openQA Infrastructure - action #115547

openqaworker20 fails to boot, broken hardware size:M

2022-08-19 08:32 - favogt

Status: Feedback
Priority: High
Assignee: nicksinger
Category: Estimated time: 0.00 hour
Target version: Ready

Description

Today I noticed that openqaworker20 was MIA. It didn't respond to ping and the BMC revealed that it got stuck really early during boot, on the BIOS splash screen!

The phase it got stuck in was "DXE--SB Initialization". A reset helped, but after loading the kernel and initrd the system crashed again and got stuck on the BIOS splash, this time during "PCI resource allocation". I did a power cycle and turned off "Quiet boot" in the BIOS settings for good measure and added verbose debug to the kernel cmdline.

Unfortunately it crashes in a rather bad place:

```
[ 2.794996] T1] smpboot: CPU0: AMD EPYC 7543P 32-Core Processor (family: 0x19, model: 0x1,
    stepping: 0x1)
[ 2.802523] T1] ... version: 0
[ 2.806522] T1] ... bit width: 48
[ 2.810522] T1] ... generic registers: 6
[ 2.814522] T1] ... value mask: 0000ffffffffffff
[ 2.818522] T1] ... max period: 00007fffffffffff
[ 2.822522] T1] ... fixed-purpose events: 0
[ 2.826522] T1] ... event mask: 000000000000003f
[ 2.838797] T1] smp: Bringing up secondary CPUs ...
[ 2.842584] T1] x86: Booting SMP configuration:
(stuck here for ~10s, then reset)
```

Booting the older 5.14.21-150400.22-default kernel doesn't work either.

Suggestions

- Disabling some of the CPU's seems to work around the issue, so maybe this is a hardware fault in one or more of the CPU's
- Run a memory test
- Consider updating the BIOS/ firmware
- Visit the servere room, or find someone with access who can examine the machine
- Contact the vendor, likely Delta Computers to take the machine back

Related issues:

Related to openQA Infrastructure - action #111473: Get replacements for imagetester and openqaworker1 size:M added
Resolved 2022-05-23 2022-09-02

Related to openQA Infrastructure - action #115418: Setup ow19+20 to be able t... Workable 2022-08-17

History

#1 - 2022-08-19 08:44 - favogt
- Related to action #111473: Get replacements for imagetester and openqaworker1 size:M added

#2 - 2022-08-19 08:45 - favogt
I also noticed that the BMC Web UI is unable to report sensor readings and reports the fan states as "critical". On ow19, temperature and fan speed readings are available.

#3 - 2022-08-19 08:50 - cdywan
- Related to action #115418: Setup ow19+20 to be able to run MM tests size:M added

2022-09-05
With maxcpus=0 1 the system manages to boot, and then the sensor readings are available as well and look fine. It fails with maxcups=1 already, so it appears to be a severe HW issue with either CPU or RAM.

I also tried booting the 15.3 GA kernel+initrd with GRUB's http module, but that also crashes.

The sensor issues are a red herring, the values probably only show up if the system is fully booted.

Only the fans FAN1, FAN3 and FANA have “health status” in the BMC and it is critical. Maybe just wrong sensor reading (as most sensor data is in fact missing) but it could also be the culprit. That disabling CPUs helped for a while would be in accordance with and overheating system due to fan issues.

Note that I tried powering on the system again. However, the system got stuck even before reaching the bootloader. Then I tried it again and it could reach the bootloader and I could select the UEFI firmware menu from there. While being in the firmware menu, the sensor readings on the BMC interface were unchanged. Unfortunately the UEFI menu itself doesn't show any interesting information (which would e.g. confirm that fans are broken). I only found a few entries in the event log which I've attached.

I could not reach the OS anymore so I could not check any sensors from there. I’ve shut the machine down again as I don't know what I could do next. I suppose as a next step someone with access to SRV1 needs to do a physical check.

The only thing I can think of we could do from our side is a firmware update. However, if it is really a fan that might not be a good idea. In general, if it is really a fan we should likely refrain from doing too much.

Only the fans FAN1, FAN3 and FANA have “health status” in the BMC and it is critical. Maybe just wrong sensor reading (as most sensor data is in fact missing) but it could also be the culprit. That disabling CPUs helped for a while would be in accordance with and overheating system due to fan issues.

Note that I tried powering on the system again. However, the system got stuck even before reaching the bootloader. Then I tried it again and it could reach the bootloader and I could select the UEFI firmware menu from there. While being in the firmware menu, the sensor readings on the BMC interface were unchanged. Unfortunately the UEFI menu itself doesn't show any interesting information (which would e.g. confirm that fans are broken). I only found a few entries in the event log which I've attached.

When I booted the system with maxcpus=0, I got perfectly fine and healthy sensor readings while it was "up".

mkittler wrote:

Only the fans FAN1, FAN3 and FANA have “health status” in the BMC and it is critical. Maybe just wrong sensor reading (as most sensor data is in fact missing) but it could also be the culprit. That disabling CPUs helped for a while would be in accordance with and overheating system due to fan issues.

Note that I tried powering on the system again. However, the system got stuck even before reaching the bootloader. Then I tried it again and it could reach the bootloader and I could select the UEFI firmware menu from there. While being in the firmware menu, the sensor readings on the BMC interface were unchanged. Unfortunately the UEFI menu itself doesn't show any interesting information (which would e.g. confirm that fans are broken). I only found a few entries in the event log which I've attached.

When I booted the system with maxcpus=0, I got perfectly fine and healthy sensor readings while it was "up".

mkittler wrote:
I've just created an infra ticket: https://sd.suse.com/servicedesk/customer/portal/1/SD-96734

Can you add me to CC?

#13 - 2022-08-25 13:47 - tinita
favogt wrote:

Can you add me to CC?

I added you

#14 - 2022-08-26 08:59 - mkittler
We can rule out the broken fan theory. Gerhard checked and fans are running. When accessing the system today, I found it again stuck at the boot screen (SXE--CPU Initialization ... 63). Normally that state during the boot screen is very fast so it was clearly stuck again. Not sure why Gerhard said he could reliably reach GRUB. However, after a power reset I could reach GRUB again and tried also maxcpus=0. Then sensors also shows normal temperature values and all sensors are showing up green in the BMC as @fvogt mentioned (including the ones under the Fan tab).

Note that @fvogt attempted a network installation and it crashes there as well. So a re-installation of the system doesn't make much sense.

#15 - 2022-08-26 10:41 - mkittler
Btw, that's the hardware we have:

```bash
openqaworker20:~ # inxi -F
System: Host: openqaworker20 Kernel: 5.14.21-150400.24.18-default x86_64 bits: 64 Console: pty pts/0
   Distro: openSUSE Leap 15.4
   Mobo: Supermicro model: H12SSL-i v: 1.02 serial: ZM224S601286 UEFI: American Megatrends v: 2.3 date: 10/20/2021
   CPU: Info: Single Core model: AMD EPYC 7543P bits: 64 type: UP cache: L2: 512 KiB Speed: 1801 MHz min/max: 1500/2800 MHz Core speed (MHz): 1: 1801
          Display: server: No display server data found. Headless machine? tty: 316x82
          Message: Advanced graphics data unavailable in console for root.
   Audio: Message: No device data found.
   Network: Device-1: Intel Ethernet X710 for 10GbE SFP+ driver: i40e
          IF: eth0 state: down mac: 40:a6:b7:3f:b7:74
          Device-2: Intel Ethernet X710 for 10GbE SFP+ driver: i40e
          IF: eth1 state: down mac: 40:a6:b7:3f:b7:75
          Device-3: Broadcom NetXtreme BCM5752 Gigabit Ethernet PCIe driver: tg3
          IF: eth2 state: up speed: 1000 Mbps duplex: full mac: 3c:ec:ef:93:aa:22
         Device-4: Broadcom NetXtreme BCM5752 Gigabit Ethernet PCIe driver: tg3
          IF: eth3 state: down mac: 3c:ec:ef:93:aa:23
          IF-ID-1: usb0 state: down mac: 92:1b:4c:c1:45:13
   Bluetooth: Device-1: Insyde RNDIS/Ethernet Gadget type: USB driver: rndis_host
   Drives: Local Storage: total: 6.99 TiB used: 433.5 GiB (6.1%) ID-1: /dev/nvme0n1 vendor: Samsung model: MZQL23T8HCLS-00A07 size: 3.49 TiB
          ID-2: /dev/nvme0n1 vendor: Samsung model: MZQL23T8HCLS-00A07 size: 3.49 TiB
   Partition: ID-1: / size: 100 GiB used: 5.97 GiB (6.0%) fs: btrfs dev: /dev/nvme0n1p2
          ID-2: /boot/efi size: 511 MiB used: 5 MiB (1.0%) fs: vfat dev: /dev/nvme0n1p1
          ID-3: /home size: 100 GiB used: 5.97 GiB (6.0%) fs: btrfs dev: /dev/nvme0n1p2
          ID-4: /opt size: 100 GiB used: 5.97 GiB (6.0%) fs: btrfs dev: /dev/nvme0n1p2
          ID-5: /tmp size: 100 GiB used: 5.97 GiB (6.0%) fs: btrfs dev: /dev/nvme0n1p2
          ID-6: /var size: 100 GiB used: 5.97 GiB (6.0%) fs: btrfs dev: /dev/nvme0n1p2
   Swap: Alert: No swap data was found.
   Sensors: Message: No sensor data found. Is lm-sensors configured?
   Info: Processes: 190 Uptime: 1h 37m Memory: 503.57 GiB used: 6.19 GiB (1.2%) Init: systemd runlevel: 3 Shell: Bash
inxi: 3.3.0
```

And that's the chassis we have: SuperMicro 825BTQC-R1K23LPB (from order)

#16 - 2022-08-26 11:43 - mkittler
- Assignee changed from mkittler to nicksinger

Looks like we have to contact delta computers.

#17 - 2022-08-31 08:40 - nicksinger
I contacted Delta. They offered us to send back the server so they can do further debugging in their Lab. I just wrote Gerhard if he can clarify who can handle the package handling back to delta.

**Files**

<table>
<thead>
<tr>
<th>Name</th>
<th>Size</th>
<th>Date</th>
<th>User</th>
</tr>
</thead>
<tbody>
<tr>
<td>ow20-events.webm</td>
<td>390 KB</td>
<td>2022-08-25</td>
<td>mkittler</td>
</tr>
</tbody>
</table>