

# VMware ESX Server 网络设置

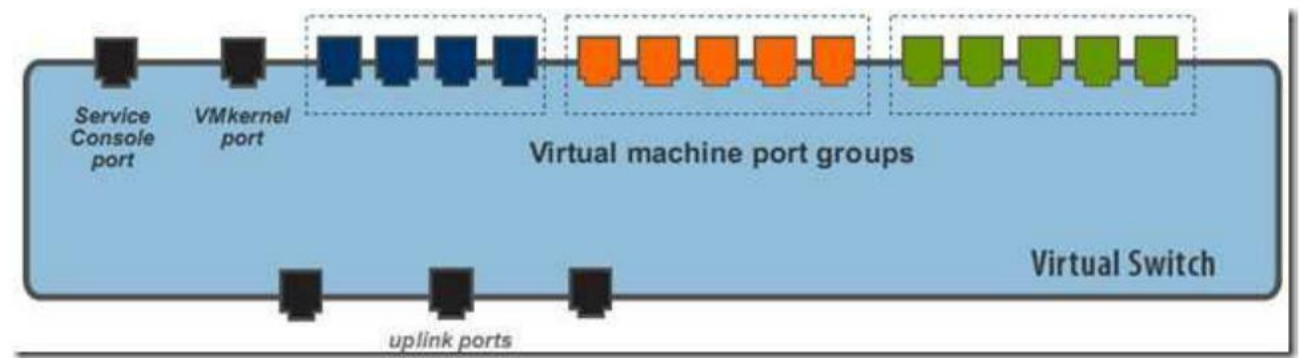
时间: 2011-02-14 分类: [VMware](#), [推荐主题](#)

标签: [ESX](#), [VMware](#), [vSwitch](#) / 4,748 次浏览 0 评论

## 1. [ESX](#) 网络基础

[ESX](#) server 总共有三种形式的网络:

- Service console port: [ESX](#) Server 的管理网络, 如: vSphere Client 和 vCenter 就是使用此网络管理 ESX。
- VMkernel port: ESX Server 内核接口, 用来 VMotion, iSCSI 和 NFS/NAS networks
- Virtual machine port group: 虚拟机网络连接。



**ESX 界面:**



## 2. ESX 的 [vSwitch](#)

VMWare 的官方解释是这样的:

A virtual switch, [vSwitch](#), works much like a physical Ethernet switch. It detects which virtual machines are logically connected to each of its virtual ports and uses that information to forward traffic to the correct virtual machines. A vSwitch can be connected to physical switches using physical Ethernet adapters, also referred to as uplink adapters, to join virtual networks with physical networks. This type of connection is similar to connecting physical switches together to create a larger network. Even though a vSwitch works much like a physical switch, it does not have some of the advanced functionality of a physical switch. For more information on vSwitches, see [Virtual Switches](#).

其实 vSwitch 就是上图左边端口的组合, 默认端口数为 56(最大可以到 1016), 我们可以为一个 ESX server 创建最多 248 个 vSwitch。在 VM 不断增多的情况下, 出于冗余和流量或安全考虑, 可以增加多个 vSwitch, 并设置相应规则。

### 3. 增加 vSwitch

#### 1) 列出系统的网卡

```
[root@VM1 ~]# esxcfg-nics -l
```

Name	PCI Address	Driver	Link	Speed	Duplex	MAC
vmnic0	01:00.00	bnx2	Up	1000Mbps	Full	a4:ba:db:45:c8:b2
1500	Broadcom Corporation	Broadcom NetXtreme II	BCM5709	1000Base-T		
vmnic1	01:00.01	bnx2	Up	1000Mbps	Full	a4:ba:db:45:c8:b4
1500	Broadcom Corporation	Broadcom NetXtreme II	BCM5709	1000Base-T		
vmnic2	02:00.00	bnx2	Down	0Mbps	Half	a4:ba:db:45:c8:b6
1500	Broadcom Corporation	Broadcom NetXtreme II	BCM5709	1000Base-T		
vmnic3	02:00.01	bnx2	Down	0Mbps	Half	a4:ba:db:45:c8:b8
1500	Broadcom Corporation	Broadcom NetXtreme II	BCM5709	1000Base-T		

#### 2) 列出你当前的虚拟交换机

```
[root@VM1 ~]# esxcfg-vswitch -l
```

Switch Name	Num Ports	Used Ports	Configured Ports	MTU	Uplinks
vSwitch0	32	10	32	1500	vmnic0,vmnic1

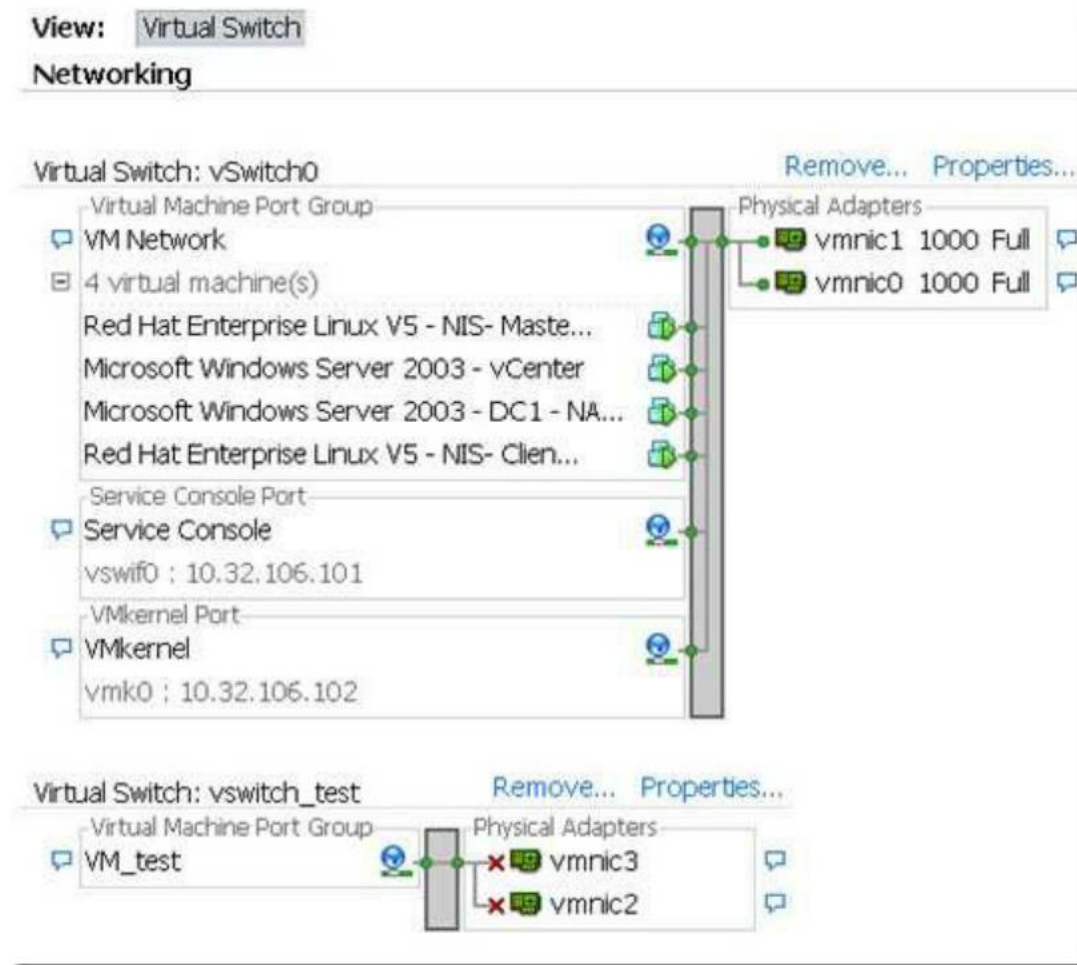
  

PortGroup Name	VLAN ID	Used Ports	Uplinks
VM Network	0	5	vmnic0,vmnic1
Service Console	0	1	vmnic0,vmnic1
VMkernel	0	1	vmnic0,vmnic1

3) 添加一个虚拟交换机，名字叫 (**vswitch\_test**) 连接到两块不用的物理网卡(目前没有插网卡)

```
[root@VM1 ~]# esxcfg-vswitch -a vswitch_test
[root@VM1 ~]# esxcfg-vswitch -A VM_test vswitch_test
[root@VM1 ~]# esxcfg-vswitch -L vmnic2 vswitch_test
[root@VM1 ~]# esxcfg-vswitch -L vmnic3 vswitch_test
```

然后就可以在 GUI 里修改相应配置(修改 Port Group 或其他)。



## 4. vSwitch 流量分隔及冗余切换

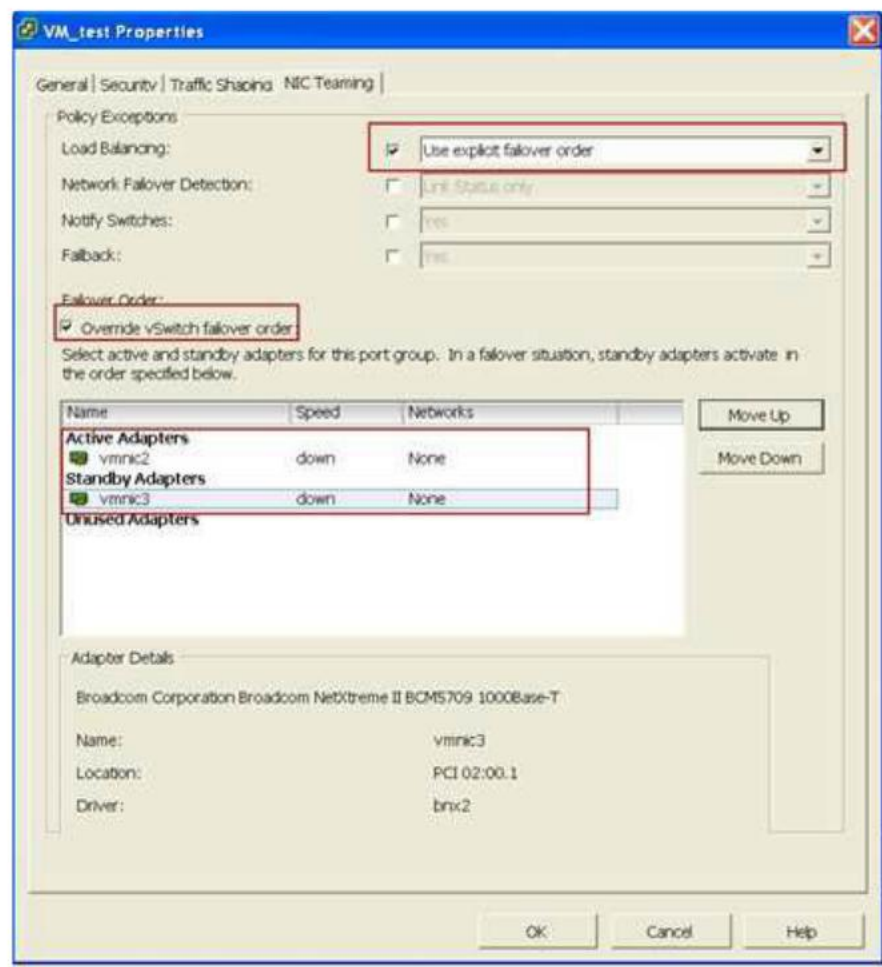
**VMware** 建议 ESX 中的三种网络类型最好配置不同的 Vlan，使用不同的上行链路以分隔流量，并且为了保证故障冗余，每个网络类型至少关联 2 块网卡，我这里用 vSwitch\_test 为例：

配置：

VMotion 使用 vmnic2 作为活动适配器，vmnic3 作为待机适配器，VLAN ID=201；



虚拟机网络 VM\_test 使用 vmnic2 作为活动适配器,vmnic3 作为待机适配器,VLAN ID=200;  
虚拟机网络 VM\_test2 使用 vmnic3 作为活动适配器,vmnic2 作为待机适配器,VLAN ID=202。



## 5. vSwitch 网络负载均衡

### VMware ESX Server 网络负载均衡配置

时间: 2011-02-14 分类: VMware

标签: ESX, VMware, vSwitch / 3,687 次浏览 2 评论

#### 简介:

- 1) 3 种类型的 [VMware](#) 网络都支持 NIC Teaming , 详细请参考《[VMWare ESX Server 中的网络设置](#)》(VMkernel, Service Console 和 VM port group)
- 2) uplink 连接到那些物理交换机的端口都必须在同一个广播域中。(也就是必须在同一个 VLAN 中, 不能跨路由)
- 3) 如果 uplink 要配置 VLAN, 则每个 uplink 必须都配置成 VLAN Trunk 并且具有相同的 VLAN

配置。

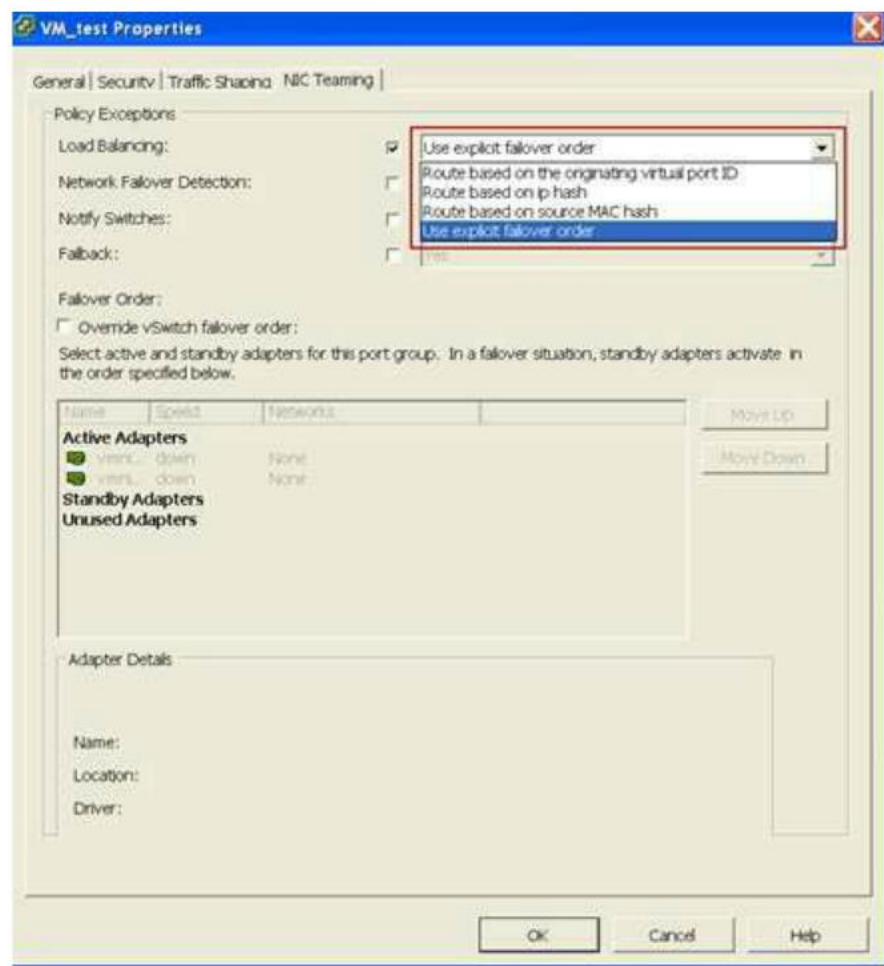
4) **VMware** 的负载均衡(Load Balancing)只是出站(Outbound)的负载均衡，因此要做到 Inbound 的负载均衡，必须在物理交换机上做同样 IP Hash 方式的配置。

5) NIC Teaming 的 Load Balancing 和一些高级路由算法的 Load Balancing 不同，它不是按照 Teaming 中网卡上通过的数据流量来负载均衡，而是根据网卡上的连接(connection)来进行负载均衡。

## VMware 的 3 种负载均衡

VMware 的 NIC Teaming Load Balancing 策略有 3 种。

- (1) 基于端口的负载均衡(默认)
- (2) 基于源 MAC 的负载均衡
- (3) 基于 IP hash 的负载均衡



## 1. 基于端口的负载均衡 (Route based on the originating virtual port ID)

这种方式下，负载均衡是基于 vPort ID 的。一个 vPort 和 Host 上的一个 pNIC(从 [vSwitch](#) 角度看就是某个 uplink)捆绑在一起，只有当这个 pNIC 失效的时候，才切到另外的 pNIC 链路上。这种方式的负载均衡只有在 vPort 数量大于 pNIC 的数量时才生效。

什么是 vport? 一个 VM 上的 vNIC 或者某一个 VMKernel 或者 Service Console 的某个 vswif。用一个图来直观的表述，vPort 在下图中显示为 [vSwitch](#) 上左侧的那些绿点。而 pNIC 在图中显示为右边的 vmnicX。



对于 VM 来说，因为某台 VM 的 vNIC 是捆绑在某一个 pNIC 上的，也就是说这台 VM(如果只有一个 vNIC 的话)对外的数据流量将固定在某一个 pNIC 上。这种负载均衡是在 VM 之间的均衡，对于某一台 VM 而言，其 uplink 的速率不可能大于单个 pNIC 的速率。此外，只有当 VM 的数量足够多，并且这些 VM 之间的数据流量基本一致的情况下，Host 上的 NIC Teaming 的 Load Balancing 才较为有效。对于以下这些极端情况，基于端口方式的负载均衡根本不起作用或者效果很差，充其量只能说是一种端口冗余。

(1) Host 上只有一台只具有单 vNIC 的 VM (此时完全没有 Load balancing)

(2) Host 上的 VM 数量比 pNIC 少(比如有 4 台 VM 但是 Teaming 中有 5 块 pNIC，此时有一块 pNIC 完全没用上，其他每个 vNIC 各自使用一块 pNIC，此时也没有任何负载均衡实现)

(3) Host 上虽然有多台 VM，但是 99% 的网络流量都是某一台 VM 产生的



## 2. 基于源 MAC 地址的负载均衡 Route based on source MAC hash

---

这种方式下，负载均衡的实现是基于源 MAC 地址的。因为每个 vNIC 总是具有一个固定的 MAC 地址，因此这种方式的负载均衡同基于端口的负载均衡具有同样的缺点。同样也是要求 vPort 数量大于 pNIC 的时候才会有效。同样是 vNIC 的速率不会大于单个 pNIC 的速率。

## 3. 基于 IP Hash 的负载均衡 Route based on IP hash

---

这种方式下，负载均衡的实现是根据源 IP 地址和目的 IP 地址的。因此同一台 VM(源 IP 地址总是固定的)到不同目的的数据流，就会因为目的 IP 的不同，走不同的 pNIC。只有这种方式下，VM 对外的流量的负载均衡才能真正实现。

不要忘记，VMware 是不关心对端物理交换机的配置的，VMware 的负载均衡只负责从 Host 出站的流量(outbound)，因此要做到 Inbound 的负载均衡，必须在物理交换机上做同样 IP Hash 方式的配置。此时，pNIC 必须连接到同一个物理交换机上。

需要注意的是，VMware 不支持动态链路聚合协议(例如 802.3ad LACP 或者 Cisco 的 PAgP)，因此只能实现静态的链路聚合。(类似于 HP 的 SLB)。不仅如此，对端的交换机设置静态链路聚合的时候也要设置成 IP Hash 的算法。否则这种方式的负载均衡将无法实现。

这种方式的缺点是，因为 pNIC 是连接到同一台物理交换机的，因此存在交换机的单点失败问题。此外，在点对点的链路中(比如 VMotion)，2 端地址总是固定的，所以基于 IP Hash 的链路选择算法就失去了意义。

## 4. 总结

---

不管采用以上哪一种方法的 Load Balancing，它会增加总聚合带宽，但不会提升某单个连接所获的带宽。为啥会这样？同一个 Session 中的数据包为啥不能做到 Load Balancing？这是因为网络的 7 层模型中，一个 Session 在传输过程中会被拆分成多个数据包，并且到目的之后再重组，他们必须具有一定的顺序，如果这个顺序弄乱了，那么到达目的重组出来的信息就是一堆无意义的乱码。这就要求同一个 session 的数据包必须在同一个物理链路中按照顺序传输过去。所以，10 条 1Gb 链路组成的 10Gb 的聚合链路，一定不如单条 10Gb 链路来的高速和有效。