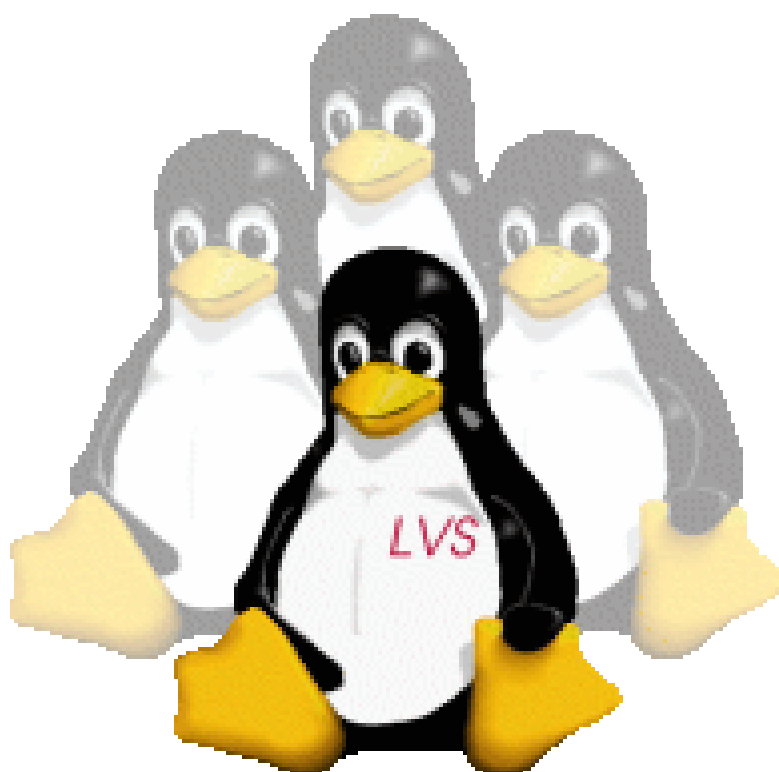


LVS 在 Fedora Core5 上的配置指南



目 录

1 简介	3
2 实现 VS/NAT	5
2.1 网络拓扑图	5
2.2 Director 的配置	5
2.3 Realserver 的配置	6
2.3.1 Linux Realserver 的配置	6
2.3.2 FreeBSD Realserver 的配置	6
2.3.3 Solaris x86 Realserver 的配置	6
3 实现 VS/DR	8
3.1 网络拓扑图	8
3.2 Director 的配置	9
3.3 Realserver 的配置	9
3.3.1 Linux Realserver 的配置	9
3.3.2 FreeBSD Realserver 的配置	10
3.3.3 Solaris x86 Realserver 的配置	11
4 后记	11

§1 简介

[Linux 服务器集群系统](#)(以下简称 LVS), 是一个由章文嵩博士发起的自由软件项目, 现在已经是 Linux 标准内核的一部分。

LVS 从 Linux 内核还在 2.0 版本的时代开始, 就得到广泛的应用, 目前 Linux 的内核已经是 2.6 版, 完全内置了 LVS 的各个模块, 已经不需要在给内核打任何补丁。

Fedora 项目是由 [Redhat 公司](#) 赞助的开源 Linux 发行版本, 由于 Fedora 系列界面简单易用, 升级方便而且免费, 得到笔者和很多 Linux 爱好者的追捧。

Fedora Core5(代号 Bordeaux, 内核为 2.6.15-1, 以下简称 FC5) 是 Fedora 最新的稳定版本, 界面豪华美观, 中文支持极佳。是笔者平常的办公学习用系统。

本文 FC5 以作为负载均衡服务器 (以下简称 Director), 分别以 FC5、Sun Solaris10 x86、FreeBSD5.4 作为真实服务器 (以下简称 RealServer), 实现了 LVS 中的两种负载均衡方式 VS/NAT、VS/DR。调度算法采用了比较简单的加权轮转算法。

笔者在 VS/TUN 方式下得出了新的结论, 但正在完善之中, 本文暂不使用 VS/TUN 方式。对于地理分布的系统而言, 这几乎是唯一的方法。

三种 IP 负载均衡技术的优缺点比较:

杂项	VS/NAT	VS/TUN	VS/DR
服务器 (OS)	任意	支持隧道	多数 (支持 Non-arp)
服务器网络	私有网络	局域网/广域网	局域网
服务器数目 (100M 网络)	10-20	100	多 (100)
服务器网关	负载均衡器	自己的路由	自己的路由
效率	一般	高	最高

如何配置 Director 和 RealServer 是本文关注的焦点, 读者可以将本文当作一篇 Howto 文档, 文中给出了实现的详细方法、配置、和步骤。本文对 LVS 的原理阐述的非常少, 如有需要, 请参考章文嵩先生的博士论文和相关文章。

性能也不是本文的重点, 事实上文中的很多机器都是在 VMware 上运行的。

本文中的所有例子, 都是采用的 Apache 服务, 这也是 LVS 经常应用的场合。三种操作系统的 ip 的最后一位, Linux 总是 160, FreeBSD 总是 170, Solaris 总是 180。测试的手段是在客户机 (Linux) 下运行 links 或 elinks 命令, 用法如下:

`links -source http://VIP`

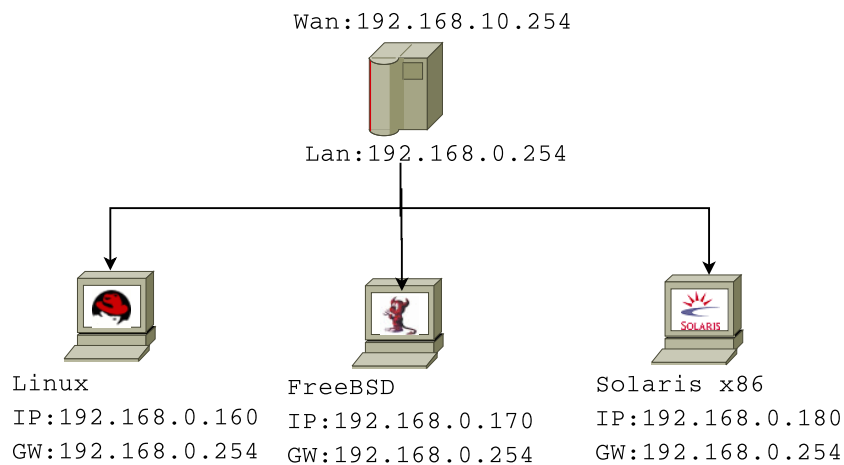
这个手段非常简单，而且快速，笔者实验系统上的每个 RealServer 上 Apache 服务器上的缺省首页都不同，其中 Linux 上会显示“This is 160”，连续执行此命令，会明显的看到 ipvs 在起作用。

本文原稿是用 vim 编辑的 T_EX 文件，使用了 CJK 宏包。文中的拓扑图用 Gnome Desktop Project 旗下的 dia 软件软件绘制。用多种 L^AT_EX 工具生成 dvi、pdf 和 html 版本的文件。笔者相信，即使在排版领域，自由软件也是最出色的。

§2 实现 VS/NAT

VS/NAT 是一种最简单的方式，所有的 RealServer 只需要将自己的网关指向 Director 即可。客户端可以是任意操作系统，但此方式下，一个 Director 能够带动的 RealServer 比较有限。在 VS/NAT 的方式下，Director 也可以兼为一台 RealServer。

§2.1 网络拓扑图



§2.2 Director 的配置

Director 的配置如下：外部地址为 192.168.10.254 内部地址为 192.168.0.254
LVS 在 3 种方式下均需要打开 ip_forward 功能，这可以配置/etc/sysctl.conf 文件，保证其中有如下一行：

```
net.ipv4.ip_forward = 1
```

执行：

```
sysctl -p
```

可以使配置立即生效，并且重启电脑或 network 服务后，此设置仍然保存
ipvs 的脚本如下：

Listing 1 Script of VS/NAT

```
#!/bin/bash
ipvsadm -C
ipvsadm -A -t 192.168.10.254:80 -s wlc
ipvsadm -a -t 192.168.10.254:80 -r 192.168.0.160:80 -m
ipvsadm -a -t 192.168.10.254:80 -r 192.168.0.170:80 -m
ipvsadm -a -t 192.168.10.254:80 -r 192.168.0.180:80 -m
ipvsadm -a -t 192.168.10.254:80 -r 192.168.0.254:80 -m
```

```
ipvsadm -Ln
```

§2.3 Realserver 的配置

3 台 RealServer 的配置如下。

§2.3.1 Linux Realserver 的配置

/etc/sysconfig/network 文件的内容如下:

Listing 2 /etc/sysconfig/network

```
NETWORKING=yes  
HOSTNAME=localhost.localdomain  
GATEWAY=192.168.0.254
```

Listing 3 /etc/sysconfig/network-scripts/ifcfg-eth0

```
DEVICE=eth0  
ONBOOT=yes  
BOOTPROTO=static  
IPADDR=192.168.0.160  
NETMASK=255.255.255.0  
BROADCAST=192.168.0.255
```

§2.3.2 FreeBSD Realserver 的配置

Listing 4 /etc/rc.conf

```
sshd_enable="YES"  
ifconfig_lnc0="inet 192.168.0.170 netmask 255.255.255.0"  
defaultrouter="192.168.0.254"  
hostname="localhost.localdomain"  
apache_enable="YES"
```

§2.3.3 Solaris x86 Realserver 的配置

在 Solaris 上配置主机名称、ip 地址、子网以及网关, 要比 Linux 和 FreeBSD 要复杂, 需要配置多个文件。我机器上网卡名为 pcn0, 如果您的网卡名称不同, 就请取代相应的部分。/etc/hosts 和/etc/hostname.pcn0 联合作用的结果是配置 ip 地址和子网掩码, 在 hostname.pcn0 中填写了子网掩码, /etc/netmasks 文件就可以留空了。/etc/defaultrouter 文件的作用是配置网关。只需要写一个做为网关的 ip 地址即可。/etc/nodename 文件的内容为 hostname。

Listing 5 /etc/hosts

```
127.0.0.1      localhost loghost localhost.localdomain
192.168.0.180 solaris
```

Listing 6 /etc/hostname.pcn0

```
192.168.0.180/24
```

Listing 7 /etc/defaultrouter

```
192.168.0.254
```

Listing 8 /etc/nodename

```
localhost.localdomain
```

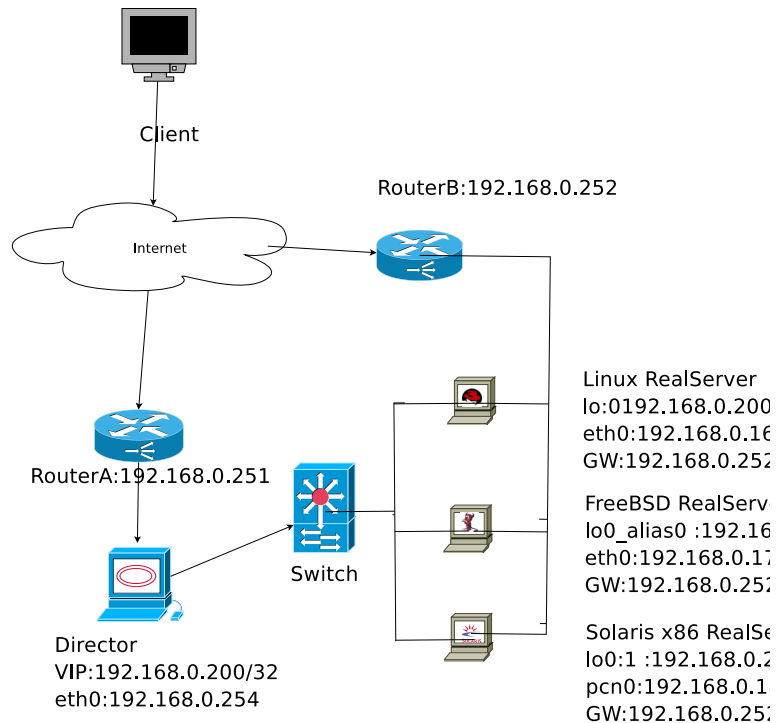
§3 实现 VS/DR

在很多 Internet 服务中, 例如 ftp 和 http, 请求的流量往往远远小于回应的流量, VS/DR 方式利用这种非对称的特点, 只负责调度请求, 而 RealServer 直接将相应的报文返回给客户机。

VS/DR 方式是通过改写请求报文中的 MAC 地址部分, 来实现的 Director 和 RealServer 必需在物理上有一个网卡通过不间断的局域网相连。RealServer 上绑定 VIP 配置在各自 Non-ARP 的网络设备上 (如 lo 或 tunl), Director 的 VIP 地址对外可见, 而 RealServer 的 VIP 对外是不可见的。RealServer 的地址即可以是内部地址, 也可以是真实地址。

在实现 VS/DR 的方式中, 我们配置了两台 Router 来模拟这种非对称路由的情况, RouterA 地址为 192.168.0.251, RouterB 地址为 192.168.0.252。当客户机访问 Director 时, 由 RouterA 进入, 当 RealServer 返回给客户机信息时, 由 RouterB 出去, 根据实际情况您也可以只使用一台或使用多台 Router。只要能确保客户发送信息到 Director, 而且 RealServer 可以发送信息到客户即可。

§3.1 网络拓扑图



§3.2 Director 的配置

Director 为单网卡, 其中 eth0 的地址为 192.168.0.254/32, eth0:0 为 192.168.0.200/32。网关为 192.168.0.251。这里的 VIP 地址就是 192.168.0.200, 当客户访问此地址, 根据相应的调度算法, 会访问到不同的 RealServer。

Listing 9 Script of VS/DR

```
#!/bin/bash
ipvsadm -C
#VS/DR
ipvsadm -A -t 192.168.0.200:80 -s wlc
ipvsadm -a -t 192.168.0.200:80 -r 192.168.0.160 -g
ipvsadm -a -t 192.168.0.200:80 -r 192.168.0.170 -g
ipvsadm -a -t 192.168.0.200:80 -r 192.168.0.180 -g
ipvsadm -Ln
```

Listing 10 /etc/sysconfig/network-scripts/ifcfg-eth0:0

```
DEVICE=eth0:0
ONBOOT=yes
BOOTPROTO=static
IPADDR=192.168.0.200
NETMASK=255.255.255.255
BROADCAST=192.168.0.200
```

Listing 11 /etc/sysconfig/network

```
NETWORKING=yes
HOSTNAME=localhost.localdomain
GATEWAY=192.168.0.251
```

§3.3 Realserver 的配置

3 台 RealServer 的网关均指向另一路由器, 地址为 192.168.0.252。

§3.3.1 Linux Realserver 的配置

Linux Realserver 配置的关键之一在于关闭 arp, 这要配置/etc/sysctl.conf 文件。

Listing 12 /etc/sysctl.conf

```
...
net.ipv4.ip_forward = 1
```

```
#for ipvs
net.ipv4.conf.lo.arp_ignore = 1
net.ipv4.conf.lo.arp_announce = 2
net.ipv4.conf.all.arp_ignore = 1
net.ipv4.conf.all.arp_announce = 2
net.ipv4.conf.default.rp_filter = 1
#end of ipvs
..
```

在/etc/sysconfig/network-scripts/ifcfg-eth0 中配置 ip 地址和网关。

Listing 13 /etc/sysconfig/ifcfg-eth0

```
DEVICE=eth0
ONBOOT=yes
BOOTPROTO=static
IPADDR=192.168.0.160
NETMASK=255.255.255.0
BROADCAST=192.168.0.255
GATEWAY=192.168.0.252
```

配置 VIP 的网卡为/etc/sysconfig/network-scripts/ifcfg-lo:0。

Listing 14 /etc/sysconfig/network-scripts/ifcfg-lo:0

```
DEVICE=lo:0
ONBOOT=yes
IPADDR=192.168.0.200
NETMASK=255.255.255.255
```

§3.3.2 FreeBSD Realserver 的配置

FreeBSD 上的配置最为简单,只需要配置一个/etc/rc.conf 文件即可。lnc0 为真实网卡,lo0 的别名上配置的网卡为 VIP。

Listing 15 /etc/rc.conf

```
....
sshd_enable="YES"
ifconfig_lnc0="inet 192.168.0.170 netmask 0xfffff00"
ifconfig_lo0_alias0="inet 192.168.0.200 netmask 0xffffffff"
defaultrouter="192.168.0.252"
hostname="localhost.localdomain"
```

```
apache_enable="YES"
....
```

§3.3.3 Solaris x86 Realserver 的配置

Solaris 上的网卡 ip 为 192.168.0.180/24。lo0:1 为 VIP 192.168.0.200/32，另外需配置 Solaris 的缺省网关为 192.168.0.252。

Listing 16 /etc/hosts

```
127.0.0.1      localhost localhost.localdomain
192.168.0.200  vip
192.168.0.180  realserver
```

Listing 17 /etc/hostname.pcn0

```
192.168.0.180/24
```

Listing 18 /etc/hostname.lo0:1

```
192.168.0.200/32
```

Listing 19 /etc/nodename

```
localhost.localdomain
```

Listing 20 /etc/defaultrouter

```
192.168.0.252
```

§4 后记

最近突然对 LVS 感兴趣了，根据 www.linuxvirtualserver.org 上的文档，在 Fedora Core5 Linux 上成功配通了 VS/NAT、VS/DR 和 VS/TUN 方式，顺手又用 FreeBSD 5.4 和 Solaris 10 x86 实验了一下，也就完成了本文档。

如果文中如有任何错误，无论是技术错误还是笔误，都欢迎 email 给作者

本文及其插图使用 [GPL 第二版](#)。保留一切权利。