Subject: megaraid_mbox: garbage in file
Posted by vaverin on Thu, 04 May 2006 18:46:04 GMT
View Forum Message <> Reply to Message

Hello all,

I've investigated customers claim on the unstable work of their node and found a strange effect: reading from some files leads to the
"attempt to access beyond end of device" messages.

I've checked filesystem, memory on the node, motherboard BIOS version, but it does not help and issue still has been reproduced by simple file reading.

Reproducer is simple:

```
echo 0xffffffff >/proc/sys/dev/scsi/logging_level ;
cat /vz/private/101/root/etc/ld.so.cache >/tmp/ttt  ;
echo 0 >/proc/sys/dev/scsi/logging
```

It leads to the following messages in dmesg

```
sd_init_command: disk=sda, block=871769260, count=26
sda : block=871769260
sda : reading 26/26 512 byte blocks.
scsi_add_timer: scmd: f79ed980, time: 7500, (c02b1420)
sd 0:1:0:0: send 0xf79ed980                sd 0:1:0:0:
        command: Read (10): 28 00 33 f6 24 ac 00 00 1a 00
buffer = 0xf7cfb540, bufflen = 13312, done = 0xc0366b40, queuecommand 0xc0344010
leaving scsi_dispatch_cmnd()
scsi_delete_timer: scmd: f79ed980, rtn: 1
sd 0:1:0:0: done 0xf79ed980 SUCCESS        0 sd 0:1:0:0:
        command: Read (10): 28 00 33 f6 24 ac 00 00 1a 00
scsi host busy 1 failed 0
sd 0:1:0:0: Notifying upper driver of completion (result 0)
sd_rw_intr: sda: res=0x0
26 sectors total, 13312 bytes done.
use_sg is 4
attempt to access beyond end of device
sda6: rw=0, want=1044134458, limit=951401367
Buffer I/O error on device sda6, logical block 522067228
attempt to access beyond end of device
sda6: rw=0, want=1178878530, limit=951401367
Buffer I/O error on device sda6, logical block 589439264
...
```

As far as I see first read operation has finished without errors, but when we read the rest of file we get an access to beyond end of device.

Originally it was found on Virtuozzo kernels (2.6.8.1-based x86 32-bit),
reproduced on RHEL4 kernels 2.6.9-22.EL and 2.6.9-34.EL,
on FC5 (2.6.16-1.2096_FC5) and on vanilla 2.6.16 kernels.

However, when I first read these blocks by using dd with bs=512 or 1024 it works
without any troubles. Then I can cat this file, copy it, map it and so on -- and
get correct content without any errors. Moreover, this issue may be workarounded
by memory limitation: it helps to use mem=4G in kernel commandline or kernels
without PAE support.

I've noticed that we attempt to access to the blocks with a strange numbers:

522067228 = 0x1f1e1d1c
589439264 = 0x23222120 and so on.

Then I've found that I've read strange garbage from file:

```
# hexdump /tmp/ttt
0000000 0100 0302 0504 0706 0908 0b0a 0d0c 0f0e
0000010 1110 1312 1514 1716 1918 1b1a 1d1c 1f1e
0000020 2120 2322 2524 2726 2928 2b2a 2d2c 2f2e
0000030 3130 3332 3534 3736 3938 3b3a 3d3c 3f3e
0000040 4140 4342 4544 4746 4948 4b4a 4d4c 4f4e
0000050 5150 5352 5554 5756 5958 5b5a 5d5c 5f5e
0000060 6160 6362 6564 6766 6968 6b6a 6d6c 6f6e
0000070 7170 7372 7574 7776 7978 7b7a 7d7c 7f7e
0000080 0100 0302 0504 0706 0908 0b0a 0d0c 0f0e
0000090 1110 1312 1514 1716 1918 1b1a 1d1c 1f1e
00000a0 2120 2322 2524 2726 2928 2b2a 2d2c 2f2e
...
00000f0 7170 7372 7574 7776 7978 7b7a 7d7c 7f7e
0000100 0100 0302 0504 0706 0908 0b0a 0d0c 0f0e
...
```

Then I've discovered that "access beyond end of device" occurs due reading of
the same garbage from the 13-th (Indirect) block of the file.

I've tried to understand where we got this garbage and found that it is present
in the data buffers beginning at megaraid_mbox driver functions.

Could somebody explain me what is the strange garbage: repeated 0...127?
Seokmann, Atul, could you please tell me if it is a known issue?
James, from my point of view it is not looks like a driver bug, but probably I'm
wrong?

I suppose it is MegaRAID SATA 150-4 firmware issue. I've seen similar firmware
fixes for MegaRAID SATA 300 controllers ("Support PAE mode fixed" and "Fixed the
operating systems using more than 4 gig of memory"). Is it probably the same

issues are present in SATA 150-4 firmware? Or may be I use broken controller?

Hardware Environment:
Tyan B2881
2 x Opteron 246
8G RAM
LSI MegaRAID SATA 150-4
/vz partition formatted as ext3 with 1Kb blocksize

megaraid cmm: 2.20.2.6 (Release Date: Mon Mar 7 00:01:03 EST 2005)
megaraid: 2.20.4.7 (Release Date: Mon Nov 14 12:27:22 EST 2005)
megaraid: probe new device 0x1000:0x1960:0x1000:0x4523: bus 1:slot 4:func 0
ACPI: PCI Interrupt 0000:01:04.0[A] -> GSI 29 (level, low) -> IRQ 16
megaraid: fw version:[713N] bios version:[G119]
scsi0 : LSI Logic MegaRAID driver
scsi[0]: scanning scsi channel 0 [Phy 0] for non-raid devices
scsi[0]: scanning scsi channel 1 [virtual] for logical drives
  Vendor: MegaRAID Model: LD 0 RAID1 476G Rev: 713N
  Type: Direct-Access   ANSI SCSI revision: 02

Also I would note that from my point of view this issue looks similar to
http://bugzilla.kernel.org/show_bug.cgi?id=6052

It seems for me both of our cases may have the same cause.

Thank you,
 Vasily Averin

SWsoft Virtuozzo/OpenVZ Linux kernel team