Debugging Embedded Linux Systems: Understand Kernel Oops Logs

Debugging Embedded Linux Training Series [Part 6]

Debugging Embedded Linux Training Series

- Part 1: Linux/Kernel Overview
- Part 2: Kernel Logging System Overview
- Part 3: printk and Variations
- Part 4: Dynamic Debug
- Part 5: Locate Device Driver Source Code
- **Part 6: Understand Kernel Oops Logs**

Agenda

- What is kernel oops?
- Kernel oops log structure
- Tools to locate errors in source code
- Oops log examples

What is kernel Oops?

- Deviation from correct behavior of the Linux kernel
- Produces certain error messages in kernel logs
- Why does kernel generate oops logs?
- Severity varies:
  - `panic() --> WARN()`

Kernel oops log structure

- Error Summary
- Error Type
- CPU#/PID#/Kernel-Version
- Hardware
- CPU Register Dump
  - PC/LR
- Stack Dump
- Backtrace

Kernel oops log structure example

Error Summary ->

```
Unable to handle kernel NULL pointer dereference at virtual address 00000000
pgd = eeda0000
[00000000] *pgd=aedb8831, *pte=00000000, *ppte=00000000
```

Error Type ->

```
Internal error: Oops: 817 [#1] PREEMPT ARM
Modules linked in: musb_am335x(+) rtc_omap omap_wdt ti_am335x_tscadc matrix_keypad matrix_keymap
```

CPU#/PID#/kernel-Version ->

```
CPU: 0 PID: 135 Comm: udevd Not tainted 4.4.48-02799-g2f0993afde90-dirty #440
```

Hardware ->

```
Hardware name: Generic AM33XX (Flattened Device Tree)
task: eeeaa400 ti: eeeda000 task.ti: eeeda000
```

CPU Register Dump ->

```
PC is at am335x_child_probe+0x2c/0x58 [musb_am335x]
LR is at am335x_child_probe+0x24/0x58 [musb_am335x]
pc : [<bf01902c>]    lr : [<bf019024>]    psr: 600b0013
sp : eeedbcb8  ip : eeedbcb8  fp : eeedbccc
r10: 00000000  r9 : 0000000e  r8 : bf019230
r7 : fffffdfb  r6 : bf019230  r5 : ee99aa00  r4 : ee99aa10
r3 : ee99aa10  r2 : 00000000  r1 : 00000001  r0 : ee99aa10
Flags: nZCv  IRQs on  FIQs on  Mode SVC_32  ISA ARM  Segment none
Control: 10c5387d  Table: aeda0019  DAC: 00000051
Process udevd (pid: 135, stack limit = 0xeeeda210)
```

Stack Dump ->

```
Stack: (0xeeedbcb8 to 0xeeedc000)
bca0:                                                       ee99aa10 ee99aa10
bcc0: eeedbcec eeedbcd0 c03772b4 bf01900c ee99aa10 c089f530 00000000 c08686b0
...
```

Backtrace ->

```
Backtrace:
[<bf019000>] (am335x_child_probe [musb_am335x]) from [<c03772b4>] (platform_drv_probe+0x5c/0xc0)
[<c0377258>] (platform_drv_probe) from [<c037501c>] (driver_probe_device+0x228/0x484)
[<c0374df4>] (driver_probe_device) from [<c0375314>] (__driver_attach+0x9c/0xa0)
[<c0375278>] (__driver_attach) from [<c0372dbc>] (bus_for_each_dev+0x7c/0xb0)
[<c0372d40>] (bus_for_each_dev) from [<c0374934>] (driver_attach+0x28/0x30)
```

Tools for locating errors in source code

- gdb

  list command

- addr2line

  -fe option

- objdump

  -dS option

Locate errors example 1: Kernel (1)

```
Unable to handle kernel NULL pointer dereference at virtual address 00000000
pgd = eeecc000
[00000000] *pgd=aedb8831, *pte=00000000, *ppte=00000000
Internal error: Oops: 817 [#1] PREEMPT ARM
Modules linked in: musb_am335x(+) rtc_omap omap_wdt ti_am335x_tscadc
matrix_keypad matrix_keymap
CPU: 0 PID: 125 Comm: udevd Not tainted 4.4.48-02799-g2f0993afde90-dirty #446
Hardware name: Generic AM33XX (Flattened Device Tree)
task: eeddc000 ti: eeec6000 task.ti: eeec6000
PC is at cppi41_dma_probe+0x2c4/0x52c
LR is at 0x0
pc : [<c0328614>]    lr : [<00000000>]    psr: 60000013
...
```

Locate errors example 1: Kernel (2)

```
$ gdb vmlinux
(gdb) list *(cppi41_dma_probe+0x2c4)
```

Locate errors example 1: Kernel (2)

```
$ gdb vmlinux
(gdb) list *(cppi41_dma_probe+0x2c4)
```

```
0xc0328614 is in cppi41_dma_probe (drivers/dma/cppi41.c:679).
674                     cchan = kzalloc(sizeof(*cchan), GFP_KERNEL);
675                     if (!cchan)
676                             goto err;
677
678                     cchan = 0;
679                     cchan->cdd = cdd;
680                     if (i & 1) {
681                             cchan->gcr_reg = cdd->ctrl_mem + DMA_TXGCR(i >> 1);
682                             cchan->is_tx = 1;
683                     } else {
```

Locate errors example 1: Kernel (3)

```
...
PC is at cppi41_dma_probe+0x2c4/0x52c
LR is at 0x0
pc : [<c0328614>]    lr : [<00000000>]    psr: 60000013
...
```

Locate errors example 1: Kernel (3)

```
...
PC is at cppi41_dma_probe+0x2c4/0x52c
LR is at 0x0
pc : [<c0328614>]    lr : [<00000000>]    psr: 60000013
...
```

```
$ arm-linux-gnueabihf-addr2line -fe vmlinux c0328614
```

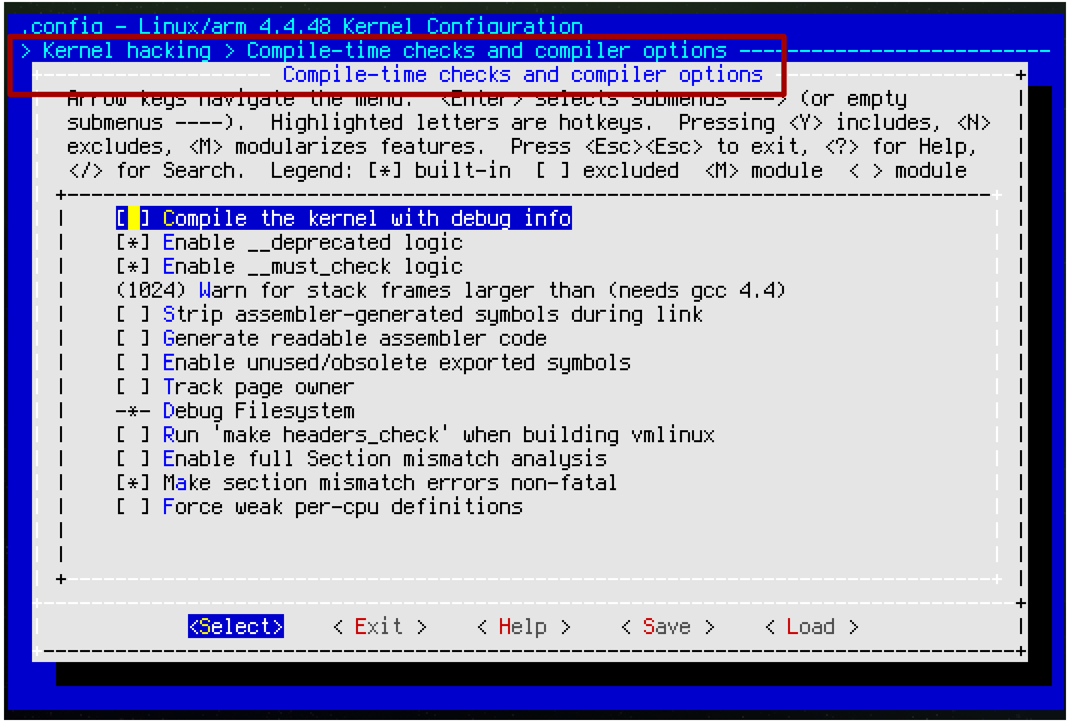
Locate errors example 1: Kernel (3)

```
...
PC is at cppi41_dma_probe+0x2c4/0x52c
LR is at 0x0
pc : [<c0328614>]    lr : [<00000000>]    psr: 60000013
...
```

$ arm-linux-gnueabihf-addr2line -fe vmlinux c0328614

```
cppi41_add_chans
drivers/dma/cppi41.c:679
```

Locate errors: Kernel config



Locate errors example 2: Module (1)

```
Unable to handle kernel NULL pointer dereference at virtual address 00000000
pgd = eeda0000
[00000000] *pgd=aedb8831, *pte=00000000, *ppte=00000000
Internal error: Oops: 817 [#1] PREEMPT ARM
Modules linked in: musb_am335x(+) rtc_omap omap_wdt ti_am335x_tscadc
matrix_keypad matrix_keymap
CPU: 0 PID: 135 Comm: udevd Not tainted 4.4.48-02799-g2f0993afde90-dirty #440
Hardware name: Generic AM33XX (Flattened Device Tree)
task: eeeaa400 ti: eeeda000 task.ti: eeeda000
PC is at am335x_child_probe+0x2c/0x58 [musb_am335x]
LR is at am335x_child_probe+0x24/0x58 [musb_am335x]
...
```

Locate errors example 2: Module (2)

```
$ gdb drivers/usb/musb/musb_am335x.ko
(gdb) list *(am335x_child_probe+0x2c)
```

Locate errors example 2: Module (2)

```
$ gdb drivers/usb/musb/musb_am335x.ko
(gdb) list *(am335x_child_probe+0x2c)
```

```
0x2c is in am335x_child_probe (drivers/usb/musb/musb_am335x.c:12).
7       {
8               int ret;
9
10              pm_runtime_enable(&pdev->dev);
11
12              *(int*)0 = 0;
13              ret = of_platform_populate(pdev->dev.of_node, NULL, NULL, &pdev->dev);
14              if (ret)
15                      goto err;
```

Locate errors example 3: NULL pointer in workqueue (1)

```
Unable to handle kernel NULL pointer dereference at virtual address 00000000
Internal error: Oops: 207 [#1] PREEMPT SMP ARM
CPU: 0 PID: 21548 Comm: kworker/u4:1 Not tainted 4.1.18-rt17-yocto-standard #1
Hardware name: Generic DRA74X (Flattened Device Tree)
PC is at pwq_activate_delayed_work+0x38/0xe8
LR is at pwq_dec_nr_in_flight+0x84/0xe4
pc : [<c004ee14>]    lr : [<c0051a4c>]    psr: 60000013
sp : edff5ea0  ip : edff5ec0  fp : edff5ebc
Flags: nZCv  IRQs on  FIQs on  Mode SVC_32  ISA ARM  Segment kernel
Control: 30c5387d  Table: adc1d700  DAC: fffffffd
Process kworker/u4:1 (pid: 21548, stack limit = 0xedff4218)
Backtrace:
[<c004eddc>] (pwq_activate_delayed_work) from [<c0051a4c>] (pwq_dec_nr_in_flight+0x84/0xe4)
[<c00519c8>] (pwq_dec_nr_in_flight) from [<c005223c>] (process_one_work+0x1e0/0x478)
[<c005205c>] (process_one_work) from [<c0052528>] (worker_thread+0x54/0x510)
[<c00524d4>] (worker_thread) from [<c0057844>] (kthread+0xdc/0xf4)
[<c0057768>] (kthread) from [<c000fdb8>] (ret_from_fork+0x14/0x3c)
Code: e34c30ae e5932004 e3520000 ca00000b (e5961000)
---[ end trace 0000000000000002 ]---
Kernel panic - not syncing: Fatal exception
```

Locate errors example 3: NULL pointer in workqueue (2)

kernel/workqueue.c:

```
static void pwq_activate_delayed_work(struct work_struct *work)
{
        struct pool_workqueue *pwq = get_work_pwq(work);

        trace_workqueue_activate_work(work);
        move_linked_works(work, &pwq->pool->worklist, NULL);
        __clear_bit(WORK_STRUCT_DELAYED_BIT, work_data_bits(work));
        pwq->nr_active++;
}
```

The work is cancelled.
So pwq is NULL here.

pwq->pool: illegal access

L3 custom error

```
[0.210757] omap_l3_noc 44000000.ocp: L3 debug error: target 8 mod:0 (unclearable)
[0.210778] omap_l3_noc 44000000.ocp: L3 debug error: target 13 mod:1 (unclearable)
[0.210817] omap_l3_noc 44000000.ocp: L3 application error: target 8 mod:0 (unclearable)
[0.210830] ------------[ cut here ]------------
[0.210848] WARNING: CPU: 0 PID: 1 at drivers/bus/omap_l3_noc.c:147
           l3_interrupt_handler+0x220/0x34c()
[0.210856] 44000000.ocp:L3 Custom Error: MASTER M2 (64-bit) TARGET L4_WKUP (Read):
           Data Access in User mode during Functional access
```

- Due to access clock-gated module:
  - Improper dts change
  - Driver bug, mainly in runtime PM
- Or due to wrong access address:
  - Incorrect hardware design. For example, DDR size, RTC, etc.
  - Driver bug

Locate errors example 4: Access clock gated module - AES

```
WARNING: CPU: 0 PID: 0 at drivers/bus/omap_l3_noc.c:147 l3_interrupt_handler+0x25c/0x368()
44000000.ocp:L3 Standard Error: MASTER MPU TARGET AES1 (Read): At Address: 0x00100054 :
           Data Access in Supervisor mode during Functional access
Modules linked in: algif_skcipher af_alg bc_example(O) rpmsg_proto rpmsg_pru rpmsg_rpc
CPU: 0 PID: 0 Comm: swapper/0 Tainted: G           O    4.4.12-g3639bea54a #1
Hardware name: Generic DRA74X (Flattened Device Tree)
Backtrace:
[<c00130c0>] (dump_backtrace) from [<c00132bc>] (show_stack+0x18/0x1c)
...
[<c02d05fc>] (l3_interrupt_handler) from [<c0079228>] (handle_irq_event_percpu+0xb4/0x160)
...
[<c000944c>] (gic_handle_irq) from [<c0013d80>] (__irq_svc+0x40/0x74)
...
[<bf1d5000>] (omap_aes_dma_trigger_omap2 [omap_aes_driver]) from [<bf1d5094>]
           (omap_aes_dma_trigger_omap4+0x34/0x38 [omap_aes_driver])
[<bf1d5060>] (omap_aes_dma_trigger_omap4 [omap_aes_driver]) from [<bf1d62a4>]
           (omap_aes_crypt_dma_start+0x2d4/0x498 [omap_aes_driver])
[<bf1d5fd0>] (omap_aes_crypt_dma_start [omap_aes_driver]) from [<bf1d6860>]
           (omap_aes_handle_queue+0x324/0x390 [omap_aes_driver])
```

Locate errors example 4: Access clock gated module - USB

```
omap_l3_noc 44000000.ocp: L3 debug error: target 5 mod:1 (unclearable)
WARNING: CPU: 0 PID: 823 at drivers/bus/omap_l3_noc.c:147 l3_interrupt_handler+0x24c/0x350()
44000000.ocp:L3 Custom Error: MASTER USB3 TARGET GPMC (Idle): Data Access in User mode during
Functional access
Modules linked in: g_mass_storage usb_f_mass_storage usb_f_ss_lb libcomposite configfs dwc3 CPU: 0
PID: 823 Comm: sh Not tainted 4.4.12-00004-gfb912bf-dirty #34
Hardware name: Generic DRA74X (Flattened Device Tree)
Backtrace:
[<c0012ffc>] (dump_backtrace) from [<c00131f8>] (show_stack+0x18/0x1c)
...
[<c02cd5f4>] (l3_interrupt_handler) from [<c0078d28>] (handle_irq_event_percpu+0x90/0x148)
...
[<c0009434>] (gic_handle_irq) from [<c0013cc0>] (__irq_svc+0x40/0x74)
[<c00807e0>] (resume_irqs) from [<c0080900>] (resume_device_irqs+0x14/0x18)
[<c00808ec>] (resume_device_irqs) from [<c03dce70>] (dpm_resume_noirq+0x210/0x22c)
[<c03dcc60>] (dpm_resume_noirq) from [<c00743e0>] (suspend_devices_and_enter+0x21c/0x508)
[<c00741c4>] (suspend_devices_and_enter) from [<c0074960>] (pm_suspend+0x294/0x310)
[<c00746cc>] (pm_suspend) from [<c00734c4>] (state_store+0x70/0xc0)
...
[<c0115814>] (SyS_write) from [<c000f9e0>] (ret_fast_syscall+0x0/0x34)
```

Locate errors example 6: Spinlock dead lock (1)

```
=============================================
[ INFO: possible recursive locking detected ]
4.6.0-08691-g7f3db9a #37 Not tainted
---------------------------------------------
usb/733 is trying to acquire lock:
 (&(&dev->lock)->rlock){-.....}, at: [<bf129288>] ep0_complete+0x18/0xdc [gadgetfs]


but task is already holding lock:
 (&(&dev->lock)->rlock){-.....}, at: [<bf12a420>] ep0_read+0x20/0x5e0 [gadgetfs]

 *** DEADLOCK ***


 May be due to missing lock nesting notation


2 locks held by usb/733:
 #0:  (&f->f_pos_lock){+.+.+.}, at: [<c02a6114>] __fdget_pos+0x40/0x48
 #1:  (&(&dev->lock)->rlock){-.....}, at: [<bf12a420>] ep0_read+0x20/0x5e0 [gadgetfs]
```

Locate errors example 6: Spinlock dead lock (2)

stack backtrace:

```
CPU: 0 PID: 733 Comm: usb Not tainted 4.6.0-08691-g7f3db9a #37
Hardware name: Generic AM33XX (Flattened Device Tree)
[<c010ffbc>] (unwind_backtrace) from [<c010c1bc>] (show_stack+0x10/0x14)
[<c010c1bc>] (show_stack) from [<c04207fc>] (dump_stack+0xb0/0xe4)
[<c04207fc>] (dump_stack) from [<c01886ec>] (__lock_acquire+0xf68/0x1994)
[<c01886ec>] (__lock_acquire) from [<c0189528>] (lock_acquire+0xd8/0x238)
[<c0189528>] (lock_acquire) from [<c06ad6b4>] (_raw_spin_lock_irqsave+0x38/0x4c)
[<c06ad6b4>] (_raw_spin_lock_irqsave) from [<bf129288>] (ep0_complete+0x18/0xdc [gadgetfs])
[<bf129288>] (ep0_complete [gadgetfs]) from [<bf10a728>] (musb_g_giveback+0x118/0x1b0 [musb_hdrc])
[<bf10a728>] (musb_g_giveback [musb_hdrc]) from [<bf108768>] (musb_g_ep0_queue+0x16c/0x188
[musb_hdrc])
[<bf108768>] (musb_g_ep0_queue [musb_hdrc]) from [<bf12a944>] (ep0_read+0x544/0x5e0 [gadgetfs])
[<bf12a944>] (ep0_read [gadgetfs]) from [<c0284470>] (__vfs_read+0x20/0x110)
[<c0284470>] (__vfs_read) from [<c0285324>] (vfs_read+0x88/0x114)
[<c0285324>] (vfs_read) from [<c0286150>] (SyS_read+0x44/0x9c)
[<c0286150>] (SyS_read) from [<c0107820>] (ret_fast_syscall+0x0/0x1c)
```

Summary

- Kernel oops log has sufficient information.
- *gdb* or *addr2line* helps to locate the error in source code when kernel debug info is enabled in kernel config.

For more information

- Processor SDK Training Series: http://training.ti.com/processor-sdk-training-series
- Debugging Embedded Linux Training Series: http://training.ti.com/debug-embedded-linux-training-series
- Processor SDK Linux Getting Started Guide: http://processors.wiki.ti.com/index.php/Processor_SDK_Linux_Getting_Started_Guide
- Download Processor SDK Linux for Embedded Processors: http://www.ti.com/processorsdk
- For questions about this training, refer to the E2E Embedded Linux Community Forum: http://e2e.ti.com/support/embedded/linux/f/354