

FreeNAS - Bug #2056

ESXi with LSI Cards

02/21/2013 05:09 PM - Paul Bucher

Status:	Closed	
Priority:	Expected	
Assignee:		
Category:		
Target version:	9.1.0-RELEASE	
Seen in:		Needs Merging: Yes
Severity:	New	Needs Automation: No
Reason for Closing:		Support Suite Ticket: n/a
Reason for Blocked:		Hardware Configuration:
Needs QA:	Yes	ChangeLog Required: No
Needs Doc:	Yes	

Description

Out of the box you can not boot [[FreeNAS]] in a VMware ESXi VM if you have a physical LSI card passed through to the VM. This is mainly ESXi's fault with having some kind of limitation related to msix interrupts(I don't understand it but it's been that way through multiple ESXi releases so I'm guessing it's either impossible to fix or really low on the todo list).

Anyways if you simple bring up [[FreeNAS]] before adding the LSI card and add the tuning variable "hw.pci.enable_msix=0" life is great and everything is good.

The bug/problem is when you upgrade [[FreeNAS]] is, it boots up with a generic setup to do the db upgrade and finalize things and it doesn't have the customized loader.conf file in place yet and you get a hung system.

Work Around:

1. Install update, catch it at the boot prompt and pause boot cycle at the boot prompt.
2. Power Off VM and remove LSI card from VM.
3. Power Up VM, after the 1st boot does the database upgrade and reboots pause it again at the boot prompt.
4. Power Off VM add back LSI card to VM.
5. Power Up and enjoy your upgraded [[FreeNAS]] server.

Ideas/Thoughts:

1. If it's possible to detect your booting in a VM, automatically add "hw.pci.enable_msix=0" to rc.conf with the other VMware lines.
2. I'm guessing it would be considered undesirable to add that line across the board.
3. Produce a VMWare only upgrade that has that automatically added to rc.conf or loader.conf.
4. Some how save the working loader.conf.local file from the existing setup and boot up with that for the upgrade and then overwrite it.

Note:

I'm been running a slight mod of this this(hw.mps.disable_msix="1") in production since the Beta2 days of 8.3 with great results. I'm able to hit 500MB per sec writes to disk(not OS cached or ZIL cached) on my ZFS vol without a problem(3x5 drive Raid1Z) inside the VM. I'm going to switch my test SAN over to "hw.pci.enable_msix=0" to validate it before I move it to production.

History

#1 - 02/21/2013 05:29 PM - Julien Bénic

Yes it's really good idea because when I update my [[FreeNAS]] I have the same error.

#2 - 02/21/2013 06:12 PM - Anonymous

What version(s)/build(s) of ESXi are you seeing these issues on?

#3 - 02/21/2013 06:29 PM - Julien Bénic

Replying to [comment:2 dwhite]:

What version(s)/build(s) of ESXi are you seeing these issues on?

My Software:

ESXi 5.1 build 914609 (the last)

[[FreeNAS]]-8.3.2-ALPHA-r13323-x64

My Hardware:

CPU: i5-3570T

RAM: 32Go

RAID ADAPTER: LSI 9207-8i in passthrough for VM [[FreeNAS]]

HDD: 7* 3To RAIDZ1

The error is say in the forum:

<http://forums.freenas.org/archive/index.php/t-6166.html>

When update [[FreeNAS]] , vm reboot and crash then turn off the vm.

And delete the adapter LSI, and turn on for finalize the update and turn off and add the adapter.

That's all and it's annoying :)

#4 - 02/21/2013 06:38 PM - Paul Bucher

I've had the problem under ESXi 5.0(various updates/builds), 5.1 shipped version, and 5.1 build 914609 with several different LSI cards and servers. It's pretty well known that [[FreeBSD]], LSI cards using the mps driver, and ESXi don't do anything but hang you up at boot time.

#5 - 02/21/2013 10:26 PM - paleoN -

Not a solution, but have you tried this workaround instead:

1. Install update, choose [6. Escape to loader prompt](#) at the boot prompt.

1. set hw.pci.enable_msix=0

1. boot

You may need to do this twice depending on when `/boot/loader.conf` is actually updated.

#6 - 02/22/2013 12:10 AM - Julien Bénic

Replying to [comment:5 paleoN]:

Not a solution, but have you tried this workaround instead:

1. Install update, choose [6. Escape to loader prompt](#) at the boot prompt.

1. set hw.pci.enable_msix=0

1. boot

You may need to do this twice depending on when `/boot/loader.conf` is actually updated.

I haven't think to this solution. Great!

It's possible to add "set hw.pci.enable_msix=0" in Init\Shutdown Scripts?

Thanks

#7 - 02/22/2013 01:06 AM - paleoN -

Replying to [comment:6 [[MotorSport]]]:

It's possible to add "set hw.pci.enable_msix=0" in Init\Shutdown Scripts?

Like I said it's a workaround. At this point you are at [Stage Three](#) which reads `/boot/loader.conf/.local` among other things. Which would put you back to the original request.

#8 - 02/22/2013 01:08 AM - Paul Bucher

Replying to [comment:6 [[MotorSport]]]:

It's possible to add "set hw.pci.enable_msix=0" in Init\Shutdown Scripts?

That's what loader.conf is more or less.

Also 1 important comment I forgot, if this settings could be made standard for VMed installs it would save a ton of tech support posts along with folks who just walk away from [[FreeNAS]] because it won't boot. There are also variations of the fix which cause other problems, myself and others have spent untold hours to find this 1 simple line that makes [[FreeNAS]] rock under ESXi and based on the # of posts I've turned up googling there are a lot of folks trying this. Esp in when used in conjunction with the vmxnet3 driver you now a SAN connected by 10G ethernet to the other VMs on the ESXi box.

#9 - 05/21/2013 06:42 PM - Anonymous

This ticket is important, but it will have to wait until post-9.1

#10 - 06/19/2013 01:19 AM - Paul Bucher

See Ticket [#1894](#) for some discussion about this issue.

Here's some new bread crumbs from 9.1, where I'm having interrupt storms when ever I add more then 1 CPU core to the VM.

I diffed the dmesg log between a verbose boot with 1 CPU and 2 CPUs and this is the only item that was different that seemed to point to the issue.

```
pass36 at mps0 bus 0 scbus2 target 43 lun 0
pass36: <SEAGATE ST31000424SS 0006> Fixed Direct Access SCSI-5 device
pass36: Serial Number 9WK3P6W40000C132EFCP
pass36: 600.000MB/s transfers
pass36: Command Queueing enabled
SMP: AP CPU #1 Launched!
cpul AP:
  ID: 0x01000000  VER: 0x00050014  LDR: 0x00000000  DFR: 0xffffffff
  lint0: 0x00010700  lint1: 0x00000400  TPR: 0x00000000  SVR: 0x000001ff
  timer: 0x000100ef  therm: 0x00010000  err: 0x000000f0  pmc: 0x00010400
ioapic0: routing intpin 1 (ISA IRQ 1) to lapic 1 vector 48
ioapic0: routing intpin 12 (ISA IRQ 12) to lapic 1 vector 49
ioapic0: routing intpin 16 (PCI IRQ 16) to lapic 1 vector 50
ioapic0: routing intpin 18 (PCI IRQ 18) to lapic 1 vector 51
TSC timecounter discards lower 1 bit(s)
Timecounter "TSC-low" frequency 1250000000 Hz quality -100
interrupt storm detected on "irq18: "; throttling interrupt source
interrupt storm detected on "irq18: "; throttling interrupt source
```

#11 - 06/19/2013 08:16 PM - Paul Bucher

I'm going to open a new ticket since my interrupt issue is something different then the original request of this ticket. I don't want to loose the original request asked for in this ticket.

#12 - 06/28/2013 05:57 PM - Paul Bucher

I've got my fingers crossed that this issue will be resolved by some fixes in [[FreeBSD]] kernel that should just make everything auto-magically work out of the box for ESXi. See ticket [#2293](#) for more info.

#13 - 07/15/2013 03:37 PM - William Grzybowski

- Status changed from Unscreened to Closed

Fix committed as [1bf3929c] in [[TrueOS]]. See [#2293](#) for more info.