

# Linux kernel debugging for sysadmins

---

Minto Joseph  
August, 2017

# Agenda

- Investigation
- Common issues
- Tools

# Why?

- To learn more about the systems we manage
- For debugging efficiently
- For better Root cause Analysis

**Help !**  
**My system is not responding**

# Investigation

## Before reboot

- System or Service
- Try to see the screen before rebooting. (HP iLO, Dell DRAC, VM console, Monitor, KVM etc..)
- Network connectivity. Is only networking affected? (ethtool, ifconfig, logs)
- **SysRq**

# Investigation

## After reboot

- Check syslog
- Check sysstat
- **Check Vmcore, if available**

# Panics

```
Dec 13 10:34:53 server1 kernel: Kernel BUG at exit.c:904
Dec 13 10:34:53 server1 kernel: invalid operand: 0000 [1] SMP
Dec 13 10:34:53 server1 kernel: CPU 1
Dec 13 10:34:53 server1 kernel: Modules linked in: nfs nfsd exportfs lockd nfs_acl parport_pc lp parport autofs4
i2c_dev i2c_core ipmi_devintf ipmi_si ipmi_msghandler sunrpc ds yenta_so
cket pcmcia_core cpufreq_powersave mptctl ide_dump scsi_dump diskdump zlib deflate dm_mirror dm_mod button
battery ac joydev usbnet mii md5 ipv6 uhci_hcd ehci_hcd hw_random bnx2 bonding(U)
ext3 jbd mptscsih mptsas mptspi mptscsi mptbase sd_mod scsi_mod
Dec 13 10:34:53 server1 kernel: Pid: 5441, comm: httpd Not tainted 2.6.9-78.ELsmp
Dec 13 10:34:53 server1 kernel: RIP: 0010:[<ffffffff8013a844>] <ffffffff8013a844>{next_thread+12}
Dec 13 10:34:53 server1 kernel: RSP: 0018:000001004bca7f00 EFLAGS: 00010046
Dec 13 10:34:53 server1 kernel: RAX: 0000000000000000 RBX: 00000101b0232030 RCX: 0000003248b34b28
Dec 13 10:34:53 server1 kernel: RDX: 0000000000000064 RSI: 000000000068cbf4 RDI: 000001004f5b77f0
Dec 13 10:34:53 server1 kernel: RBP: 0000000000000003 R08: 0000003248a0f080 R09: 0000000000000002
Dec 13 10:34:53 server1 kernel: R10: 0000003248b34b28 R11: 0000000000000202 R12: 0000000000000004
Dec 13 10:34:53 server1 kernel: R13: 000000000066a0d8 R14: 0000000000000000 R15: 000000004e610490
Dec 13 10:34:53 server1 kernel: FS: 000000004e616960(005b) GS:ffffffff8050d300(0000) knlGS:0000000000000000
Dec 13 10:34:53 server1 kernel: CS: 0010 DS: 0000 ES: 0000 CR0: 000000008005003b
Dec 13 10:34:53 server1 kernel: CR2: 0000000082635000 CR3: 0000000006968000 CR4: 000000000000006e0
Dec 13 10:34:53 server1 kernel: Process httpd (pid: 5441, threadinfo 000001004bca6000, task 00000101b0232030)
Dec 13 10:34:53 server1 kernel: Stack: ffffffff80146a6d 0000000000000000 0000000000000000 000001036fa29880
Dec 13 10:34:53 server1 kernel: 0000000000000007 00000101a67efbc0 0000002a958b4e60 ffffffff80000000
Dec 13 10:34:53 server1 kernel: 0000000000000008 000000000006b7520
Dec 13 10:34:53 server1 kernel: Call Trace:<ffffffff80146a6d>{sys_times+103} <ffffffff801102f6>{system_call+126}
Dec 13 10:34:53 server1 kernel:
Dec 13 10:34:53 server1 kernel: Code: 0f 0b 08 1f 33 80 ff ff ff ff 88 03 8b 80 08 08 00 00 85 c0
Dec 13 10:34:53 server1 kernel: RIP <ffffffff8013a844>{next_thread+12} RSP <000001004bca7f00>
```

# Lockups

```
BUG: soft lockup - CPU#0 stuck for 10s! [swapper:0]
Pid: 0, comm:          swapper
EIP: 0060:[<c061dc4e>] CPU: 0EIP is at _spin_lock_bh+0xf/0x18
EFLAGS: 00000286   Not tainted (2.6.18-194.3.1.el5 #1)
EAX: c0748000 EBX: f69fe550 ECX: f6f9a52c EDX: f69fe000
ESI: f69fe4fc EDI: f69fe550 EBP: f69418c0 DS: 007b ES: 007b
CR0: 8005003b CR2: 0069c270 CR3: 00742000 CR4: 000006d0
[<f8b240ea>] rlb_arp_recv+0x98/0x11d [bonding]
[<c05c0aa8>] netif_receive_skb+0x3ac/0x401
[<f8a33bf8>] bnx2_poll_work+0xc3b/0xd45 [bnx2]
[<c041000c>] mtrr_bp_init+0x1f7/0x21a
[<c041db0c>] kmap_atomic_to_page+0x34/0x54
[<c041f79d>] try_to_wake_up+0x3e8/0x3f2
[<c043887c>] hrtimer_run_queues+0xef/0x176
[<c042d5f5>] lock_timer_base+0x15/0x2f
[<f8a3406f>] bnx2_poll+0xbd/0x1ce [bnx2]
[<c05c2995>] net_rx_action+0x9c/0x1a7
[<c042a377>] __do_softirq+0x87/0x114
[<c04073cf>] do_softirq+0x52/0x9c
[<c044f158>] __do_IRQ+0x0/0xd6
[<c04074ce>] do_IRQ+0xb5/0xc3
[<c0405946>] common_interrupt+0x1a/0x20
[<c0403ce7>] mwait_idle+0x25/0x38
[<c0403ca8>] cpu_idle+0x9f/0xb9
[<c07099fa>] start_kernel+0x37b/0x383
```



# hung\_task

```
Dec 21 09:56:53 kernel: INFO: task oracleasm-read-:6138 blocked for more than 120 seconds.
Dec 21 09:56:53 kernel: "echo 0 > /proc/sys/kernel/hung_task_timeout_secs" disables this
message.
Dec 21 09:56:53 kernel: oracleasm-rea D ffffffff80151248 0 6138 6137 (NOTLB)
Dec 21 09:56:53 kernel: ffff8102793c3cb8 0000000000000082 0000000000000400
ffffffff8001c211
Dec 21 09:56:53 botein kernel: ffff8102793c3c98 0000000000000004 ffff81027c7b00c0
ffff810108d22080
Dec 21 09:56:53 botein kernel: 0000002d96f8074c 00000000000050d4f ffff81027c7b02a8
0000000200000003
Dec 21 09:56:53 botein kernel: Call Trace:
Dec 21 09:56:53 botein kernel: [<ffffffff8001c211>] generic_make_request+0x211/0x228
Dec 21 09:56:53 botein kernel: [<ffffffff8006f1f5>] do_gettimeofday+0x40/0x90
Dec 21 09:56:53 botein kernel: [<ffffffff800647ea>] io_schedule+0x3f/0x67
Dec 21 09:56:53 botein kernel: [<ffffffff800f5824>] __blockdev_direct_IO+0x8da/0xa80
Dec 21 09:56:53 botein kernel: [<ffffffff800e6859>] blkdev_direct_IO+0x32/0x37
Dec 21 09:56:53 botein kernel: [<ffffffff800e6791>] blkdev_get_blocks+0x0/0x96
Dec 21 09:56:53 botein kernel: [<ffffffff8000c514>] __generic_file_aio_read+0xb8/0x198
Dec 21 09:56:53 botein kernel: [<ffffffff800c78fb>] generic_file_read+0xac/0xc5
Dec 21 09:56:53 botein kernel: [<ffffffff800a1ba4>] autoremove_wake_function+0x0/0x2e
Dec 21 09:56:53 botein kernel: [<ffffffff8002a6d0>] __vma_link+0x42/0x4b
Dec 21 09:56:53 botein kernel: [<ffffffff8001cca6>] vma_link+0x70/0xfd
Dec 21 09:56:53 botein kernel: [<ffffffff800b878c>] audit_syscall_entry+0x180/0x1b3
Dec 21 09:56:53 botein kernel: [<ffffffff8000b6b0>] vfs_read+0xcb/0x171
Dec 21 09:56:53 botein kernel: [<ffffffff80013626>] sys_pread64+0x50/0x70
Dec 21 09:56:53 botein kernel: [<ffffffff8005e229>] tracesys+0x71/0xe0
Dec 21 09:56:53 botein kernel: [<ffffffff8005e28d>] tracesys+0xd5/0xe0
```

# Out of memory

May 4 17:15:21 tsp2850a kernel: Free pages: 663984kB (650752kB HighMem)  
**May 4 17:15:21 tsp2850a kernel: Active:969716 inactive:730309 dirty:53518 writeback:56 unstable:0 free:165996 slab:203048 mapped:213066 pagetables:1341**  
May 4 17:15:21 tsp2850a kernel: DMA free:12576kB min:16kB low:32kB high:48kB active:0kB inactive:0kB present:16384kB pages\_scanned:17312 all\_unreclaimable? yes  
May 4 17:15:21 tsp2850a kernel: protections[]: 0 0 0  
**May 4 17:15:22 tsp2850a kernel: Normal free:656kB min:928kB low:1856kB high:2784kB active:104kB inactive:1072kB present:901120kB pages\_scanned:5016 all\_unreclaimable? yes**  
May 4 17:15:22 tsp2850a kernel: protections[]: 0 0 0  
May 4 17:15:22 tsp2850a kernel: HighMem free:650752kB min:512kB low:1024kB high:1536kB active:3878760kB inactive:2920164kB present:8519680kB pages\_scanned:0 all\_unreclaimable? no  
May 4 17:15:22 tsp2850a kernel: protections[]: 0 0 0 May 4 17:15:22 tsp2850a kernel: DMA: 2\*4kB 5\*8kB 3\*16kB 4\*32kB 3\*64kB 1\*128kB 1\*256kB 1\*512kB 1\*1024kB 1\*2048kB 2\*4096kB = 12576kB  
**May 4 17:15:22 tsp2850a kernel: Normal: 0\*4kB 28\*8kB 27\*16kB 0\*32kB 0\*64kB 0\*128kB 0\*256kB 0\*512kB 0\*1024kB 0\*2048kB 0\*4096kB = 656kB**  
May 4 17:15:22 tsp2850a kernel: HighMem: 0\*4kB 0\*8kB 0\*16kB 0\*32kB 0\*64kB 3330\*128kB 523\*256kB 117\*512kB 30\*1024kB 0\*2048kB 0\*4096kB = 650752kB  
May 4 17:15:22 tsp2850a kernel: Swap cache: add 153154, delete 151772, find 11197/16185, race 0+0 May 4 17:15:22 tsp2850a kernel: 0 bounce buffer pages  
**May 4 17:15:22 tsp2850a kernel: Free swap: 8120648kB**  
May 4 17:15:22 tsp2850a kernel: 2359296 pages of RAM  
May 4 17:15:22 tsp2850a kernel: 1867710 pages of HIGHMEM  
May 4 17:15:22 tsp2850a kernel: 282147 reserved pages  
May 4 17:15:22 tsp2850a kernel: 1254777 pages shared  
May 4 17:15:22 tsp2850a kernel: 1382 pages swap cached  
**May 4 17:15:22 tsp2850a kernel: Out of Memory: Killed process 7944 (mysqld).**

# Hangs

- Hardware issues (MCE - Machine Check Exceptions, EDAC - Error Detection And Correction)

*Aug 30 15:13:51 dxv9bw1 kernel: [Hardware Error]: Machine check events logged*

*Check /var/log/mcelog*

*Check Hypervisor for Vms.*

- CPU, Memory or I/O utilization

# Vmcore

**crashkernel** parameter in grub.conf

## **Panic**

netdump, diskdump

## **kdump**

To cause panic:

## **Soft lockup**

kernel.softlockup\_panic=1

## **Out of Memory**

vm.panic\_on\_oom = 1

## **Hung\_task**

kernel.hung\_task\_panic = 1

## **Hangs**

alt+sysrq+c

# Thank you - Q&A



joseph@pythian.com