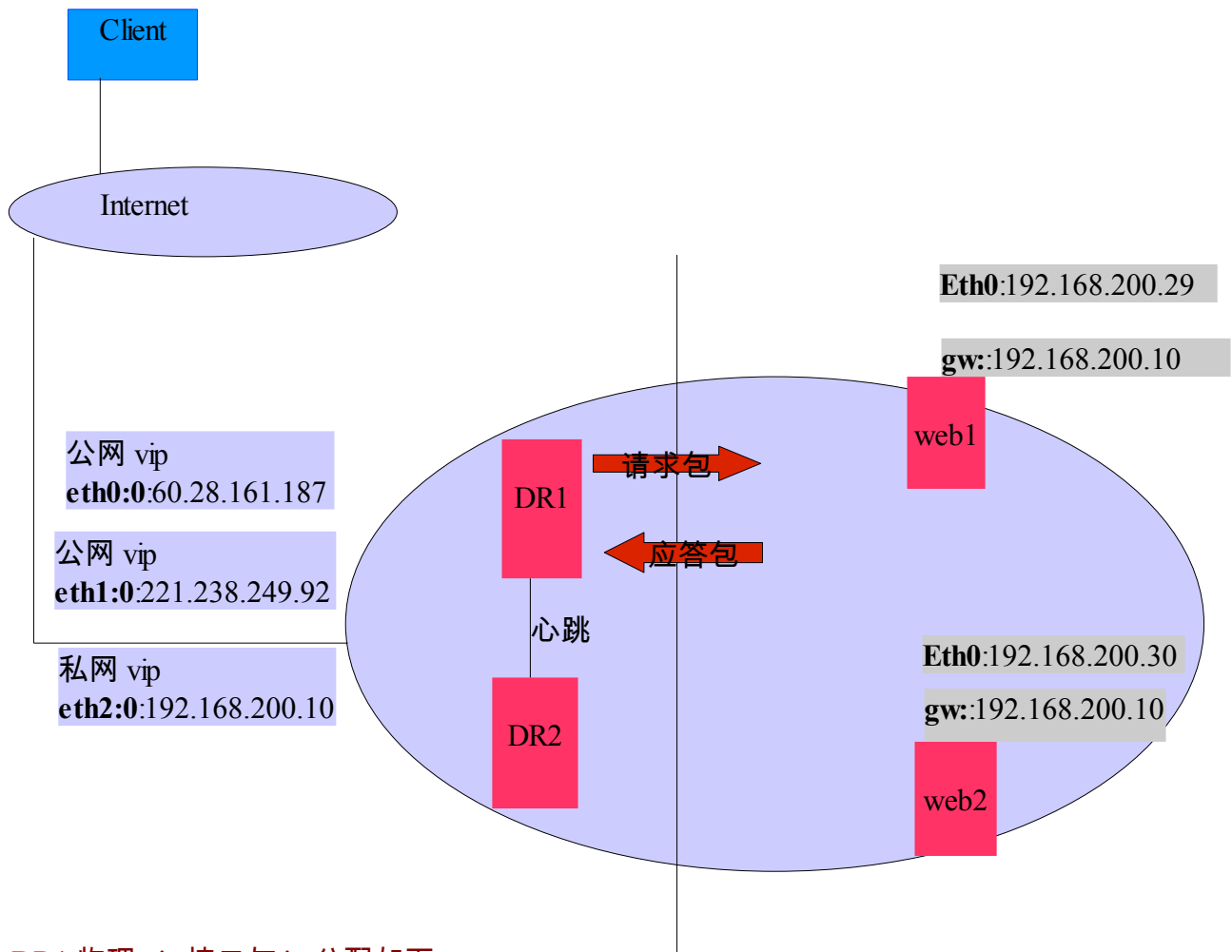


一. lvs-nat HA 结构图原理

图一



DR1 物理 nic 接口与 ip 分配如下:

Eth0:60.28.161.188 netmask:255.255.255.192 gw60.28.161.129 连接网通

Eth1:221.238.249.93 netmask:255.255.255.224 gw221.238.249.65 连接电信

Eth2:192.168.200.3/24 gw:为空

Eth3:192.168.200.13/24 gw:为空, 用于心跳检测

DR2物理nic接口与ip分配如下:

Eth0:60.28.161.189 netmask:255.255.255.192 gw60.28.161.129 连接网通

Eth1:221.238.249.94 netmask:255.255.255.224 gw221.238.249.65 连接电信

Eth2:192.168.200.4/24 gw:为空

Eth3:192.168.200.14/24 gw:为空, 用于心跳检测

web1 eth0:192.168.200.29/24 gw: 192.168.200.10

web2 eth0:192.168.200.30/24 gw: 192.168.200.10

二. 目的:

1. 实现 Server 集群服务的高可用性, 做到 Director 零故障转移, 同时对 realserver 的实时监管. 配置成 lvs-Nat+heartbeat+ldirectord 架构

三. 测试环境 与 架构要求

1. host 环境 Linux 2.6.9-55.ELsmp i686 i686 i386 GNU/Linux 双网卡为:eth0,eth1
2. vmwareServer1.0.4 上新建四台 guest 各 guest 机器名分别为:
 - DR1(主 director)4块网卡 eth0 连网通 eth1 连电信 eth2 私网 eth3 私网用于心跳
 - DR2(备 director)4块网卡 eth0 连网通 eth1 连电信 eth2 私网 eth3 私网用于心跳
 - web1 单网卡 eth0 接私网
 - web2 单网卡 eth0 接私网

四. DR1(主 Director)配置

4.1 配置 DR1 公网 ip 私网 ip 及心跳接口 ip

注: 浮动 IP 即: vip 是不写入系统网口配置文件的, 它一般由 HA 软件来控制

eth0 配置

```
[[root@DR1 network-scripts]# cat ifcfg-eth0
BOOTPROTO=none
GATEWAY=60.28.161.129
HWADDR=00:0C:29:FA:32:D5
DEVICE=eth0
MTU=""
NETMASK=255.255.255.192
BROADCAST=""
IPADDR=60.28.161.188
NETWORK=60.28.161.128
ONBOOT=yes
```

eth1 配置

```
[root@DR1 network-scripts]# cat ifcfg-eth1
BOOTPROTO=none
GATEWAY=221.238.249.65
HWADDR=00:0C:29:FA:32:DF
DEVICE=eth1
MTU=""
NETMASK=255.255.255.224
BROADCAST=""
IPADDR=221.238.249.93
ONBOOT=yes
```

eth2 配置

```
[root@DR1 network-scripts]# cat ifcfg-eth2
# Advanced Micro Devices [AMD] 79c970 [PCnet32 LANCE]
DEVICE=eth2
ONBOOT=yes
BOOTPROTO=static
```

```
IPADDR=192.168.200.3
NETMASK=255.255.255.0
```

eth3 配置

```
[root@DR1 network-scripts]# cat ifcfg-eth3
# Advanced Micro Devices [AMD] 79c970 [PCnet32 LANCE]
DEVICE=eth3
ONBOOT=yes
BOOTPROTO=static
IPADDR=192.168.200.13
NETMASK=255.255.255.0
```

4.2 设置系统转发相关参数

```
vi /etc/sysctl.conf
net.ipv4.ip_forward = 1
再运行命令 sysctl -p
#turn OFF icmp redirects(1 on , 0 off)
    echo 0 > /proc/sys/net/ipv4/conf/all/send_redirects
    echo 0 > /proc/sys/net/ipv4/conf/default/send_redirects
    echo 0 > /proc/sys/net/ipv4/conf/eth0/send_redirects
```

将 dr1 dr2 的机器名与 ip 对应起来分别写入各自的/etc/hosts 文件

```
127.0.0.1          DR1 localhost.localdomain localhost
60.28.161.188      DR1
221.238.249.93     DR1
60.28.161.189      DR2
221.238.249.94     DR2
```

4.3 配置 iptables 做成 nat 形式对包处理

```
/sbin/modprobe ip_tables
/sbin/modprobe iptable_nat
/sbin/modprobe ip_conntrack
/sbin/modprobe iptable_filter
/sbin/modprobe iptable_mangle
/sbin/modprobe ipt_LOG
/sbin/modprobe ipt_limit
/sbin/modprobe ipt_state
iptables -t nat -F
iptables -t nat -X
iptables -t nat -A POSTROUTING -s 192.168.200.0/255.255.255.0 -o eth0 -j MASQUERADE
iptables -t nat -A POSTROUTING -s 192.168.200.0/255.255.255.0 -o eth1 -j MASQUERADE
mkdir /etc/joyconfig
iptables-save>/etc/joyconfig/myiptables
```

4.4 所需软件列表用 yum 安装:

1. yum -y install ipvsadm //用于 lvs-nat 架构
2. yum -y install libnet
3. yum -y install heartbeat-* //用于 director HA, 可以得 heartbeat 所需的工具包:

```
ipvsadm-1.24-6
heartbeat-stonith-2.1.2-3.el4.centos
heartbeat-gui-2.1.2-3.el4.centos
```

heartbeat-pils-2.1.2-3.el4.centos

heartbeat-ldirectord-2.1.2-3.el4.centos //用于监控 realServer 的运行状况

heartbeat-2.1.2-3.el4.centos

4.5 配置 heartbeat, 用到/etc/ha.d/目录下的 4 个配置文件

(1)authkeys //认证文件, 里面有不同的效验算法让你选, 这个文件的权限必须是 600

(2)ha.cf //heartbeat 的主要配置文件, 可以对 heartbeat 的多数性能和状态进行配置。大部分选项的取值可以采用默认

(3)haresources //此文件是节点调用的资源文件如可 heartbeat 去监管/etc/init.d/或 /etc/ha.d/resource.d 目录下的脚本

(4)ldirectord.cf //监控 RealServer 的运行状况

4.6 将相关配置文件复制到/etc/ha.d 目录下

cp /usr/share/doc/heartbeat-2.1.2/ha.cf /etc/ha.d/ha.cf

cp /usr/share/doc/heartbeat-2.1.2/authkeys /etc/ha.d/authkeys

cp /usr/share/doc/heartbeat-2.1.2/haresources /etc/ha.d/haresources

cp /usr/share/doc/heartbeat-ldirectord-2.1.2/ldirectord.cf /etc/ha.d/ldirectord.cf

4.7 开始配置 authkeys 文件

chmod 600 /etc/ha.d/authkeys

vi /etc/ha.d/authkeys 内容如下:

auth 1 //打开 auth 并选择一个用于主从节点的效验算法

1 crc //本例中使用 crc 算法

#2 sha1 HI!

#3 md5 Hello!

4.8 配置 ha.cf 文件

[root@DR1 ha.d]# cat ha.cf

debugfile /var/log/ha-debug

#crm yes

logfile /var/log/ha-log

logfacility local0

keepalive 2

deadtime 20

warntime 10

initdead 50

udpport 694

ucast eth3 192.168.200.14

#auto_failback on

#watchdog /dev/watchdog

node DR1

node DR2

#ping_group group1 60.28.161.129 221.238.249.65

ping 60.28.161.129

ping 221.238.249.65

#apiauth ipfail gid=root uid=root

#use_logd yes/no

#respawn root /usr/lib/heartbeat/ipfail

如果是 os 是 64 位的话下面这样写, 这里都注销采用默认值

#respawn root /usr/lib64/heartbeat/ipfail

uuidfrom nodename //uuid 的生成的方法之一, 当你用 clone DR1 方法生成 DR2 时可解决 uuid 重复的问题

4.9 配置 haresources 文件

```
[root@DR1 ha.d]# cat haresources
#DR1 60.28.161.187/26/eth0:0 221.238.249.92/27/eth1:0 ipvsadm ldirectord::ldirectord.cf
DR1 IPAddr2::60.28.161.187/26/eth0:0 IPAddr2::221.238.249.92/27/eth1:0 ldirectord::ldirectord.cf
DR1 IPAddr2::192.168.200.10/24/eth2:0
```

注:

ldirectord 监管 realserver, 在 haresources 文件中是以 **ldirectord::配置文件**, 的形式出现, 所以::符号后跟的配置文件名是可以任意的, 只要它在/etc/ha.d/目录下存在!

4.10 配置 directord.cf 文件

```
[root@DR1 ha.d]# cat ldirectord.cf
checktimeout=3
checkinterval=2
autoreload=yes
quiescent=yes
logfile="/var/log/ldirectord.log"
virtual=60.28.161.187:80
    real=192.168.200.29:80 masq 100
    real=192.168.200.30:80 masq 100
    fallback=127.0.0.1:80
    request="index.html"
    receive="TestPage"
    protocol=tcp
    checktype=negotiate
    checkport=80
    scheduler=wrr
virtual=221.238.249.92:80
    real=192.168.200.29:80 masq 100
    real=192.168.200.30:80 masq 100
    fallback=127.0.0.1:80
    request="index.html"
    receive="TestPage"
    protocol=tcp
    checktype=negotiate
    checkport=80
    scheduler=wrr
```

关于 heartbeat 配置完成!

4.11 配置 ipvsadm

```
[root@DR1 ha.d]# cat /etc/sysconfig/ipvsadm
ipvsadm -A -t 60.28.161.187:80 -s wrr
ipvsadm -a -t 60.28.161.187:80 -r 192.168.200.29:80 -m -w 100
ipvsadm -a -t 60.28.161.187:80 -r 192.168.200.30:80 -m -w 100
ipvsadm -A -t 221.238.249.92:80 -s wrr
ipvsadm -a -t 221.238.249.92:80 -r 192.168.200.29:80 -m -w 100
ipvsadm -a -t 221.238.249.92:80 -r 192.168.200.30:80 -m -w 100
```

4.12#为了能在系统启动时, 相关服务也启动起来, 我们将命令写入/etc/rc.local 文件

```
[root@DR1 ha.d]# vi /etc/rc.local
echo 0 > /proc/sys/net/ipv4/conf/all/send_redirects
echo 0 > /proc/sys/net/ipv4/conf/default/send_redirects
echo 0 > /proc/sys/net/ipv4/conf/eth0/send_redirects
iptables-restore</etc/joyconfig/myiptables
/etc/init.d/heartbeat start
到此主 lvs-nat + heartbeat+ldirectord 配置完成
```

五 DR2(备用 director)配置:

注: DR2 除了 ip 和将 **ha.cf** 文件中的 **ucast eth3 192.168.200.14** 改成 **ucast eth3 192.168.200.1 3** 之外其它配置要与 DR1 一样。

注: 在对外提供多个 vip 的 slb-HA 环境中, 建意在 ha.cf 文件中用指命 **ucast** 不用 **bcast**

5. DR2 公网 ip 私网 ip 及心跳接口 ip

eth0 配置

```
[[root@DR2 network-scripts]# cat ifcfg-eth0
BOOTPROTO=none
GATEWAY=60.28.161.129
DEVICE=eth0
MTU=""
NETMASK=255.255.255.192
BROADCAST=""
IPADDR=60.28.161.189
NETWORK=60.28.161.128
ONBOOT=yes
```

eth1 配置

```
[root@DR2 network-scripts]# cat ifcfg-eth1
BOOTPROTO=none
DEVICE=eth1
MTU=""
BROADCAST=""
IPADDR=221.238.249.94
NETMASK=255.255.255.224
GATEWAY=221.238.249.65
ONBOOT=yes
```

eth2 配置

```
[root@DR2 network-scripts]# cat ifcfg-eth2
# Advanced Micro Devices [AMD] 79c970 [PCnet32 LANCE]
DEVICE=eth2
ONBOOT=yes
BOOTPROTO=static
IPADDR=192.168.200.4
NETMASK=255.255.255.0
```

eth3 配置

```
[root@DR2 network-scripts]# cat ifcfg-eth3
# Advanced Micro Devices [AMD] 79c970 [PCnet32 LANCE]
DEVICE=eth3
ONBOOT=yes
BOOTPROTO=static
IPADDR=192.168.200.14
NETMASK=255.255.255.0
```

六. 两个 RealServer Web1, Web2 的配置

web1 和 web2 都采用 rpm 包安装版为 httpd-2.0.52 在两台 webserver 上都运行 `yum -y install httpd` 命令来安装 apache

6.1 配置 web1 ip 和禁止 ip 转发

配置 eth0

```
[root@web1 network-scripts]# cat ifcfg-eth0
# Advanced Micro Devices [AMD] 79c970 [PCnet32 LANCE]
DEVICE=eth0
ONBOOT=yes
BOOTPROTO=static
IPADDR=192.168.200.29
NETMASK=255.255.255.0
GATEWAY=192.168.200.10
```

```
echo "0" > /proc/sys/net/ipv4/ip_forward //禁止 ip 转发
```

在 Real Server1 即 Web1 中添加监控页:

```
echo "TestPage111111" >> /var/www/html/index.html
```

6.2 配置 web2 ip 和禁止转发, web2 的设定与 web1 配置方法相同!

配置 eth0

```
[root@web2 network-scripts]# cat ifcfg-eth0
# Advanced Micro Devices [AMD] 79c970 [PCnet32 LANCE]
DEVICE=eth0
ONBOOT=yes
BOOTPROTO=static
IPADDR=192.168.200.30
NETMASK=255.255.255.0
GATEWAY=192.168.200.10
```

```
echo "0" > /proc/sys/net/ipv4/ip_forward //禁止 ip 转发
```

在 Real Server2 即 Web2 中添加监控页:

```
echo "TestPage22222" >> /var/www/html/index.html
```

七. 测试步骤:

1. 首先测试 lvs 能正常工作, 关闭 heartbeat ldirectord 及其它与 lvs 不相关的服务.

2. heartbeat 的 HA 功能:

<2.1>在主 DR1 和 DR2 开启在正常工作的情况下, 将 DR1 机器关闭或将 heartbeat

服务关闭, 关闭用 如下命令/etc/init.d/heartbeat stop 或 service heartbeat stop

<2.2>去 DR2 上用 ifconfig 看一下应该看到 eth0:0 eth1:0 eth2:0 三个浮动接口自动启

用了, 而 DR1 上的 eth0:0 eth1:0 eth2:0 三个浮动接口被自动关闭了, 说明 DR2 成功将 DR1 的工作接管过来了

3. 再将 DR1 开启 heartbeat 服务, 这时 DR1 上的而 DR1 上的 eth0:0 eth1:0 eth2:0 三个浮动接口自动起来了, 而 DR2 上 eth0:0 eth1:0 eth2:0 三个浮动接口也自动关闭了, 说明 DR1 故障恢复之后成功将之前 DR2 接管的工作又重新接任下来了。

4. 以上测试通过, 从而达到了 HA 真正的效果

八. 测试 ldirectord 监管 RealServer 效果的步骤:

1. 在 DR1 DR2 web1 web2 都正常运作的环境中, 先将 web1 或 web2 的 http 服务关闭, 然到调度 lvs 的 director 机器上观察 lvs 的调度情况, 可用命令:

```
watch -n 1 ipvsadm -lnc 动态查看
```

当然也可以在 client 端去访问 vip，看此时访问的结果，可用以下方法：

<http://60.28.161.187>

<http://221.238.249.92>