

luci web界面搭建HA

2019年12月29日 15:17

(1) 登录luci
<https://cluster01.test.com:8084> 或<https://20.20.20.7:8084>，使用root/123qweQ登录

(2) 创建集群

Create New Cluster

Cluster Name

test_cluster

☐ Use the Same Password for All Nodes

| Node Name | Password | Ricci Hostname | Ricci Port |
|--------------------|----------|----------------|------------|
| cluster01.test.com | | 20.20.20.7 | 11111 |
| cluster02.test.com | | 20.20.20.8 | 11111 |

Add Another Node

☐ Download Packages
☒ Use Locally Installed Packages

☐ Reboot Nodes Before Joining Cluster
☒ Enable Shared Storage Support

Create Cluster

Cancel

(3) 集群配置(按需配置，本文只配置仲裁QDisk)
「常规」、「Fence守护进程」、「网络」、「冗余环」、「QDisk」、「日志」

「常规」

- 1 集群名称：不支持修改
- 2 配置版本：初始值为1，每次修改配置后自动增加1，可修改。

General

Fence Daemon

Network

Redundant Ring

QDisk

Logging

General Properties

Cluster Name

test_cluster

Configuration Version

14

Apply

「Fence守护进程」

- 1 失败后延迟(Post Fail Delay)：参数为在节点失败之后，执行节点(fence 域中的成员)fencing 前，fence 守护进程(fenced)要等待的秒数。默认值为 0。
- 2 后加入延迟(Post Join Delay)：参数是该节点加入 fence 守护进程(fenced)后，该守护进程 fence 该节点前要等待的秒数。默认值为 6，一般在 20-30 秒之间。

General

Fence Daemon

Network

Redundant Ring

QDisk

Logging

Fence Daemon Properties

Post Fail Delay (seconds)

0

Post Join Delay (seconds)

3

Apply

「网络」：配置网络传输类型

- 1 UDP多播并让集群选择多播地址：根据集群ID自动创建一个多播地址，生成地址的后16字节，并根据IP协议将其附加到地址迁移部分的后面。
 - i. 对于 IPv4 — 该地址格式为 239.192 加上红帽高可用性附加组件软件生成的后 16 字节。
 - ii. 对于 IPv6 — 该地址格式为 FF15:: 加上红帽高可用性附加组件软件生成的后 16 字节。

```
[root@cluster01 ~]# cman_tool status #查看集群ID等信息
Version: 6.2.0
Config Version: 5
Cluster Name: test_cluster
Cluster Id: 249
Cluster Member: Yes
Cluster Generation: 68
Membership state: Cluster-Member
Nodes: 2
Expected votes: 1
```

Total votes: 2
Node votes: 1
Quorum: 1
Active subsystems: 9
Flags: 2node
Ports Bound: 0 11 177
Node name: cluster01.test.com
Node ID: 1
Multicast addresses: 239.192.0.249
Node addresses: 192.168.55.7

- 2 UDP多播并手动指定多播地址：应该使用 cman 采用的 239.192.x.x 序列（IPv6 使用 FF15::），修改以后必须重启该集群以便其生效
- 3 UDP单播(UDPU)：6.2以后，集群节点可以使用UDP单播传输机制进行沟通。但建议您在集群网络中使用IP多播。UDP单播是IP多播不可用时的备用方法。不建议在GFS2部署中使用UDP单播。

GeneralFence DaemonNetworkRedundant RingQDiskLogging

Network Configuration

Network Transport Type

☒ UDP Multicast and Let Cluster Choose the Multicast Address

☐ UDP Multicast and Specify the Multicast Address Manually

Multicast Address

☐ UDP Unicast (UDPU)

Apply

「冗余环」：≥RHEL 6.4

「备用环多播地址」、「备用环 CMAN 端口」和「备用环多播数据包 TTL」

- 1 如果为第二环指定多播地址，则备用多播地址或备用端口必须与第一环的多播地址不同。
- 2 如果指定了备用端口，则第一个环和第二个环的端口号必须至少相差两个，因为系统本身使用port和port-1来执行操作。
- 3 如果不指定备用多播地址，系统将自动为第二个环使用不同的多播地址。

GeneralFence DaemonNetworkRedundant RingQDiskLogging

Redundant Ring Protocol Configuration

Alternate Ring Multicast Address

Alternate Ring CMAN Port

Alternate Ring Multicast Packet TTL

Redundant Ring Cluster Node Configuration

Cluster Node

cluster01.test.com

cluster02.test.com

Alternate Name

Apply

「QDisk」

- 1 指定物理设备: 根据设备标签(By Device Label)、根据文件系统路径(不推荐)

i. 指定 mkqdisk 程序生成的仲裁磁盘标签。如果使用该字段，则仲裁守护进程会读取 /proc/partitions 文件，并在每个找到的块设备中检查仲裁磁盘签名，与指定的标签进行对比。这在节点间使用不同仲裁设备名称时很有用。
- 2 试探法：即仲裁判定设置

i. 到程序的路径——用来决定这个试探是否可用的程序(必填项)。它是 /bin/sh -c 可执行的任意程序, 返回值为 0 表示成功; 其他则表示失败。

ii. 间隔——interval 读取/写入循环的频率，单位为秒。。每次试探间的默认间隔为 2 秒。

iii. 分数——score 仲裁守护进程告知 cman 它有足够分数时的票数。每个试探的默认分数为 1。

iv. TKO——宣布节点死亡时节点必须错过的循环数。
- 3 最小总分：

i. 视节点为“活跃”所需的最小分数。如果忽略或者将其设定为 0，则使用默认功能 floor((n+1)/2)，其中 n 为试探法分数之和。

ii. 最小总分数值必须永远小于试探法分数之和，否则将无法使用仲裁磁盘。

信息记录在/etc/cluster/cluster.conf，如果要让仲裁磁盘生效，或要对仲裁磁盘参数的修改生效，必须重启集群以保证在每个节点中都重启 qdiskd 守护进程。

General
Fence Daemon
Network
Redundant Ring
QDisk
Logging

Quorum Disk Configuration

☐ Do Not Use a Quorum Disk
☒ Use a Quorum Disk

Specify Physical Device

☒ By Device Label

cluster_qdisk

☐ By Filesystem Path to Device (deprecated)

Heuristics

| Path to Program | Interval | Score | TKO |
|-------------------------------|----------|-------|-----|
| /bin/ping -c1 -t2 20.20.20.20 | 1 | 1 | 2 |

Add Another Heuristic

Minimum Total Score

Reset values to defaults

Apply

「日志」

略

*** (4) 安装配置 Fence ***

待完善

(4.1) 安装 Fence

```
# yum install fence-virt.x86_64 fence-virt-multicast.x86_64 fence-virt-libvirt.x86_64
# dd if=/dev/urandom of=/etc/cluster/fence_xvm.key bs=4k count=1
1+0 records in
1+0 records out
4096 bytes (4.1 kB) copied, 0.000867824 s, 4.7 MB/s
```

(4.2) Fence 配置

(5) 创建文件系统

本文使用 gfs 文件系统，在其中任意一个节点执行：

```
1 # service clvmd start
2 # cknconfig clvmd on
# lvs
LV VG Attr LSize Pool Origin Data% Move Log Cpy% Sync Convert
LogVol00 vg_cluster01 -wi-ao--- 19.75g
3 # pvcreate /dev/sdb
Physical volume "/dev/sdb" successfully created
# pvs
PV VG Fmt Attr PSize PFree
/dev/sda2 vg_cluster01 lvm2 a-- 19.75g 0
/dev/sdb lvm2 a-- 20.00g 20.00g
4 # vgcreate vg_ftp /dev/sdb
Clustered volume group "vg_ftp" successfully created
# vgs
VG #PV #LV #SN Attr VSize VFree
vg_cluster01 1 1 0 wz--n- 19.75g 0
vg_ftp 1 0 0 wz--nc 19.99g 19.99g
5 # lvcreate -n lv_ftp -l 100%FREE vg_ftp
Logical volume "lv_ftp" created
# lvs
LV VG Attr LSize Pool Origin Data% Move Log Cpy% Sync Convert
LogVol00 vg_cluster01 -wi-ao--- 19.75g
lv_ftp vg_ftp -wi-a---- 19.99g
6 # mkfs.gfs2 -t test_cluster:lv_ftp -p lock_dlm -j 2 /dev/vg_ftp/lv_ftp
This will destroy any data on /dev/vg_ftp/lv_ftp.
It appears to contain: symbolic link to `../dm-1'
```

Are you sure you want to proceed? [y/n] y

```
Device: /dev/vg_ftp/lv_ftp
Blocksize: 4096
Device Size: 19.99 GB (5240832 blocks)
Filesystem Size: 19.99 GB (5240831 blocks)
Journals: 2
Resource Groups: 80
```

Locking Protocol: "lock_dlm"
Lock Table: "test_cluster:lv_ftp"
UUID: b21d0cc3-e3bb-b292-6628-2b7050a6bea5

(6) 故障切换域

- 输入故障切换域的名字
- Prioritized:是否启用故障切换域成员间设置故障切换优先权，如果启用了，则可以为每个节点配置优先级
- Restricted:是否在失败转移域成员中启用服务故障切换限制
- No Failback:当最高优先级恢复时，不自动回切服务到最高优先级节点
- Member:勾选，则表示节点属于本故障切换域
- Priority:指定优先级，priority值越小，表示优先级越高，为了以后添加节点方便，建议节点间优先级最好相差5左右

Add Failover Domain to Cluster

Name

ftp_failover

☒

Prioritized

Order the nodes to which services failover.

☒

Restricted

Service can run only on nodes specified.

☐

No Failback

Do not send service back to 1st priority node when it becomes available again.

| | Member | Priority |
|--------------------|-------------------------------------|--------------|
| cluster01.test.com | <input checked="" type="checkbox"/> | <div>1</div> |
| cluster02.test.com | <input checked="" type="checkbox"/> | <div>2</div> |

CreateCancel

(7) 添加资源

添加IP资源

Add Resource to Cluster

IP Address

IP Address

192.168.55.6

Netmask Bits (optional)

24

Monitor Link

☒

Disable Updates to Static Routes

☐

Number of Seconds to Sleep After Removing an IP Address

10

SubmitCancel

- 1 IP Address: IP地址，通常作为集群浮动IP;
- 2 Netmask Bits: 子网掩码，本项目中是255.255.255.0，即24;
- 3 其他默认即可
- 4 点击“Submit”提交。

文件系统添加

由于之前我们将使用的共享存储做成了lvm逻辑卷系统，因此在添加文件系统时需要先添加HA LVM，注意启用HA LVM需要确保lvm2-cluster服务启动。

Add Resource to Cluster

HA LVM

HA LVM

Name

test_lvm

Volume Group Name

vg_ftp

Logical Volume Name

lv_ftp

Fence the Node if It is Unable to Clean Up LVM Tags

☐

SubmitCancel

- 1 Name: 给HA LVM取个名字;
- 2 Volume Group Name: 共享存储的卷组名;
- 3 Logical Volume Name: 需要添加的逻辑卷名;
- 4 Fence the Node if It is Unable to Clean Up LVM Tags: 当节点LVM标记无法清除时认为节点脑裂，这里不启用。
- 5 点击“Submit”提交。

然后再添加文件系统:

注: 若填写UUID服务启动失败, 需要更换成路径/dev/mapper/vg_ftp-lv_ftp

Add Resource to Cluster

GFS2

GFS2

Name

test_gfs2

Mount Point

/data

Device, FS Label, or UUID

fc52c9af-7c34-c828-027C

Filesystem Type

GFS2

Mount Options

Filesystem ID (optional)

Force Unmount

☒

Enable NFS daemon and lockd workaround

☐

Reboot Host Node if Unmount Fails

☒

Submit

Cancel

- 1 Name: 给文件系统取个名字;
- 2 Filesystem Type: 指定文件系统类型, 之前创建逻辑卷时将其格式化为ext4格式了, 可以选择ext4, 也可以选Autodetect自动检测;
- 3 Mount Point: 设备挂载位置;
- 4 Device, FS Label, or UUID: 需要挂载的设备, 本项目使用的是LVM逻辑卷, 所以这里填写逻辑卷路径即可: /dev/oraclevg/oraclelv;
- 5 Mount Options和Filesystem ID无视即可。
- 6 Force Unmount: 是否强制卸载, 集群切换时挂载的设备有时候会因为资源占用等问题无法卸载, 导致集群切换失败, 这里启用强制卸载, 可以防止此问题发生;
- 7 Force fsck: 是否强制文件系统检查, 文件系统检查会拖慢集群切换时间, 这里不启用;
- 8 Enable NFS daemon and lockd workaround: 启用NFS守护进程和上锁的解决方法, 无视;
- 9 Use Quick Status Checks: 使用快速状态检查, 无视;
- 10 Reboot Host Node if Unmount Fails: 设备卸载失败时是否重启该节点, 同样可以防止设备卸载失败导致集群切换失败的问题, 可以勾选, 这里无视。
- 11 点击“Submit”提交。

脚本添加

本实验添加的脚本主要用于控制集群切换时vsftpd实例的起停。RHCS集群对于脚本有专门的格式要求, 通常需要有start、stop等函数进行自动调用。脚本应当先分别保存在/etc/init.d目录下并赋予可执行权限

Add Resource to Cluster

Script

Script

Name

test_vsftpd_script

Full Path to Script File

/etc/init.d/vsftpd

Submit

Cancel

至此资源全部添加完成, 如下图所示。

| Nodes | Fence Devices | Failover Domains | Resources | Service Groups | Configure |
|--|--------------------|---------------------------|-----------|----------------|-----------|
| <div><div>+ Add</div><div>✕ Delete</div></div> | | | | | |
| Name/IP | | Type | | In Use | |
| <input type="checkbox"/> | 192.168.55.6/24 | IP Address | | ✓ | |
| <input type="checkbox"/> | test_lvm | Logical Volume Management | | ✓ | |
| <input type="checkbox"/> | test_vsftpd_script | Script | | ✓ | |
| <input type="checkbox"/> | test_gfs2 | GFS2 | | ✓ | |

Select an item to view details

(8) 创建资源组

集群资源需要加入资源组中才能真正被集群中的节点所使用。

Add Service Group to Cluster

| | |
|----------------------------------|---|
| Service Name | <input type="text" value="test_ftp_service"/> |
| Automatically Start This Service | <input checked="" type="checkbox"/> |
| Run Exclusive | <input type="checkbox"/> |
| Failover Domain | <input type="text" value="ftp_failover"/> |
| Recovery Policy | <input type="text" value="Relocate"/> |

Restart Options

| | |
|---|----------------------|
| Maximum Number of Restart Failures Before Relocating | <input type="text"/> |
| Length of Time in Seconds After Which to Forget a Restart | <input type="text"/> |

Add Resource

Submit

Cancel

- 1 Service Name: 给资源组取个名字;
- 2 Automatically Start This Service: 自动启动集群资源组服务, 勾选;
- 3 Run Exclusive: 集群服务独占方式运行, 拒绝其他集群服务运行在节点服务器上, 这里是否勾选都行, 建议不勾选;
- 4 Failover Domain: 指定刚才配置的失败转移域;
- 5 Recovery Policy: 恢复策略, 选择Relocate迁移, 一旦节点失效立刻迁移到其他节点。
- 6 Restart Options: 节点重启策略, 默认即可。

然后我们需要为资源组添加资源, 依次添加之前已经配置好的IP Address, HA LVM和Script资源即可。

查询结果如下:

```
[root@cluster01 ~]# clustat
Cluster Status for test_cluster @ Sat Dec 28 20:54:22 2019
Member Status: Quorate

Member Name                                ID   Status
-----
cluster01.test.com                        1   online, Local, rgmanager
cluster02.test.com                        2   online, rgmanager
/dev/block/8:32                           0   online, Quorum Disk

Service Name                                owner (Last)                                State
-----
service:test_ftp_service                    cluster01.test.com                          started
```