Kylin高可用集群配置与管理

## 一．建立集群

1. 关闭两台主机防火墙，关闭selinux

systemctl disable firewalld

systemctl stop firewalld

setenforce 0

2. 修改两台主机名

hostnamectl set-hostname zdatadkydb1 （zdatadkydb2）

3. 配置两台主机网络

vi /etc/sysconfig/network-scripts/ifcfg-eno1

vi /etc/sysconfig/network-scripts/ifcfg-bond0

4. 确定两台机器时间一致

5. 编辑/etc/hosts文件，添加两台主机地址

10.55.3.82 zdatadkydb1 #节点zdatadkydb1的业务网口bond0(由eno3与ens6f0做主备绑定)

10.55.3.83 zdatadkydb2 #节点zdatadkydb2的业务网口bond0(由eno3与ens6f0做主备绑定)

192.168.101.9 dkydb1 #节点zdatadkydb1的心跳网口eno1

192.168.101.10 dkydb2 #节点zdatadkydb2的心跳网口eno1

6. 安装 psc corossync pacemaker crmsh

将附件 ha.iso 上传到 /opt 目录，并使用 root 用户进行挂载：

mount /opt/ha.iso /mnt

将附件 ha.repo 文件上传到 /etc/yum.repos.d/ 目录

yum安装：

yum clean all

yum makecache

yum install pcs corosync corosynclib pacemaker pacemaker-cli crmsh\*

7. 在两个node上修改管理密码

echo "kylin.2020" |passwd --stdin hacluster

8. 两个node上启动pscd服务

systemctl start pcsd.service

systemctl status pcsd.service

systemctl enable pcsd.service

一下在一台节点配置即可

9. 集群节点认证

pcs cluster auth zdatadkydb1 zdatadkydb2 -u hacluster -p kylin.2020

10 创建集群

pcs cluster setup --name my\_cluster zdatadkydb1 zdatadkydb2

11. 集群启动

pcs cluster start --all

pcs status

设置开机自启动：

pcs cluster enable --all(两台机器都执行)

12. 设置资源默认粘性（防止资源回切），设置优先级，防止资源回切，为正数表示倾向，负数则会离开，-inf表示负无穷，inf表示正无穷

pcs resource defaults resource-stickiness=100

pcs resource defaults

13. 设置资源超时时间

pcs resource op defaults timeout=90s

pcs resource op defaults

14. 二个节点时，忽略节点quorum功能,正常集群需要quorum需要半数以上的票数，是双节点的话，会存在法定票数不足导致的资源不转移，因此忽略集群没有法定票数时直接忽略，在加一个ping node节点，当ping不通ping node资源时，节点会释放资源，3个节点以上不需要

法定票数不足也就是只有两个节点时，会导致资源不转移，解决办法如下：

1. 可以增加一个ping node节点
2. 可以增加一个仲裁磁盘
3. 让集群中的节点数成奇数个，也就是增加节点
4. 直接忽略当集群没有法定票数时直接忽略

pcs property set no-quorum-policy=ignore

15. 没有 Fencing设备时，禁用STONITH 组件功能，

1）STONITH：节点级别的隔离，通过断开一个节点的电源或者重新启动节点

2）fencing：资源级别的隔离，类似通过向交换机发送隔离信号，特意让数据无法通过此接口

当集群分裂，即分裂后的一个集群的法定票数小于总票数的一半时采取对资源的控制策略

pcs property set stonith-enabled=false

16. 验证群集配置信息（没有输出为正常）

crm\_verify -L -V

## 二．配置集群资源

1. 创建vip资源

pcs resource create vip ocf:heartbeat:IPaddr2 ip=**192.168.8.130** cidr\_netmask=24 nic=ens33 op monitor interval=30s

2. 创建lvm资源和lv资源（设备名可用UUID） 使用UUID前面加上-U

pcs resource create zdata\_lvm LVM volgrpname=SSD\_POOL01\_VG exclusive=true

pcs resource create dbdata\_lv Filesystem device="/dev/mapper/SSD\_POOL01\_VG-dbdata" directory="/dbdata" fstype="xfs"

pcs resource create dbarch\_lv Filesystem device="/dev/mapper/SSD\_POOL01\_VG-dbarch" directory="/dbarch" fstype="xfs"

pcs resource create dbbak\_lv Filesystem device="/dev/mapper/SSD\_POOL01\_VG-dbbak" directory="/dbbak" fstype="xfs"

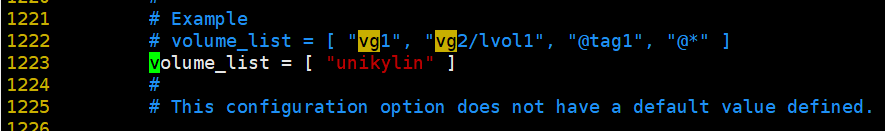
**开放halvm功能**

添加lvm资源报错

1)lvmconf --enable-halvm --services --startstopservices

2)cp /etc/lvm/lvm.conf{,.bak}

3)编辑lvm.conf文件



查看现有vg，放开volume\_list的注释，将除了集群共享存储的vg加入到括号呢内

注意：主机后续添加新的非集群管理vg，必须将vg名加入到volume\_list内

4）备份boot/initramfs-$(uname -r).img文件

cp /boo/initramfs-$(uname -r).img{,.bak}

5）dracut -H -f /boot/initramfs-$(uname -r).img $(uname -r)

6)reboot 重启

[root@node1 lvm]# lvscan

inactive '/dev/vgdbbak/lvdbbak' [<19.99 GiB] inherit

ACTIVE '/dev/unikylin/swap' [2.00 GiB] inherit

ACTIVE '/dev/unikylin/root' [17.80 GiB] inherit

lvdbbak 的状态为ACTIVE为成功

两个节点，一个为ACTIVE，一个为inactive,当前在哪个节点哪个节点为ACTIVE

3. 创建达梦服务

pcs resource create dmserverd systemd:httpd op monitor interval=30s

pcs resource create DmAPService systemd:DmAPService op monitor interval=90s

4. 资源设置排列约束(资源同时启动，关闭，重启)，将资源绑定在一起，始终保持在一个节点上

pcs constraint colocation add vip zdata\_lvm dbdata\_lv dbarch\_lv dbbak\_lv dmserverd DmAPService INFINITY

INFINITY：无穷大：q

5. 设置顺序约束(资源启动关闭顺序)

pcs constraint order (start）vip then zdata\_lvm

pcs constraint order zdata\_lvm then dbdata\_lv

pcs constraint order dbdata\_lv then dbarch\_lv

pcs constraint order dbarch\_lv then dbbak\_lv

pcs constraint order dbbak\_lv then dmserverd

pcs constraint order dmserverd then DmAPService

6. 创建资源组：

pcs resource group add mygroup vip zdata\_lvm dbdata\_lv dbarch\_lv dbbak\_lv dmserverd DmAPService

@@7. 配置网络ping监控，当服务器指定的ip（10.55.3.254）不通时，节点释放资源组

配置ping node主资源:

(pcs resource create PingCheck ocf:pacemaker:ping multiplier=100 host\_list="192.168.8.108 router" op monitor interval=30s timeout=10s

)

crm

configure

primitive myping ocf:pacemaker:ping params host\_list=**192.168.142.150** multiplier=100 op monitor interval=10s timeout=60s op start timeout=60s

commit

配置myping的克隆资源

crm

configure

clone cl\_myping myping meta globally-unique="false"

commit

添加判断规则，绑定资源组,配置ping node 失败的时候，资源的故障转移

crm

configure

location loc\_group\_on\_myping  **mygroup** rule $id="loc\_ip\_cluster\_on\_ping\_rule" -inf: not\_defined pingd or pingd lte 0

commit

添加规则，将pingd的状态和需要保护的资源进行绑定。

* loc\_group\_on\_myping：约束名称，自定义
* my\_res：需要保护的资源或资源组名称
* loc\_ip\_cluster\_on\_ping\_rule：规则名称，自定义
* -inf: not\_defined pingd or pingd lte 0：ping的规则，大意是如果节点上没有pingd或者是pingd失效，将放弃对资源的管理，交给其它节点。

## 三．集群管理

集群配置了服务pcsd，pacemaker,corosync开机默认自启动，常见的集群管理如下：

### 查看集群状态

查看集群整体状态：pcs status

[root@zdatadkydb2 ~]# pcs status

Cluster name: my\_cluster

Stack: corosync

Current DC: zdatadkydb1 (version 1.1.19-8.ky3.4-c3c624ea3d) - partition with quorum

Last updated: Thu Apr 23 17:14:39 2020

Last change: Thu Apr 23 17:07:33 2020 by root via cibadmin on zdatadkydb1

2 nodes configured

9 resources configured

**Online: [ zdatadkydb1 zdatadkydb2 ]**

Full list of resources:

Resource Group: mygroup

vip (ocf::heartbeat:IPaddr2): Started zdatadkydb2

zdata\_lvm (ocf::heartbeat:LVM): Started zdatadkydb2

dbdata\_lv (ocf::heartbeat:Filesystem): Started zdatadkydb2

dbarch\_lv (ocf::heartbeat:Filesystem): Started zdatadkydb2

dbbak\_lv (ocf::heartbeat:Filesystem): Started zdatadkydb2

dmserverd (systemd:dmserverd): Started zdatadkydb2

DmAPService (systemd:DmAPService): Started zdatadkydb2

Clone Set: cl\_myping [myping]

Started: [ zdatadkydb1 zdatadkydb2 ]

Daemon Status:

corosync: active/enabled

pacemaker: active/enabled

pcsd: active/enabled

查看集群状态pcs cluster status

[root@zdatadkydb2 opt]# pcs cluster status

Cluster Status:

Stack: corosync

Current DC: zdatadkydb1 (version 1.1.19-8.ky3.4-c3c624ea3d) - partition with quorum

Last updated: Thu Apr 23 17:09:24 2020

Last change: Thu Apr 23 17:07:42 2020 by root via cibadmin on zdatadkydb1

2 nodes configured

9 resources configured

PCSD Status:

zdatadkydb2: Online

zdatadkydb1: Online

查看集群实时刷新状态：crm\_mon –i1

Cluster name: my\_cluster

Stack: corosync

Current DC: zdatadkydb1 (version 1.1.19-8.ky3.4-c3c624ea3d) - partition with quorum

Last updated: Thu Apr 23 17:14:39 2020

Last change: Thu Apr 23 17:07:33 2020 by root via cibadmin on zdatadkydb1

2 nodes configured

9 resources configured

**Online: [ zdatadkydb1 zdatadkydb2 ]**

Full list of resources:

Resource Group: mygroup

vip (ocf::heartbeat:IPaddr2): Started zdatadkydb2

zdata\_lvm (ocf::heartbeat:LVM): Started zdatadkydb2

dbdata\_lv (ocf::heartbeat:Filesystem): Started zdatadkydb2

dbarch\_lv (ocf::heartbeat:Filesystem): Started zdatadkydb2

dbbak\_lv (ocf::heartbeat:Filesystem): Started zdatadkydb2

dmserverd (systemd:dmserverd): Started zdatadkydb2

DmAPService (systemd:DmAPService): Started zdatadkydb2

Clone Set: cl\_myping [myping]

Started: [ zdatadkydb1 zdatadkydb2 ]

### 2. 启动集群

启动集群的所有节点（2个）：pcs cluster start –all

[root@zdatadkydb2]# pcs cluster start --all

zdatadkydb1: Starting Cluster (corosync)...

zdatadkydb2: Starting Cluster (corosync)...

zdatadkydb2: Starting Cluster (pacemaker)...

zdatadkydb1: Starting Cluster (pacemaker)...

启动集群的单个节点（在zdatadkydb2执行，则仅启动zdatadkydb2的集群服务）：pcs cluster start

[root@zdatadkydb2]# pcs cluster start

Starting Cluster (corosync)...

Starting Cluster (pacemaker)...

启动集群的指定的单个节点（在zdatadkydb 2上启动zdatadkydb 1）: pcs cluster start zdatadkydb1

[root@**zdatadkydb2**]# pcs cluster start zdatadkydb 1

Zdatadkydb1：Starting Cluster (corosync)...

Zdatadkydb1：Starting Cluster (pacemaker)...

### 3. 停止集群

停止集群的所有节点（2个）：pcs cluster stop --all

[root@zdatadkydb2 koji11]# pcs cluster stop --all

zdatadkydb2: Stopping Cluster (pacemaker)...

zdatadkydb1: Stopping Cluster (pacemaker)...

zdatadkydb2: Stopping Cluster (corosync)...

zdatadkydb1: Stopping Cluster (corosync)...

停止集群的单个节点，如果对方节点正常，则会发生资源切换：pcs cluster stop

[root@zdatadkydb2 corosync]# pcs cluster stop

Stopping Cluster (pacemaker)...

Stopping Cluster (corosync)...

在仅有一个节点运行时，停止集群的单个节点：pcs cluster stop --force

[root@zdatadkydb1 ~]# pcs cluster stop --force

Stopping Cluster (pacemaker)...

Stopping Cluster (corosync)...

停止集群的指定的单个节点（在zdatadkydb 2上停止zdatadkydb 1）: pcs cluster start zdatadkydb1

[root@zdatadkydb2 ~]# pcs cluster stop zdatadkydb1

zdatadkydb1:Stopping Cluster (pacemaker)...

zdatadkydb1:Stopping Cluster (corosync)...

### 4. 其他集群操作

##把指定节点置于后备standby状态

pcs cluster standby zdatadkydb1

##把处于后备standby状态的节点恢复到offline

pcs cluster unstandby zdatadkydb1

## 清除指定资源（vip）的状态与错误计数

pcs resource cleanup vip

### 5. 查看群集配置

## 检查配置有无错误

crm\_verify -L -V

## 查看群集属性

pcs property

## 查看stonith

pcs stonith

## 查看资源约束

pcs constraint

pcs config

## 查看群集资源配置

pcs cluster cib

## 以XML格式显示群集配置