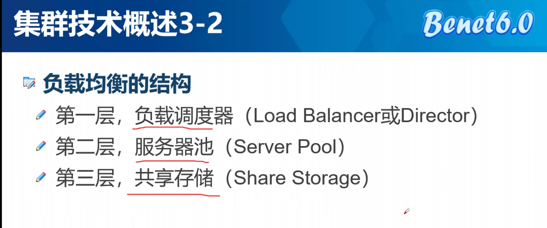
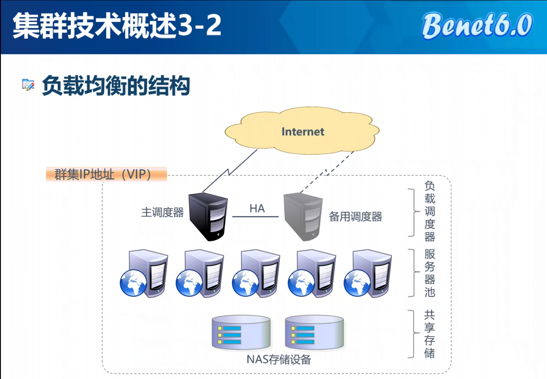
# 集群基础知识点：





**集群类型：LB、HA、HP、DS**

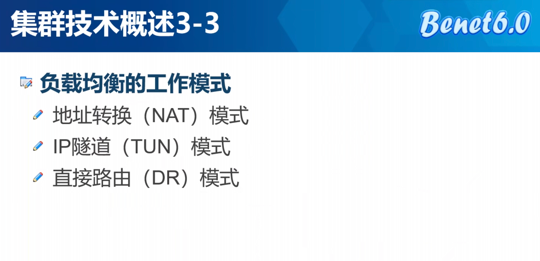
**LB：load balance，负载均衡集群，director：负载均衡器或者调度器，上游服务（upstream server），后端服务器称为real server；**

**HA：high avalilability,高可用集群，评估高可用性能Availability = 平均无故障时间/（平均无故障时间+平均修复时间）**

**HP：high performance 超级计算机类型**

**LB集群的实现：硬件、软件（lvs: Linux Virtual Server、haproxy、nginx、ats (apache traffic server)、perlbal、**

**基于工作的协议层次划分：**





**传输层：(工作于内核)可以突破socket 可用端口的限制一共6W多个；lvs, haproxy (mode tcp)**

**应用层：（用户空间）：haproxy (mode http), nginx, ats, perlbal**

**lvs的实现基于传输层：类型有四种nat、dr(direct routing)、tun、fullnat**

**lvs-nat：多目标的DNAT：通过将请求报文的目标地址和目标端口修改为挑选出某RS的RIP和PORT来实现；**

(1) RIP和DIP应该使用私网地址，RS的网络应该指向DIP；

(2) 请求和响应报文都要经由director转发；极高负载的场景中，Director可能会成为系统瓶颈；

(3) 支持端口映射；

(4) VS必须为Linux，RS可以是任意的OS；

(5) RS的RIP与Director的DIP必须在同一IP网络；

**lvs-dr：direct routing通过修改请求报文的MAC地址进行转发；IP首部不会发生变化（源IP为CIP，目标IP始终为VIP）；**

(1) 确保前端路由器将目标IP为VIP的请求报文一定会发送给Director；

解决方案：

静态绑定；在路由器上静态绑定VIP和director的Mac地址,禁止RS响应VIP的ARP请求；

     (a) arptables；

     (b) 修改各RS的内核参数，并把VIP配置在特定的接口上实现禁止其响应；

(2) RS的RIP可以使用私有地址，也可以使用公网地址；

(3) RS跟Director必须在同一物理网络中；

(4) 请求报文必须由Director调度，但响应报文必须不能经由Director；

(5) 不支持端口映射；

(6) 各RS可以使用大多数的OS；

**lvs-tun：ip tunneling，ip隧道；转发方式：不修改请求报文的IP首部（源IP为CIP，目标IP为VIP），而是在原有的IP首部这外再次封装一个IP首部（源IP为DIP，目标IP为RIP）；**

(1) RIP，DIP，VIP全得是公网地址；

(2) RS的网关不能也不可能指向DIP；

(3) 请求报文经由Director调度，但响应报文将直接发给CIP；

(4) 不支持端口映射；

(5) RS的OS必须支持IP隧道功能；

**lvs-fullnat：（非标准模型）通过同时修改请求报文的源IP地址（cip-->dip）和目标IP地址（vip --> rip）实现转发；**

(1) VIP是公网地址；RIP和DIP是私网地址，且可以不在同一IP网络中，但需要通过路由互相通信；

(2) RS收到的请求报文的源IP为DIP，因此其响应报文将发送给DIP；

(3) 请求报文和响应报文都必须经由director；

(4) 支持端口映射；

(5) RS可使用任意OS；

**LVS相关术语**

**1、 DS：Director Server。指的是前端负载均衡器节点。**

**2、 RS：Real Server。后端真实的工作服务器。**

**3、 VIP：向外部直接面向用户请求，作为用户请求的目标的IP地址。**

**4、 DIP：Director Server IP，主要用于和内部主机通讯的IP地址。**

**5、 RIP：Real Server IP，后端服务器的IP地址。**

**6、 CIP：Client IP，访问客户端的IP地址**

# 1.实验：搭建LVS集群

**LVS Linux virtual server**

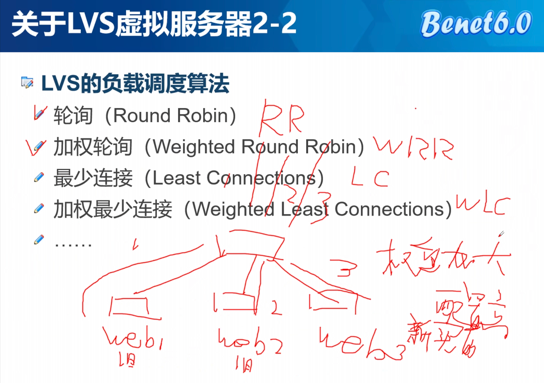
**lb1 192.168.7.10**

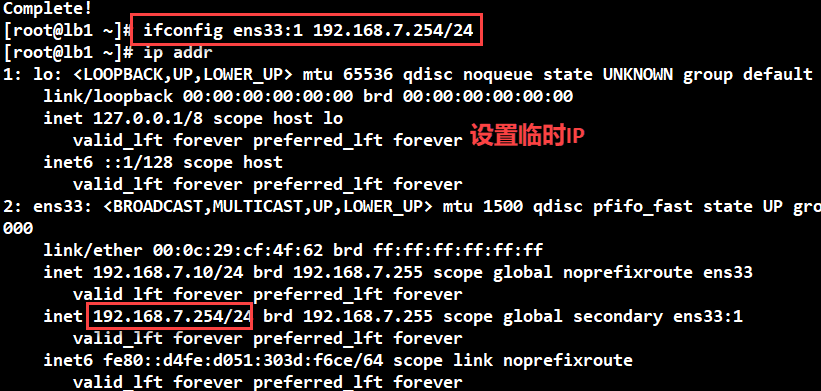
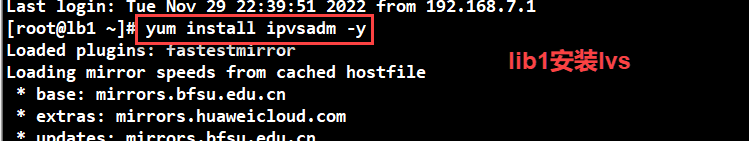
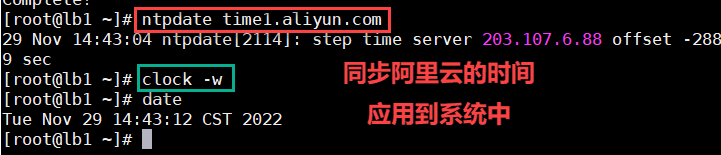
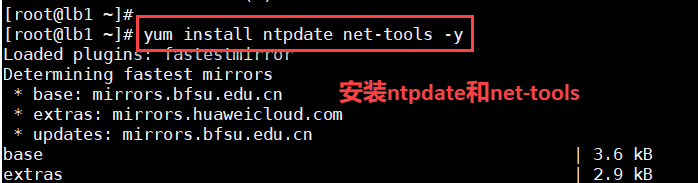
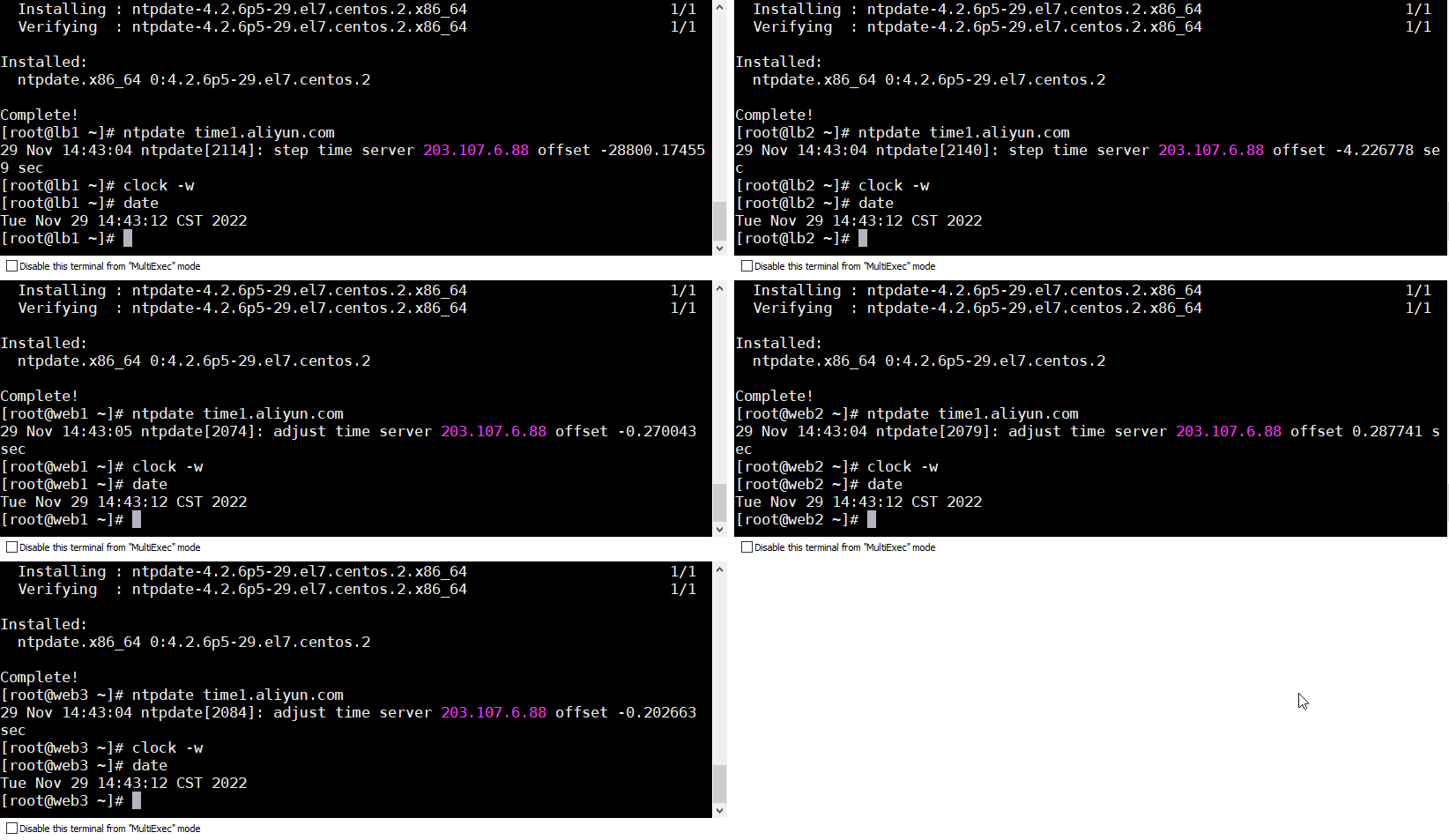
**lb2 192.168.7.11**

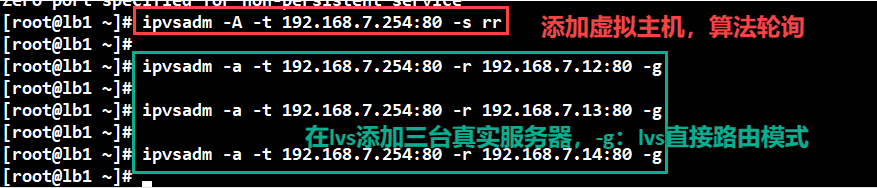
**web1 192.168.7.12**

**web2 192.168.7.13**

**web3 192.168.7.14**



****

****

**-A 添加虚拟服务**

**-a 添加一个真实的主机到虚拟服务**

**-S 保存**

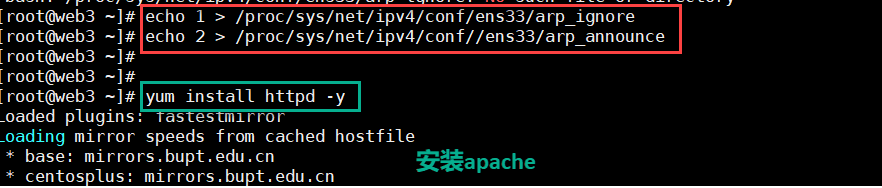
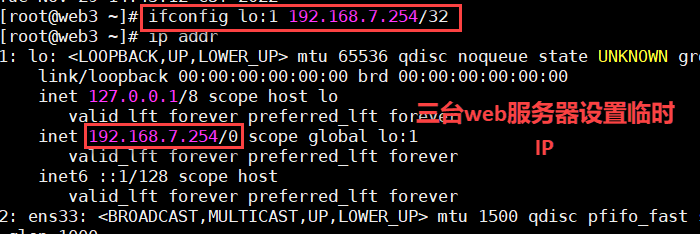
**-s 选择调度方法**

**rr 轮训调度**

**-g --gatewaying 指定LVS 的工作模式为直接路由模式（也是LVS 默认的模式）**

**-m 网络地址转换NAT**

**-t | -u TCP/UDP协议的虚拟服务**

****

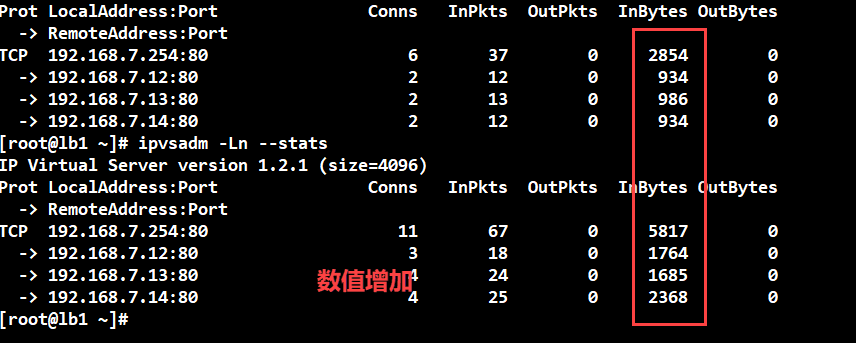
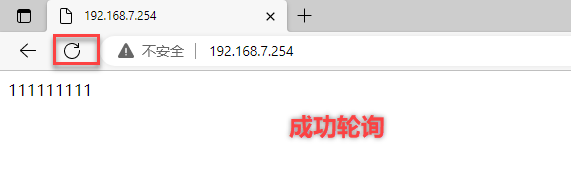
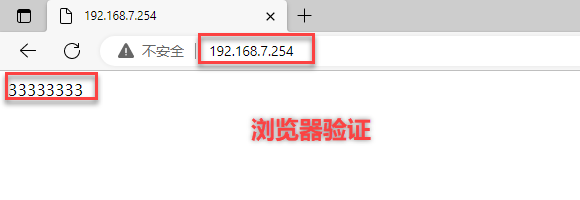
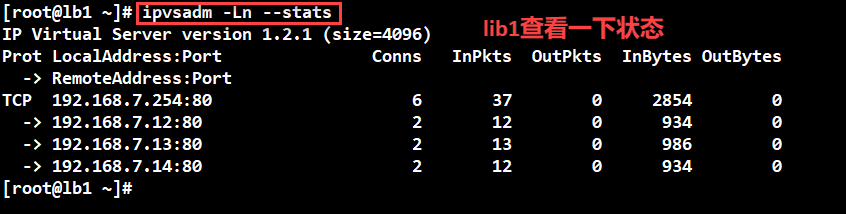
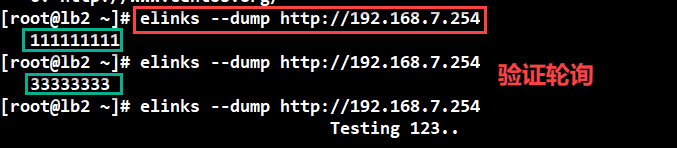
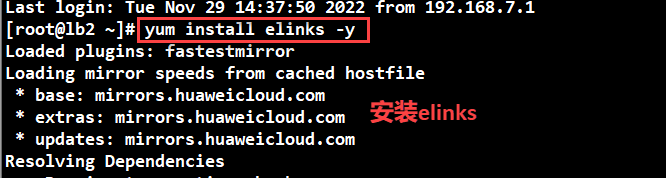
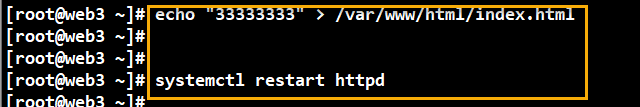
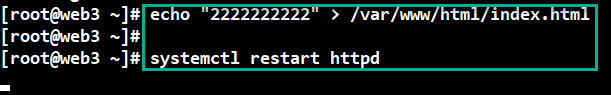
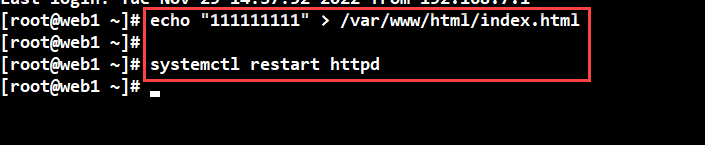
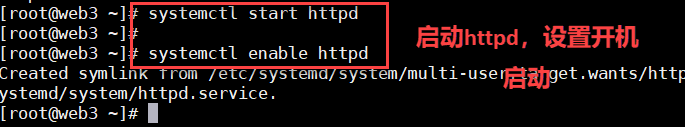
**参数说明：**

**arp\_ignore 为：1：**

**只回答目标 IP 地址是访问本网络接口（ens33）的 ARP 查询请求。 在设置参数的时候将 arp\_ignore 设置为 1，意味着当别人的 arp 请求过来的时候，如果接收的网卡设备上面没有这个 ip，就不做出响应，默认是0，只要这台机器上面任何一个网卡设备上面有这个 ip，就响应arp 请求，并发送mac 地址。**

**arp\_announce（宣告） 为2：**

**对查询目标使用最适当的本地地址。例如，如果在 ens33 接口上接受到了一个 VIP 的 arp 请求包。内核判断这个 VIP 地址是不是 ens33 接口上的IP 一样。如果一样，则回复这个包。如果不一样，就丢弃不回应。**

****

**LVS（ipvsadm）命令参数详解**

**Commands:**

**-A --add-service 在内核的虚拟服务器表中添加一条新的虚拟服务器记录。也就是增加一台新的虚拟服务器。**

**-E --edit-service 编辑内核虚拟服务器表中的一条虚拟服务器记录。**

**-D --delete-service 删除内核虚拟服务器表中的一条虚拟服务器记录。**

**-C --clear 清除内核虚拟服务器表中的所有记录。**

**-R --restore 恢复虚拟服务器规则**

**-S --save 保存虚拟服务器规则，输出为-R 选项可读的格式**

**-L|-l –list 显示内核虚拟服务器表**

**-Z --zero 虚拟服务表计数器清零（清空当前的连接数量等）**

**-a --add-server 在内核虚拟服务器表的一条记录里添加一条新的真实服务器记录。也就是在一个虚拟服务器中增加一台新的真实服务器**

**-e --edit-server 编辑一条虚拟服务器记录中的某条真实服务器记录**

**-d --delete-server 删除一条虚拟服务器记录中的某条真实服务器记录**

**--set tcp tcpfin udp 设置连接超时值**

**--start-daemon 启动同步守护进程。他后面可以是master或backup，用来说明LVS Router 是master 或是backup。在这个功能上也可以采用 keepalived的VRRP 功能。**

**--stop-daemon 停止同步守护进程**

**-h --help 显示帮助信息**

**Options:**

**-t --tcp-service service-address 说明虚拟服务器提供的是tcp的服务[vip:port] or [real-server-ip:port]**

**-u --udp-service service-address 说明虚拟服务器提供的是udp的服务[vip:port] or [real-server-ip:port]**

**-s --scheduler scheduler 使用的调度算法，有这样几个选项rr|wrr|lc|wlc|lblc|lblcr|dh|sh|sed|nq,默认的调度算法是： wlc.**

**-p --persistent [timeout] 持久稳固的服务。这个选项的意思是来自同一个客户的多次请求，将被同一台真实的服务器处理。timeout 的默认值为300 秒。**

**-r --real-server server-address 真实的服务器[Real-Server:port]**

**-g --gatewaying 指定LVS 的工作模式为直接路由模式（也是LVS 默认的模式）**

**-i --ipip 指定LVS 的工作模式为隧道模式**

**-m --masquerading 指定LVS 的工作模式为NAT 模式**

**-w --weight weight 真实服务器的权值**

**--mcast-interface interface 指定组播的同步接口**

**-c --connection 显示LVS 目前的连接 如：ipvsadm -L –c**

**-f --fwmark-service fwmark 说明是经过iptables标记过的服务类型。**

**-M --netmask netmask persistent**

**--timeout 显示tcp tcpfin udp 的timeout 值 如：ipvsadm -L –timeout**

**--daemon 显示同步守护进程状态**

**--stats 显示统计信息**

**--rate 显示速率信息**

**--sort 对虚拟服务器和真实服务器排序输出**

**--numeric -n 输出IP 地址和端口的数字形式**