

keepalived配置

前提：两台主机形成互为主从的配置、主机ip固定

ps: 如果在容器里配置的话，要确保生成容器时使用了--privileged=true 选项。

一、安装keepalived

1. 下载keepalived所依赖的组件（如果没有，后边安装的时候会报各种错误、解决措施就是装这些组件）

```
yum install -y gcc openssl openssl-devel popt popt-devel libnl libnl-devel
yum install -y ipvsadm
yum install -y libnfnetlink-devel
yum install -y rsyslog
```

2. 下载keepalived 包

```
wget -P /usr/local/etc/ http://www.keepalived.org/software/keepalived-1.4.2.tar.gz
cd /usr/local/etc/
tar -xzf keepalived-1.4.2.tar.gz
mv keepalived-1.4.2 keepalived
cd keepalived
./configure
make
make install
```

3. 创建链接（keepalived安装位置与运行需要的默认位置不一致，所以要做链接或拷贝）

```
mkdir /etc/keepalived
## 备份keepalived主体文件
test -f /usr/local/etc/keepalived/keepalived.conf.bak || cp /usr/local/etc/keepalived/keepalived.conf /usr/local/etc/keepalived/keepalived.conf.bak
## 创建链接
test -f /etc/keepalived/keepalived.conf || ln -s /usr/local/etc/keepalived/keepalived.conf /etc/keepalived/keepalived.conf

chmod 664 /etc/keepalived/keepalived.conf

test -f /etc/sysconfig/keepalived || ln -s /usr/local/etc/sysconfig/keepalived /etc/sysconfig/keepalived

test -f /sbin/keepalived || ln -s /usr/local/sbin/keepalived /sbin/keepalived
```

4. 设置开机自启

```
## 复制keepalived 服务启动脚本到开机启动脚本目录中
test -f /etc/rc.d/init.d/keepalived || cp
/usr/local/etc/keepalived/keepalived/etc/init.d/keepalived
/etc/rc.d/init.d/keepalived
## 授权
chmod +x /etc/rc.d/init.d/keepalived
## 下面两条指令都是设置开机自启的方式，一般来说第二条会执行成功，第一条会报错。不用管
kconfig keepalived on
systemctl enable keepalived.service
service keepalived start
```

注：上述所有指令可以一条条执行，也可以写一个shell脚本放在一起按顺序批量执行。

5. 修改keepalived日志存放位置

默认日志存放在/var/log/messages中。更改为/var/log/keepalived.log

```
vim /etc/sysconfig/keepalived
## 将最后一行 KEEPALIVED_OPTIONS="-D" 注释掉
## 写入
KEEPALIVED_OPTIONS="-D -S 0"

vim /etc/rsyslog.conf
## 找到 # Save boot messages also to boot.log 字样，在下边追加（或者直接在末尾追加）
local0.* /var/log/keepalived.log

## 重启rsyslog
systemctl restart rsyslog.service
## 重启keepalived
service keepalived restart
```

至此、keepalived安装完成、可以使用 `service keepalived status` 命令查看运行情况，如果未能正常运行，查看日志寻找原因。（只要按步骤安装，一般到这里不会出错）

二、下面开始配置keepalived

我这里使用的两台主机ip分别为 172.30.0.2、172.30.0.3，网卡均为eth0。想要配置的虚拟ip为 172.30.0.88。

ps：网上关于配置文件keepalived.conf的写法五花八门，有的会做邮箱通知、心跳检测、切换到各种状态要执行的功能。有的实现了防止主机竞争、防脑裂。一开始的时候都想试试，最终都没成功。所以最后还是回到最基础的功能：参照着keepalived安装时产生的keepalived.conf做类似的最简单的配置（在安装keepalived的第3步创建链接时，曾备份主体文件，将原有的keepalived.conf备份成了 keepalived.conf.bak）

修改/usr/local/etc/keepalived/keepalived.conf的内容如下

```
! Configuration File for keepalived

global_defs {
    notification_email {
        tangshiheng@mastercom.cn ##这里的邮箱就是写着玩，实际上并不会收到邮件
    }
    notification_email_from HAproxy01@one.com
```

```

smtp_server 127.0.0.1 ## 邮箱服务器，如果配置正确的话可以收到邮件通知
smtp_connect_timeout 30

router_id mysql_HA
vrrp_skip_check_adv_addr
## 此项注释掉，否则虚拟ip无法ping通
## vrrp_strict
vrrp_garp_interval 0
vrrp_gna_interval 0
}

vrrp_instance VI_1 {
    ## 定义：实例角色，都写为BACKUP，通过优先级来竞争MASTER
    state BACKUP
    ## 定义：承载VIP地址的物理接口
    interface eth0
    ## 定义：承载VIP地址的物理接口
    virtual_router_id 51
    ## 定义：优先级初始值、多个节点应不同
    priority 100
    ## 定义：VRRP通知报文的时间间隔
    advert_int 1
    ## 设置：验证信息（两个节点必须一致）
    authentication {
        auth_type PASS
        auth_pass 1111
    }
    virtual_ipaddress {
        172.30.0.88 ##虚拟ip
    }
}

virtual_server 172.30.0.88 3306 {
    delay_loop 6
    ## 这两项注释掉，是看别人说的，想知道效果的话取消注释尝试一下
    ## lb_algo rr
    ## 这一项不能注释，否则无法登陆端口
    lb_kind NAT
    persistence_timeout 50
    protocol TCP

    real_server 172.30.0.2 3306 {
        ## 权重
        weight 1
        ## 自杀脚本，出现问题关掉本节点的keepalived服务
        notify_down /etc/keepalived/mysql_down.sh

        TCP_CHECK {
            connect_timeout 10
            nb_get_retry 3
            delay_before_retry 3
            connect_port 3306
        }
    }
}

real_server 172.30.0.3 3306 {
    weight 1
    notify_down /etc/keepalived/mysql_down.sh
}

```

```

TCP_CHECK {
    connect_timeout 10
    nb_get_retry 3
    delay_before_retry 3
    connect_port 3306
}

}

}

```

ps：以上内容可以根据自己的实际情况更改相应ip。可以在windows下编辑好。通过docker命令：

docker cp 文件地址 容器id:文件地址 直接拷贝到容器中（同样可以从容器拷贝文件出来，不管容器是否开启都可以）例：`docker cp E:/Docker/keepalived.conf d0:/usr/local/etc/keepalived/`

上面的脚本拷贝到容器后应修改 priority 100 为不同的值，对应的real_server也要修改。然后创建“自杀”脚本

```

## 创建配置文件里所需要的mysql_down脚本
cat > /etc/keepalived/mysql_down.sh <<EOF
#!/bin/bash
service keepalived stop
EOF
chmod +x /etc/keepalived/mysql_down.sh

```

配置完成后重启keepalived.此时可能会重启失败，查看日志如果是报如下错误：

Configuration file '/etc/keepalived/keepalived.conf' is not a regular non-executable file.

是由于keepalived.conf权限问题造成的，授权即可解决（链接文件在每次重建指向文件后都要重新授权）

```
chmod 644 /etc/keepalived/keepalived.conf
```

```

Aug 20 14:19:41 408a38fe6011 systemd[1]: Started System Logging Service.
[root@408a38fe6011 log]# service keepalived restart
Restarting keepalived (via systemctl): Job for keepalived.service failed. See "systemctl status keepalived.service" and "journalctl -xe" for details.
[FAILED]
[root@408a38fe6011 log]# tail -f /var/log/keepalived.log
Aug 20 14:19:41 408a38fe6011 Keepalived[5227]: Starting Keepalived v1.4.2 (02/24,2018), git commit v1.4.1-41-g6a2987e+
Aug 20 14:19:41 408a38fe6011 Keepalived[5227]: Running on Linux 4.19.76-linuxkit #1 SMP Tue May 26 11:42:35 UTC 2020 (built for Linux 3.10.0)
Aug 20 14:19:41 408a38fe6011 Keepalived[5227]: Configuration file '/etc/keepalived/keepalived.conf' is not a regular non-executable file
Aug 20 14:19:41 408a38fe6011 Keepalived[5227]: Stopped Keepalived v1.4.2 (02/24,2018), git commit v1.4.1-41-g6a2987e+
Aug 20 14:26:04 408a38fe6011 Keepalived[5316]: Starting Keepalived v1.4.2 (02/24,2018), git commit v1.4.1-41-g6a2987e+
Aug 20 14:26:04 408a38fe6011 Keepalived[5316]: Running on Linux 4.19.76-linuxkit #1 SMP Tue May 26 11:42:35 UTC 2020 (built for Linux 3.10.0)
Aug 20 14:26:04 408a38fe6011 Keepalived[5316]: Configuration file '/etc/keepalived/keepalived.conf' is not a regular non-executable file
Aug 20 14:26:04 408a38fe6011 Keepalived[5316]: Stopped Keepalived v1.4.2 (02/24,2018), git commit v1.4.1-41-g6a2987e+
^C
[root@408a38fe6011 log]# cd /etc/keepalived/
[root@408a38fe6011 keepalived]# ls
keepalived.conf  mysql_down.sh
[root@408a38fe6011 keepalived]# chmod 644 keepalived.conf

```

三、检查是否配置成功

如果前边的步骤顺利完成，两个节点均能正常启动keepalived，那其中一台节点（优先级更高的，我这里是172.30.0.3）会链接到虚拟ip上。查看网卡配置。`ip add | grep eth0` (eth0为配置文件配置的网卡)

```

[root@408a38fe6011 log]# ip add | grep eth0
17: eth0@if18: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state UP group default
    inet 172.30.0.3/16 brd 172.30.255.255 scope global eth0
       inet 172.30.0.88/32 scope global eth0
[root@408a38fe6011 log]#

```

可以看到, 172.30.0.88已经挂载到网卡上了。

使用第三台主机连接到虚拟ip(宿主机不行, 只能是容器),可以正常登陆。

```
[root@985c5b093f33 ~]# ifconfig
eth0: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
    inet 172.30.0.4 netmask 255.255.0.0 broadcast 172.30.255.255
    ether 02:42:ac:1e:00:04 txqueuelen 0 (Ethernet)
    RX packets 171316 bytes 10659828 (10.1 MiB)
    RX errors 0 dropped 0 overruns 0 frame 0
    TX packets 12836 bytes 735540 (718.3 KiB)
    TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0

lo: flags=73<UP,LOOPBACK,RUNNING> mtu 65536
    inet 127.0.0.1 netmask 255.0.0.0
    loop txqueuelen 1000 (Local Loopback)
    RX packets 24 bytes 1336 (1.3 KiB)
    RX errors 0 dropped 0 overruns 0 frame 0
    TX packets 24 bytes 1336 (1.3 KiB)
    TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0

[root@985c5b093f33 ~]# mysql -h 172.30.0.88 -P 3306 -udtauser -pdtauser
mysql: [Warning] Using a password on the command line interface can be insecure.
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 205
Server version: 8.0.21 MySQL Community Server - GPL

Copyright (c) 2000, 2020, Oracle and/or its affiliates. All rights reserved.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql>
```

接下来模拟一下故障发生:

初始时刻, 172.30.0.3的优先级更高, 虚拟ip在此节点挂载。手动执行自杀脚本。可以看到keepalived服务被关闭, mysql服务未关闭(因为自杀脚本关掉的是keepalived服务), 查看ip挂载发现虚拟ip消失。

```
bash: ./mysql_down.sh: Command not found
[root@408a38fe6011 keepalived]# ./mysql_down.sh
Stopping keepalived (via systemctl): [ OK ]
[root@408a38fe6011 keepalived]# service mysql status
SUCCESS! MySQL running (441)
[root@408a38fe6011 keepalived]# service keepalived status
Usage: /etc/init.d/keepalived {start|stop|reload|restart|condrestart|status}
[root@408a38fe6011 keepalived]# service keepalived status
● keepalived.service - LVS and VRRP High Availability Monitor
   Loaded: loaded (/usr/lib/systemd/system/keepalived.service; enabled; vendor preset: disabled)
   Active: inactive (dead) since Thu 2020-08-20 15:54:53 CST; 35s ago
     Process: 5386 ExecStart=/usr/local/sbin/keepalived $KEEPALIVED_OPTIONS (code=exited, status=0/SUCCESS)
    Main PID: 5387 (code=exited, status=0/SUCCESS)

Aug 20 14:31:44 408a38fe6011 Keepalived_vrrp[5389]: Sending gratuitous ARP on eth0 for 172.30.0.88
Aug 20 14:31:44 408a38fe6011 Keepalived_vrrp[5389]: Sending gratuitous ARP on eth0 for 172.30.0.88
Aug 20 14:31:44 408a38fe6011 Keepalived_vrrp[5389]: Sending gratuitous ARP on eth0 for 172.30.0.88
Aug 20 15:54:52 408a38fe6011 Keepalived[5387]: Stopping
Aug 20 15:54:52 408a38fe6011 Keepalived_vrrp[5389]: VRRP_Instance(VI_1) sent 0 priority
Aug 20 15:54:52 408a38fe6011 Keepalived_vrrp[5389]: VRRP_Instance(VI_1) removing protocol VIPs.
Aug 20 15:54:52 408a38fe6011 Keepalived_vrrp[5389]: VRRP_Instance(VI_1) removing protocol iptable drop rule
Aug 20 15:54:52 408a38fe6011 systemd[1]: Stopping LVS and VRRP High Availability Monitor...
Aug 20 15:54:53 408a38fe6011 Keepalived[5387]: Stopped Keepalived v1.4.2 (02/24,2018), git commit v1.4.1-41-g6a2987e-
Aug 20 15:54:53 408a38fe6011 systemd[1]: Stopped LVS and VRRP High Availability Monitor.
[root@408a38fe6011 keepalived]# ip addr | grep eth0
17: eth0@if18: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state UP group default
    inet 172.30.0.3/16 brd 172.30.255.255 scope global eth0
[root@408a38fe6011 keepalived]#
```

到172.30.0.2节点查看ip挂载, 发现虚拟ip挂载到了此节点上。

```

valid_crt forever preferred_crt forever
[root@d0c4f7dac21b log]# ip addr | grep eth0
15: eth0@if16: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state UP group default
    inet 172.30.0.2/16 brd 172.30.255.255 scope global eth0
    inet 172.30.0.88/32 scope global eth0
[root@d0c4f7dac21b log]#

```

假设故障解决，重启0.3节点的keepalived服务，再次查看ip，虚拟ip重新回到172.30.0.3节点。

```

[root@408a38fe6011 keepalived]# service keepalived restart
Restarting keepalived (via systemctl): [ OK ]
[root@408a38fe6011 keepalived]# ip addr | grep eth0
17: eth0@if18: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state UP group default
    inet 172.30.0.3/16 brd 172.30.255.255 scope global eth0
    inet 172.30.0.88/32 scope global eth0
[root@408a38fe6011 keepalived]#

```

再模拟mysql服务挂掉的情况:关掉0.3节点的mysql服务，发现keepalived也被关闭（自杀脚本），ip挂载再次消失。到0.2节点可以看到虚拟ip重新挂载。

```

[root@408a38fe6011 keepalived]# service mysql stop
Shutting down MySQL..... SUCCESS!
[root@408a38fe6011 keepalived]# service keepalived status
● keepalived.service - LVS and VRRP High Availability Monitor
   Loaded: loaded (/usr/lib/systemd/system/keepalived.service; enabled; vendor preset: disabled)
   Active: inactive (dead) since Thu 2020-08-20 16:01:47 CST; 12s ago
     Process: 5559 ExecStart=/usr/local/sbin/keepalived $KEEPALIVED_OPTIONS (code=exited, status=0/SUCCESS)
    Main PID: 5560 (code=exited, status=0/SUCCESS)
   CGroup: /docker/408a38fe601169e572cc5e0375ad3d56b547400f4637a69205b0448a3bbbf98/system.slice/keepalived.service

Aug 20 16:01:46 408a38fe6011 Keepalived_healthcheckers[5561]: Removing service [172.30.0.3]:tcp:3306 to VS [172.30.0.88]:tcp:3306
Aug 20 16:01:46 408a38fe6011 Keepalived_healthcheckers[5561]: Executing [/etc/keepalived/mysql_down.sh] for service [172.30.0.3]:tcp:3306 in VS [172.30.0.88]:tcp:3306
Aug 20 16:01:46 408a38fe6011 Keepalived_healthcheckers[5561]: Lost quorum 1-0-1 > 0 for VS [172.30.0.88]:tcp:3306
Aug 20 16:01:46 408a38fe6011 Keepalived[5560]: SMTP connection ERROR to [127.0.0.1]:25.
Aug 20 16:01:46 408a38fe6011 Keepalived[5560]: Stopping
Aug 20 16:01:46 408a38fe6011 systemd[1]: Stopping LVS and VRRP High Availability Monitor...
Aug 20 16:01:46 408a38fe6011 Keepalived_vrrp[5562]: VRRP_Instance(VI_1) sent 0 priority
Aug 20 16:01:46 408a38fe6011 Keepalived_vrrp[5562]: VRRP_Instance(VI_1) removing protocol VIPs.
Aug 20 16:01:46 408a38fe6011 Keepalived_vrrp[5562]: VRRP_Instance(VI_1) removing protocol iptable drop rule
Aug 20 16:01:47 408a38fe6011 systemd[1]: Stopped LVS and VRRP High Availability Monitor.
Hint: Some lines were ellipsized, use -l to show in full.
[root@408a38fe6011 keepalived]# ip addr | grep eth0
17: eth0@if18: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state UP group default
    inet 172.30.0.3/16 brd 172.30.255.255 scope global eth0
[root@408a38fe6011 keepalived]#

```

至此，keepalived配置完成，且可以正常运行。

四、问题

1. 不知道是否是docker容器的原因，由于主机和容器无法通信（之前的通信是基于暴露容器端口进行的，这里虚拟ip无法按照之前的方法形成端口映射，也可能是暂时没找到方法）。此虚拟ip只能在容器内访问，所以这里的虚拟ip实用性很小，但熟悉了整个配置过程，如果在真正的服务器上配置的话，应该不存在这个问题。
2. 前边也提到了，有很多配置项没有尝试，包括邮件发送、心跳检测、防脑裂。这几个都是比较重要的，有兴趣可以再尝试一下。