High Availability: Fail-over, Clustered FileSystem and Load Balancing

New ISV Support for Red Hat Enterprise Linux for IBM System z
Sine Nomine Associates

SNA HAO - High Availability Option for Red Hat Enterprise Linux offers:

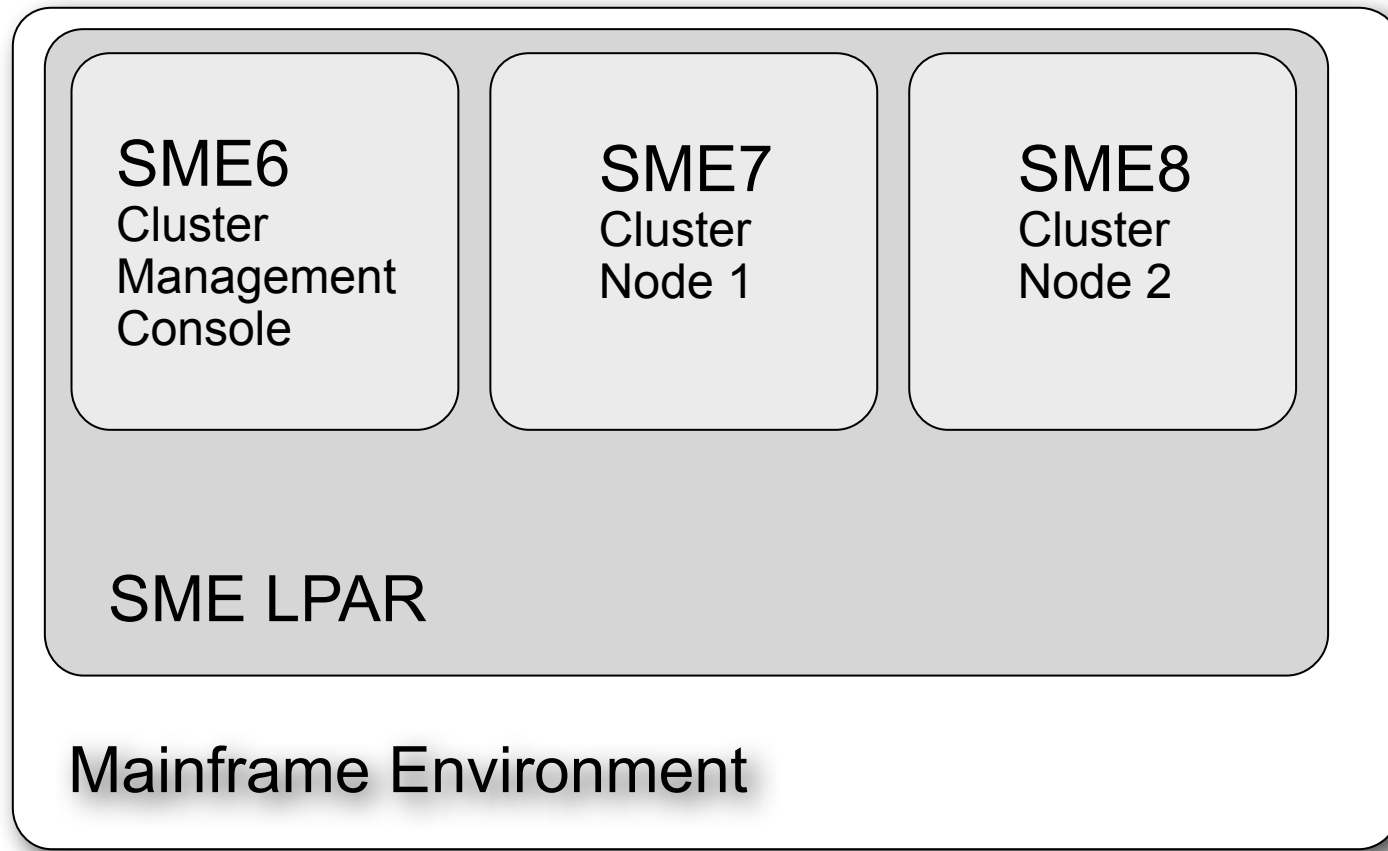- Fail-over
- Clustered File System GFS2
- Load Balancing

David Boyes
Sine Nomine Associates
+1 703 723 6673
info@sinenomine.net

Homebase

Login

Username   root

Password   ••••••

Login

SINE NOMINE
ASSOCIATES

High Availability
management

About   Admin   Preferences   Logout

Homebase   Clusters   SNA_HA

Homebase

Manage Clusters

● SNA_HA

Nodes	Fence Devices	Failover Domains	Resources	Service Groups	Configure

● Add   ⚙ Reboot   § Join Cluster   ⚲ Leave Cluster   ⊗ Delete

!	Node Name	Node ID	Votes	Status	Uptime	Hostname
☐	sme7.s390.bos.redhat.com	1	1	Cluster Member	01:13:30:55	sme7.s390.bos.redhat.com
☐	sme8.s390.bos.redhat.com	2	1	Cluster Member	00:13:23:47	sme8.s390.bos.redhat.com

Select an item to view details

Homebase   Clusters   SNA_HA

Homebase
Manage Clusters

● SNA_HA

Nodes   Fence Devices   Failover Domains   Resources   Service Groups   Configure

● Add   ● Reboot   ● Join Cluster   ● Leave Cluster   ● Delete

!	Node Name	Node ID	Votes	Status	Uptime	Hostname
☐	sme7.s390.bos.redhat.com	1	1	Cluster Member	01:13:35:01	sme7.s390.bos.redhat.com
☐	sme8.s390.bos.redhat.com	2	1	Cluster Member	00:13:27:53	sme8.s390.bos.redhat.com

sme8.s390.bos.redhat.com
Status Cluster Member

Properties

Update Properties

Number of votes	1
ricci host	sme8.s390.bos.redha
ricci port	11111

Services

Failover Domains
                                              Priority
Failover_domain_SNA

Fence Devices
Method
zVM_TCP   Remove

Name	Type/Values	
fence_zvm_2	IBM z/VM – SSI	⊗
	target : SME8	
Add Fence Instance		

Add Fence Method

Cluster Daemons
                                                          Status

cman	Running
rgmanager	Running
ricci	Running
modclusterd	Running

High Availability
management

Homebase 〉 Clusters 〉 SNA_HA

Homebase

Manage Clusters

● SNA_HA

| Nodes | Fence Devices | Failover Domains | **Resources** | Service Groups | Configure |

⊕ Add   ⊗ Delete

Name/IP	Type	In Use
☐ 10.16.106.20/21	IP Address	✔
☐ MyHttpd	Apache Server	✔

10.16.106.20/21   ⊗

IP Address

IP Address	10.16.106.20
Netmask Bits (optional)	21
Monitor Link	☑
Disable Updates to Static Routes	☑
Number of Seconds to Sleep After Removing an IP Address	10

Apply

SINE NOMINE
ASSOCIATES

High Availability
management

About Admin Preferences Logout

Homebase > Clusters > SNA_HA

Homebase

Manage Clusters

● SNA_HA

Nodes | Fence Devices | Failover Domains | **Resources** | Service Groups | Configure

● Add   ⊗ Delete

Name/IP	Type	In Use
☐ 10.16.106.20/21	IP Address	✔
☐ MyHttpd	Apache Server	✔

MyHttpd   ⊗

Apache

Name	MyHttpd
Server Root	/etc/httpd
Config File	conf/httpd.conf
httpd Options	
Shutdown Wait (seconds)	0

Apply

High Availability
management

About   Admin   Preferences   Logout

Homebase   >   Clusters   >   SNA_HA

Homebase
Manage Clusters

● SNA_HA

Nodes	Fence Devices	Failover Domains	Resources	**Service Groups**	Configure

⊕ Add   ⊙ Start   ⟳ Restart   ⊝ Disable   ⊗ Delete

!	Name	Status		Autostart	Failover Domain
☐	MyWeb	Running on sme7.s390.bos.redhat.com		✓	Failover_domain_SNA

MyWeb
Status   Running on sme7.s390.bos.redhat.com   [Start on node... ‡]

Edit service

Service Name	MyWeb
Automatically Start This Service	☑
Run Exclusive	☐
Failover Domain	Failover_domain_SNA ‡
Recovery Policy	Relocate ‡

--- **Restart Options** ---

Maximum Number of Restart Failures Before Relocating	[ ]
Length of Time in Seconds After Which to Forget a Restart	[ ]

Remove

IP Address

IP Address	10.16.106.20
Netmask Bits (optional)	21
Monitor Link	✓
Disable Updates to Static Routes	✓
Number of Seconds to Sleep After Removing an IP Address	10
Independent Subtree	☐
Non-Critical Resource	☐

--- **Independent Subtree/Non-Critical Options** ---

Maximum Number of Failures	[ ]
Failure Expire Time (seconds)	[ ]
Maximum Number of Restarts	[ ]
Restart Expire Time (seconds)	[ ]

[Add Child Resource]

```
[root@sme8 ~]# clustat
Cluster Status for SNA_HA @ Mon Mar 11 03:54:09 2013
Member Status: Quorate

 Member Name                              ID   Status
 ------ ----                              ---- ------
 sme7.s390.bos.redhat.com                   1 Online, rgmanager
 sme8.s390.bos.redhat.com                   2 Online, Local, rgmanager

 Service Name                 Owner (Last)                       State
 ------- ----                 ----- ------                       -----
 service:MyWeb                sme7.s390.bos.redhat.com           started
[root@sme8 ~]#
```

Simulating a application outage and failover action

```
[root@sme7 ~]# killall -9 httpd
[root@sme7 ~]#
```

```
Mar 11 03:56:04 sme8 rgmanager[2118]: Recovering failed service service:MyWeb
Mar 11 03:56:04 sme8 rgmanager[28383]: [ip] Adding IPv4 address 10.16.106.20/21 to
eth0
Mar 11 03:56:08 sme8 rgmanager[28579]: [apache] Checking Existence Of File /var/run
/cluster/apache/apache:MyHttpd.pid [apache:MyHttpd] > Failed
Mar 11 03:56:09 sme8 rgmanager[28601]: [apache] Monitoring Service apache:MyHttpd >
 Service Is Not Running
Mar 11 03:56:09 sme8 rgmanager[28623]: [apache] Starting Service apache:MyHttpd
Mar 11 03:56:10 sme8 rgmanager[2118]: Service service:MyWeb started
```

```
[root@sme7 ~]# clustat
Cluster Status for SNA_HA @ Mon Mar 11 03:56:58 2013
Member Status: Quorate

 Member Name                                  ID   Status
 ------ ----                                  ---- ------
 sme7.s390.bos.redhat.com                        1 Online, Local, rgmanager
 sme8.s390.bos.redhat.com                        2 Online, rgmanager

 Service Name               Owner (Last)                        State
 ------- ----               ----- ------                        -----
 service:MyWeb              sme8.s390.bos.redhat.com            started
[root@sme7 ~]#
```

Simulating virtual machine outage

```
Mar 11 03:59:57 sme7 corosync[1452]:     [TOTEM ] A processor failed, forming new configuration.
Mar 11 03:59:59 sme7 corosync[1452]:     [QUORUM] Members[1]: 1
Mar 11 03:59:59 sme7 corosync[1452]:     [TOTEM ] A processor joined or left the membership and a new mem
bership was formed.
Mar 11 03:59:59 sme7 corosync[1452]:     [CPG   ] chosen downlist: sender r(0) ip(10.16.106.7) ; members(
old:2 left:1)
Mar 11 03:59:59 sme7 corosync[1452]:     [MAIN  ] Completed service synchronization, ready to provide ser
vice.
Mar 11 03:59:59 sme7 corosync[1452]:     [CPG   ] chosen downlist: sender r(0) ip(10.16.106.7) ; members(
old:2 left:1)
Mar 11 03:59:59 sme7 corosync[1452]:     [MAIN  ] Completed service synchronization, ready to provide ser
vice.
Mar 11 03:59:59 sme7 rgmanager[1948]: State change: sme8.s390.bos.redhat.com DOWN
Mar 11 03:59:59 sme7 fenced[1513]: fencing node sme8.s390.bos.redhat.com
Mar 11 03:59:59 sme7 kernel: dlm: closing connection to node 2
Mar 11 04:00:01 sme7 fence_zvm[15358]: Recycling of SME8 successful
Mar 11 04:00:01 sme7 fenced[1513]: fence sme8.s390.bos.redhat.com success
Mar 11 04:00:02 sme7 rgmanager[1948]: Taking over service service:MyWeb from down member sme8.s390.bos.
redhat.com
Mar 11 04:00:02 sme7 rgmanager[15420]: [ip] Adding IPv4 address 10.16.106.20/21 to eth0
Mar 11 04:00:06 sme7 rgmanager[15613]: [apache] Monitoring Service apache:MyHttpd > Service Is Not Runn
ing
Mar 11 04:00:06 sme7 rgmanager[15635]: [apache] Starting Service apache:MyHttpd
Mar 11 04:00:08 sme7 rgmanager[1948]: Service service:MyWeb started
```

```
[root@sme7 ~]# clustat
Cluster Status for SNA_HA @ Mon Mar 11 04:00:55 2013
Member Status: Quorate

 Member Name                                    ID   Status
 ------ ----                                    ---- ------
 sme7.s390.bos.redhat.com                          1 Online, Local, rgmanager
 sme8.s390.bos.redhat.com                          2 Online, rgmanager

 Service Name                          Owner (Last)                              State
 ------- ----                          ----- ------                              -----
 service:MyWeb                         sme7.s390.bos.redhat.com                  started
[root@sme7 ~]#
```

Customer Success case, Industry:
Government/Health Services (Australia)

Case Study: HA Websphere MQ Broker Deployment for RHEL on System z

Problem:

A national health services single payer organization had created an application based on Web sphere MQ message queuing streams from health care providers to the health services administration for records and controlled substance management. The application was based on RHEL on Intel. IBM was contracted to host and move the application to a System z-based host. The MQ high-availability code requires a cluster management solution, a cluster aware file system and a workload distribution utility to ensure that work units and message states are preserved if a cluster node is unavailable.

Solution:

IBM employed SNA's High-Availability Option (HAO) for RHEL on System z to configure a set of four-image clusters sharing a common file system. The individual images are deployed on RHEL 6.3 running as guests in 10 z/VM 6.2 LPARs on five physical System z cabinets. The underlying z/VM systems are configured in HA pairs, using the z/VM Single System Image capability to complement the clustering capabilities of the HAO for the RHEL systems. Each LPAR contains 4 RHEL systems running Websphere MQ brokers, with additional clusters deployed on demand. The HAO cluster management code permits concurrent maintenance of individual Linux nodes, and the DASD and FCP disk storage components of the cluster management code manage storage availability to individual LPARs and coordination of write access to all nodes in the cluster. The workload management component of the HAO distributes requests across the clusters of Linux systems, and manages failover and maintenance of individual nodes. The GFS2 cluster file system coordinates data integrity between nodes within single clusters, and across multiple clusters on different machines.

More Information:
Please contact:
David Boyes
Sine Nomine Associates
+1 703 723 6673
info@sinenomine.net