistoryServer**
- **WebAppProxy**

However, you can still monitor these processes under JVM report, in the same way as **SecondaryNameNode** can be monitored in **MRv1** setups. You can

also define any alerts which are based on JVM metrics which should be good enough for most situations.

Integration

- Instructions: https://apps.sematext.com/ui/howto/Hadoop/overview

Metrics

Metric Name	Key	Agg	Type	Description
num allocated containers	hadoop.yarn.containers.allocated	Avg	Double	containers.allocated
containers launched	hadoop.yarn.num.containers.launched	Sum	Long	containers.launched
containers killed	hadoop.yarn.num.containers.killed	Sum	Long	containers.killed
containers completed	hadoop.num.shuffle	Sum	Long	shuffle.output.bytes
containers completed	hadoop.yarn.num.containers.completed	Sum	Long	containers.completed
available GB	hadoop.yarn.available	Avg	Double	available
shuffle connections	hadoop.num.shuffle	Sum	Long	shuffle.connections

Metric Name	Key	Agg	Type	Description
shuffle outputs failed	hadoop...	Sum	Long	shuffle.output.failed
containers running	hadoop.yarn...	Avg	Double	containers.running
containers failed	hadoop...	Sum	Long	containers.failed
containers inited	hadoop.yarn...	Avg	Double	containers.initiating
shuffle outputs ok	hadoop...	Sum	Long	shuffle.output.ok
allocated GB	hadoop.yarn...	Avg	Double	allocated
applications killed	hadoop...	Sum	Long	apps.killed
running 300	hadoop.yarn...	Avg	Double	running.300
available MB	hadoop.yarn...	Avg	Double	memory.available
applications failed	hadoop...	Sum	Long	apps.failed
applications submitted	hadoop...	Sum	Long	apps.submitted
running 60	hadoop.yarn...	Avg	Double	running.60

Metric Name	Key	Agg	Type	Description
pending MB	hadoop...yarn.memory...	Avg	Double	pending
applications pending	hadoop...yarn.apps...	Avg	Double	pending
active applications	hadoop...yarn.apps...	Avg	Double	active
containers allocated	hadoop...yarn.containers.alloc	Sum	Long	containers.alloc
running 0	hadoop...yarn.running0	Avg	Double	0
containers pending	hadoop...yarn.containers.pending	Avg	Double	containers.pending
applications completed	hadoop...yarn.apps...	Sum	Long	completed
allocated MB	hadoop...yarn.memory.alloc	Avg	Double	alloc
containers reserved	hadoop...yarn.containers.reserved	Avg	Double	containers.reserved
active users	hadoop...yarn.users...	Avg	Double	active
containers released	hadoop...yarn.containers.released	Sum	Long	containers.released

Metric Name	Key	Agg	Type	Description
reserved MB	hadoop...mem...	Avg	Double	reserved
applications running	hadoop...apps	Avg	Double	running
running 1440	hadoop...running	Avg	Double	running 1440
active NMs	hadoop...nm.active	Avg	Double	active
decom NMs	hadoop...nm.active	Avg	Double	active.decom
rebooted NMs	hadoop...nm.active	Avg	Double	active.rebooted
lost NMs	hadoop...nm.active	Avg	Double	active.lost
unhealthy NMs	hadoop...nm.active	Avg	Double	active.unhealthy
blocks replicated	hadoop...blocks	Sum	Long	blocks.replicated
blocks read	hadoop...blocks	Sum	Long	blocks.read
writes from local client	hadoop...io.write	Sum	Long	io.write.local

Metric Name	Key	Agg	Type	Description
copy block op avg time	hadoop	Avg	Double	hadoop.blockop.copies.time
heartbeats num ops	hadoop.dfsio.write	Sum	Long	heartbeats.hartbeats
reads from remote client	hadoop.dfsio.read	Sum	Long	read.remote
read from local client	hadoop.dfsio.read	Sum	Long	read.local
write block num ops	hadoop.blocks	Sum	Long	blockop.writes
blocks removed	hadoop.blocks	Sum	Long	blocks.removed
blocks verified	hadoop.blocks	Sum	Long	blocks.verified
replace block op avg time	hadoop	Avg	Double	hadoop.blockop.replaces.time

Metric Name	Key	Agg	Type	Description
writes from remote client	hadoop.dfs.io.write.remote	Sum	Long	
replace block num ops	hadoop.dfs.blockop.replaces	Sum	Long	
copy block num ops	hadoop.dfs.blockop.copies	Sum	Long	
write block op avg time	hadoop.dfs.blockop.writes.time	Avg	Double	
blocks written	hadoop.dfs.blockwrite	Sum	Long	
read block op avg time	hadoop.dfs.blockop.reads.time	Avg	Double	
heartbeats avg time	hadoop.dfs.io.heartbeats.time	Avg	Double	
bytes read	hadoop.dfs.io.read	Sum	Long	

Metric Name	Key	Agg	Type	Description
read block num ops	hadoop...blocks	Sum	Long	hadoop.reads
bytes written	hadoop...io.write	Sum	Long	
map task slots	hadoop...mapslots	Avg	Double	
reduces running	hadoop...reduces	Avg	Double	running
tasks failed timeout	hadoop...tasksfailed	Sum	Long	.timeout
tasks failed ping	hadoop...tasksfailed	Sum	Long	.ping
tasks completed	hadoop...taskscom	Sum	Long	pleted
reduce task slots	hadoop...reduceslots	Avg	Double	
maps running	hadoop...mapsrunning	Avg	Double	
waiting reduces	hadoop...waiting	Avg	Double	waiting

Metric Name	Key	Agg	Type	Description
waiting maps	hadoop.jt.map	Avg	Double	waiting maps
jobs killed	hadoop.jt.jobs.killed	Sum	Long	jobs killed
jobs submitted	hadoop.jt.jobs.sub	Sum	Long	jobs submitted
maps failed	hadoop.jt.maps.fail	Sum	Long	maps failed
running 300	hadoop.jt.running	Avg	Double	running 300
maps killed	hadoop.jt.maps.kill	Sum	Long	maps killed
reduce slots	hadoop.jt.reduce.slots	Avg	Double	reduce slots
running 60	hadoop.jt.running	Avg	Double	running 60
maps launched	hadoop.jt.maps.launch	Sum	Long	maps launched
jobs preparing	hadoop.jt.jobs.prep	Avg	Double	jobs preparing
reduces completed	hadoop.jt.reduces	Sum	Long	reduces completed
jobs failed	hadoop.jt.jobs.fail	Sum	Long	jobs failed

Metric Name	Key	Agg	Type	Description
running 1440	hadoop.jt	Avg	Double	running 1440
reduces launched	hadoop.jt	Sum	Long	reduces launched
map slots	hadoop.jt	Avg	Double	map slots
reduces killed	hadoop.jt	Sum	Long	reduces killed
maps completed	hadoop.jt	Sum	Long	maps completed
jobs running	hadoop.jt	Avg	Double	jobs running
reduces failed	hadoop.jt	Sum	Long	reduces failed
jobs completed	hadoop.jt	Sum	Long	jobs completed
running 0	hadoop.jt	Avg	Double	running 0
get listing ops	hadoop	Sum	Long	ops.listing
renamed files	hadoop	Sum	Long	renamed
decom nodes	hadoop	Avg	Double	decom

Metric Name	Key	Agg	Type	Description
dead nodes	hadoop.nodes.dead	Avg	Double	
create file ops	hadoop.fs.ops.create	Sum	Long	
blocks pending replication	hadoop.blocks.pending.replication	Avg	Long	
missing blocks	hadoop.blocks.missing	Avg	Long	
under replicated blocks	hadoop.blocks.underreplicated	Avg	Long	
file info ops	hadoop.fs.ops.info	Sum	Long	
blocks total	hadoop.blocks	Avg	Long	
blocks pending deletion	hadoop.blocks.pending.deletion	Avg	Long	
corrupt blocks	hadoop.blocks.corrupt	Avg	Long	

Metric Name	Key	Agg	Type	Description
capacity remaining	hadoopAvg	Long	hadoop.capacity.remaining	
total files	hadoopAvg	Long	hadoop.files	
appended files	hadoopSum	Long	hadoop.files.appended	
excess blocks	hadoopAvg	Long	hadoop.blocks.excess	
capacity used	hadoopAvg	Long	hadoop.capacity.used	
deleted files	hadoopSum	Long	hadoop.files.deleted	
delete file ops	hadoopSum	Long	hadoop.files.ops.delete	
capacity total	hadoopAvg	Long	hadoop.capacity	
created files	hadoopSum	Long	hadoop.files.created	
live nodes	hadoopAvg	Double	hadoop.nodes.live	
scheduled replication blocks	hadoopAvg	Long	hadoop.blocks.scheduled.replication	
heartbeats	hadoopSum	Long	hadoop.heartbeats	

Metric Name	Key	Agg	Type	Description
blacklisted trackers	hadoop.jt...	SUM	Long	trackers.blacklisted
running reduces	hadoop.jt...	AVG	Double	running
occupied map slots	hadoop.jt...	AVG	Double	map.occupied
running maps	hadoop.jt...	AVG	Double	running
blacklisted maps	hadoop.jt...	SUM	Long	blacklisted
blacklisted reduces	hadoop.jt...	SUM	Long	blacklisted
occupied reduce slots	hadoop.jt...	AVG	Double	reduce.occupied
decommissioned trackers	hadoop.jt...	SUM	Long	trackers.decommissioned
trackers	hadoop.jt...	SUM	Long	trackers
waiting reduces	hadoop.jt...	AVG	Double	waiting reduces
waiting maps	hadoop.jt...	AVG	Double	waiting maps

Metric Name	Key	Agg	Type	Description
map slots	hadoop.jt.map.slots	Avg	Double	map
graylisted trackers	hadoop.jt.reduce	Sum	Long	trackers.graylisted
reduce slots	hadoop.jt.reduce.slots	Avg	Double	reduce