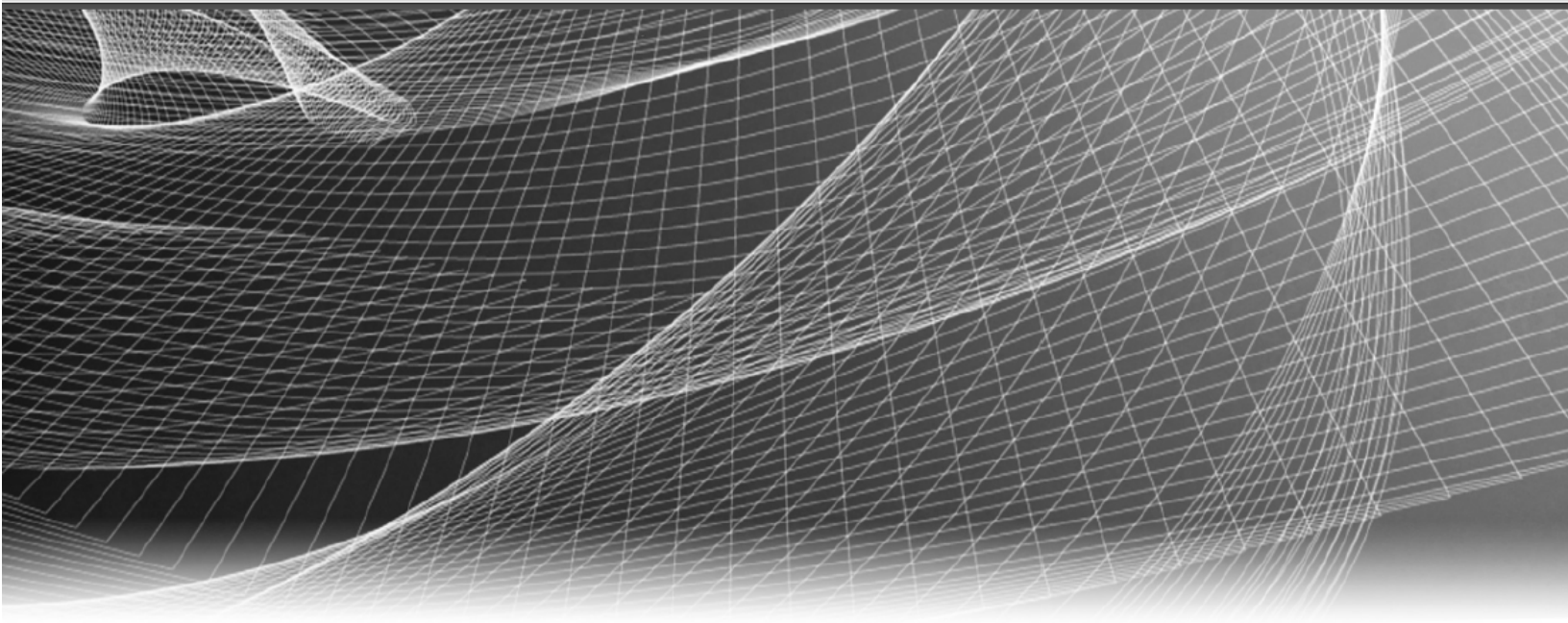


RSA Security Analytics

Warehouse (MapR) Configuration Guide

for Version 10.6.5

RSA EMC²

Contact Information

RSA Link at https://community.rsa.com contains a knowledgebase that answers common questions and provides solutions to known problems, product documentation, community discussions, and case management.

Trademarks

For a list of RSA trademarks, go to www.emc.com/legal/emc-corporation-trademarks.htm#rsa.

License Agreement

This software and the associated documentation are proprietary and confidential to EMC, are furnished under license, and may be used and copied only in accordance with the terms of such license and with the inclusion of the copyright notice below. This software and the documentation, and any copies thereof, may not be provided or otherwise made available to any other person.

No title to or ownership of the software or documentation or any intellectual property rights thereto is hereby transferred. Any unauthorized use or reproduction of this software and the documentation may be subject to civil and/or criminal liability.

This software is subject to change without notice and should not be construed as a commitment by EMC.

Third-Party Licenses

This product may include software developed by parties other than RSA. The text of the license agreements applicable to third-party software in this product may be viewed on the product documentation page on RSA Link. By using this product, a user of this product agrees to be fully bound by terms of the license agreements.

Note on Encryption Technologies

This product may contain encryption technology. Many countries prohibit or restrict the use, import, or export of encryption technologies, and current use, import, and export regulations should be followed when using, importing or exporting this product.

Distribution

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

October 2017

Contents

RSA Analytics Warehouse (MapR) Configuration Guide

This guide provides an overview of RSA Analytics Warehouse (MapR) and detailed instructions on how to configure the Warehouse appliance in your network. This guide provides the configuration procedures only for RSA Analytics Warehouse instances running MapR.

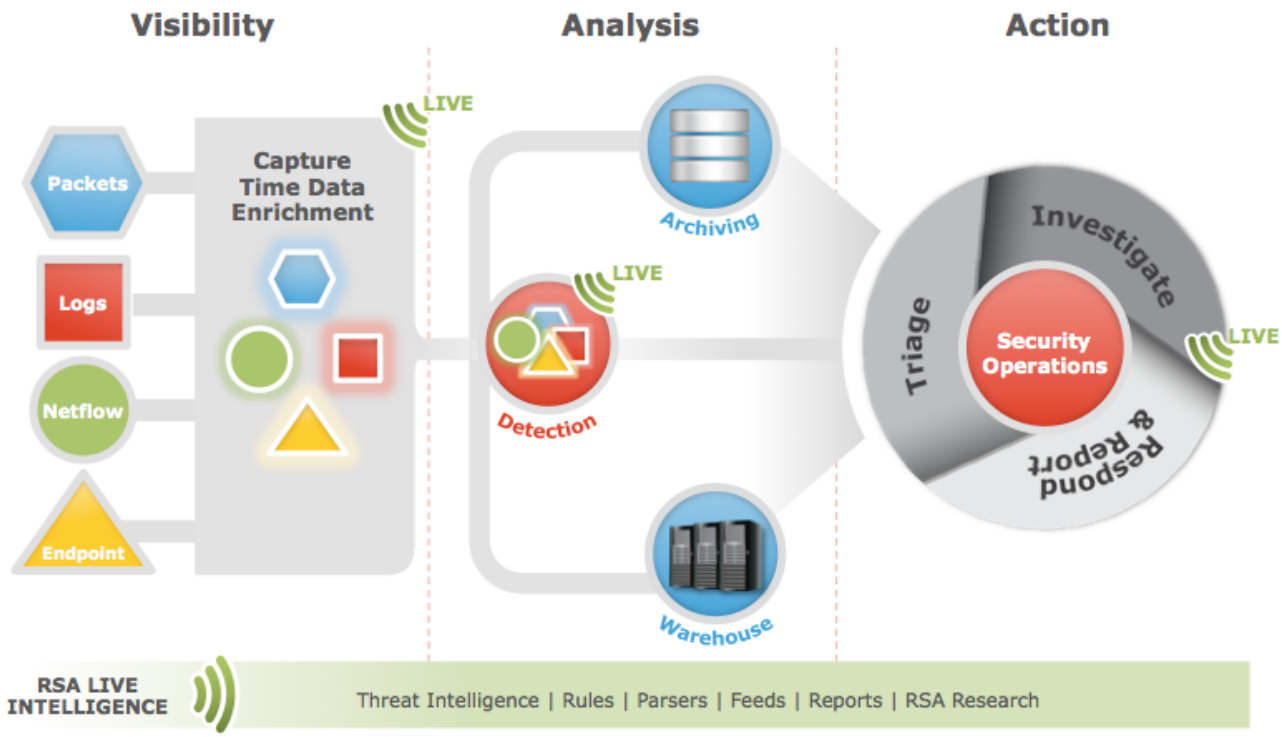
Topics

RSA Analytics Warehouse Overview

This topic provides an overview of the RSA Analytics Warehouse. RSA Analytics Warehouse provides the capacity to process large amounts of current and longer term data through a Hadoop-based distributed computing system that collects, manages, and enables advanced analytics and reporting on Security Analytics data. RSA Analytics Warehouse requires a service called Warehouse Connector to collect metadata and events from Decoder and Log Decoder and write them in Avro format into a Hadoop-based distributed computing system. For more information on the Warehouse Connector, see the **Warehouse Connector Overview** topic in the *Warehouse Connector Configuration Guide*.

The Warehouse can be made up of three or more nodes depending on the organization's analytic, archiving, and resiliency requirements.

The following diagram depicts the architecture of a Security Analytics network that implements the RSA Analytics Warehouse component.



This guide provides the configuration procedures only for RSA Analytics Warehouse instances running MapR.

Configure RSA Analytics Warehouse (MapR)

This topic provides instructions for configuring the nodes for the RSA Analytics Warehouse (MapR). It only applies to RSA Analytics Warehouse instances running MapR.

Prerequisites

Make sure that you have:

- Installed the RSA Analytics Warehouse (Warehouse) appliance in your network environment. For more information, see the *RSA Analytics Warehouse (MapR) Setup Guide* in the Hardware Setup guides.
- Configured the network interface of the Warehouse appliance.

Procedure

Perform the following procedures in the order shown to set up Warehouse. If you are planning to have a cluster of Warehouse appliances, make sure that you perform the following tasks on all the appliances in the cluster.

Caution: Prerequisites are mandatory. Your installation will fail if you have not set the network configuration as described in the *RSA Analytics Warehouse (MapR) Setup Guide* or *Virtual Host Setup Guide* depending on your deployment.

Process	Tasks/Instructions
1. Generate and Update the Default UUID in the Appliances.	Refer to Step 1. Generate and Update the Default UUID in Appliances.
2. Update the configuration template file in the Warehouse Appliance.	Refer to Step 2. Update the Configuration Template File.
3. Upgrade the RSA Analytics Warehouse cluster.	Refer to Step 3. Upgrade the Warehouse Cluster.
4. Install the Warehouse license file.	Refer to Step 4. Install the Warehouse License File.

Process	Tasks/Instructions
5. Generate the Virtual IP address.	Refer to Step 5. Generate the Virtual IP Address for Primary Appliance.
6. Configure Warehouse Connector to write to Warehouse.	Refer to Step 6. Configure the Connector to Write to Warehouse.
7. Configure other Security Analytics Services for the Warehouse.	Refer to Step 7. Configure other Security Analytics Services.

Post Configuration Steps

The following table lists the optional tasks that you can perform after configuring the Warehouse:

Process	Tasks/Instructions
1. Generate and Update the Default UUID in the Appliances.	Refer to Step 1. Generate and Update the Default UUID in Appliances.
2. Update the configuration template file in the Warehouse Appliance.	Refer to Step 2. Update the Configuration Template File.
3. Upgrade the RSA Analytics Warehouse cluster.	Refer to Step 3. Upgrade the Warehouse Cluster.
4. Install the Warehouse license file.	Refer to Step 4. Install the Warehouse License File.
5. Generate the Virtual IP address.	Refer toStep 5. Generate the Virtual IP Address for Primary Appliance.
6. Configure Warehouse Connector to write to Warehouse.	Refer to Step 6. Configure the Connector to Write to Warehouse.
7. Configure other Security Analytics Services for the Warehouse.	Refer to Step 7. Configure other Security Analytics Services.

Process	Tasks/Instructions
Tasks	**Reference**
1. Access MapR Control System user interface for Warehouse cluster administration.	Refer to Access MapR Control System UI for Cluster Administration.
2. Enable MapR Metrics on the Warehouse Cluster.	Refer to Enable MapR Metrics on RSA Analytics Warehouse Cluster.

Step 1. Generate and Update the Default UUID in Appliances

This topic provides instructions on how to manually generate and update the default UUID on the Appliances in the cluster.

Procedure

You need to manually generate and update the default UUID on the Appliances in the cluster. The UUID should be unique to the Appliance in the cluster.

To generate and update the default UUID in the Appliance:

1. Log on to the Appliance as root user.
2. To generate the UUID and copy it in the proper files, type the following commands:
   - `/opt/mapr/server/mruuidgen > /opt/mapr/hostid`
   - `cp /opt/mapr/hostid /opt/mapr/server/hostid.xxxxx`

   Where, xxxxx refers to the 5-digit number randomly assigned to the existing file.

   Note: Review **/opt/mapr/server** for the full name of this file.
3. To restart the appliance, enter the following command:

   ```
   reboot
   ```

Step 2. Update the Configuration Template File

This topic provides instructions on how to update the configuration template file in the RSA Analytics Warehouse Appliance. The configuration template file in the RSA Analytics Warehouse appliance must include the following parameters:

- nodes
- Internalnetworks
- clustername
- disks

By default, a configuration template is provided with the RSA Analytics Warehouse appliance. The configuration template is located on the RSA Analytics Warehouse appliance at **/opt/rsa/saw/install**.

Prerequisites

Make sure that you have validated the volume in the server to identify available drive space for Warehouse to store data. The total drive space of the additional volume is considered as a single drive by the HDFS. In Warehouse, the AVRO files are stored in this drive space.

Note: The server contains additional volumes of identical size other than the operation system volume.

To identify free volume, enter the command `fdisk -l | grep /dev/s |sort` in the Warehouse node. You will get a list of disks that are not partitioned for usage. You need to list the identified disks in the configuration template file so that Warehouse utilizes this space for the Hadoop Cluster.

Procedure

To update the configuration template file in the RSA Analytics Warehouse Appliance:

1. Log on to the appliance as the root user.
2. Navigate to **/opt/rsa/saw/install**, enter the following command:

   ```
   cd /opt/rsa/saw/install
   ```
3. Create a copy of the configuration template, enter the following command:

   ```
   cp conf.template conf.template-<name>
   ```

   where `<name>` is the custom name of the configuration template file.
4. Edit the configuration template file, enter the following command:

```
vi conf.template-<name>
```

Parameter	Description
nodes	List the IP addresses of the appliances in the cluster separated by spaces. All the appliances in the cluster must be listed in the same order in every configuration file for every RSA Analytics Warehouse appliance.
Internalnetworks	List the network addresses in CIDR format separated by spaces. This Warehouse appliance cluster communication is limited to the provided network addresses. **Note:** RSA recommends that you do not leave this parameter blank.
clustername	Name of the cluster. The cluster name is used to identify the Network File System (NFS) share.
disks	Displays the list of disks recognized by the operation system, and these disks will be formatted in HDFS for the Warehouse when this configuration script is executed.

The following figure shows a sample configuration template file:

```
[root@saw-node2 install]# vi conf.template-test
[global]

# nodes: List of the first 5 node IP addresses in the cluster, separated by
#        spaces. Use addresses on internal network if restricting network traffic

nodes=xxx.108.x.25 xxx.108.x.27 xxx.108.x.33

# internalnetworks: List of network addresses, in CIDR format separated by
#                   spaces, that cluster communication will be limited to.
#                   Leave blank to allow communication over any network

internalnetworks=xxx.108.0/24

# clustername: Name of cluster. NFS share will be /mapr/<clustername>

clustername=saw

# Internal settings - changing these may result in unsupported behavior

[internal]

disks=/dev/sdb /dev/sdc /dev/sdd /dev/sde /dev/sdf /dev/sdg /dev/sdh /dev/sdi /dev/sdj
```

5. Execute the configuration template file, enter the following command:

```
./configure.py conf.template-<name>
```

6. Restart the appliance, enter the following command:

```
reboot
```

Step 3. Upgrade the Warehouse Cluster

This topic provides instructions on how to upgrade the warehouse cluster. This procedure is required after updating the configuration template file and rebooting the RSA Analytics Warehouse appliance.

Procedure

1. Follow the "Upgrade MapR to Latest Components" instructions in the *RSA Security Analytics v10.6 Upgrade Instructions* located on RSA SecurCare Online (SCOL): https://knowledge.rsasecurity.com
2. Manually open Hiveserver port 10000, which is not opened by default:
   a. Get the line number where the REJECT statement appears in the Iptable.
   b. Ensure that the Iptables service is running and enter the following command:

      ```
      NUM=$(iptables -L INPUT -n --line-numbers |grep 'reject-
      with' |awk ' {print $1}')
      ```

      Note: The ACCEPT statements that follow the REJECT statement in the Iptables will not take effect. You can incorporate the line number of the REJECT statement in the command to ensure that the ACCEPT statements preceed the REJECT statement.

   c. Add the firewall exception for port 10000 to the Iptables. Enter the following command:

      ```
      iptables -I INPUT $NUM -m state --state NEW -p tcp --
      dport 10000 -j ACCEPT
      ```

   d. Save the Iptables. Enter the following command:

      ```
      /etc/init.d/iptables save
      ```

   e. Restart the Iptables. Enter the following command:

      ```
      /etc/init.d/iptables restart
      ```

   f. Verify if the firewall exceptions for the ports are added. Enter the following command:

      ```
      Service iptables status | grep 10000
      ```

      The following output should be displayed:

      ```
      ACCEPT tcp -- 0.0.0.0/0 0.0.0.0/0 state NEW tcp dpt:10000
      ```

Step 4. Install the Warehouse License File

This topic provides instructions on how to manually install the RSA Analytics Warehouse (Warehouse) license file in the Warehouse Appliance. You need to manually install the Warehouse license file on the Warehouse appliance. If you have a cluster of Warehouse appliances, you need to install the license file on the first Warehouse appliance in the cluster.

Prerequisites

Make sure that you have:

- Obtained the Warehouse license file.
- Copied the license file to **/root/** on the first Warehouse appliance in the cluster using a USB drive or through SCP.

Procedure

To install the Warehouse license file:

1. Log on to the appliance as the root user.
2. To install the license file, enter the following command:

   ```
   maprcli license add -is_file true -license <license_filename>
   ```

   where `<license_filename>` is filename of the RSA Analytics Warehouse license file. The license file is installed without any output messages. If you included a network range in the `internalnetworks` parameter in the configuration template file, a warning message appears suggesting that the Warehouse is configured only to communicate with the network entered in the configuration template file. You can ignore this warning as this does not have any functional issue.
3. To confirm the license file installation, enter the following command:

   ```
   maprcli license list
   ```

   Output messages appear on the console screen. The last two lines of the output message should be similar to the following sample:

   ```
   hash: "b8xO1flW8EMNSqq7zztn8D2BXnQ="
    3 May 14, 2013
   ```
4. To retrieve a list of directories, run the following command:

   ```
   hadoop fs -ls /
   ```

Step 5. Generate the Virtual IP Address for Primary Appliance

This topic provides instructions to generate a virtual IP address for the primary RSA Analytics Warehouse (Warehouse) appliance.

Prerequisites

Make sure that you note the MAC addresses of all the Warehouse appliances in the cluster. Use the following command on the appliance to view the MAC address of appliance:

```
ifconfig <interface> | grep HWaddr
```

where `<interface>` is the network interface.

Procedure

To generate a virtual IP address for the primary Warehouse appliance:

1. Log on to the primary appliance as root user.
2. Create the virtual IP address. Enter the following command:
   ```
   maprcli virtualip add -virtualip <VIP_address> -netmask
   <netmask> -macs <mac_node1> <mac_node2> <mac_node3> ……< mac_
   node n>
   ```
   where:
   - `<VIP_address>` is the virtual IP address for the primary Warehouse appliance.
   - `<netmask>` is the netmask address of the primary Warehouse appliance.
   - `<mac_node1>` is the MAC address of the first node in the Warehouse cluster.
   - `<mac_node2>` is the MAC address of the second node in the Warehouse cluster.

   For example, if the MAC address for node 1 is `01:Z1:1X:00:20:Y1` and node 2 is `32:Y2:4Z:40:10:X3`, and the IP address is `192.168.100.10`, then enter the command as following:
   ```
   maprcli virtualip add -virtualip 192.168.100.10 -
   netmask <netmask> -macs 01:Z1:1X:00:20:Y1 32:Y2:4Z:40:10:X3
   ```
3. Verify the virtual IP address. Enter the following command:
   ```
   maprcli virtualip list
   ```

To add or remove virtual IP addresses, you can use the command line or the MapR Control System. Edit and Remove Virtual IP Addresses (Command Line) and Add and Remove a Virtual IP Address (MapR UI) provide additional information.

Step 6. Configure the Connector to Write to Warehouse

This topic provides instructions for enabling the services to write to RSA Analytics Warehouse.

Procedure

To configure Warehouse Connector to write to the Warehouse, perform the following tasks on the Log Decoders and Decoders where the Warehouse Connectors are installed:

Note: If you are configuring on a virtual environment, perform these tasks on a standalone Warehouse Connector server.

Tasks	References
1. Verify the Network File System (NFS) Services status.	Refer to Verify the Network File System (NFS) Services Status
2. Install the NFS services package.	Refer to Install the Network File System Packages.
3. Mount RSA Analytics Warehouse on the appliance where you have installed the Warehouse Connector service.	Refer to Mount the Warehouse on the Warehouse Connector.

Verify the Network File System (NFS) Services Status

Perform the following to verify the status of the Network File System (NFS) on the appliance where you have installed the Warehouse Connector service.

To verify the NFS services status:

1. Log on to the Warehouse Connector appliance where you have installed the Warehouse Connector service.
2. Enter the following command:
   ```
   rpm -qa |grep nfs
   ```
   The NFS package names appear in the output message. For example:
   ```
   nfs-utils-lib-1.1.5-6.el6.x86_64
   nfs-utils-1.2.3-36.el6.x86_64
   ```
3. If the output message is empty, install the NFS packages.

Install the Network File System Packages

This topic provides instructions to install the Network File System packages.

Prerequisites

If the NFS packages are already downloaded on the appliances manually, install the packages and mount RSA Analytics Warehouse. You need to have internet access to complete this task. If internet access is not available, you must download the RPM packages offline and copy them to this machine for installation.

Procedure

Perform the following to install the Network File System (NFS) packages on the appliance where you have installed the Warehouse Connector service.

Note: Install the NFS packages only if the NFS packages are not displayed when you verify the status of NFS in the Warehouse Connector appliance or on the appliance where you have installed the Warehouse Connector service.

To install NFS packages:

1. Log on to the Warehouse Connector appliance or on the appliance where you have installed the Warehouse Connector service.
2. To verify the NFS status, enter the following command:

   ```
   rpm -qa |grep nfs
   ```

   The NFS package names appear in the output message. For example:

   ```
   nfs-utils-lib-1.1.5-6.el6.x86_64
    nfs-utils-1.2.3-36.el6.x86_64
   ```

   If the `nfs-utils` and `nfs-utils-lib` are properly identified, you can skip the remaining steps in this procedure (*Install the Network File System Packages*).
3. To search for NFS package, enter the following command:

   ```
   yum search nfs-utils
   ```

   The output ends with the following message:

   ```
   "name and summary matches only, use "search all" for
   everything."
   ```

   Note: Contact RSA Customer Support if the output ends with the following message: `"no matches found"`
4. To install the NFS programs, enter the following command:

```
yum install nfs-utils nfs-utils-lib
```

The output prompts for **y** or **n**. Type **y** and press **ENTER**.

The NFS packages are successfully installed.

Mount the Warehouse on the Warehouse Connector

This topic provides instructions to mount RSA Analytics Warehouse on the Warehouse Connector.

Prerequisites

Make sure that the Network File System packages are on the appliance where you have installed the Warehouse Connector service.

Procedure

Perform the following steps to mount RSA Analytics Warehouse on the appliance where you have installed the Warehouse Connector service to write to RSA Analytics Warehouse:

1. To create a new directory named **/saw**, enter the following command:
   `mkdir /saw`
2. Enter the following command:
   `ll /`
   The new directory is displayed.
3. To mount the Warehouse, enter the following command:
   `mount -t nfs -o nolock,tcp,hard,intr <IP_Address_for_SAW>:/mapr/<cluster-name> /saw`
   Where `<IP_Address_for_SAW>` is the IP address of the primary Warehouse appliance in the cluster and `<cluster-name>` is the name provided in the template file.

   Note: If a virtual IP address is configured for the Warehouse, you have to use it as the IP address in `<IP_Address_for_SAW>`.
4. To verify if the Warehouse is mounted successfully, enter the following command
   `mount`
   The IP address of the primary Warehouse appliance and other details you have provided in **step 3** appear in the last line of the output message.
5. To list the content in the newly created directory, **/saw**, enter the following command:
   `ll /saw`
   The following directories are displayed:
   hbase
   index-scratch
   jars
   logs

user

var

6. To add NFS to the Auto-mount options. Do the following:

   a. To check if the IP address of the primary Warehouse appliance and other details you have provided while mounting Warehouse appears in **/etc/fstab**, enter the following command:

      ```
      cat /etc/fstab
      ```

      If the detail does not appear in the /etc/fstab file, perform the following steps.

   b. Enter the following command:

      ```
      tail -n 1 /etc/mtab
      ```

      The IP address of the primary Warehouse appliance and other details you provided while mounting Warehouse appear in the last line of the output message.

   c. Enter the following command:

      ```
      tail -n 1 /etc/mtab >> /etc/fstab
      ```

   d. Edit the /etc/fstab file to add the word 'auto' at the end of the file. Enter the following command:

      ```
      vi /etc/fstab
      ```

      For example, `10.11.111.11:/mapr/saw /saw nfs rw,nolock,tcp,auto,addr=10.11.111.11 0 0`

Step 7. Configure other Security Analytics Services

This topic provides additional instructions for configuring Security Analytics services for the RSA Analytics Warehouse (MapR).

Procedure

1. If you are not using Vulnerability Response Management (VRM), disable the Hbase services to return the configured memory so that it is available for use elsewhere in the cluster. To stop the Hbase services, you can use the command line or the MapR Control System. Stop the Hbase Services Using the Command Line and Stop the Hbase Services Using the MapR Control System provide additional information.
2. Add Warehouse data sources to the Reporting Engine. For the detailed procedure, see the **Add Warehouse as Data Source to Reporting Engine** topic in the *Reporting Engine Configuration Guide*.

Stop the Hbase Services Using the Command Line

This topic provides the steps to stop the Hbase services using the command line. If you are not using Vulnerability Response Management (VRM), stop the Hbase services to return the configured memory so that it is available for use elsewhere in the cluster.

Procedure

To stop the Hbase services using the command line:

1. To stop the **Hbase RegionServer** service on *all of the appliances*, enter the following command:
   ```
   maprcli node services -hbregionserver stop -filter "[hn==*]"
   ```

2. To stop the **Hbase RegionServer** service on *a specific node*:
   ```
   maprcli node services -hbregionserver stop -filter "
   [hn==<Hostname>]"
   ```
   Where `<Hostname>` is the specific node hostname.

3. To stop the **Hbase Master** service on *all of the appliances*, enter the following command:
   ```
   maprcli node services -hbmaster stop -filter "[hn==*]"
   ```

4. To stop the **Hbase Master** service on *a specific node*:
   ```
   maprcli node services -hbmaster stop -filter "
   [hn==<Hostname>]"
   ```
   Where `<Hostname>` is the specific node hostname.

Hbase Services Stop and Start Commands Summary

The following tables summarize the commands used to stop and start the Hbase services for the **HBase RegionServer** and **HBase Master** services.

HBase RegionServer	Command to run using the Command Line
Stop on All the Appliances	`maprcli node services -hbregionserver stop -filter "[hn==*]"`

HBase RegionServer	Command to run using the Command Line
Start on All the Appliances	`maprcli node services -hbregionserver start -filter "[hn==*]"`
Stop on Specific node	`maprcli node services -hbregionserver stop -filter "[hn==<Hostname>]"`
Start on Specific node	`maprcli node services -hbregionserver start -filter "[hn==<Hostname>]"`

HBase Master	Command to run using the Command Line
Stop on All the Appliances	`maprcli node services -hbmaster stop -filter "[hn==*]"`
Start on All the Appliances	`maprcli node services -hbmaster start -filter "[hn==*]"`
Stop on Specific node	`maprcli node services -hbmaster stop -filter "[hn==<Hostname>]"`
Start on Specific node	`maprcli node services -hbmaster start -filter "[hn==<Hostname>]"`

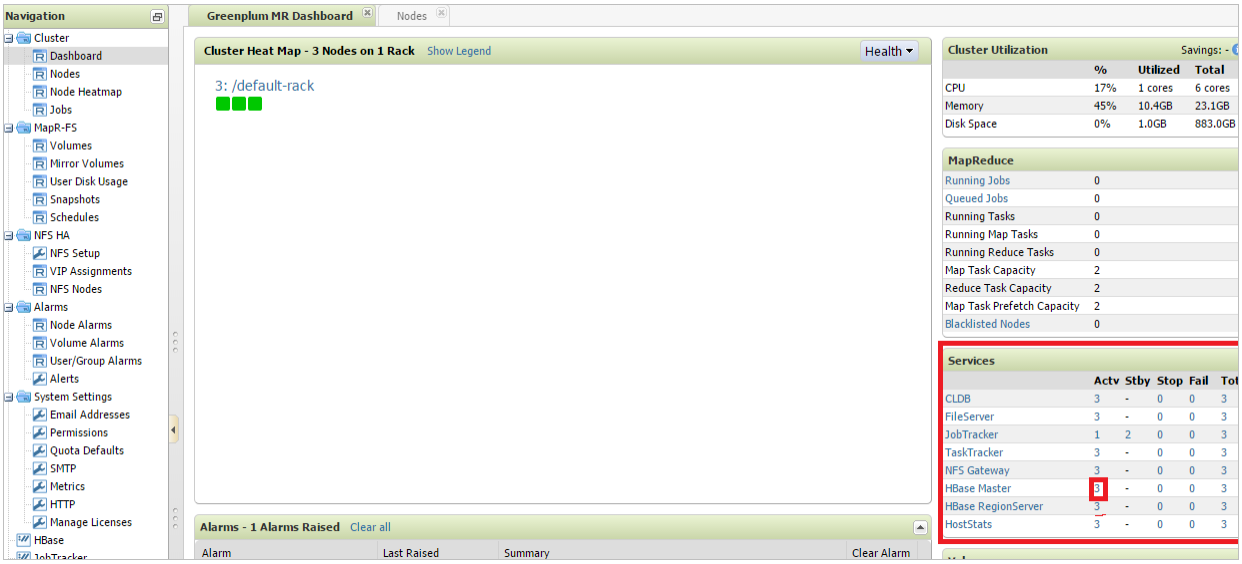
Where `<Hostname>` is the specific node hostname.

Stop the Hbase Services Using the MapR Control System

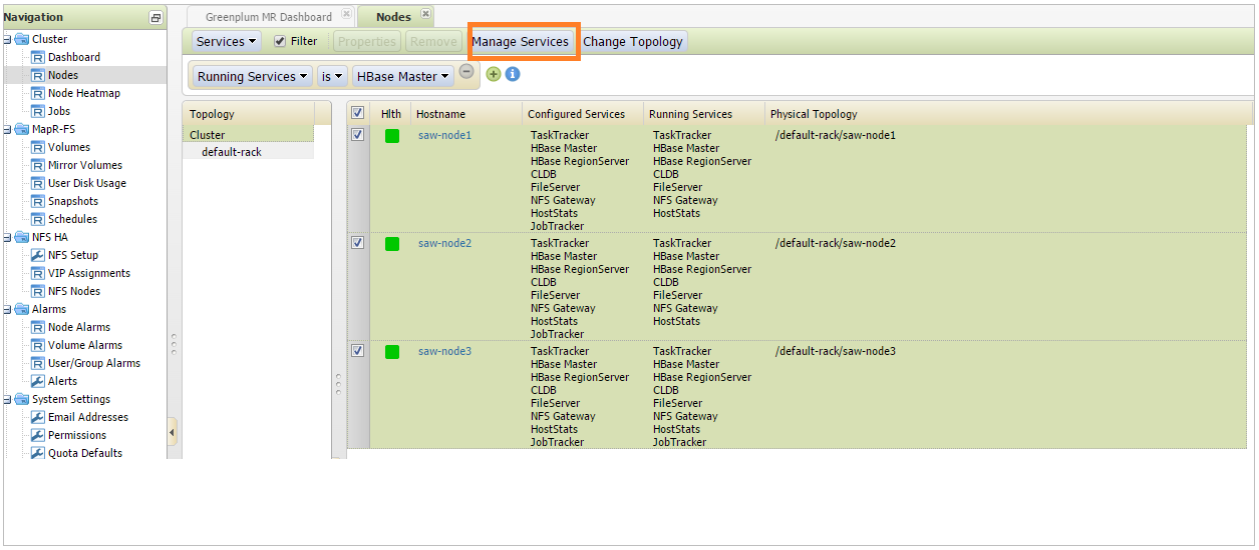
This topic provides the steps to stop the Hbase services using the MapR Control System. If you are not using Vulnerability Response Management (VRM), stop the Hbase services to return the configured memory so that it is available for use elsewhere in the cluster.

Procedure

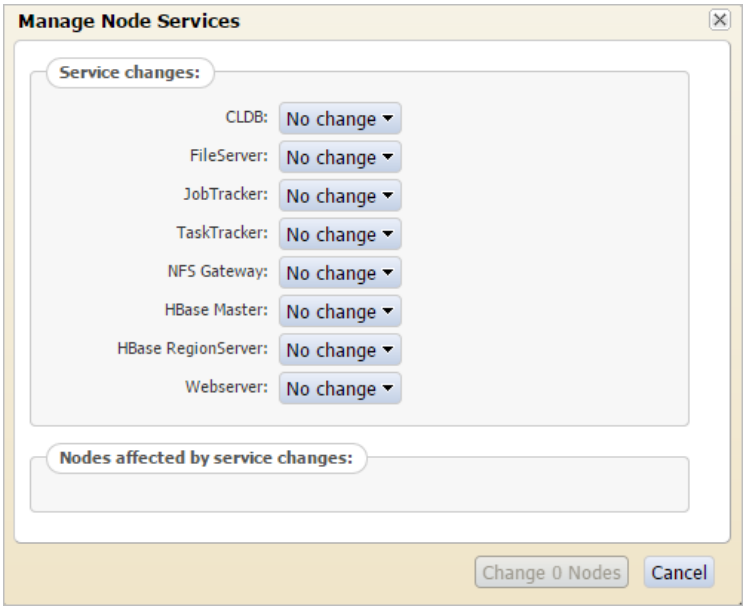
1. Log on to the MapR Control System user interface following the instructions in Access MapR Control System UI for Cluster Administration.
2. To stop the **HBase Master** services, in the **Services** section of the dashboard, click the number in the **Actv** column for the **HBase Master** service. This is the number of active services for the **HBase Master** service.



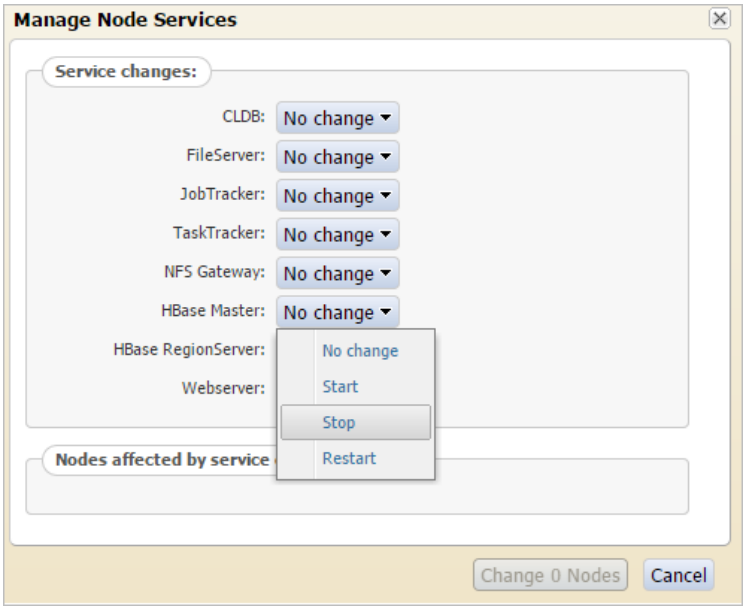
3. On the **Cluster Nodes** tab, click **Manage Services**.



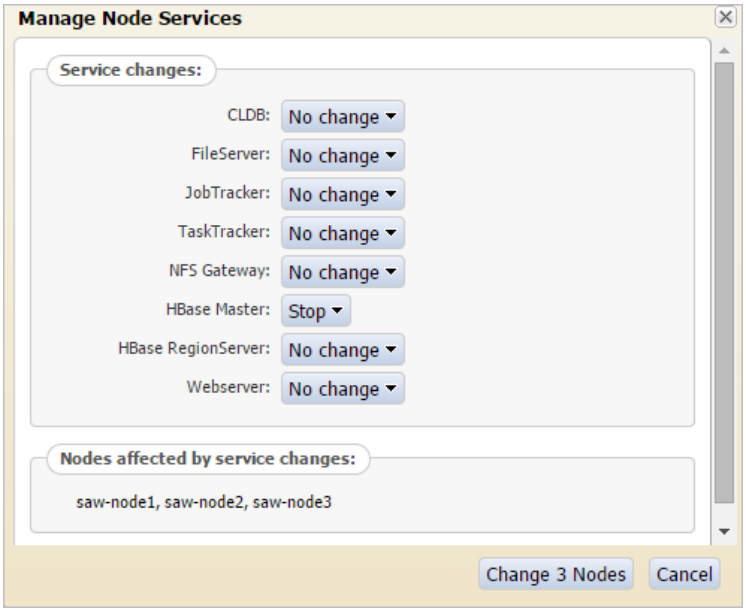
The **Manage Node Services** dialog is displayed.



4. In the **HBase Master** field, select **Stop**.



5. Click **Change <number_of_nodes> Nodes.**
   Where <number_of_nodes> is the number of active nodes selected.
   For example, click **Change 3 Nodes**.



   The **Hbase Master** service on the selected nodes should be in a stopped state.

6. To stop the **Hbase RegionServer** services, repeat steps 2 to 5 for the **Hbase RegionServer** services.

Additional Procedures

This topic is a collection of individual procedures, which an Administrator may perform at any time, and they are not required to complete the initial setup of Warehouse. These procedures are presented in alphabetical order.

Topics

Access MapR Control System UI for Cluster Administration

This topic provides instructions for accessing the MapR Control System user interface for RSA Analytics Warehouse cluster administration. MapR Control System user interface enables you to administer the RSA Analytics Warehouse cluster. The MapR Control System user interface provides details of the following:

- Nodes
- Node Heatmap
- Jobs
- MapR Tables
- Volumes
- Mirrors
- User Disk Usage
- Snapshots
- Schedules
- NFS Setup
- Virtual IP Assignments
- NFS Nodes
- Node Alarms
- Volume Alarms
- User/Group Alarms
- HBase
- JobTracker
- CLDB

To access the MapR Control System user interface:

1. Log on to one of the appliances in the RSA Analytics Warehouse cluster.
2. Start the webserver. Enter the following command:

   ```
   /opt/mapr/adminuiapp/webserver start
   ```

   Note: The default port used by the webserver is **8443**.

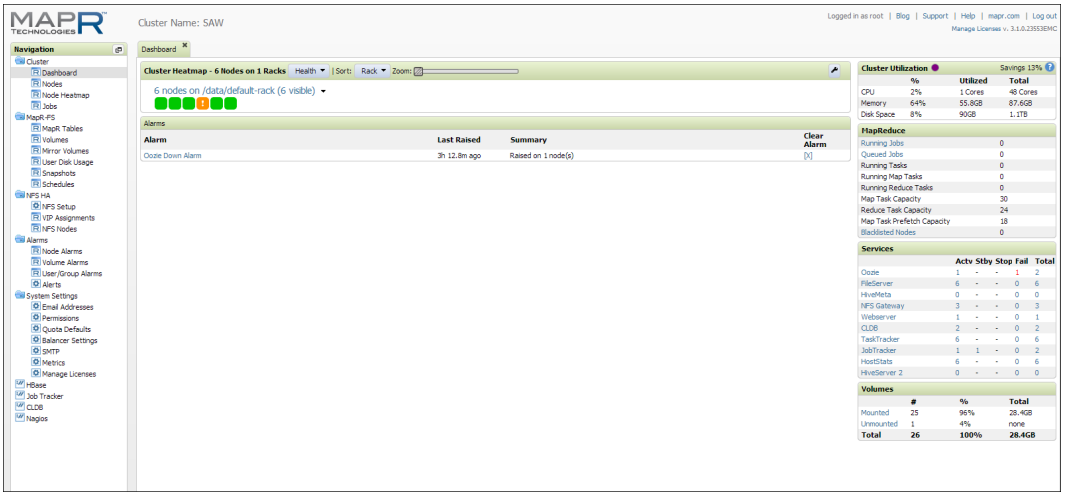
Note: If you receive the error `/opt/mapr/conf/ssl_keystore (No such file or directory)` in the **/opt/mapr/logs/adminuiapp.log** after executing the command `/opt/mapr/adminuiapp/webserver start`, enter the following commands:

```
./configure.sh -R -genkeys
service mapr-warden restart
```

3. Using a web browser to access the MapR Control System, type the following url:

   `https://<NODE-IP-OR-HOSTNAME>:8443`

   The MapR Control System user interface is displayed.



Enable MapR Metrics on RSA Analytics Warehouse Cluster

This topic provides links to instructions on how to enable MapR Metrics on the RSA Analytics Warehouse cluster. This optional procedure enables Administrators to see job details in the MapR Control System UI rather than going to the JobTracker for details.

Prerequisites

Make sure that you have the following MapR Metrics dependencies installed in your environment:

- MySQL Server installed and configured.
- Libraries hosted on the EPEL Repository.
- Libraries hosted on the CentOS base repositories.

Procedure

To enable MapR Metrics on the RSA Analytics Warehouse cluster, following the instructions at the following links:

- http://doc.mapr.com/display/MapR/Setting+up+the+MapR+Metrics+Database
- http://doc.mapr.com/display/MapR/MapR+Metrics+and+Job+Performance

Note: Make sure that you install MapR Mertics on the nodes in your RSA Analytics Warehouse Cluster where Job Tracker or Web Server is running.

Edit and Remove Virtual IP Addresses (Command Line)

This topic provides instructions on how to edit and remove virtual IP addresses in the Warehouse cluster using the command line. This procedure is optional and used when you want to change the virtual IP addresses in the Warehouse cluster.

Prerequisites

Make sure that you note the MAC addresses of all the Warehouse appliances in the cluster. Use the following command on the appliance to view the MAC address of the appliance:

```
ifconfig <interface> | grep HWaddr
```

where `<interface>` is the network interface.

Also note the MAC addresses of the Warehouse appliances that you want to add.

Procedure

Adding and removing Warehouse appliances to and from a virtual IP group is accomplished by executing an **edit** command. This is the same as the add command, except that ALL of the MAC addresses are replaced with ONLY the MAC addresses that you enter.

To add or remove a virtual IP address in the primary Warehouse appliance:

1. Log on to the primary appliance as root user.
2. Edit the virtual IP address. Enter the following command:

   ```
   maprcli virtualip edit -virtualip <VIP_address> -netmask
   <netmask> -macs <mac_node1> <mac_node2> <mac_node3> ……< mac_
   node n>
   ```

   where:

   - `<VIP_address>` is the virtual IP address for the primary Warehouse appliance.
   - `<netmask>` is the netmask address of the primary Warehouse appliance.
   - `<mac_node1>` is the MAC address of the first node in the Warehouse cluster.
   - `<mac_node2>` is the MAC address of the second node in the Warehouse cluster.

   For example, if the IP address of the primary warehouse is `192.168.100.10` and the MAC address for node 1 is `01:Z1:1X:00:20:Y1`, node 2 is `32:Y2:4Z:40:10:X3`, and you want to add node 3, which is `20:Y2:4Z:20:10:X3`, then enter the following:

   ```
   maprcli virtualip edit -virtualip 192.168.100.10 -
   netmask <netmask> -
   macs 01:Z1:1X:00:20:Y1 32:Y2:4Z:40:10:X3 20:Y2:4Z:20:10:X3
   ```

3. Verify the virtual IP addresses. Enter the following command:

   ```
   maprcli virtualip list
   ```

To remove the virtual IP address of the primary Warehouse appliance group entirely:
Enter the following command:

```
maprcli virtualip remove -virtualip 192.168.100.10
```

Add and Remove a Virtual IP Address (MapR UI)

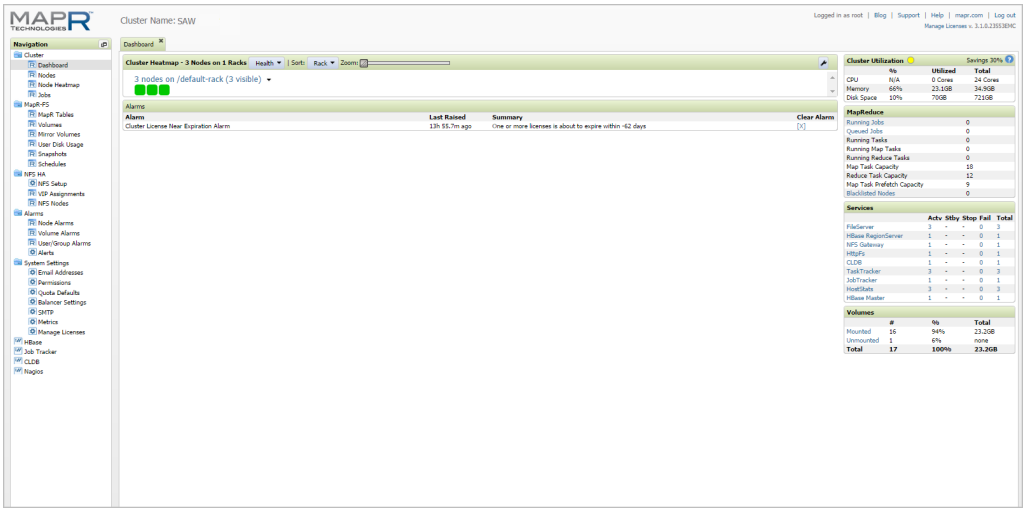
This topic provides instructions on how to add a virtual IP address in the Warehouse cluster using the MapR Control System. This procedure is optional and used when you want to add a virtual IP address (VIP) in the Warehouse cluster.

Prerequisites

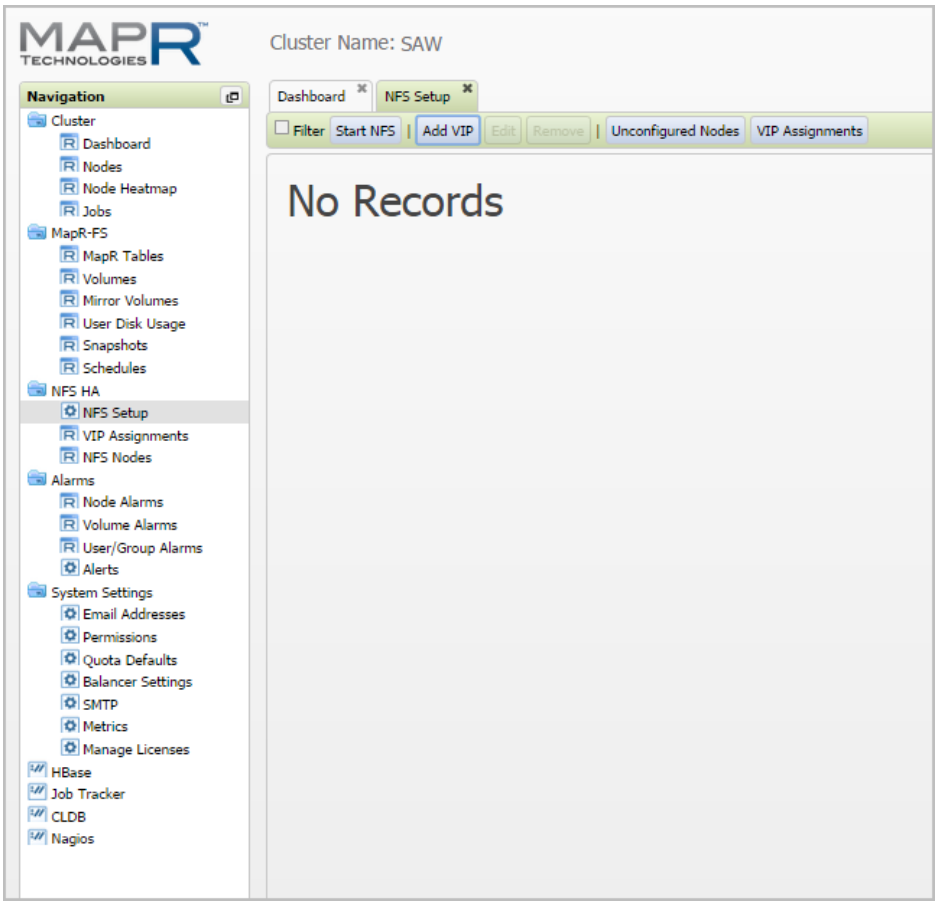
Follow the instructions in Access MapR Control System UI for Cluster Administration before completing this procedure.

Procedure

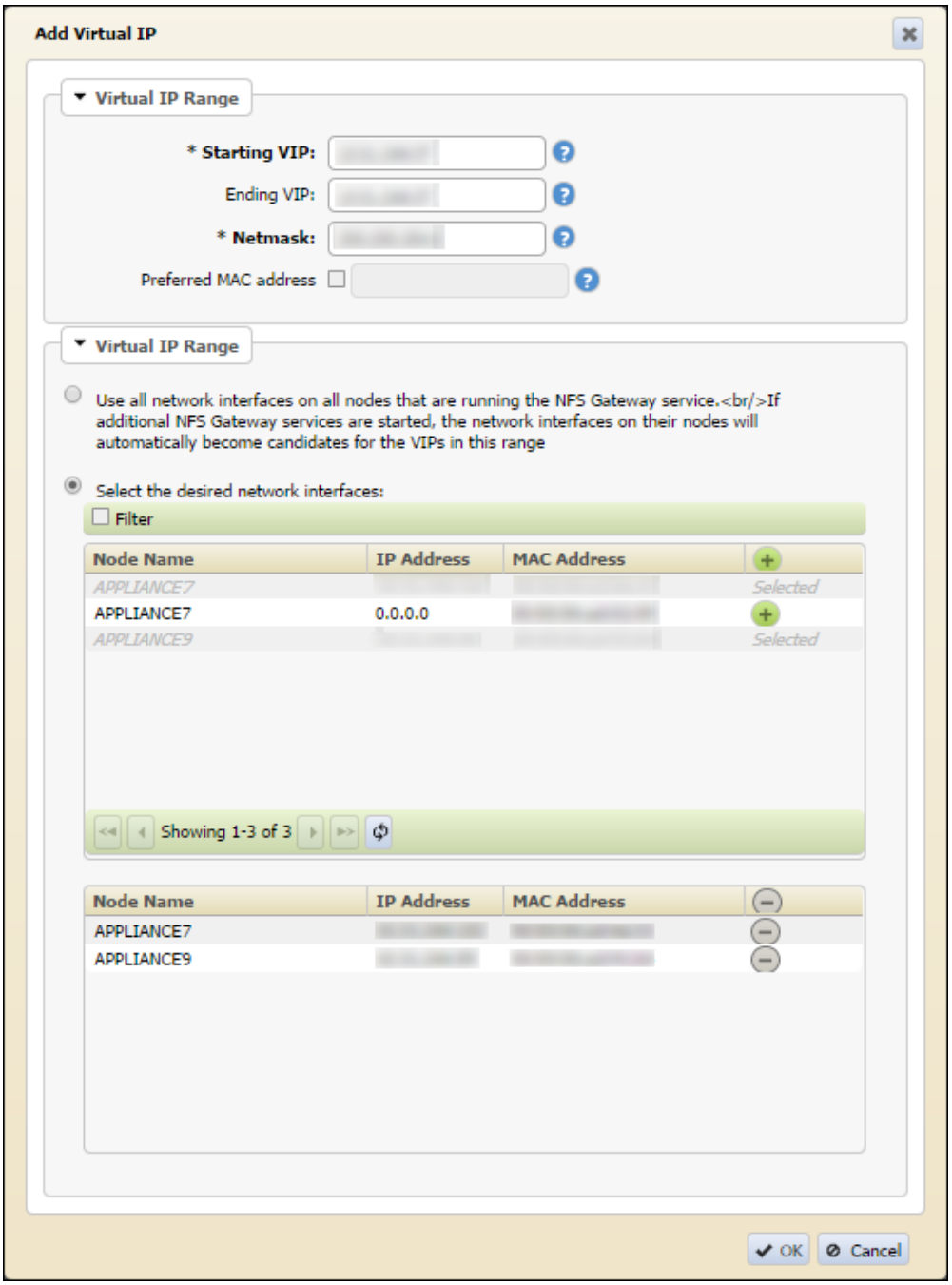
1. Log on to the MapR Control System.



2. In the Navigation panel, select **NFS HA > NFS Setup**.
   The NFS Setup tab is displayed. The NFS Setup tab enables you to edit, remove or add VIPs in the Warehouse cluster.
3. On the **NFS Setup** tab, click the **Add VIP** button.



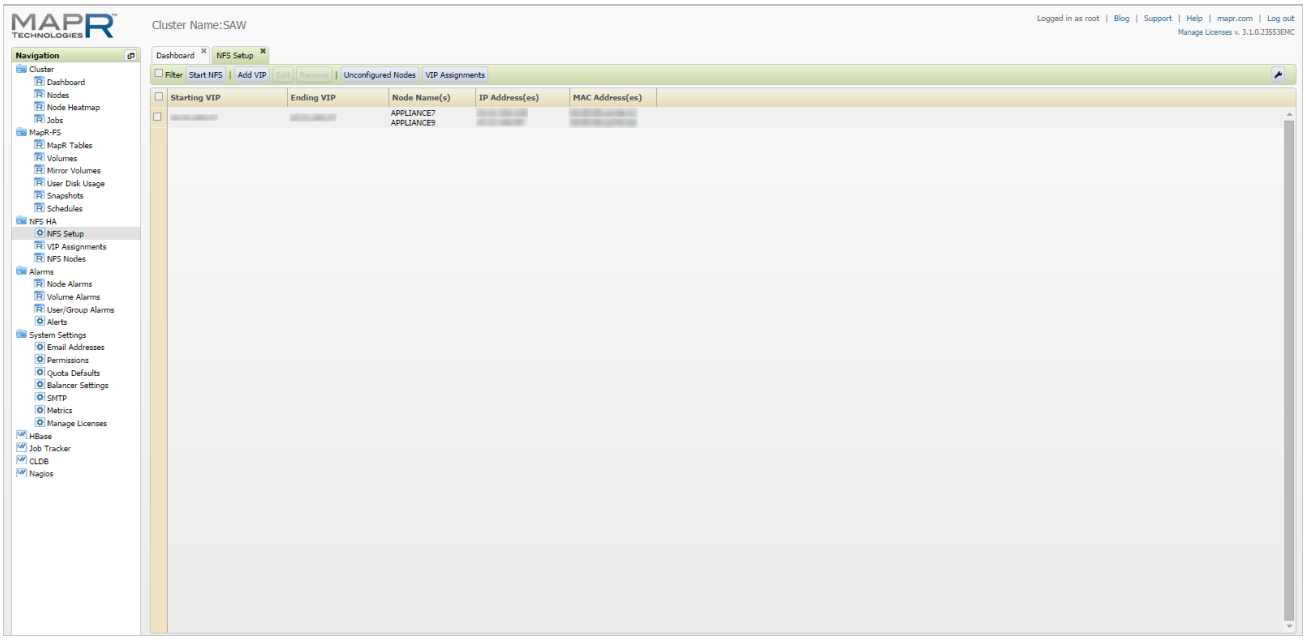
The **Add Virtual IP** dialog is displayed.



4. In the **Add Virtual IP** dialog, do the following:
   a. In the **Starting VIP** field, type the starting IP Address for VIP.
   b. In the **Ending VIP** field, type the ending IP Address for VIP. If this field is left blank, only one IP address is used for VIP allocation.
   c. In the **Netmask** field, type the Netmask for the deployment.

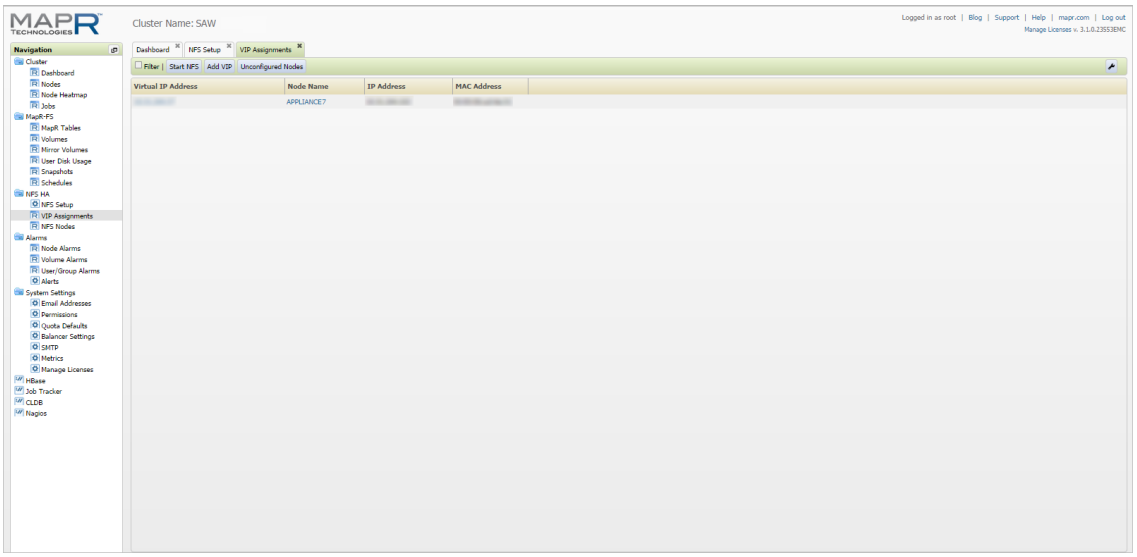
d. Select **Select Desired Network Interfaces** to choose the available Network Interfaces that need to be used for VIP assignment. Select all of the external Interfaces from the list of available nodes by clicking the plus button next to the interface entry. Selected Interfaces will appear in the bottom list.

e. Click **OK** to add the VIP.
The newly added VIP appears in the list on the **NFS Setup** tab.



Note: VIP allocation can also be removed or edited from the **NFS HA > NFS Setup** tab by selecting a VIP and clicking the **Edit** or **Remove** button.

5. In the Navigation panel, select **NFS HA > VIP Assignment** to view the node that is assigned to the newly added VIP.


MAPR
TECHNOLOGIES
Cluster Name: SAW
Logged in as root | Blog | Support | Help | mapr.com | Log out
Manage Licenses v. 3.1.0.23553EMC
Navigation
Cluster
Dashboard
Nodes
Node Heatmap
Jobs
MapR-FS
MapR Tables
Volumes
Mirror Volumes
User Disk Usage
Snapshots
Schedules
NFS HA
NFS Setup
VIP Assignments
NFS Nodes
Alarms
Node Alarms
Volume Alarms
User/Group Alarms
Alerts
System Settings
Email Addresses
Permissions
Quota Defaults
Balancer Settings
SMTP
Metrics
Manage Licenses
HBase
Job Tracker
CLDB
Nagios
Dashboard
NFS Setup
VIP Assignments
Filter
Start NFS
Add VIP
Unconfigured Nodes
Virtual IP Address
Node Name
IP Address
MAC Address
APPLIANCE7

Add a Virtual IP Address with Multiple Nodes (MapR UI)

This topic provides instructions on how to add a virtual IP address (VIP) with multiple nodes. Virtual IP (VIP) is a technique used to load balance data access into HDFS by using a floating IP Address among the cluster nodes. This technique is mostly used by the MapR Hadoop Distribution along with the MapR-NFS Service. VIP can provide High Availability and Load Balancing by dynamically allocating the Floating IP among the nodes.

Optimal VIP Configuration

We recommend using one VIP for every three Nodes, because the replication factor for HDFS is 3 by default. This also helps in optimizing the performance of the cluster.

In the case of High Data Load (>20K EPS), a single NFS might overload while replicating the file into the cluster. If the NFS Server crashes before the data is replicated, you could have data loss.

Multiple NFS Servers also allow more distributed data locality which helps in High Availability and Fault Tolerance.

Prerequisites

Calculate how many VIPs you can afford.

- We suggest **One VIP per 3 Nodes**.
- In case the number of nodes that you have is not a multiple of three, you can allocate multiple VIPs to more than three nodes. For example, two VIPs among five Nodes.

The steps to add the VIP are the same as adding any other VIP, but instead of choosing “all nodes” for VIP, you choose a subset of nodes to participate in the VIP.

- A node can participate in Multiple VIPs.
- For more information, see http://doc.mapr.com/display/MapR/Setting+Up+VIPs+for+NFS

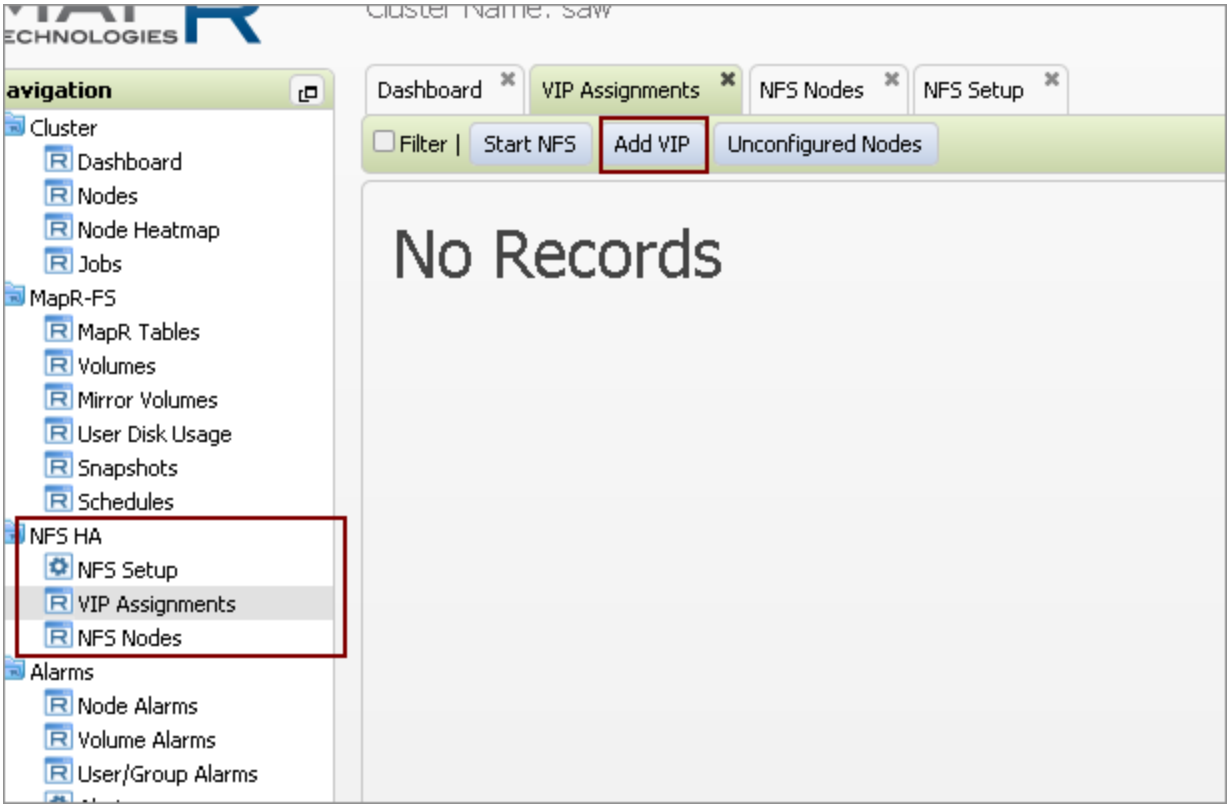
Optimal Configuration with the Warehouse Connector

The best configuration is to have one VIP per Warehouse Connector. In cases where Warehouse Connector numbers are higher than VIPs, configure multiple Warehouse Connectors to write to a VIP in a way so that traffic on VIPs can be normalized.

Add a Virtual IP Address that has Multiple Nodes

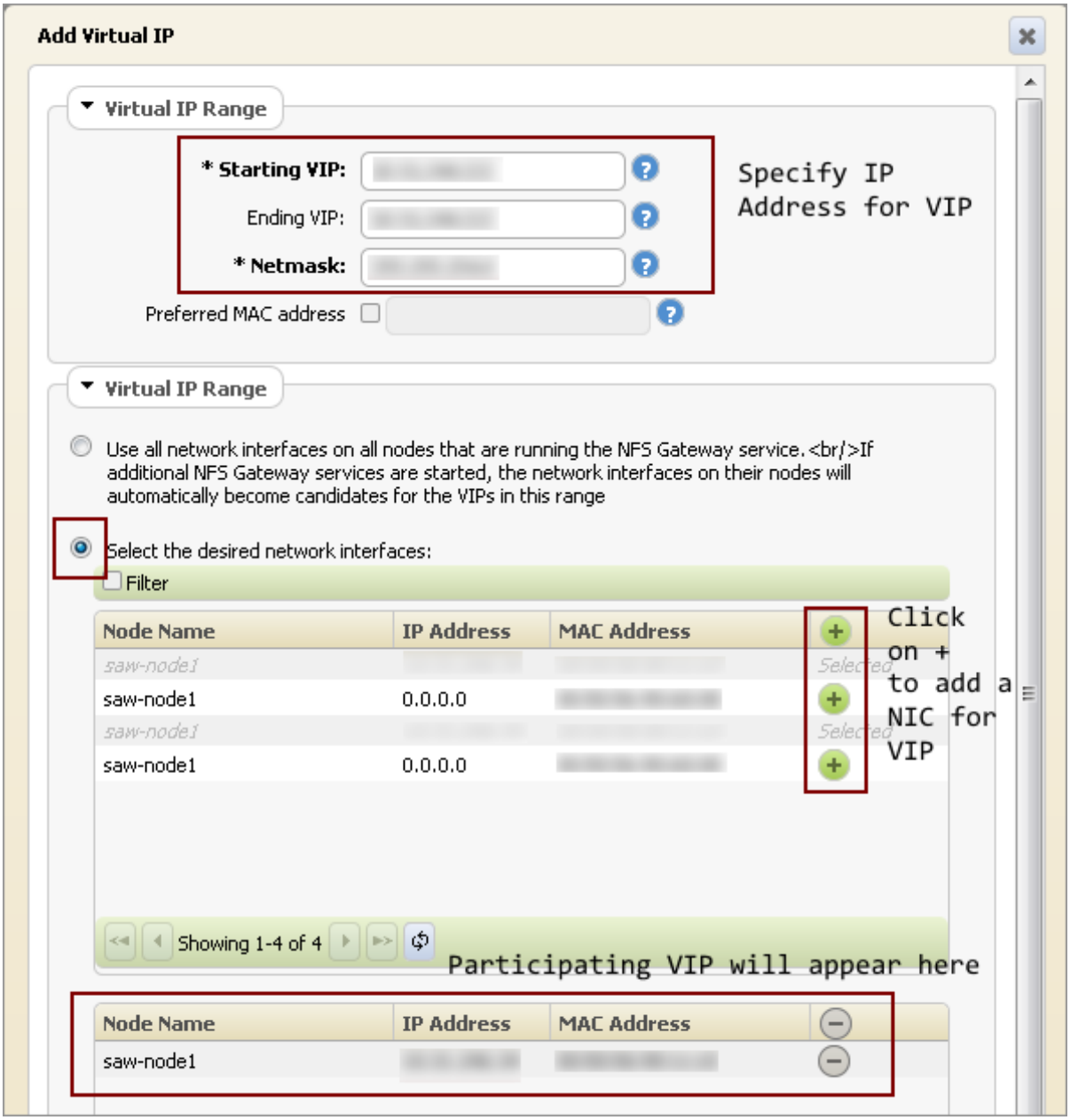
1. Log on to the MapR Control System.
2. In the Navigation panel, select **NFS-HA > VIP Assignments**.

3. On the **NFS Setup** tab, click the **Add VIP** button.



4. In the **Add Virtual IP** dialog, do the following:

a. Specify the Starting and Ending VIP as the same IP address.



b. Select **Select the Desired Network Interfaces** to choose the available Network Interfaces that need to be used for the VIP assignment. Select the NIC Cards that you want to participate in the VIP. A node can have multiple NICs, so depending on the Network Configuration you can select them.

c. Click **OK** to add the VIP.

Example VIP Configurations

The following table shows example configurations of virtual IP addresses (VIPs) with different numbers of nodes in the cluster.

Number of Nodes in Cluster	Number of VIPs
3 Nodes	1 VIP
5 Nodes	2 VIPs (3 Nodes each, 1 Common Node)
7 Nodes	2 VIPs (3 Nodes each, 1 Free Node)
8 Nodes	3 VIPs (3 Nodes each, 1 Common Node among 2 VIPs)
11 Nodes	4 VIPs (3 Nodes each, 1 Common Node among 2 VIPs)
11 Nodes	3 VIPs (3 Nodes each, 2 Free Nodes)