

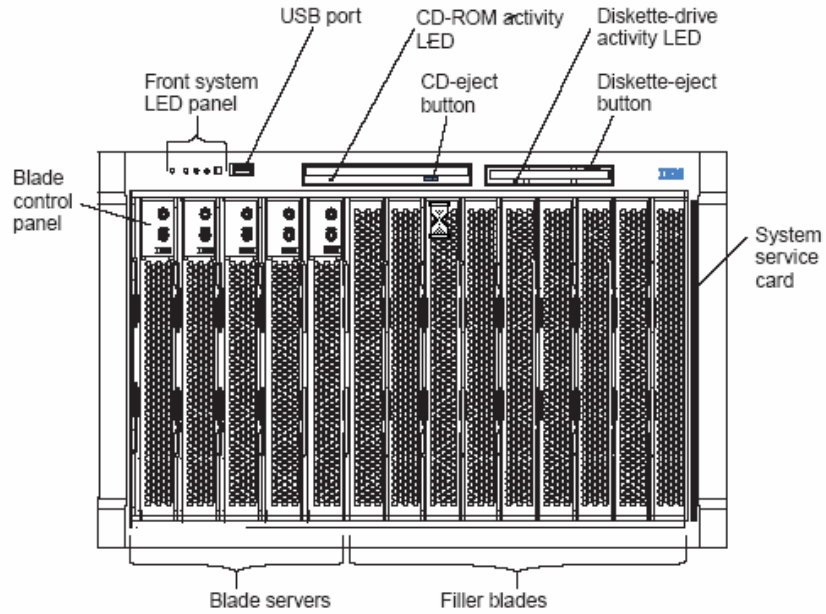
VMware ESX 2.1 NIC Bonding and VLANs On BladeCenter HS20

A BladeCenter.....	3
A Management Module	4
Two Ethernet Switches	5
SCSI blade Server Storage required	6
Spec, docs, white-papers, support.....	7
VMware	7
IBM.....	7
How to config Bonding and VLANs on BladeCenter	8
1.0 Install ESX 2.1 using the graphical installer and default options.....	8
2.0 Reassign NICs, create BOND & enable vmxnet_console.....	8
2.1 Log into COS and run from the cmd line: “vmkpcidivy -i”.....	8
2.2 Create the bond.....	9
2.3 Use vmxnet_console module to give network access to COS,.....	9
2.3.1 Creating a VLAN with NIC Teaming for the Service Console.....	10
2.3.2 Configure the physical ports on your external switch	10
2.4 REBOOT Server	10
Load-balancing	11
Dynamic commands to set load balancing.....	11
Static commands to set load balancing	11
Dynamic command to view bond N	11
Fail-over only mode.....	12
Dynamic command to designate top NIC, bond0, as home.....	12
Dynamic command to designate bottom NIC, bond0, as home	12
Dynamic command to REMOVE bottom NIC, bond0, as home.....	12
Static command to designate top NIC, bond0, as home	12
Static command to designate bottom NIC, bond0, as home	12
Dynamic command to view bond0	12
Switch failover mode	13
Dynamic command to set beacon interval on bond0 to 3 secs	13
Dynamic command to disable beacon interval on bond0	13
Static command to set beacon interval on bond0 to 3 secs.....	13
Dynamic command to view bond0	13
Example config with SAN --- preferred solution	14
Example config without SAN – must use SCSI sidecar on blades.....	14
Minimum hw config possible - use SCSI sidecar, vmnic1= home.....	15
Appendix A --- Example output of step 2.1 “vmkpcidivy -i ”.....	16
Appendix B --- Debugging	18

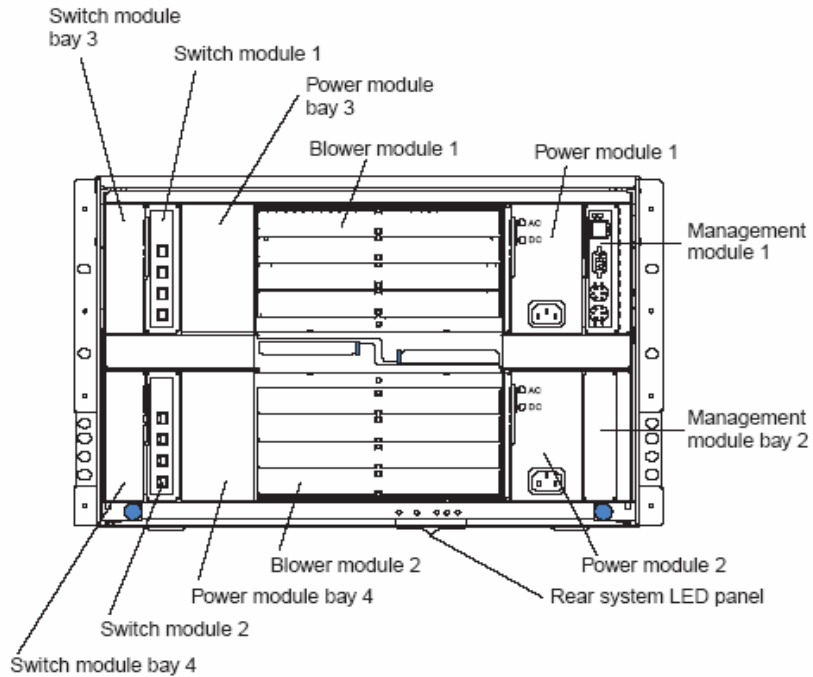
A BladeCenter

Here's a BladeCenter:

Front view

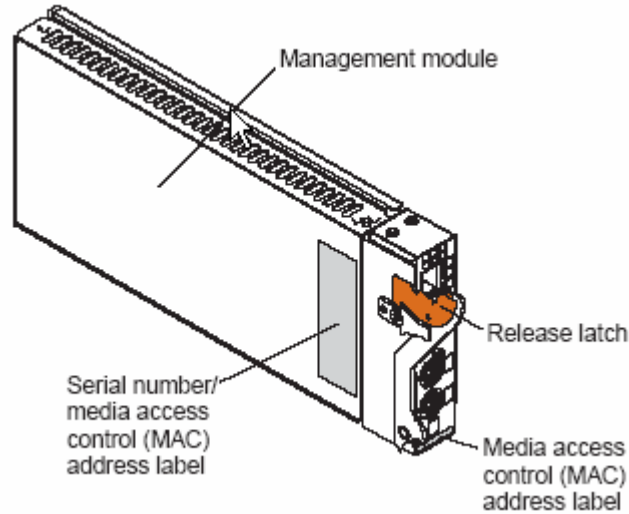


Rear view



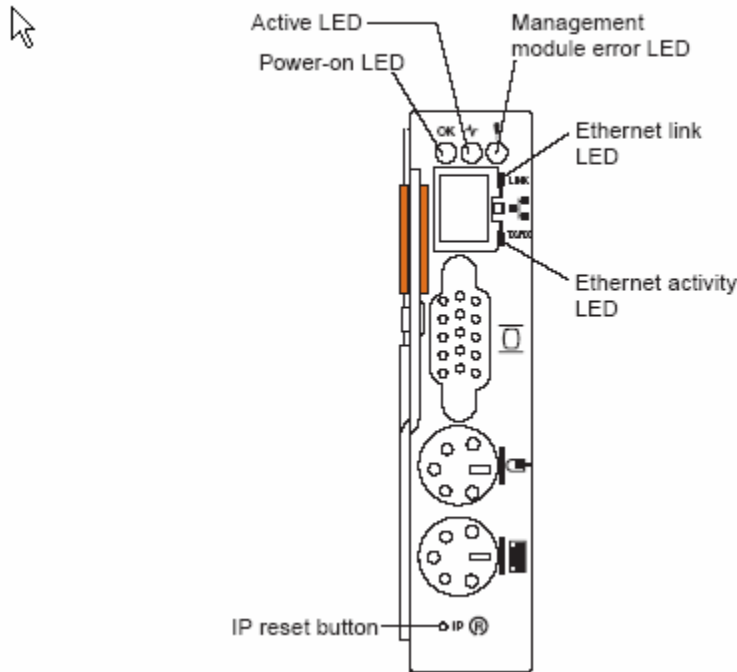
A Management Module

BladeCenter requires one Management Module. (A second, for failover, is optional.)



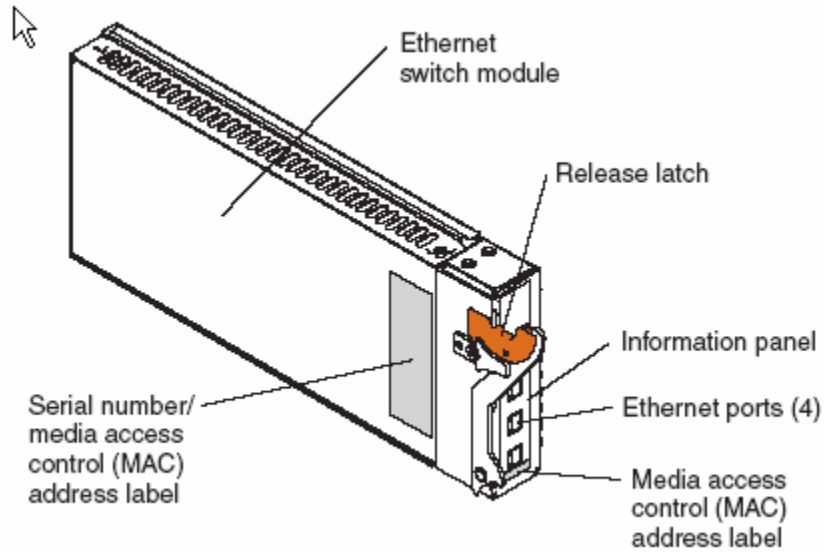
The MM has a physical IP reset button, shown below, pushed via paper clip. It resets the MM to its default IP address of 192.168.70.125, netmask 255.255.255.0.

Management module controls and indicators

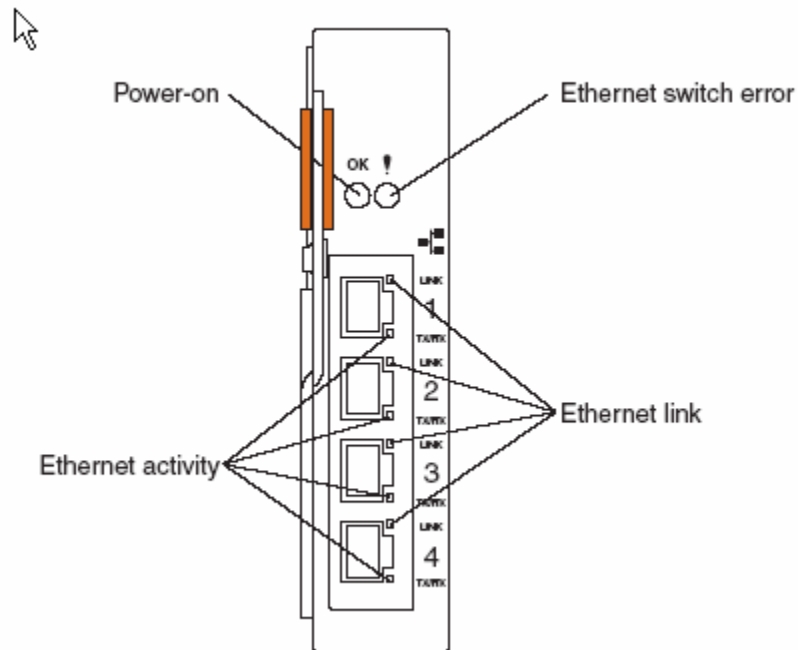


Two Ethernet Switches

Two Ethernet Switch modules *are recommended* for VMware. (A single NIC solution can be configured for a low-end solution.)



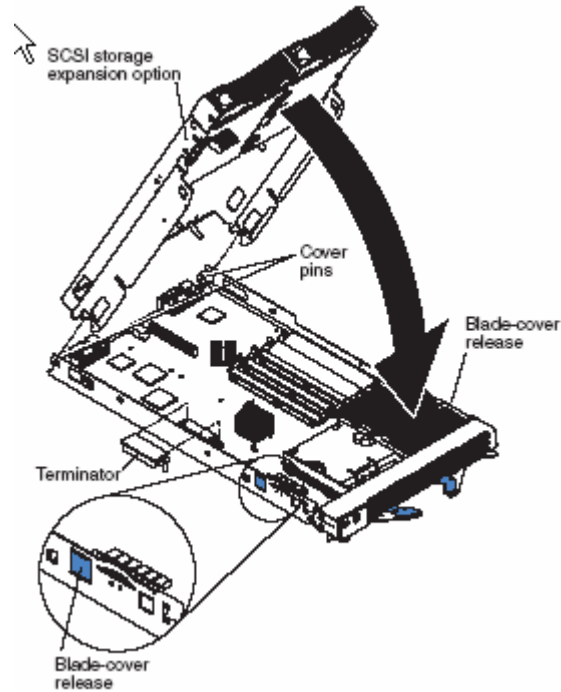
An example faceplate is:



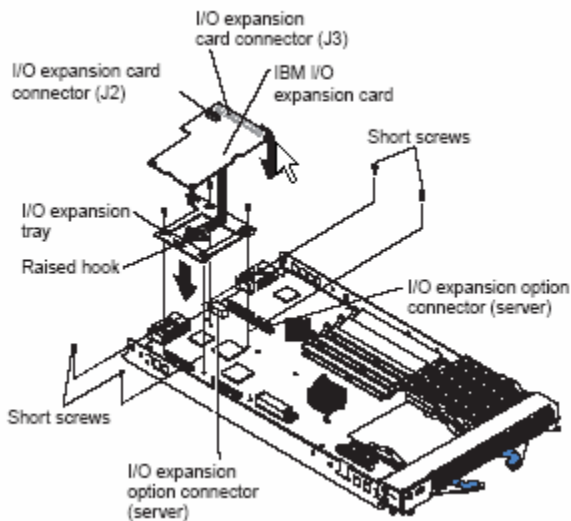
SCSI blade Server Storage required

You can install the Vmkernel/Service Console on an IDE drive. However, VMware ESX Server ***requires SCSI blade server storage***, either local or SAN, for virtual machines.

If local, then one must install a SCSI expansion option on each blade.



If SAN, then one must have a FC expansion card in each blade (shown below), plus Fibre Channel Switch modules in bays 3 and 4 (not shown). SAN is the preferred media as it enables VirtualCenter to dynamically move VMs from blade to blade via Vmotion.



Spec, docs, white-papers, support**VMware**

- 1) VMware's white-paper --- "Configuring and Installing IBM BladeCenter"
http://www.vmware.com/pdf/esx21_IBM_blade.pdf
- 2) VMware's CD-ROM install steps ---
http://www.vmware.com/support/esx21/doc/esx21install_text-teps_install.html
- 3) VMware's ESX Server 2 specs, docs, and white papers ---
http://www.vmware.com/products/server/esx_features.html

IBM

- 1) Personal computing support ---
<http://www.ibm.com/pc/support>
- 2) Blade servers ---
<http://www.ibm.com/servers/eserver/bladecenter>

How to config Bonding and VLANs on BladeCenter

NOTE: Do not connect cables to switch trunk ports yet.
(Trunk ports are on the GbE switches external to BladeCenter).

1.0 Install ESX 2.1 using the graphical installer and default options.

By default 1 NIC will be dedicated to the Service Console's Operating System (COS).
Make sure and type in all static IP information for the COS during installation.
System will reboot after installation is complete.

2.0 Reassign NICs, create BOND & enable vmxnet_console

2.1 Log into COS and run from the cmd line: "vmkpcidiv -i"

to reassign the COS's NIC to the virtual machines. Change the first NIC labeled with a [c] to shared mode by typing "s" so it will be shared between the COS and the VMs. Ensure the remaining NIC is dedicated to the VMs by typing "v". Both HS20 blade NICs will now be dedicated to the virtual machines. ((Refer to Appendix A for a screenshot of "vmkpcidiv -i")).
DO NOT REBOOT YET.

After step 2.4, you can review the result of step 2.1 via the MUI. Click on the Options tab, then click on Startup Profile, and you see the following:

Hardware Profile	
Ethernet controller: BROADCOM Corporation NetXtreme BCM5704 Gigabit Ethernet (rev 02)	
Dedicated To	Shared
Driver	bcm5700.o
PCI Bus:Device.Function	1:0.0
Ethernet controller: BROADCOM Corporation NetXtreme BCM5704 Gigabit Ethernet (rev 02)	
Dedicated To	Virtual Machines
Driver	bcm5700.o
PCI Bus:Device.Function	1:0.1

2.2 Create the bond.

Edit /etc/vmware/hwconfig and add the following two lines to the end of the file, where *x* is a bond number from 0 to 9.

```
nicteam.vmnic0.team = "bond<x>"
nicteam.vmnic1.team = "bond<x>"
```

For example:

```
nicteam.vmnic0.team = "bond0"
nicteam.vmnic1.team = "bond0"
```

NOTE1:

In teaming mode, both NICs on the blade need to be in the same broadcast domain, that is, they have to be layer 2 connected.

NOTE2:

As it stands now, both NICs are used for traffic in the bond. They will use MAC addresses to load-balance, which is the default.

If you want to enable Fail-over mode, so that data traffic only flows over a single "home" or "primary" NIC, with the other NIC reserved strictly for backup, then the line to enable this also goes here in /etc/vmware/hwconfig.

For example, if we want the top NIC on the HS20 (which is vmnic1 and connects to GbESw1) to be the "home" or "primary" NIC on bond0, then we add the following line also:

```
nicteam.bond0.home_link = "vmnic1"
```

2.3 Use vmxnet_console module to give network access to COS,

thereby creating COS eth0. (See pg 337 of ESX 2.1 Admin Manual).

Edit /etc/rc.local to add the following at the end of the file:

```
# vmxnet_console through bond<x>
/etc/rc.d/init.d/network stop
rmmod vmxnet_console
insmod vmxnet_console devName="bond<x>"
/etc/init.d/network start
mount -a
```

For example:

```
# vmxnet_console through bond0
/etc/rc.d/init.d/network stop
rmmod vmxnet_console
insmod vmxnet_console devName="bond0"
/etc/init.d/network start
mount -a
```

A MINIMAL example:

```
# vmxnet_console through bond0
insmod vmxnet_console devName="bond0"
ifup eth0
```

2.3.1 Creating a VLAN with NIC Teaming for the Service Console

If you are going to use a VLAN with the COS you must include the VLAN ID with the devName for the vmxnet_console. The /etc/rc.local file will have this entry for VLAN 10

```
# vmxnet_console through bond0 VLAN 10
```

Edit /etc/rc.local and add the following lines to the end of the file, where x is the bond number (0-9) and y is the VLAN tag number (1-4095). These lines allow you to give network access through the bond to the COS.

```
# vmxnet_console through bond<x> VLAN<y>
/etc/rc.d/init.d/network stop
rmmod vmxnet_console
insmod vmxnet_console devName="bond<x>.<y>"
/etc/init.d/network start
mount -a
```

A MINIMAL example:

```
insmod vmxnet_console devName="bond0.10"
ifup eth0
```

NOTE:

In the MUI there will be no reference to a Port Group with VLAN 10.

The vmkernel reads the devName="bond0.10" and knows to tag all packets from the COS NIC with VLAN ID 10.

If you want to use VLAN 10 for virtual machines you will need to create a Port Group in the MUI associated with bond0 with VLAN ID 10.

2.3.2 Configure the physical ports on your external switch

to be trunk ports that will support the VLAN IDs you chose. I refer to the external switch NOT the switch in the BladeCenter.

2.4 REBOOT Server

NOTE:

Unless you have specified vmnic1 as the "home" NIC in /etc/vmware/hwconfig as discussed in Step 2.2, **after reboot, the COS uses the bottom NIC on the HS20.**

```
vmnic1 ---- top NIC
vmnic0 ---- bottom NIC.
```

Load-balancing

Dynamic commands to set load balancing

Load balancing is now based by default on source MAC, not source IP. This gives better interaction with multi-switch configurations.

You can dynamically switch between “out-mac” and “out-ip” load balancing via the following commands, where “N” is the number assigned to the bond.

```
echo "nicateaming load-balance out-mac" > /proc/vmware/net/bondN/config
echo "nicateaming load-balance out-ip" > /proc/vmware/net/bondN/config
```

Static commands to set load balancing

Add a line to /etc/vmware/hwconfig as so:

```
nicateam.bondN.load_balance_mode = "out-ip"
nicateam.bondN.load_balance_mode = "out-mac"
```

Dynamic command to view bond N

```
cat /proc/vmware/net/bondN/config
```

```
VlanHwTxAccel          Yes
VlanHwRxAccel          Yes
VlanSwTagging           Yes
PromiscuousAllowed     No
InterruptClustering    No
Link state:            Up
Speed:                 1000 Mbps, full duplex
PCI (bus:slot.func):   -1:-1.-1
Minimum Capabilities   0x0
Device Capabilities    0x74b
Maximum Capabilities   0x76b
NICteamingSlaves:
Name                   LinkUp   BeaconOkay
vmnic0                 Yes      N/A
vmnic1                 Yes      N/A
NICteamingLoadBalance: On:      out-ip
NICteamingSwitchFailover: Off
Interrupt vector       0xfffffffffe
DebugSocket            Closed
```

Fail-over only mode

In fail-over only mode, one NIC is the home or primary NIC, and all others simply stand by. ((Designating a home NIC overrides the value of load-balance mode.))

To force a fail-over for testing purposes, power-off one of the GbE Switch modules in the BladeCenter chassis.

Dynamic command to designate top NIC, bond0, as home

```
echo "nicateaming home vmnic1" > /proc/vmware/net/bond0/config
```

Dynamic command to designate bottom NIC, bond0, as home

```
echo "nicateaming home vmnic0" > /proc/vmware/net/bond0/config
```

Dynamic command to REMOVE bottom NIC, bond0, as home

```
echo "nicateaming unhome vmnic0" > /proc/vmware/net/bond0/config
```

Static command to designate top NIC, bond0, as home

Add the following line to /etc/vmware/hwconfig as so:

```
nicteam.bond0.home_link = "vmnic1"
```

Static command to designate bottom NIC, bond0, as home

Add the following line to /etc/vmware/hwconfig as so:

```
nicteam.bond0.home_link = "vmnic0"
```

Dynamic command to view bond0

```
cat /proc/vmware/net/bond0/config
VlanHwTxAccel          Yes
VlanHwRxAccel          Yes
VlanSwTagging          Yes
PromiscuousAllowed     No
InterruptClustering    No
Link state:            Up
Speed:                 1000 Mbps, full duplex
PCI (bus:slot.func):   -1:-1.-1
Minimum Capabilities   0x0
Device Capabilities    0x74b
Maximum Capabilities   0x76b
NICteamingSlaves:
Name      LinkUp   BeaconOkay
vmnic0    Yes      N/A
vmnic1    Yes      N/A
NICteamingLoadBalance: Off (HomeLink: vmnic1)
NICteamingSwitchFailover: Off
Interrupt vector       0xfffffffffe
DebugSocket            Closed
```

Switch failover mode

In switch-failover mode, bonds watch for an upstream switch failure using a beacon. ((The upstream switch is external to the BladeCenter chassis.))

Dynamic command to set beacon interval on bond0 to 3 secs

```
echo "nict teaming switch-failover 3" > /proc/vmware/net/bond0/config
```

Dynamic command to disable beacon interval on bond0

```
echo "nict teaming switch-failover 0" > /proc/vmware/net/bond0/config
```

Static command to set beacon interval on bond0 to 3 secs

Add the following line to /etc/vmware/hwconfig as so:

```
nicteam.bond0.switch_failover_threshold = "3"
```

Dynamic command to view bond0

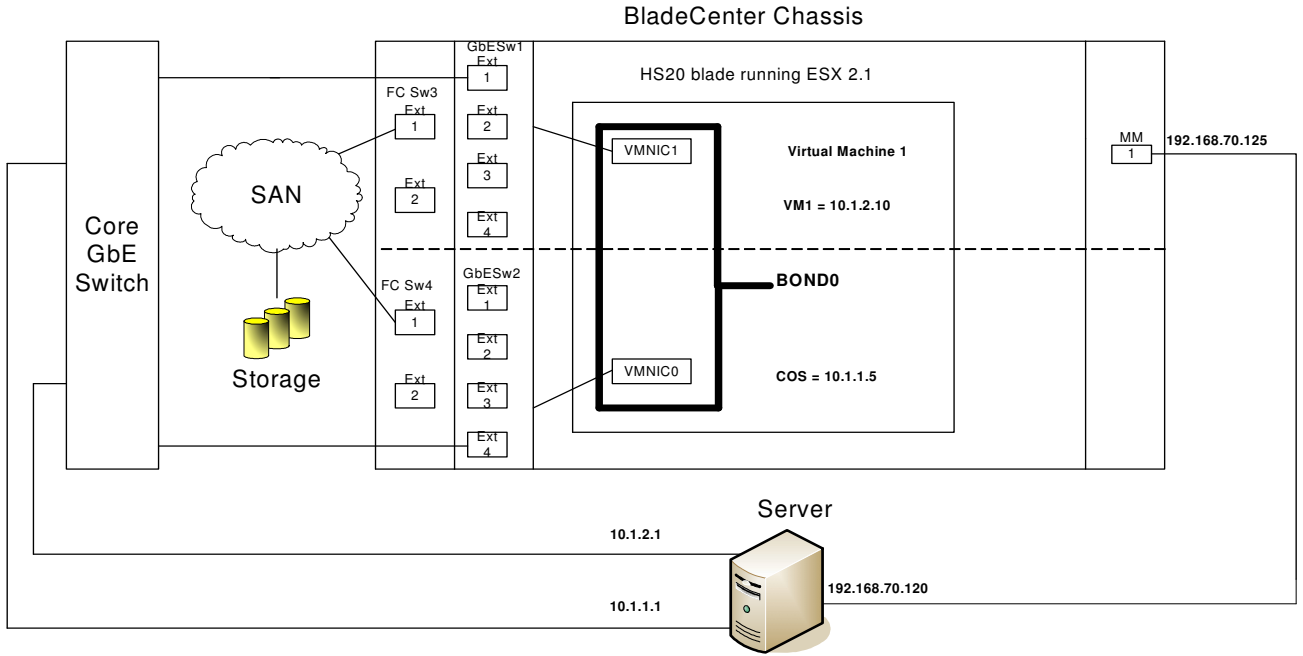
```
cat /proc/vmware/net/bond0/config
```

```

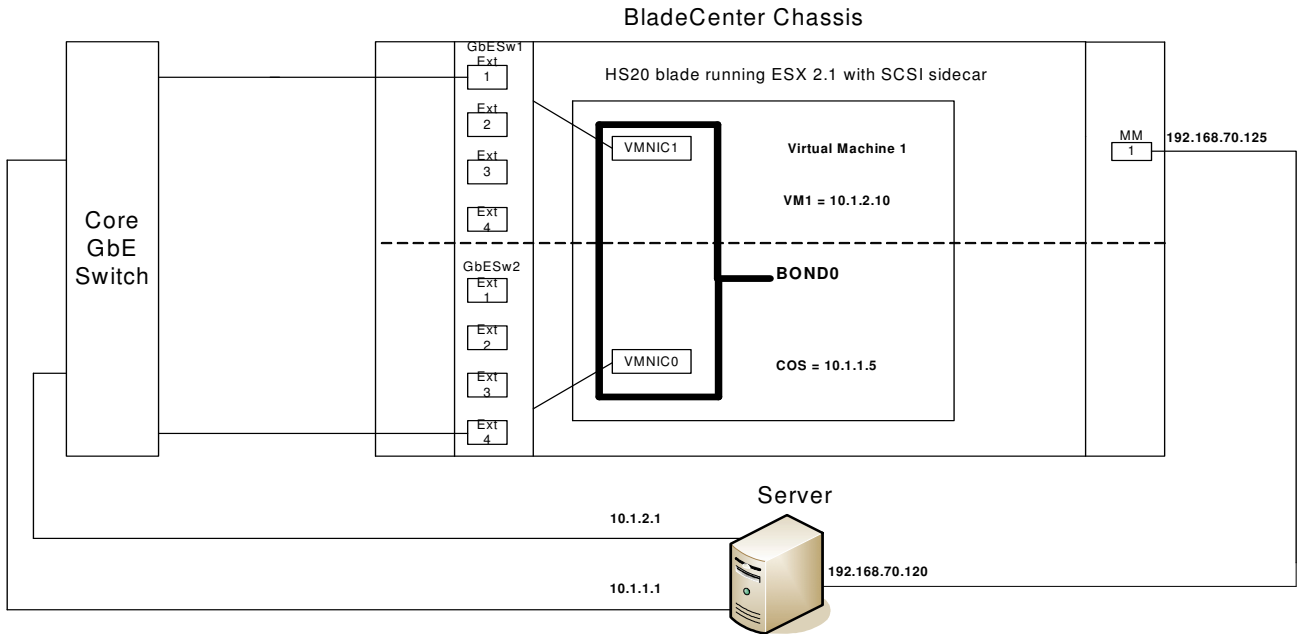
VlanHwTxAccel          Yes
VlanHwRxAccel          Yes
VlanSwTagging           Yes
PromiscuousAllowed     No
InterruptClustering    No
Link state:            Up
Speed:                 1000 Mbps, full duplex
PCI (bus:slot.func):   -1:-1.-1
Minimum Capabilities   0x0
Device Capabilities    0x74b
Maximum Capabilities   0x76b
NICTeamingSlaves:
Name      LinkUp   BeaconOkay
vmnic0    Yes      No
vmnic1    Yes      No
NICTeamingLoadBalance: Off (HomeLink: vmnic1)
NICTeamingSwitchFailover: On (3), NO recent beacons (261)
Interrupt vector       0xfffffffffe
DebugSocket            Closed

```

Example config with SAN --- preferred solution



Example config without SAN – must use SCSI sidecar on blades



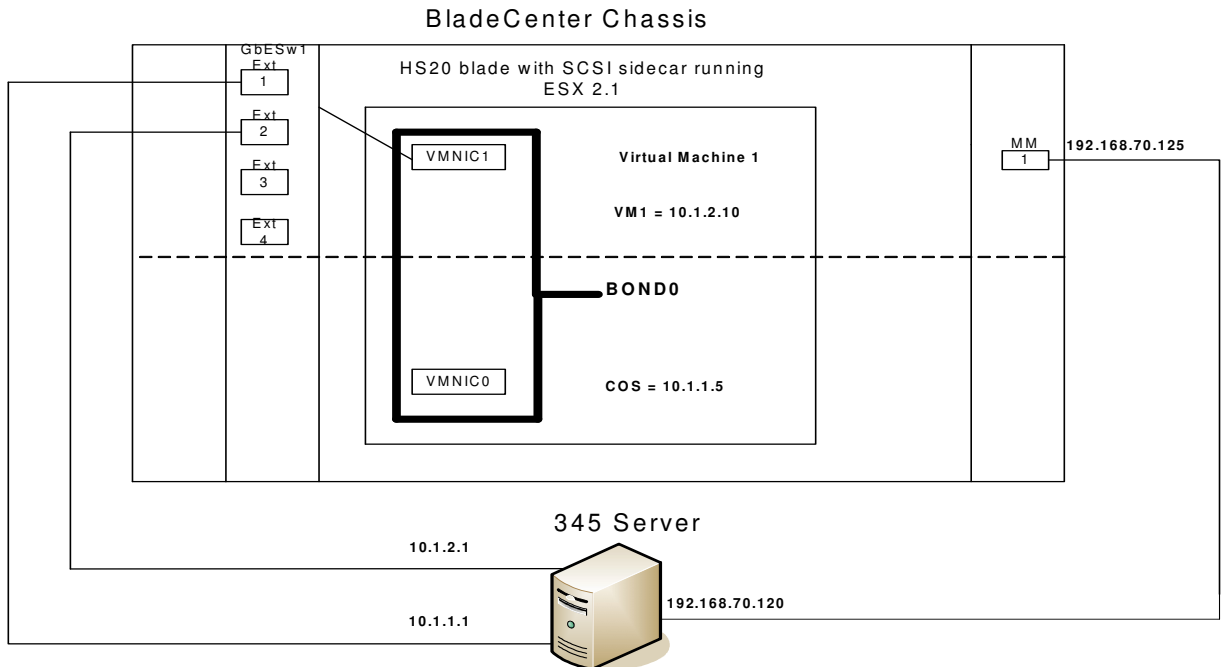
Minimum hw config possible - use SCSI sidecar, vmnic1= home

(no failover possible)

This is the minimum hardware configuration possible. Customers save the price of the second GbE Switch at the cost of wasting vmnic0 on each blade and losing fail-over.

Notice that in this configuration, “vmkpcidivv -i” did assign both NICs to the Service Console. However, since vmnic1 in this configuration must be the “home” NIC, the Service Console uses vmnic1 instead of vmnic0. Recall lines must be inserted at the bottom of /etc/vmware/hwconfig as follows:

```
nicteam.vmnic0.team = "bond0"
nicteam.vmnic1.team = "bond0"
nicteam.bond0.home_link = "vmnic1"
```



Appendix A --- Example output of step 2.1 “vmpcidivy -i ”

```
[root@BC7 root]# vmpcidivy -i
Checking for existing VMnix Boot Configurations.

The following VMnix kernel images are defined on your system:

Boot image configuration: esx
  Image file: /boot/vmlinuz-2.4.9-vmnix2
  Memory: 192M
  Service Console devices:
    Ethernet controller: BROADCOM Corporation NetXtreme BCM5704 Gigabit Ethernet (rev 02)
    SCSI storage controller: Symbios Logic Inc. (formerly NCR) LSI Logic Fusion MPT
53C1030 (rev 07) (shared)
  VM devices:
    Ethernet controller: BROADCOM Corporation NetXtreme BCM5704 Gigabit Ethernet (rev 02)
    Fiber storage controller: QLogic Corp QLA231x/2340 (rev 02)
    SCSI storage controller: Symbios Logic Inc. (formerly NCR) LSI Logic Fusion MPT
53C1030 (rev 07) (shared)

Type in the name of the boot image configuration you wish to configure
or type "new" to create a new image [esx]: new

Creating new VMnix boot configuration.

Configuring your setup.  Enter in the options that you want.  Defaults have
been provided.

Name of this configuration [esx2]: test

Now allocating memory to the Service Console.  The memory that is not allocated
to the Service Console will be allocated to the VMs.  Be sure to indicate a
number smaller than you actually have on your system or you may not be able to
run any VMs.

Your system is reporting approximately 2,047 MB of memory.
Note: you cannot allocate more than 800 MB to the Service Console.

If you plan to run 4 virtual machines or fewer, 128 MB should suffice.
Set this to 192 MB for 8 virtual machines, 272 MB for 16 virtual
machines, 384 MB for 32 virtual machines, or 512 MB for more than
32 virtual machines.

Amount of memory (in megs) to allocate to the Service Console [272]:
Enable hyperthreading? (Only works with 8 or less physical processors) [y]:

Now to divide the SCSI controllers and network adapters.  In this process,
you will select whether to give the previous devices to the Service Console or
the virtual machines that you will be running.  In order to ensure optimal
performance for your virtual machines, we suggest you allocate as many devices
as possible to the virtual machines, leaving only the ones necessary to
configure and maintain the Service Console.

Type 'c' to allocate to the Service Console or 'v' to allocate to the virtual
machines, or 's' to share between both.

Ethernet controller BROADCOM Corporation NetXtreme BCM5704 Gigabit Ethernet (rev 02)
running on bus 1 device 0 function 0 (bcm5700.o) [] [c]: s
Ethernet controller BROADCOM Corporation NetXtreme BCM5704 Gigabit Ethernet (rev 02)
running on bus 1 device 0 function 1 (bcm5700.o) [vmnic0] [c]: v
Fiber storage controller QLogic Corp QLA231x/2340 (rev 02) running on bus 2 device 2
(qla2300_604.o) [vmhba0] [c]: v
SCSI storage controller Symbios Logic Inc. (formerly NCR) LSI Logic Fusion MPT 53C1030
(rev 07) running on bus 3 device 1 function 0 (mptscsi.o) [vmhba2] [c]: s
```

```
Current configuration is not the default configuration.
Use this configuration as the default (y/n)? [n]: y

Configuration complete.
Commit changes (y/n)? [y]: y

Committing changes...VMware::ExtHelpers::System[485]: '/bin/mv' '-f'
'/tmp/exthelpers.tmpfile.2293' '/etc/lilo.conf'
VMware::ExtHelpers::System[1725]: '/sbin/lilo'
VMware::ExtHelpers::System[485]: '/bin/mv' '-f' '/tmp/exthelpers.tmpfile.2293'
'/etc/vmware/vmkmodule.conf'
Successful.
[root@BC7 root]#
```

Appendix B --- Debugging

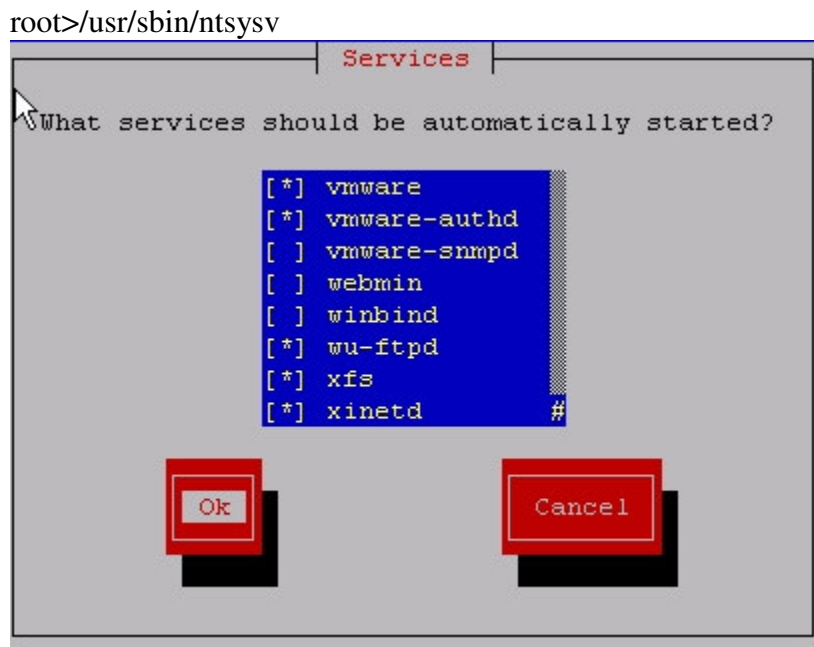
VMware has a service support script named `/usr/bin/vm-support`, which gathers up all the relevant files required to debug. For example, from the Service Console, type:

```
root>/usr/bin/vm-support
root>ls -lt esx*
-rw-r--r--  1 root   root   393426 May  5 07:55 esx-2004-05-05.6327.tgz
```

You have several options to get the above `esx*.tgz` file off the system.

A ssh server runs on the COS, so from a Windows box you can use PuTTY's `pscp.exe` to do a secure copy. Or you can log into the COS and initiate an ftp session or `scp` to a remote server.

Or, on the COS, you can use `/usr/sbin/ntsysv` to start an FTP server on the next reboot via `wu-ftpd`. (You can also enable telnet from here.)



If you need to change the IP address of the Service Console, edit this file:

```
/etc/sysconfig/network-scripts/ifcfg-eth0
```

If you need to change the gateway address of the Service Console, edit this file:

```
/etc/sysconfig/network
```