双主介绍：

MySQL主从复制架构可以在很大程度保证MySQL的高可用，在一主多从的架构中还可以利用读写分离将读操作分配到从库中，减轻主库压力。但是在这种架构中，主库出现故障时需要手动将一台从库提升为主库。在对写操作要求较高的环境中，主库故障在主从架构中会成为单点故障。因此需要主主互备架构，避免主节点故障造成写操作失效。

基于MySQL 5.7多源复制+Keepalived搭建高可用

##说明 本内容来源于[知数堂](http://zhishuedu.com/) 公开课 : 《MySQL 5.7 高可用新玩法》--吴炳锡 相关视频推荐：<https://ke.qq.com/course/172600>

##基本环境准备 使用Centos 6.X 64位系统 MySQL 使用 MySQL-5.7.17-x86\_64 版本，去官方下载mysql-5.7.17-linux-glibc2.5-x86\_64.tar.gz 版本

|  |  |  |
| --- | --- | --- |
| **机器名** | **操作系统** | **IP** |
| node1 | centos-6.6 | 172.21.109.67 |
| node2 | centos-6.6 | 172.21.109.68 |
|  |  |  |

对应的VIP： 192.168.11.110

特别提示: 关闭iptables chkconfig --del iptables /etc/init.d/iptables stop

关闭:selinux setenforce 0 vim /etc/sysconfig/selinux SELINUX=permissive 更改为： SELINUX=disabled

###下载MySQL ：

mkdir /data/Soft cd /data/Soft wget https://dev.mysql.com/get/Downloads/MySQL-5.7/mysql-5.7.17-linux-glibc2.5-x86\_64.tar.gz

###MySQL部署约定 二进制文件放置： /opt/mysql/ 下面对应的目录 数据文件全部放置到 /data/mysql/ 下面对应的目录 原始二进制文件下载到/data/Soft/

##MySQL基本安装 以下安装步骤需要在node1, node2, node3上分别执行。

1. mkdir /opt/mysql
2. cd /opt/mysql
3. tar zxvf /data/Soft/mysql-5.7.17-linux-glibc2.5-x86\_64.tar.gz
4. ln -s /opt/mysql/mysql-5.7.17-linux-glibc2.5-x86\_64 /usr/local/mysql
5. mkdir /data/mysql/mysql3309/{data,logs,tmp} -p
6. groupadd mysql
7. useradd -g mysql -s /sbin/nologin -d /usr/local/mysql -M mysql
8. chown -R mysql:mysql /data/mysql/
9. chown -R mysql:mysql /usr/local/mysql
10. cd /usr/local/mysql/
11. ./bin/mysqld --defaults-file=/data/mysql/mysql3309/my3309.cnf --initialize (配置读写分离，主库read-only=off,从库read-only=on);
12. cat /data/mysql/mysql3309/data/error.log |grep password
13. /usr/local/mysql/bin/mysqld --defaults-file=/data/mysql/mysql3309/my3309.cnf &
14. echo "export PATH=$PATH:/usr/local/mysql/bin" >>/etc/profile
15. source /etc/profile
16. mysql -S /tmp/mysql3309.sock -p #输才查到密码进入MySQL
17. mysql>alter user user() identified by '\*\*\*\*\*\*'
18. mysql>grant replication slave on *.* to 'repl'@'%' identified by 'repl4slave';
19. mysql>grant all privilegs on *\*.\** to 'zhaixiaona'@'%' identified by '\*\*\*\*\*' # 一会测试使用的帐号
20. mysql>reset master

在node1和node2上执行change master之后，形成双主模式

之后配置keepalived；

mount /dev/sro /mnt

**挂载yum源之后，yum install keepalived**

**安装python依赖模块：**

**yum install MySQL-python.x86\_64 yum install python2-filelock.noarch**

**将checkmysql.py检测数据库是否存活脚本放在/etc/keepalived目录下，看是否可执行**

**chmod +x checkMySQL.py**

**python checkmysql.py**

**包可以根据下载之后上传；**

安装之后，主库修改配置文件

keepalived.conf

vrrp\_script vs\_mysql\_82 {

script "python /etc/keepalived/checkMYSQL.py -h 172.21.109.67 -P 3306"

interval 60

}

vrrp\_instance VI\_82 {

state BACKUP

nopreempt

interface eth0

virtual\_router\_id 82

priority 100

advert\_int 5

authentication {

auth\_type PASS

auth\_pass 1111

}

track\_script {

vs\_mysql\_82

}

virtual\_ipaddress {

172.21.109.200

}

}

**--------------------------------------------------**

**解释配置文件：**

vrrp\_script vs\_mysql\_82 {

script "python /etc/keepalived/checkMYSQL.py -h 172.21.109.67 -P 3306"

interval 60

}

这部分代表，检测不到数据库存活，则keepalived出现fault报错，不会切换

；

vrrp\_instance VI\_82 {

state BACKUP #状态备库

nopreempt #不抢占模式

interface eth0

virtual\_router\_id 82

priority 100 #优先级

advert\_int 5

authentication {

auth\_type PASS

auth\_pass 1111

}

track\_script {

vs\_mysql\_82

}

virtual\_ipaddress { #虚拟vip

172.21.109.200

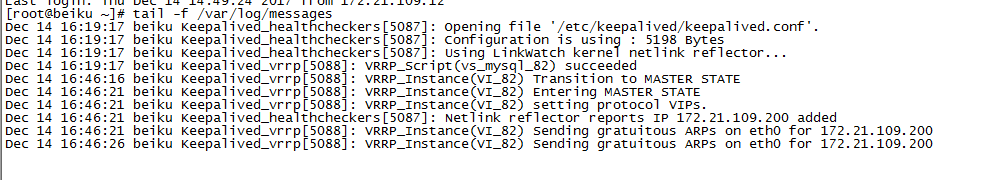
}

}

启动keepalived

使用yum 安装，则可使用/etc/init.d/keepalived start即可

之后查看cat /var/log/messages 启动日志



会出现成功状态，大概1分钟，之后绑定vip；

可使用ip addr show 检测，vip是否可在

ps -aux |grep keepalived

--------------------------------------------------

在节点2配置文件

vrrp\_script vs\_mysql\_82 {

script "python /etc/keepalived/checkMYSQL.py -h 172.21.109.68 -P 3306"

interval 15

}

vrrp\_instance VI\_82 {

state BACKUP

nopreempt

interface eth0

virtual\_router\_id 82

priority 50 #优先级小于主库

advert\_int 5

authentication {

auth\_type PASS

auth\_pass 1111

}

track\_script {

vs\_mysql\_82

}

notify /etc/keepalived/notify.py

virtual\_ipaddress {

172.21.109.200

}

}

/etc/init.d/keepalived start

tail -f /var/log/messages

------------------------------------------------------------------------

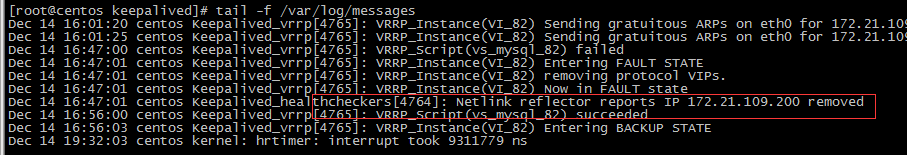
测试，在主库上

ps -ef |grep mysql

kill 进程号

停掉mysql

tail -f /var/log/messages



之后 查看备库

tail -f /var/log/messages

发现从库已经绑定上vip

实现自动切换

mysql -uzhaixiaona -h 172.21.109.200 -P 3306 -p

select @@hostname

show variables like '%read%';

查看读写分离继续

其中checkmysql.py

大概总结脚本如下：

#！/bin/bash

count "ps -aux |grep -v |grep mysql |wc -l"

if { $count >0 }; then

exit 0

else

exit 1

fi

------------------------

**#!/bin/bash**

counter=$(netstat -na|grep "LISTEN"|grep "3306"|wc -l)

if [ "${counter}" -eq 0 ]; then

    /etc/init.d/keepalived stop

fi

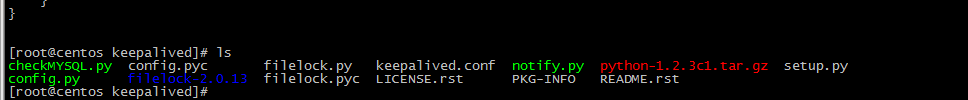
ps -ef | egrep -i \"mysqld\" | grep %s | egrep -iv \"mysqld\_safe\" | grep -v grep | wc -l

原理分析：

healthcheck

vrrp

先进入backup state 状态，运行一次vrrp\_scripts，成功后，发现没有主 》master> 拉起vip完成



**------------------------------------温馨提示（Keepalived的抢占和非抢占模式）---------------------------------------**

|  |  |
| --- | --- |
|  |  |

keepalive是基于vrrp协议在linux主机上以守护进程方式，根据配置文件实现健康检查。

VRRP是一种选择协议，它可以把一个虚拟路由器的责任动态分配到局域网上的VRRP路由器中的一台。

控制虚拟路由器IP地址的VRRP路由器称为主路由器，它负责转发数据包到这些虚拟IP地址。

一旦主路由器不可用，这种选择过程就提供了动态的故障转移机制，这就允许虚拟路由器的IP地址可以作为终端主机的默认第一跳路由器。

keepalive通过组播，单播等方式（自定义），实现keepalive主备推选。工作模式分为抢占和非抢占（通过参数nopreempt来控制）。

1）抢占模式：

主服务正常工作时，虚拟IP会在主上，备不提供服务，当主服务优先级低于备的时候，备会自动抢占虚拟IP，这时，主不提供服务，备提供服务。

也就是说，工作在抢占模式下，不分主备，只管优先级。

如上配置，不管keepalived.conf里的state配置成master还是backup，只看谁的priority优先级高（一般而言，state为MASTER的优先级要高于BACKUP）。

priority优先级高的那一个在故障恢复后，会自动将VIP资源再次抢占回来！！

2）非抢占模式：

这种方式通过参数nopreempt（一般设置在advert\_int的那一行下面）来控制。不管priority优先级，只要MASTER机器发生故障，VIP资源就会被切换到BACKUP上。

并且当MASTER机器恢复后，也不会去将VIP资源抢占回来，直至BACKUP机器发生故障时，才能自动切换回来。

千万注意：

nopreempt这个参数只能用于state为backup的情况，所以在配置的时候要把master和backup的state都设置成backup，这样才会实现keepalived的非抢占模式！

也就是说：

a）当state状态一个为master，一个为backup的时候，加不加nopreempt这个参数都是一样的效果。即都是根据priority优先级来决定谁抢占vip资源的，是抢占模式！

b）当state状态都设置成backup，如果不配置nopreempt参数，那么也是看priority优先级决定谁抢占vip资源，即也是抢占模式。

c）当state状态都设置成backup，如果配置nopreempt参数，那么就不会去考虑priority优先级了，是非抢占模式！即只有vip当前所在机器发生故障，另一台机器才能接管vip。即使优先级高的那一台机器恢复  后也不会主动抢回vip，只能等到对方发生故障，才会将vip切回来。