

freebsd-current: Re: pci powerstate related: aac(4) broken on Perc 3/Di on -CURRENT

Re: pci powerstate related: aac(4) broken on Perc 3/Di on -CURRENT

Source: <http://unix.derkeiler.com/Mailing-Lists/FreeBSD/current/2005-01/0181.html>

From: Scott Long (scottl_at_freebsd.org)

Date: 01/06/05

Date: Thu, 06 Jan 2005 15:58:39 -0700
To: Warner Losh <imp@rover.village.org>

Warner Losh wrote:

```
> From: "Simon L. Nielsen" <simon@nitro.dk>
> Subject: Re: pci powerstate related: aac(4) broken on Perc 3/Di on -CURRENT
> Date: Thu, 6 Jan 2005 14:13:28 +0100
>
>
>> On 2004.12.23 07:48:44 -0700, Scott Long wrote:
>>
>>> Simon L. Nielsen wrote:
>>>
>>>> Hello
>>>>
>>>> Recent -CURRENT seems to have broken aac(4) on a Dell Perc 4/Di. The
>>>> system is a Dell PowerEdge 2650 with 4 36GB IBM disks in a RAID0+1
>>>> configuration.
>>>>
>>>> It runs fine on a 5-STABLE kernel, but when booting -CURRENT it prints
>>>> a lot of errors from the RAID controller and then fails to mount the
>>>> root file-system.
>>>>
>>>> I have attached dmesg from 6-CURRENT and 5-STABLE, but the main
>>>> interesting parts from -CURRENT are:
>>>>
>>>> aac0: <Dell PERC 3/Di> mem 0xf0000000-0xf7ffffff irq 30 at device 8.1 on
>>>> pci4
>>>> aac0: [FAST]
>>>> aacd0: <RAID 0/1> on aac0
>>>> aacd0: 69425MB (142182912 sectors)
>>>> SMP: AP CPU #3 Launched!
>>>> SMP: AP CPU #1 Launched!
>>>> SMP: AP CPU #2 Launched!
>>>> aac0: **Monitor** NMI ISR: NMI_SECONDARY_ATU_ERROR
>>>> aac0: **Monitor** NMI ISR: NMI_SECONDARY_ATU_ERROR
>>>> aac0: COMMAND 0xc2409438 TIMEOUT AFTER 41 SECONDS
```

Re: pci powerstate related: aac(4) broken on Perc 3/Di on -CURRENT

freebsd-current: Re: pci powerstate related: aac(4) broken on Perc 3/Di on -CURRENT

>>>
>>> *There are very few differences between the driver in 6-CURRENT and*
>>> *5-STABLE, and none of the differences look like ones that could*
>>> *cause problems. Would you get able to step the source backwards until*
>>> *you find the point where it starts working again?*
>>
>> *After several rounds of backstepping I found that the problem is*
>> *caused by sys/dev/pci/pci.c v. 1.268 which sets hw.pci.do_powerstate=1*
>> *by default. If I add hw.pci.do_powerstate="0" to loader.conf the*
>> *system boots fine. I have no idea why this only manifests itself as*
>> *an aac(4) error.*
>>
>> *This system has a Dell remote management card and I remember that*
>> *Lukas Ertl, some time ago, reported some problem with the power state*
>> *change and a (HP?) remote management card, so perhaps this is a*
>> *similar issue.*
>
>
> *Interesting. This is even after my changes to current to make it not*
> *power down system devices? Can you send me a complete pciconf -lv for*
> *this system?*
>
> Warner

One thing to keep in mind with the Dell PERC systems is that the RAID CPU is an i960 with a transparent PCI-PCI bridge. The i960 device (which the driver attaches to) sits before the bridge, while a SCSI chip sits behind it. Anywhere from 0 - 2 devices of this SCSI chip are exposed through the bridge, depending on how the RAID BIOS is configured. It 'hides' the other devices by changing the pci id of them to something that the ahc driver will not attach to. I thought that it also swizzled the INTx and IDSEL lines, but that appears not to be the case; maybe it only does the INTx lines. For a refresher, this is what it looks like in the dmesg:

```
pci4: <ACPI PCI bus> on pcib1
pcib2: <ACPI PCI-PCI bridge> at device 8.0 on pci4
pci5: <ACPI PCI bus> on pcib2
pci5: <mass storage, SCSI> at device 6.0 (no driver attached)
pci5: <mass storage, SCSI> at device 6.1 (no driver attached)
aac0: <Dell PERC 3/Di> mem 0xf0000000-0xf7ffffff irq 30 at device 8.1 on
pci4
aac0: [FAST]
aac0: i960RX 100MHz, 118MB cache memory, optional battery present
aac0: Kernel 2.7-1, Build 3170, S/N f810d3
aac0: Supported
Options=75c<WCACHE,DATA64,HOSTTIME,WINDOW4GB,SOFTERR,NORECOND,SGMAP64>
```

So why is the aac firmware getting mad? Because Warner powered down the SCSI devices that it was using.

Re: pci powerstate related: aac(4) broken on Perc 3/Di on -CURRENT

freebsd-current: Re: pci powerstate related: aac(4) broken on Perc 3/Di on -CURRENT

This type of thing is why I've always been very nervous about the automatic power management control that was committed to the tree. The above example is completely in spec, but we are taking the liberty of assuming that all unattached devices should be powered down (modulo the exception that was made for video devices). I don't know of a generic way to fix this; you'll have to either add an exception to the PM code for these specific SCSI devices, or write a do-nothing driver to attach to it so it doesn't get spammed by the PM code. Either way it's just an exception for this particular case, and who knows how many other cases with similar needs will be broken when 6.0 is released?

It should be noted that WinXP tried to get fancy in a similar way with automatic powerdown of devices, and broke these PERC devices in a similar way. Due to restrictions of the MS driver framework, the only solution that Adaptec could use was to modify the firmware to make the bridge be opaque. This solved the issue of the OS seeing devices that belong to the firmware, but made it impossible to run the controller in split-channel mode, where one channel is for RAID and the other channel is pure SCSI. So the next layer of hacks was to force the 'non-RAID' channel to be controlled by the RAID firmware and be a child of the RAID driver. This has led to endless problems since the RAID firmware doesn't pass SCSI commands through very well. As a side note, this is exactly why I recommend PERC owners to refrain from using version 2.8 firmware. Anyways, the moral of the story is to not be like Microsoft.

Scott

freebsd-current@freebsd.org mailing list

<http://lists.freebsd.org/mailman/listinfo/freebsd-current>

To unsubscribe, send any mail to "freebsd-current-unsubscribe@freebsd.org"