

SnapScale RAINcloudOS 4.0 Split-Bond Configuration for VMware

May 2014

Overview

This Technical Information Bulletin describes how to configure a SnapScale cluster running RAINcloudOS (ROS) 4.0 for use in a VMware environment to support the two IP addresses required for a multipath failover configuration.

The interfaces to the split-bond are called bond1 and bond2 at the Ethernet level, and called Eth 1 and Eth 2 at the GUI level. Each bond receives a static IP address that is configured via a command line program executed on each node. The resulting public IP address configuration has two IP addresses on Eth 1 (one static and one from CTDB) and one IP address on Eth 2.

Some items of note:

- Some hand configuration in the shell on the cluster is needed to put the cluster nodes into a split-bond mode (reboot required).
- You may need to issue some iSCSI configuration commands from the shell as well.
- Automatic failover of multipath iSCSI targets between nodes will not be supported (though a manual method may be available).
- iSNS is automatically disabled in split-bond mode.
- Functionality of VSS/VDS is limited in split bond mode.

Start Split-Bond Mode

While in split-bond (multipath) mode the user will experience the following limitations:

- iSCSI targets will not fail over to another node if the host node fails. Targets can be launched on a different node but it will require manual editing of some configuration files. Please contact technical support to relaunch iSCSI targets on a different node.
- Changing IP addresses on the cluster (or moving to a new subnet) will require the user to hand edit the `/etc/sysconfig/network-scripts/ifcfg-bond1` and `/etc/sysconfig/network-scripts/ifcfg-bond2` files to change the IP address and subnet.

Split-bond mode instructions:

1. Log in to **each node**:
 - a. Login normally via **SSH**.
 - b. Run `osshell` to get to the Linux prompt.
 - c. Run `su` to root the user.

2. Contact your IT department to acquire more **valid IP addresses** (2 per node) to be used on your cluster.

These must be unique IP addresses on your network. It is OK to divide the addresses between two subnets. Please make sure multi-subnet configurations are wired consistently for all nodes. For example, all IP addresses for subnet 1 would be on Eth 1 and all IP addresses for subnet 2 would be on Eth 2.

3. Run this **program** on each node:

```
/usr/mfs/bin/static_ip_multipath.sh setup <IP1> <IP2> <NETMASK>
```

4. Reboot the **cluster** using the Web Management Interface.
5. Verify that the **IP addresses** are correct (**Network > Information**).

NOTE: Use iSCSI static IP addresses for accessing iSCSI only. Do not add iSCSI static IP addresses to the round-robin DNS.

Adding New Nodes

Adding new nodes to a SnapScale cluster can cause a re-shuffling of the IP addresses assigned to the nodes. When the cluster is in split-bond mode this can cause the static IP addresses configured on a node to be associated with a different assigned IP address. In addition, the newly added nodes are by default in the bonded mode and need to have the bonded interface reconfigured as two separate interfaces. Use these steps to complete the add-nodes operation:

1. Disconnect any **ESXi hosts** from their existing targets before performing the add-nodes operation.
2. At the **Storage > Nodes** page, click **Add Node** and follow the wizard.
Refer to your *SnapScale Administrator's Guide* for complete details.

3. On each **newly added node**, run:

```
/usr/mfs/bin/static_ip_multipath.sh setup <Eth_1> <Eth_2> 255.255.0.0
```

where <Eth_1> and <Eth_2> are the static IP addresses assigned to the new node.

Note that this step is identical to the one used to perform split-bond operations on the other nodes. Also, the **Storage > iSCSI** page does not show any iSCSI disks.

4. Reboot the **cluster** using the Web Management Interface.
The restart of cluster is necessary to allow the network configuration changes to take effect. After this step is done, the assigned IP address for each node (which may not be the same as before the add-node operation) does not change.
5. Verify the **target IP addresses** by using **Network > Information** (or do a rescan on the ESXi host using the node's assigned IP address to obtain its new target-portals).
6. Reconnect the **targets**.
The target-portals will have the static IP addresses for that node as portals.

New Web Management Interface Screens

Once split-bond mode is started, the IP Address columns in some Web Management Interface screens change to support it.

- Wherever IP addresses are displayed (except the iSCSI pages which display only iSCSI addresses), the Web Management Interface pages now show both CTDB and iSCSI static IP addresses.

- When creating an iSCSI disk (**Storage > iSCSI > Create iSCSI Disk button**), you can either choose a specific iSCSI IP address pair from a list or to let ROS decide using round robin.
- iSCSI IP address changes to the cluster are a manual process for the static IP addresses.

Changes to iSCSI Pages

The most obvious changes are to the iSCSI pages because they now display both IP addresses.

Default iSCSI Page (Storage > iSCSI)

The screenshot shows the SnapScale iSCSI configuration page. At the top, there are navigation tabs for SnapScale, Network, Storage, Security, Monitor, and Maintenance. Under the Storage tab, there are sub-tabs for Peer Sets, Volumes, Quotas, Snapshots, iSCSI, Nodes, and Disks. The iSCSI sub-tab is active, displaying a table of iSCSI disks. A red box highlights the IP Address column, which lists two IP addresses for each disk. Below the table are buttons for 'Create iSCSI Disk', 'VSS/VDS Access', 'Refresh', and 'Close'.

iSCSI Disk	Status	Active Clients	IP Address	Authentication	Size
iscsi0	OK	0	10.25.12.34 10.25.12.35	None	51.00 GB
iscsi1	OK	0	10.25.12.34 10.25.12.35	None	52.00 GB
iscsi2	OK	0	10.25.12.36 10.25.12.37	None	53.00 GB
iscsi3	OK	0	10.25.12.38 10.25.12.39	None	16.00 GB

Create iSCSI Page (Storage > iSCSI > Create iSCSI Disk button)

The screenshot shows the 'Create iSCSI Disk' page in SnapScale. The page has the same navigation structure as the previous screenshot. The 'iSCSI' sub-tab is active, and the 'Create iSCSI Disk' form is displayed. The 'iSCSI Disk Name' field contains 'iscsi4'. The 'IP Address' dropdown menu is open, showing the option 'Let system choose (round robin)' selected. Below it, a red box highlights three options: '10.25.12.34, 10.25.12.35 (Node2414528)', '10.25.12.36, 10.25.12.37 (Node2414532)', and '10.25.12.38, 10.25.12.39 (Node2414538)'. There are also checkboxes for 'Support Multi-Initiator' and 'Enable CHAP Logon'. At the bottom, there are buttons for 'Create iSCSI Disk' and 'Cancel'.

iSCSI Disk Properties (Storage > iSCSI > iSCSI Disk name)

The screenshot shows the 'iSCSI Disk Properties' page. The 'IP Address' field is highlighted with a red box and contains the following values:

iSCSI Disk	Status	Active Clients	IP Address	iSCSI Disk EUI	Device	Size
iscsi0	OK	0	10.25.12.34 10.25.12.35	00c0b6d7c0000001	/hd/cfs/blockstor/snapbd0	51.00 GB

Below the table, the iSCSI Disk IQN is shown as: iqn.1997-10.com.snapscale:marcscale:snapbd0

The iSCSI Disk Size is set to 51 GB (Max. size is 5.49 TB).

Options include:

- Support Multiple Initiators
Warning: Uncontrolled simultaneous access of multiple initiators to the same iSCSI target can result in data corruption. Only enable multi-initiator support if your environment or application supports it.
- Enable CHAP Logon

Buttons at the bottom: OK, Delete iSCSI Disk, Refresh, Cancel

Changes to Node Pages

Besides the iSCSI pages, three **Node** pages also reflect the dual IP addresses.

Nodes (Storage > Nodes)

The screenshot shows the 'Nodes' page with 3 nodes. The 'IP address' column is highlighted with a red box, showing dual IP addresses for each node:

Node	Description	Status	IP address	Model	Version	Type	Disks/Slots
Node2414528 (Mgmt. Node)	-	OK	10.25.12.34 10.25.12.33 10.25.12.35	X2	4.0.0.mdorsasled11	2U	6 / 12
Node2414532	-	OK	10.25.12.36 10.25.12.32 10.25.12.37	X2	4.0.0.mdorsasled11	2U	5 / 12
Node2414538	-	OK	10.25.12.38 10.25.12.31 10.25.12.39	X2	4.0.0.mdorsasled11	2U	7 / 12

Buttons at the bottom: Add Nodes, Node Identification, Refresh, Close

Node Identification (Storage > Nodes > Node Identification button)

Specify node descriptions, as well as flash node LEDs for (hardware) identification. Note that both node names and descriptions are seen only within this Web Management Interface.

3 nodes. Click this icon to identify a node by flashing its LEDs for 5 minutes. Click this icon to stop flashing a node's LEDs.

Node	Description	IP address	Model	Type	Disks/Slots
Node2414528		(SCSI) 10.25.12.34 (SCSI) 10.25.12.33 (SCSI) 10.25.12.35	X2	2U	6 / 12
Node2414532		(SCSI) 10.25.12.36 (SCSI) 10.25.12.32 (SCSI) 10.25.12.37	X2	2U	5 / 12
Node2414538		(SCSI) 10.25.12.38 (SCSI) 10.25.12.31 (SCSI) 10.25.12.39	X2	2U	7 / 12

Click here to stop flashing LEDs on all nodes.

OK Cancel

Node Properties (Storage > Nodes > Node Properties)

Node	Description	Status	IP address	Model	Version	Type	Disks/Slots
VM-Node13157556	-	OK	(SCSI) 10.25.12.36 (SCSI) 10.25.12.32 (SCSI) 10.25.12.37	VirtualNode	4.0.030	1U	4 / 4

Click here to identify this node by flashing its LEDs for 5 minutes. Click here to stop flashing this node's LEDs.

Enter a description below to help identify your node. Note that both node names and descriptions are seen only within this Web Management Interface.

Node Description (optional)

OK View Disks Remove Node from SnapScale Refresh Cancel

Changes to Network Pages

The **Network Information** page also shows the dual IP addresses.

Network Information Page (Network > Information)

SnapScale client network information.

Subnet Mask	255.255.0.0
Default Gateway	10.25.1.1
Domain Name	devnet.myoverland.net
Domain Name Servers	10.6.8.34, 10.6.8.35
WINS Servers	-
Bonding Status	Load Balance (ALB)
Management IP Address	10.25.12.30

Node-specific client network information.

Node	Ethernet Port Status	IP Address	Speed/Duplex Status	Ethernet Address
Node2414528 (Mgmt. Node)	Eth 1: OK	[sCSI] 10.25.12.34 10.25.12.33	1000 Mbps (Auto) / Full Duplex (Auto)	00:C0:B6:24:D7:C0
	Eth 2: OK	[sCSI] 10.25.12.35 10.25.12.36	1000 Mbps / Full Duplex	00:C0:B6:24:D7:C1
Node2414532	Eth 1: OK	[sCSI] 10.25.12.32 10.25.12.32	1000 Mbps (Auto) / Full Duplex (Auto)	00:C0:B6:24:D7:C4
	Eth 2: OK	[sCSI] 10.25.12.37 10.25.12.50†	1000 Mbps / Full Duplex	00:C0:B6:24:D7:C5
Node2414538	Eth 1: OK	[sCSI] 10.25.12.38 10.25.12.31	1000 Mbps (Auto) / Full Duplex (Auto)	00:C0:B6:24:D7:CA
	Eth 2: OK	[sCSI] 10.25.12.39	1000 Mbps / Full Duplex	00:C0:B6:24:D7:CB

† Utility IP address.

Changes to Monitor Pages

Some **Monitor** pages also show IP addresses.

Network Monitor (Monitor > Activity > Network Monitor)

3 nodes.

Node	Client IP	Client Network Usage (In/Out)	Storage IP	Storage Network Usage (In/Out)
Node2414528 (Mgmt. Node)	10.25.12.34	<1% 1.86 KB/s	192.0.2.251	<1% 8.11 KB/s
	10.25.12.33	<1% 1.33 KB/s		<1% 5.94 KB/s
Node2414532	10.25.12.36	<1% 1.39 KB/s	192.0.2.8	<1% 6.31 KB/s
	10.25.12.32	<1% 179 bytes/s		<1% 6.06 KB/s
Node2414538	10.25.12.38	<1% 1.86 KB/s	192.0.2.248	<1% 6.33 KB/s
	10.25.12.31	<1% 179 bytes/s		<1% 8.17 KB/s

NOTE: Only **Eth 1** is shown for the **Client Network Usage**.

Network Protocol Manager (Monitor > Protocol Manager)

The screenshot shows the SnapScale Protocol Manager interface. The top navigation bar includes 'SnapScale', 'Network', 'Storage', 'Security', 'Monitor', and 'Maintenance'. The 'Monitor' tab is active, and the 'Protocol Manager' sub-tab is selected. Below the navigation, there are links for 'System Status', 'Activity', 'Event Log', 'Protocol Manager', 'SnapScale Settings', and 'Tape'. A descriptive paragraph explains that the Protocol Manager manages networking protocols and IP address assignment across the entire SnapScale. Below this, a table lists 3 nodes:

Node	Description	Status	IP Address
Node2414528 (Mgmt. Node)	-	OK	10.25.12.34 10.25.12.33
Node2414532 (IP Manager)	-	OK	10.25.12.36 10.25.12.32
Node2414538	-	OK	10.25.12.37 10.25.12.38 10.25.12.31 10.25.12.39

Each IP address in the table is preceded by a small icon labeled 's(scsi)'. Below the table are 'Refresh' and 'Close' buttons.

Removing Split-Bond Mode

To disable the Split-Bond (Multipath) Mode:

- Log in to **each node**:
 - Login normally via **SSH**.
 - Run `osshe11` to get to the Linux prompt.
 - Run `su` to root the user.
- Run the following **command** on each node:


```
/usr/mfs/bin/static_ip_multipath.sh teardown
```
- After the command is run on all the nodes, reboot the **cluster** using the Web Management Interface.

Important Notes

Considerations while the cluster is in split-bond mode:

- Each new target will be hosted on a single node but will be accessible via either of the two new static IP addresses.
- If you configure multipath using two different subnets, the RAINcloudOS Web Management Interface will not be accessible from the Eth 2 subnet. Normally the user could use the IP address from any node and access would be rerouted to the root IP, however, this functionality is not available from the second subnet.