

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 allo- cated con- tain- ers	hadoop	Avg	Container	Containers allocated
containers launched	hadoop	Sum	Container	Containers launched
containers killed	hadoop	Sum	Container	Containers killed
containers com- pleted	hadoop	Sum	Shuffle	Shuffle output bytes completed
containers com- pleted	hadoop	Sum	Container	Containers completed
available GB	hadoop	Avg	Variable	Available
shuffle con- nec- tions	hadoop	Sum	Shuffle	Shuffle connections

Metric				
Name	Key	Agg	Type	Description
shuffle out-puts failed	hadoopSum.shuffle	Sum	Double	shuffle output.failed
containers running	hadoopAvg.cd	Avg	Double	containers.running
containers failed	hadoopSum.cd	Sum	Double	containers.failed
containers initied	hadoopAvg.cd	Avg	Double	containers.initiating
shuffle out-puts ok	hadoopSum.shuffle	Sum	Double	shuffle output.ok
allocated GB	hadoopAvg.all	Avg	Double	allocated
application killed	hadoopSum.app	Sum	Double	application.killed
running 300	hadoopAvg.ru	Avg	Double	running300
available MB	hadoopAvg.m	Avg	Double	available
application failed	hadoopSum.app	Sum	Double	application.killed
application sub-mitted	hadoopSum.app	Sum	Double	application.submitted
running 60	hadoopAvg.ru	Avg	Double	running60

<hr/>				
Metric				
Name	Key	Agg	Type	Description
<hr/>				
pendingMB	hadoop	Avg.m	Double	pending
application-pending	hadoop	Avg.app	Double	pending
active applications	hadoop	Avg.app	Double	active
containers-allocated	hadoop	Sum.container	Integer	containers.allocated
running	hadoop	Avg.run	Double	0
containers-pending	hadoop	Avg.container	Double	containers.pending
application-completed	hadoop	Sum.app	Integer	application.completed
allocatedMB	hadoop	Avg.m	Double	allocated
containers-reserved	hadoop	Avg.container	Double	containers.reserved
active users	hadoop	Avg.us	Double	active
containers-released	hadoop	Sum.container	Integer	containers.released

<hr/>				
Metric				
Name	Key	Agg	Type	Description
<hr/>				
reserved MB	hadoop	Avg	Double	reserved
application run-ning	hadoop	Avg	Double	running
running 1440	hadoop	Avg	Double	running 1440
active NMs	hadoop	Avg	Double	active NMs
decom NMs	hadoop	Avg	Double	decom NMs
rebooted NMs	hadoop	Avg	Double	rebooted NMs
lost NMs	hadoop	Avg	Double	lost NMs
unhealthy NMs	hadoop	Avg	Double	unhealthy NMs
blocks replicated	hadoop	Sum	Long	blocks replicated
blocks read	hadoop	Sum	Long	blocks read
writes from local client	hadoop	Sum	Long	writes from local client

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

<hr/>				
Metric				
Name	Key	Agg	Type	Description
<hr/>				
writes from re-mote client	hadoopSumio	write.remote	Long	
replace block num ops	hadoopSumblock	replace	Long	
copy block num ops	hadoopSumblock	copy	Long	
write block op avg time	hadoopAvgblock	write	Double	
blocks written	hadoopSumblock	write	Long	
read block op avg time	hadoopAvgblock	read	Double	
heartbeats	hadoopAvgio	heartbeat	Double	
bytes read	hadoopSumio	read	Long	

<hr/>				
Metric				
Name	Key	Agg	Type	Description
<hr/>				
read block num ops	hadoop	Sum	Double	hadoop.reads
bytes writ- ten	hadoop	Sum	Long	hadoop.io.write
map task slots	hadoop	Avg	Double	map.taskslots
reduce run- ning	hadoop	Avg	Double	reduce.taskslots
tasks failed time- out	hadoop	Sum	Long	tasks.failed.timeout
tasks failed ping	hadoop	Sum	Long	tasks.failed.ping
tasks com- pleted	hadoop	Sum	Long	tasks.completed
reduce task slots	hadoop	Avg	Double	reduce.taskslots
maps run- ning	hadoop	Avg	Double	map.taskslots
waiting re- duces	hadoop	Avg	Double	reduce.taskslots

Metric				
Name	Key	Agg	Type	Description
waiting maps	hadoop	sum	Double	Waiting maps
jobs kil- lled	hadoop	sum	Long	Jobs killed
jobs sub- mit- ted	hadoop	sum	Long	Jobs submitted
maps failed	hadoop	sum	Long	Maps failed
running 300	hadoop	avg	Double	Running 300
maps kil- lled	hadoop	sum	Long	Maps killed
reduce slots	hadoop	avg	Double	Reduce slots
running 60	hadoop	avg	Double	Running 60
maps launched	hadoop	sum	Long	Maps launched
jobs prepar- ing	hadoop	avg	Double	Jobs preparing
reduces com- pleted	hadoop	sum	Long	Reduces completed
jobs failed	hadoop	sum	Long	Jobs failed

Metric				
Name	Key	Agg	Type	Description
runninghadoop	hadoop	Avg	Double	1440
1440				
reduceshadoop	hadoop	Sum	Long	launched
launched				
map	hadoop	Avg	Double	1440
slots				
reduceshadoop	hadoop	Sum	Long	killed
kill-				
lled				
maps	hadoop	Sum	Long	completed
com-				
pleted				
jobs	hadoop	Avg	Double	1440
run-				
ning				
reduceshadoop	hadoop	Sum	Long	failed
failed				
jobs	hadoop	Sum	Long	completed
com-				
pleted				
runninghadoop	hadoop	Avg	Double	1440
0				
get	hadoop	Sum	Long	listing
list-				
ing				
ops				
renamedhadoop	hadoop	Sum	Long	renamed
files				
decom	hadoop	Avg	Double	1440
nodes				

<hr/>				
Metric				
Name	Key	Agg	Type	Description
<hr/>				
dead nodes	hadoop	Avg	double	Dead nodes
create file ops	hadoop	Sum	long	Files created
blocks pending replication	hadoop	Avg	double	Blocks pending replication
blocks missing	hadoop	Avg	double	Blocks missing
blocks under replicated	hadoop	Avg	double	Blocks under replicated
file info ops	hadoop	Sum	long	Files info ops
blocks total	hadoop	Avg	double	Blocks total
blocks pending deletion	hadoop	Avg	double	Blocks pending deletion
blocks corrupt	hadoop	Avg	double	Blocks corrupt

<hr/>				
Metric				
Name	Key	Agg	Type	Description
<hr/>				
capacity-re-main-ing	hadoop	Avg.capacity	Capacity	remaining
total files	hadoop	Avg.files	Long	total
appended files	hadoop	Sum.files	Long	appended
excess blocks	hadoop	Avg.blocks	Long	excess
capacity used	hadoop	Avg.capacity	Capacity	used
deleted files	hadoop	Sum.files	Long	deleted
delete file ops	hadoop	Sum.files	Long	delete
capacity to-tal	hadoop	Avg.capacity	Capacity	total
created files	hadoop	Sum.files	Long	created
live nodes	hadoop	Avg.nodes	Double	live
scheduled repli-ca-tion blocks	hadoop	Avg.blocks	Long	scheduled.replication
heartbeat	hadoop	Sum.heartbeat	Long	heartbeat

Metric				
Name	Key	Agg	Type	Description
blacklisted-trackers	hadoop.job. trackers	Sum	Long	trackers.blacklisted
running-reduces	hadoop.job.maps	Avg	Double	running
occupied-map-slots	hadoop.job.maps	Avg	Double	occupied
running-maps	hadoop.job.maps	Avg	Double	running
blacklisted-maps	hadoop.job.maps	Sum	Long	blacklisted
blacklisted-reduces	hadoop.job.maps	Sum	Long	blacklisted
occupied-reduce-slots	hadoop.job.maps	Avg	Double	occupied
decommissioned-trackers	hadoop.job.maps	Sum	Long	trackers.decommissioned
tracker	hadoop.job.maps	Sum	Long	trackers
waiting-reduces	hadoop.job.maps	Avg	Double	waiting
waiting-maps	hadoop.job.maps	Avg	Double	waiting

Metric				
Name	Key	Agg	Type	Description
map slots	hadoop	Agg	slot	Double
graylisted track-ers	hadoop	Sum	Double	trackers.graylisted
reduce slots	hadoop	Agg	slot	Double