

# Irq balancing problem?

**Source:** <http://linux.derkeiler.com/Mailing–Lists/Kernel/2003–12/1469.html>

---

**From:** Dan Creswell ([dan\\_at\\_dcrdev.demon.co.uk](mailto:dan_at_dcrdev.demon.co.uk))

**Date:** 12/07/03

Date: Sat, 06 Dec 2003 23:12:41 +0000  
To: [linux–kernel@vger.kernel.org](mailto:linux–kernel@vger.kernel.org)

Hi Folks,

I've been testing `–test11` here and come across an odd piece of behaviour. I'm running Fedora Core on a dual PIV Xeon and have suffered a number of stability problems. After some poking around through various archives, I found some discussion about irq balancing with multiple APIC's potentially causing problems.

With that, I first tried passing `noirqbalance` as a kernel parameter – no dice – still get lock ups.

So, then I also disabled the user–space irq balancing daemon (in addition to passing `noirqbalance`) and, voila, my system is now stable.

Even with all this balancing disabled, a `cat /proc/interrupts` shows interrupts being spread across CPU's. My next step will be to re–enable in–kernel irq balancing and see what happens.

Just wondering if anyone has any theories on what's happening/going wrong? I've attached the `dmesg` output below for `–test11`.

Cheers,

Dan.

Dmesg output:

```
a2 elevator=deadline noirqbalance
Initializing CPU#0
PID hash table entries: 4096 (order 12: 32768 bytes)
Detected 2667.013 MHz processor.
Console: colour VGA+ 80x25
Memory: 2071220k/2096576k available (1814k kernel code, 24216k reserved,
649k data, 436k init, 1179072k highmem)
Calibrating delay loop... 5259.26 BogoMIPS
Dentry cache hash table entries: 262144 (order: 8, 1048576 bytes)
Inode–cache hash table entries: 131072 (order: 7, 524288 bytes)
```

## Linux–Kernel: Irq balancing problem?

```
Mount–cache hash table entries: 512 (order: 0, 4096 bytes)
CPU: After generic identify, caps: bfebfbff 00000000 00000000 00000000
CPU: After vendor identify, caps: bfebfbff 00000000 00000000 00000000
CPU: Trace cache: 12K uops, L1 D cache: 8K
CPU: L2 cache: 512K
CPU: Physical Processor ID: 0
CPU: After all inits, caps: bfebfbff 00000000 00000000 00000080
Intel machine check architecture supported.
Intel machine check reporting enabled on CPU#0.
CPU#0: Intel P4/Xeon Extended MCE MSR(12) available
CPU#0: Thermal monitoring enabled
Enabling fast FPU save and restore... done.
Enabling unmasked SIMD FPU exception support... done.
Checking 'hlt' instruction... OK.
POSIX conformance testing by UNIFIX
CPU0: Intel(R) Xeon(TM) CPU 2.66GHz stepping 05
per–CPU timeslice cutoff: 1462.58 usecs.
task migration cache decay timeout: 2 msecs.
enabled ExtINT on CPU#0
ESR value before enabling vector: 00000000
ESR value after enabling vector: 00000000
Booting processor 1/6 eip 2000
Initializing CPU#1
masked ExtINT on CPU#1
ESR value before enabling vector: 00000000
ESR value after enabling vector: 00000000
Calibrating delay loop... 5324.80 BogoMIPS
CPU: After generic identify, caps: bfebfbff 00000000 00000000 00000000
CPU: After vendor identify, caps: bfebfbff 00000000 00000000 00000000
CPU: Trace cache: 12K uops, L1 D cache: 8K
CPU: L2 cache: 512K
CPU: Physical Processor ID: 6
CPU: After all inits, caps: bfebfbff 00000000 00000000 00000080
Intel machine check architecture supported.
Intel machine check reporting enabled on CPU#1.
CPU#1: Intel P4/Xeon Extended MCE MSR(12) available
CPU#1: Thermal monitoring enabled
CPU1: Intel(R) Xeon(TM) CPU 2.66GHz stepping 05
Total of 2 processors activated (10584.06 BogoMIPS).
WARNING: No sibling found for CPU 0.
WARNING: No sibling found for CPU 1.
ENABLING IO–APIC IRQs
Setting 2 in the phys_id_present_map
...changing IO–APIC physical APIC ID to 2 ... ok.
Setting 3 in the phys_id_present_map
...changing IO–APIC physical APIC ID to 3 ... ok.
Setting 4 in the phys_id_present_map
...changing IO–APIC physical APIC ID to 4 ... ok.
init IO_APIC IRQs
IO–APIC (apicid–pin) 2–0, 2–3, 2–5, 2–10, 2–11, 2–20, 2–21, 2–22, 3–2,
3–3, 3–4, 3–5, 3–6, 3–7, 3–8, 3–9, 3–10, 3–11, 3–12, 3–13, 3–14, 3–15,
```

Irq balancing problem?

## Linux-Kernel: Irq balancing problem?

3-16, 3-17, 3-18, 3-19, 3-20, 3-21, 3-22, 3-23, 4-0, 4-1, 4-2, 4-3, 4-4, 4-5, 4-6, 4-7, 4-8, 4-9, 4-10, 4-11, 4-12, 4-13, 4-14, 4-15, 4-16, 4-17, 4-18, 4-19, 4-20, 4-21, 4-22, 4-23 not connected.

..TIMER: vector=0x31 pin1=2 pin2=0

number of MP IRQ sources: 21.

number of IO-APIC #2 registers: 24.

number of IO-APIC #3 registers: 24.

number of IO-APIC #4 registers: 24.

testing the IO APIC.....

IO APIC #2.....

.... register #00: 02000000

..... : physical APIC id: 02

..... : Delivery Type: 0

..... : LTS : 0

.... register #01: 00178020

..... : max redirection entries: 0017

..... : PRQ implemented: 1

..... : IO APIC version: 0020

.... register #02: 00000000

..... : arbitration: 00

.... register #03: 00000001

..... : Boot DT : 1

.... IRQ redirection table:

NR Log Phy Mask Trig IRR Pol Stat Dest Deli Vect:

00 000 00 1 0 0 0 0 0 0 00

01 001 01 0 0 0 0 0 1 1 39

02 001 01 0 0 0 0 0 1 1 31

03 000 00 1 0 0 0 0 0 0 00

04 001 01 0 0 0 0 0 1 1 41

05 000 00 1 0 0 0 0 0 0 00

06 001 01 0 0 0 0 0 1 1 49

07 001 01 0 0 0 0 0 1 1 51

08 001 01 0 0 0 0 0 1 1 59

09 001 01 0 0 0 0 0 1 1 61

0a 000 00 1 0 0 0 0 0 0 00

0b 000 00 1 0 0 0 0 0 0 00

0c 001 01 0 0 0 0 0 1 1 69

0d 001 01 0 0 0 0 0 1 1 71

0e 001 01 0 0 0 0 0 1 1 79

0f 001 01 0 0 0 0 0 1 1 81

10 001 01 1 1 0 1 0 1 1 89

11 001 01 1 1 0 1 0 1 1 91

12 001 01 1 1 0 1 0 1 1 99

13 001 01 1 1 0 1 0 1 1 A1

14 000 00 1 0 0 0 0 0 0 00

15 000 00 1 0 0 0 0 0 0 00

16 000 00 1 0 0 0 0 0 0 00

17 001 01 1 1 0 1 0 1 1 A9

IO APIC #3.....

.... register #00: 03000000

..... : physical APIC id: 03

Irq balancing problem?

## Linux–Kernel: Irq balancing problem?

```
..... : Delivery Type: 0
..... : LTS : 0
.... register #01: 00178020
..... : max redirection entries: 0017
..... : PRQ implemented: 1
..... : IO APIC version: 0020
.... register #02: 03000000
..... : arbitration: 03
.... register #03: 00000001
..... : Boot DT : 1
.... IRQ redirection table:
NR Log Phy Mask Trig IRR Pol Stat Dest Deli Vect:
00 001 01 1 1 0 1 0 1 1 B1
01 001 01 1 1 0 1 0 1 1 B9
02 000 00 1 0 0 0 0 0 0 00
03 000 00 1 0 0 0 0 0 0 00
04 000 00 1 0 0 0 0 0 0 00
05 000 00 1 0 0 0 0 0 0 00
06 000 00 1 0 0 0 0 0 0 00
07 000 00 1 0 0 0 0 0 0 00
08 000 00 1 0 0 0 0 0 0 00
09 000 00 1 0 0 0 0 0 0 00
0a 000 00 1 0 0 0 0 0 0 00
0b 000 00 1 0 0 0 0 0 0 00
0c 000 00 1 0 0 0 0 0 0 00
0d 000 00 1 0 0 0 0 0 0 00
0e 000 00 1 0 0 0 0 0 0 00
0f 000 00 1 0 0 0 0 0 0 00
10 000 00 1 0 0 0 0 0 0 00
11 000 00 1 0 0 0 0 0 0 00
12 000 00 1 0 0 0 0 0 0 00
13 000 00 1 0 0 0 0 0 0 00
14 000 00 1 0 0 0 0 0 0 00
15 000 00 1 0 0 0 0 0 0 00
16 000 00 1 0 0 0 0 0 0 00
17 000 00 1 0 0 0 0 0 0 00
IO APIC #4.....
.... register #00: 04000000
..... : physical APIC id: 04
..... : Delivery Type: 0
..... : LTS : 0
.... register #01: 00178020
..... : max redirection entries: 0017
..... : PRQ implemented: 1
..... : IO APIC version: 0020
.... register #02: 04000000
..... : arbitration: 04
.... register #03: 00000001
..... : Boot DT : 1
.... IRQ redirection table:
NR Log Phy Mask Trig IRR Pol Stat Dest Deli Vect:
```

Irq balancing problem?

## Linux-Kernel: Irq balancing problem?

```
00 000 00 1 0 0 0 0 0 0 0 0
01 000 00 1 0 0 0 0 0 0 0 0
02 000 00 1 0 0 0 0 0 0 0 0
03 000 00 1 0 0 0 0 0 0 0 0
04 000 00 1 0 0 0 0 0 0 0 0
05 000 00 1 0 0 0 0 0 0 0 0
06 000 00 1 0 0 0 0 0 0 0 0
07 000 00 1 0 0 0 0 0 0 0 0
08 000 00 1 0 0 0 0 0 0 0 0
09 000 00 1 0 0 0 0 0 0 0 0
0a 000 00 1 0 0 0 0 0 0 0 0
0b 000 00 1 0 0 0 0 0 0 0 0
0c 000 00 1 0 0 0 0 0 0 0 0
0d 000 00 1 0 0 0 0 0 0 0 0
0e 000 00 1 0 0 0 0 0 0 0 0
0f 000 00 1 0 0 0 0 0 0 0 0
10 000 00 1 0 0 0 0 0 0 0 0
11 000 00 1 0 0 0 0 0 0 0 0
12 000 00 1 0 0 0 0 0 0 0 0
13 000 00 1 0 0 0 0 0 0 0 0
14 000 00 1 0 0 0 0 0 0 0 0
15 000 00 1 0 0 0 0 0 0 0 0
16 000 00 1 0 0 0 0 0 0 0 0
17 000 00 1 0 0 0 0 0 0 0 0
```

IRQ to pin mappings:

IRQ0 -> 0:2

IRQ1 -> 0:1

IRQ4 -> 0:4

IRQ6 -> 0:6

IRQ7 -> 0:7

IRQ8 -> 0:8

IRQ9 -> 0:9

IRQ12 -> 0:12

IRQ13 -> 0:13

IRQ14 -> 0:14

IRQ15 -> 0:15

IRQ16 -> 0:16

IRQ17 -> 0:17

IRQ18 -> 0:18

IRQ19 -> 0:19

IRQ23 -> 0:23

IRQ24 -> 1:0

IRQ25 -> 1:1

..... done.

Using local APIC timer interrupts.

calibrating APIC timer ...

..... CPU clock speed is 2665.0458 MHz.

..... host bus clock speed is 133.0272 MHz.

checking TSC synchronization across 2 CPUs: passed.

Starting migration thread for cpu 0

Bringing up 1

Irq balancing problem?

## Linux–Kernel: Irq balancing problem?

CPU 1 IS NOW UP!  
Starting migration thread for cpu 1  
CPUS done 8  
NET: Registered protocol family 16  
EISA bus registered  
PCI: PCI BIOS revision 2.10 entry at 0xfd8d5, last bus=5  
PCI: Using configuration type 1  
mtrr: v2.0 (20020519)  
SCSI subsystem initialized  
PCI: Probing PCI hardware  
PCI: Probing PCI hardware (bus 00)  
PCI: Ignoring BAR0–3 of IDE controller 0000:00:1f.1  
Transparent bridge – 0000:00:1e.0  
PCI: Using IRQ router PIIX/ICH [8086/24c0] at 0000:00:1f.0  
PCI→APIC IRQ transform: (B0,I29,P0) → 16  
PCI→APIC IRQ transform: (B0,I29,P1) → 19  
PCI→APIC IRQ transform: (B0,I29,P2) → 18  
PCI→APIC IRQ transform: (B0,I29,P3) → 23  
PCI→APIC IRQ transform: (B1,I0,P0) → 17  
PCI→APIC IRQ transform: (B3,I3,P0) → 24  
PCI→APIC IRQ transform: (B3,I3,P1) → 25  
PCI→APIC IRQ transform: (B5,I2,P0) → 17  
PCI→APIC IRQ transform: (B5,I3,P0) → 16  
SBF: Simple Boot Flag extension found and enabled.  
SBF: Setting boot flags 0x80  
ikconfig 0.7 with /proc/config\*  
highmem bounce pool size: 64 pages  
VFS: Disk quotas dquot\_6.5.1  
Initializing Cryptographic API  
pty: 2048 Unix98 ptys configured  
Real Time Clock Driver v1.12  
hw\_random hardware driver 1.0.0 loaded  
Linux agpgart interface v0.100 (c) Dave Jones  
agpgart: Detected an Intel E7505 Chipset.  
agpgart: Maximum main memory to use for agp memory: 1919M  
agpgart: AGP aperture is 128M @ 0xe0000000  
Serial: 8250/16550 driver \$Revision: 1.90 \$ 48 ports, IRQ sharing enabled  
ttyS0 at I/O 0x3f8 (irq = 4) is a 16550A  
Using deadline io scheduler  
Floppy drive(s): fd0 is 1.44M  
FDC 0 is a post–1991 82077  
RAMDISK driver initialized: 16 RAM disks of 4096K size 1024 blocksize  
Uniform Multi–Platform E–IDE driver Revision: 7.00alpha2  
ide: Assuming 33MHz system bus speed for PIO modes; override with idebus=xx  
ICH4: IDE controller at PCI slot 0000:00:1f.1  
ICH4: chipset revision 2  
ICH4: not 100% native mode: will probe irqs later  
  ide0: BM–DMA at 0x1460–0x1467, BIOS settings: hda:DMA, hdb:pio  
hda: HL–DT–STDVD–ROM GDR8161B, ATAPI CD/DVD–ROM drive  
ide0 at 0x1f0–0x1f7,0x3f6 on irq 14  
ide–floppy driver 0.99.newide

Irq balancing problem?

## Linux-Kernel: Irq balancing problem?

GDT: Storage RAID Controller Driver. Version: 2.08  
GDT: Found 0 PCI Storage RAID Controllers  
st: Version 20030811, fixed bufsize 32768, s/g segs 256  
Fusion MPT base driver 2.05.00.03  
Copyright (c) 1999-2002 LSI Logic Corporation  
mptbase: Initiating ioc0 bringup  
ioc0: 53C1030: Capabilities={Initiator}  
mptbase: Initiating ioc1 bringup  
ioc1: 53C1030: Capabilities={Initiator}  
mptbase: 2 MPT adapters found, 2 installed.  
Fusion MPT SCSI Host driver 2.05.00.03  
scsi0 : ioc0: LSI53C1030, FwRev=01000000h, Ports=1, MaxQ=255, IRQ=24  
Vendor: SEAGATE Model: ST336607LW Rev: 0006  
Type: Direct-Access ANSI SCSI revision: 03  
SCSI device sda: 71132960 512-byte hdwr sectors (36420 MB)  
SCSI device sda: drive cache: write through  
sda: sda1 sda2 sda3 sda4 < sda5 >  
Attached scsi disk sda at scsi0, channel 0, id 0, lun 0  
Attached scsi generic sg0 at scsi0, channel 0, id 0, lun 0, type 0  
scsi1 : ioc1: LSI53C1030, FwRev=01000000h, Ports=1, MaxQ=255, IRQ=25  
Vendor: SEAGATE Model: ST336607LW Rev: 0006  
Type: Direct-Access ANSI SCSI revision: 03  
SCSI device sdb: 71132960 512-byte hdwr sectors (36420 MB)  
SCSI device sdb: drive cache: write through  
sdb: sdb1 sdb2  
Attached scsi disk sdb at scsi1, channel 0, id 1, lun 0  
Attached scsi generic sg1 at scsi1, channel 0, id 1, lun 0, type 0  
mice: PS/2 mouse device common for all mice  
input: ImExPS/2 Generic Explorer Mouse on isa0060/serio1  
serio: i8042 AUX port at 0x60,0x64 irq 12  
input: AT Translated Set 2 keyboard on isa0060/serio0  
serio: i8042 KBD port at 0x60,0x64 irq 1  
md: md driver 0.90.0 MAX\_MD\_DEVS=256, MD\_SB\_DISKS=27  
EISA: Probing bus 0 at eisa0  
NET: Registered protocol family 2  
IP: routing cache hash table of 16384 buckets, 128Kbytes  
TCP: Hash tables configured (established 524288 bind 65536)  
NET: Registered protocol family 1  
NET: Registered protocol family 17

-

To unsubscribe from this list: send the line "unsubscribe linux-kernel" in  
the body of a message to majordomo@vger.kernel.org  
More majordomo info at <http://vger.kernel.org/majordomo-info.html>  
Please read the FAQ at <http://www.tux.org/lkml/>