

---

Subject: [BUG]: Crash with CONFIG\_FAIR\_CGROUP\_SCHED=y  
Posted by [Sukadev Bhattiprolu](#) on Thu, 08 Nov 2007 23:48:05 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

With CONFIG\_FAIR\_CGROUP\_SCHED=y, following commands on 2.6.24-rc1 crash the system.

```
$ mount -t cgroup none /cgroups  
$ ./ns_exec -cm /bin/ls
```

"ns\_exec -cm" calls clone() to clone the mount namespace and then executes the '/bin/ls' program in the cloned child.

Some observations that Serge and I made (we have been able to reproduce reliably, but crash logs have not been very useful)

- a. If we skip the 'mount' command, there is no crash.
- b. If CONFIG\_FAIR\_CGROUP\_SCHED=n again, there is no crash (even with 'mount' command).

- c. mounting just the cpu or just the ns subsystem does not lead to a crash. You can even

```
mount -t cgroup -o cpu none /mnt1  
mount -t cgroup -o ns none /mnt2
```

and you still don't get a crash

- d. mount -t cgroup -o cpu,ns none /cgroup

will always cause a crash with subsequent ns\_exec

ns\_exec.c and config file are attached. Let us know if you need more info.

Suka

---

Crash log:

```
Red Hat Enterprise Linux release 4.90 (Tikanga)  
Kernel 2.6.24-rc1 on an x86_64
```

```
elm3a241 login: Unable to handle kernel NULL pointer dereferenceUnable to handle kernel NULL  
pointer dereference at 0000000000000000 RIP:
```

```
at 0000000000000000 RIP:
```

```
[<0000000000000000>]
```

```
[<0000000000000000>]
```

PGD 102d4d067 PGD 102d4d067 PUD 102c88067 PUD 102c88067 PMD 0 PMD 0

Oops: 0000 [1] Oops: 0000 [1] SMP SMP

CPU 2 CPU 2

Modules linked in:Modules linked in:

Pid: 3273, comm: ns\_exec Not tainted 2.6.24-rc1 #9

Pid: 3273, comm: ns\_exec Not tainted 2.6.24-rc1 #9

RIP: 0010:[<0000000000000000>] RIP: 0010:[<0000000000000000>] [<0000000000000000>]  
[<0000000000000000>]

RSP: 0018:ffff8101006a6af0 EFLAGS: 00055292

RSP: 0018:ffff8101006a6af0 EFLAGS: 00055292

RAX: 0000000000000000 RBX: ffff810100d11140 RCX: ffff810101de4000

RAX: 0000000000000000 RBX: ffff810100d11140 RCX: ffff810101de4000

RDX: 0000000000000000 RSI: ffff810100d1a880 RDI: ffff810001037c00

RDX: 0000000000000000 RSI: ffff810100d1a880 RDI: ffff810001037c00

RBP: ffff810102c136c0 R08: ffff810101de4000 R09: ffff810101d31bb8

RBP: ffff810102c136c0 R08: ffff810101de4000 R09: ffff810101d31bb8

R10: 0000000000000000 R11: 00000000ffffff R12: ffff8101034075b8

R10: 0000000000000000 R11: 00000000ffffff R12: ffff8101034075b8

R13: ffff8101029d6028 R14: ffff810103407500 R15: ffffffff80869d00

R13: ffff8101029d6028 R14: ffff810103407500 R15: ffffffff80869d00

FS: 00002b80c2a396f0(0000) GS:ffff81010068f3c0(0000) knlGS:0000000000000000

FS: 00002b80c2a396f0(0000) GS:ffff81010068f3c0(0000) knlGS:0000000000000000

CS: 0010 DS: 0000 ES: 0000 CR0: 000000008005003b

CS: 0010 DS: 0000 ES: 0000 CR0: 000000008005003b

CR2: 0000000000000000 CR3: 00000001028be000 CR4: 000000000000006e0

CR2: 0000000000000000 CR3: 00000001028be000 CR4: 000000000000006e0

DR0: 0000000000000000 DR1: 0000000000000000 DR2: 0000000000000000

DR0: 0000000000000000 DR1: 0000000000000000 DR2: 0000000000000000

DR3: 0000000000000000 DR6: 00000000ffff0ff0 DR7: 00000000000000400

DR3: 0000000000000000 DR6: 00000000ffff0ff0 DR7: 00000000000000400

Process ns\_exec (pid: 3273, threadinfo ffff810101de4000, task ffff810100d11140)

Process ns\_exec (pid: 3273, threadinfo ffff810101de4000, task ffff810100d11140)

Stack: Stack: 0000000000000000 0000000000000000 0000000000000001 0000000000000001  
6c2f343662696c2f 6c2f343662696c2f 2d78756e696c2d64 2d78756e696c2d64

732e34362d363878 732e34362d363878 ffffffff00322e6f ffffffff00322e6f 762f003561736376

762f003561736376 0000357363762f63 0000357363762f63

0000000000000000 0000000000000000 0000000000000000 0000000000000000

762f73007665642f 762f73007665642f 0000317363762f63 0000317363762f63

Call Trace:

Call Trace:

Code: Code: Bad RIP value. Bad RIP value.

```
RIP RIP [<0000000000000000>]
[<0000000000000000>]
RSP <ffff8101006a6af0>
RSP <ffff8101006a6af0>
CR2: 0000000000000000
CR2: 0000000000000000
```

```
#
# Automatically generated make config: don't edit
# Linux kernel version: 2.6.24-rc1
# Thu Nov  8 13:38:29 2007
#
CONFIG_X86_64=y
CONFIG_64BIT=y
CONFIG_X86=y
CONFIG_GENERIC_TIME=y
CONFIG_GENERIC_TIME_VSYSCALL=y
CONFIG_GENERIC_CMOS_UPDATE=y
CONFIG_CLOCKSOURCE_WATCHDOG=y
CONFIG_GENERIC_CLOCKEVENTS=y
CONFIG_GENERIC_CLOCKEVENTS_BROADCAST=y
CONFIG_ZONE_DMA32=y
CONFIG_LOCKDEP_SUPPORT=y
CONFIG_STACKTRACE_SUPPORT=y
CONFIG_SEMAPHORE_SLEEPERS=y
CONFIG_MMU=y
CONFIG_ZONE_DMA=y
CONFIG_RWSEM_GENERIC_SPINLOCK=y
CONFIG_GENERIC_HWEIGHT=y
CONFIG_GENERIC_CALIBRATE_DELAY=y
CONFIG_X86_CMPXCHG=y
CONFIG_EARLY_PRINTK=y
CONFIG_GENERIC_ISA_DMA=y
CONFIG_GENERIC_IOMAP=y
CONFIG_ARCH_MAY_HAVE_PC_FDC=y
CONFIG_ARCH_POPULATES_NODE_MAP=y
CONFIG_DMI=y
CONFIG_AUDIT_ARCH=y
CONFIG_GENERIC_BUG=y
# CONFIG_ARCH_HAS_ILOG2_U32 is not set
# CONFIG_ARCH_HAS_ILOG2_U64 is not set
CONFIG_DEFCONFIG_LIST="/lib/modules/$UNAME_RELEASE/.config"
```

```
#
# General setup
#
CONFIG_EXPERIMENTAL=y
CONFIG_LOCK_KERNEL=y
CONFIG_INIT_ENV_ARG_LIMIT=32
CONFIG_LOCALVERSION=""
CONFIG_LOCALVERSION_AUTO=y
CONFIG_SWAP=y
CONFIG_SYSVIPC=y
CONFIG_SYSVIPC_SYSCTL=y
CONFIG_POSIX_MQUEUE=y
# CONFIG_BSD_PROCESS_ACCT is not set
# CONFIG_TASKSTATS is not set
CONFIG_USER_NS=y
# CONFIG_AUDIT is not set
CONFIG_IKCONFIG=y
CONFIG_IKCONFIG_PROC=y
CONFIG_LOG_BUF_SHIFT=18
CONFIG_CGROUPS=y
# CONFIG_CGROUP_DEBUG is not set
CONFIG_CGROUP_NS=y
# CONFIG_CGROUP_CPUACCT is not set
# CONFIG_CPUSETS is not set
CONFIG_FAIR_GROUP_SCHED=y
# CONFIG_FAIR_USER_SCHED is not set
CONFIG_FAIR_CGROUP_SCHED=y
CONFIG_SYSFS_DEPRECATED=y
CONFIG_RELAY=y
CONFIG_BLK_DEV_INITRD=y
CONFIG_INITRAMFS_SOURCE=""
CONFIG_CC_OPTIMIZE_FOR_SIZE=y
CONFIG_SYSCTL=y
# CONFIG_EMBEDDED is not set
CONFIG_UID16=y
CONFIG_SYSCTL_SYSCALL=y
CONFIG_KALLSYMS=y
CONFIG_KALLSYMS_ALL=y
# CONFIG_KALLSYMS_EXTRA_PASS is not set
CONFIG_HOTPLUG=y
CONFIG_PRINTK=y
CONFIG_BUG=y
CONFIG_ELF_CORE=y
CONFIG_BASE_FULL=y
CONFIG_FUTEX=y
CONFIG_ANON_INODES=y
CONFIG_EPOLL=y
```

```
CONFIG_SIGNALFD=y
CONFIG_EVENTFD=y
CONFIG_SHMEM=y
CONFIG_VM_EVENT_COUNTERS=y
CONFIG_SLAB=y
# CONFIG_SLUB is not set
# CONFIG_SLOB is not set
CONFIG_RT_MUTEXES=y
# CONFIG_TINY_SHMEM is not set
CONFIG_BASE_SMALL=0
CONFIG_MODULES=y
CONFIG_MODULE_UNLOAD=y
CONFIG_MODULE_FORCE_UNLOAD=y
# CONFIG_MODVERSIONS is not set
# CONFIG_MODULE_SRCVERSION_ALL is not set
# CONFIG_KMOD is not set
CONFIG_STOP_MACHINE=y
CONFIG_BLOCK=y
# CONFIG_BLK_DEV_IO_TRACE is not set
# CONFIG_BLK_DEV_BSG is not set
CONFIG_BLOCK_COMPAT=y
```

```
#
# IO Schedulers
#
CONFIG_IOSCHED_NOOP=y
# CONFIG_IOSCHED_AS is not set
CONFIG_IOSCHED_DEADLINE=y
CONFIG_IOSCHED_CFQ=y
# CONFIG_DEFAULT_AS is not set
# CONFIG_DEFAULT_DEADLINE is not set
CONFIG_DEFAULT_CFQ=y
# CONFIG_DEFAULT_NOOP is not set
CONFIG_DEFAULT_IOSCHED="cfq"
```

```
#
# Processor type and features
#
# CONFIG_TICK_ONESHOT is not set
# CONFIG_NO_HZ is not set
# CONFIG_HIGH_RES_TIMERS is not set
CONFIG_GENERIC_CLOCKEVENTS_BUILD=y
CONFIG_X86_PC=y
# CONFIG_X86_VSMP is not set
# CONFIG_MK8 is not set
# CONFIG_MPSC is not set
# CONFIG_MCORE2 is not set
CONFIG_GENERIC_CPU=y
```

```
CONFIG_X86_L1_CACHE_BYTES=128
CONFIG_X86_L1_CACHE_SHIFT=7
CONFIG_X86_INTERNODE_CACHE_BYTES=128
CONFIG_X86_TSC=y
CONFIG_X86_GOOD_APIC=y
# CONFIG_MICROCODE is not set
CONFIG_X86_MSR=y
CONFIG_X86_CPUID=y
CONFIG_X86_HT=y
CONFIG_X86_IO_APIC=y
CONFIG_X86_LOCAL_APIC=y
CONFIG_MTRR=y
CONFIG_SMP=y
CONFIG_SCHED_SMT=y
CONFIG_SCHED_MC=y
# CONFIG_PREEMPT_NONE is not set
CONFIG_PREEMPT_VOLUNTARY=y
# CONFIG_PREEMPT is not set
CONFIG_PREEMPT_BKL=y
CONFIG_NUMA=y
CONFIG_K8_NUMA=y
CONFIG_NODES_SHIFT=6
CONFIG_X86_64_ACPI_NUMA=y
CONFIG_NUMA_EMU=y
CONFIG_ARCH_DISCONTIGMEM_ENABLE=y
CONFIG_ARCH_DISCONTIGMEM_DEFAULT=y
CONFIG_ARCH_SPARSEMEM_ENABLE=y
CONFIG_SELECT_MEMORY_MODEL=y
# CONFIG_FLATMEM_MANUAL is not set
CONFIG_DISCONTIGMEM_MANUAL=y
# CONFIG_SPARSEMEM_MANUAL is not set
CONFIG_DISCONTIGMEM=y
CONFIG_FLAT_NODE_MEM_MAP=y
CONFIG_NEED_MULTIPLE_NODES=y
# CONFIG_SPARSEMEM_STATIC is not set
CONFIG_SPARSEMEM_VMEMMAP_ENABLE=y
CONFIG_SPLIT_PTLOCK_CPUS=4
CONFIG_MIGRATION=y
CONFIG_RESOURCES_64BIT=y
CONFIG_ZONE_DMA_FLAG=1
CONFIG_BOUNCE=y
CONFIG_VIRT_TO_BUS=y
CONFIG_HAVE_ARCH_EARLY_PFN_TO_NID=y
CONFIG_OUT_OF_LINE_PFN_TO_PAGE=y
CONFIG_NR_CPUS=32
CONFIG_PHYSICAL_ALIGN=0x200000
CONFIG_HOTPLUG_CPU=y
CONFIG_ARCH_ENABLE_MEMORY_HOTPLUG=y
```

```
CONFIG_HPET_TIMER=y
CONFIG_HPET_EMULATE_RTC=y
CONFIG_IOMMU=y
# CONFIG_CALGARY_IOMMU is not set
CONFIG_SWIOTLB=y
CONFIG_X86_MCE=y
CONFIG_X86_MCE_INTEL=y
CONFIG_X86_MCE_AMD=y
# CONFIG_KEXEC is not set
# CONFIG_CRASH_DUMP is not set
# CONFIG_RELOCATABLE is not set
CONFIG_PHYSICAL_START=0x200000
CONFIG_SECCOMP=y
# CONFIG_CC_STACKPROTECTOR is not set
# CONFIG_HZ_100 is not set
CONFIG_HZ_250=y
# CONFIG_HZ_300 is not set
# CONFIG_HZ_1000 is not set
CONFIG_HZ=250
CONFIG_K8_NB=y
CONFIG_GENERIC_HARDIRQS=y
CONFIG_GENERIC_IRQ_PROBE=y
CONFIG_ISA_DMA_API=y
CONFIG_GENERIC_PENDING_IRQ=y
```

```
#
# Power management options
#
CONFIG_PM=y
# CONFIG_PM_LEGACY is not set
# CONFIG_PM_DEBUG is not set
CONFIG_PM_SLEEP_SMP=y
CONFIG_PM_SLEEP=y
CONFIG_SUSPEND_SMP_POSSIBLE=y
CONFIG_SUSPEND=y
CONFIG_HIBERNATION_SMP_POSSIBLE=y
CONFIG_HIBERNATION=y
CONFIG_PM_STD_PARTITION=""
CONFIG_ARCH_HIBERNATION_HEADER=y
CONFIG_ACPI=y
CONFIG_ACPI_SLEEP=y
CONFIG_ACPI_PROCFS=y
CONFIG_ACPI_PROC_EVENT=y
CONFIG_ACPI_BUTTON=y
CONFIG_ACPI_FAN=y
# CONFIG_ACPI_DOCK is not set
CONFIG_ACPI_PROCESSOR=y
CONFIG_ACPI_HOTPLUG_CPU=y
```

```
CONFIG_ACPI_THERMAL=y
CONFIG_ACPI_NUMA=y
# CONFIG_ACPI_ASUS is not set
# CONFIG_ACPI_TOSHIBA is not set
CONFIG_ACPI_BLACKLIST_YEAR=0
# CONFIG_ACPI_DEBUG is not set
CONFIG_ACPI_EC=y
CONFIG_ACPI_POWER=y
CONFIG_ACPI_SYSTEM=y
CONFIG_X86_PM_TIMER=y
CONFIG_ACPI_CONTAINER=y

#
# CPU Frequency scaling
#
CONFIG_CPU_FREQ=y
CONFIG_CPU_FREQ_TABLE=y
CONFIG_CPU_FREQ_DEBUG=y
CONFIG_CPU_FREQ_STAT=y
# CONFIG_CPU_FREQ_STAT_DETAILS is not set
CONFIG_CPU_FREQ_DEFAULT_GOV_PERFORMANCE=y
# CONFIG_CPU_FREQ_DEFAULT_GOV_USERSPACE is not set
# CONFIG_CPU_FREQ_DEFAULT_GOV_ONDEMAND is not set
# CONFIG_CPU_FREQ_DEFAULT_GOV_CONSERVATIVE is not set
CONFIG_CPU_FREQ_GOV_PERFORMANCE=y
# CONFIG_CPU_FREQ_GOV_POWERSAVE is not set
CONFIG_CPU_FREQ_GOV_USERSPACE=y
CONFIG_CPU_FREQ_GOV_ONDEMAND=y
CONFIG_CPU_FREQ_GOV_CONSERVATIVE=y

#
# CPUFreq processor drivers
#
CONFIG_X86_POWERNOW_K8=y
CONFIG_X86_POWERNOW_K8_ACPI=y
# CONFIG_X86_SPEEDSTEP_CENTRINO is not set
CONFIG_X86_ACPI_CPUFREQ=y

#
# shared options
#
CONFIG_X86_ACPI_CPUFREQ_PROC_INTF=y
# CONFIG_X86_SPEEDSTEP_LIB is not set
# CONFIG_CPU_IDLE is not set

#
# Bus options (PCI etc.)
#
```

```
CONFIG_PCI=y
CONFIG_PCI_DIRECT=y
CONFIG_PCI_MMCONFIG=y
CONFIG_PCI_DOMAINS=y
CONFIG_DMAR=y
CONFIG_DMAR_GFX_WA=y
CONFIG_DMAR_FLOPPY_WA=y
CONFIG_PCIEPORTBUS=y
CONFIG_PCIEAER=y
CONFIG_ARCH_SUPPORTS_MSI=y
CONFIG_PCI_MSI=y
# CONFIG_PCI_DEBUG is not set
# CONFIG_HT_IRQ is not set
# CONFIG_PCCARD is not set
# CONFIG_HOTPLUG_PCI is not set

#
# Executable file formats / Emulations
#
CONFIG_BINFMT_ELF=y
# CONFIG_BINFMT_MISC is not set
CONFIG_IA32_EMULATION=y
CONFIG_IA32_AOUT=y
CONFIG_COMPAT=y
CONFIG_COMPAT_FOR_U64_ALIGNMENT=y
CONFIG_SYSVIPC_COMPAT=y

#
# Networking
#
CONFIG_NET=y

#
# Networking options
#
CONFIG_PACKET=y
# CONFIG_PACKET_MMAP is not set
CONFIG_UNIX=y
# CONFIG_NET_KEY is not set
CONFIG_INET=y
CONFIG_IP_MULTICAST=y
# CONFIG_IP_ADVANCED_ROUTER is not set
CONFIG_IP_FIB_HASH=y
CONFIG_IP_PNP=y
CONFIG_IP_PNP_DHCP=y
# CONFIG_IP_PNP_BOOTP is not set
# CONFIG_IP_PNP_RARP is not set
# CONFIG_NET_IPIP is not set
```

```
# CONFIG_NET_IPGRE is not set
# CONFIG_IP_MROUTE is not set
# CONFIG_ARPD is not set
# CONFIG_SYN_COOKIES is not set
# CONFIG_INET_AH is not set
# CONFIG_INET_ESP is not set
# CONFIG_INET_IPCOMP is not set
# CONFIG_INET_XFRM_TUNNEL is not set
CONFIG_INET_TUNNEL=y
# CONFIG_INET_XFRM_MODE_TRANSPORT is not set
# CONFIG_INET_XFRM_MODE_TUNNEL is not set
# CONFIG_INET_XFRM_MODE_BEET is not set
# CONFIG_INET_LRO is not set
CONFIG_INET_DIAG=y
CONFIG_INET_TCP_DIAG=y
# CONFIG_TCP_CONG_ADVANCED is not set
CONFIG_TCP_CONG_CUBIC=y
CONFIG_DEFAULT_TCP_CONG="cubic"
# CONFIG_TCP_MD5SIG is not set
CONFIG_IPV6=y
# CONFIG_IPV6_PRIVACY is not set
# CONFIG_IPV6_ROUTER_PREF is not set
# CONFIG_IPV6_OPTIMISTIC_DAD is not set
# CONFIG_INET6_AH is not set
# CONFIG_INET6_ESP is not set
# CONFIG_INET6_IPCOMP is not set
# CONFIG_IPV6_MIP6 is not set
# CONFIG_INET6_XFRM_TUNNEL is not set
# CONFIG_INET6_TUNNEL is not set
# CONFIG_INET6_XFRM_MODE_TRANSPORT is not set
# CONFIG_INET6_XFRM_MODE_TUNNEL is not set
# CONFIG_INET6_XFRM_MODE_BEET is not set
# CONFIG_INET6_XFRM_MODE_ROUTEOPTIMIZATION is not set
CONFIG_IPV6_SIT=y
# CONFIG_IPV6_TUNNEL is not set
# CONFIG_IPV6_MULTIPLE_TABLES is not set
# CONFIG_NETWORK_SECMARK is not set
# CONFIG_NETFILTER is not set
# CONFIG_IP_DCCP is not set
# CONFIG_IP_SCTP is not set
# CONFIG_TIPC is not set
# CONFIG_ATM is not set
# CONFIG_BRIDGE is not set
# CONFIG_VLAN_8021Q is not set
# CONFIG_DECNET is not set
# CONFIG_LLC2 is not set
# CONFIG_IPX is not set
# CONFIG_ATALK is not set
```

```
# CONFIG_X25 is not set
# CONFIG_LAPB is not set
# CONFIG_ECONET is not set
# CONFIG_WAN_ROUTER is not set
# CONFIG_NET_SCHED is not set

#
# Network testing
#
# CONFIG_NET_PKTGEN is not set
# CONFIG_NET_TCPPROBE is not set
# CONFIG_HAMRADIO is not set
# CONFIG_IRDA is not set
# CONFIG_BT is not set
# CONFIG_AF_RXRPC is not set

#
# Wireless
#
# CONFIG_CFG80211 is not set
# CONFIG_WIRELESS_EXT is not set
# CONFIG_MAC80211 is not set
# CONFIG_IEEE80211 is not set
# CONFIG_RFKILL is not set
# CONFIG_NET_9P is not set

#
# Device Drivers
#

#
# Generic Driver Options
#
CONFIG_UEVENT_HELPER_PATH="/sbin/hotplug"
CONFIG_STANDALONE=y
CONFIG_PREVENT_FIRMWARE_BUILD=y
CONFIG_FW_LOADER=y
# CONFIG_DEBUG_DRIVER is not set
# CONFIG_DEBUG_DEVRES is not set
# CONFIG_SYS_HYPERVISOR is not set
# CONFIG_CONNECTOR is not set
# CONFIG_MTD is not set
# CONFIG_PARPORT is not set
CONFIG_PNP=y
# CONFIG_PNP_DEBUG is not set

#
# Protocols
```

```
#
CONFIG_PNPACPI=y
CONFIG_BLK_DEV=y
CONFIG_BLK_DEV_FD=y
# CONFIG_BLK_CPQ_DA is not set
# CONFIG_BLK_CPQ_CISS_DA is not set
# CONFIG_BLK_DEV_DAC960 is not set
# CONFIG_BLK_DEV_UMEM is not set
# CONFIG_BLK_DEV_COW_COMMON is not set
CONFIG_BLK_DEV_LOOP=y
# CONFIG_BLK_DEV_CRYPTOLOOP is not set
# CONFIG_BLK_DEV_NBD is not set
# CONFIG_BLK_DEV_SX8 is not set
# CONFIG_BLK_DEV_UB is not set
CONFIG_BLK_DEV_RAM=y
CONFIG_BLK_DEV_RAM_COUNT=16
CONFIG_BLK_DEV_RAM_SIZE=4096
CONFIG_BLK_DEV_RAM_BLOCKSIZE=1024
# CONFIG_CDROM_PKTCDVD is not set
# CONFIG_ATA_OVER_ETH is not set
CONFIG_MISC_DEVICES=y
# CONFIG_IBM_ASM is not set
# CONFIG_PHANTOM is not set
# CONFIG_EEPROM_93CX6 is not set
# CONFIG_SGI_IOC4 is not set
# CONFIG_TIFM_CORE is not set
# CONFIG_SONY_LAPTOP is not set
# CONFIG_THINKPAD_ACPI is not set
CONFIG_IDE=y
CONFIG_BLK_DEV_IDE=y

#
# Please see Documentation/ide.txt for help/info on IDE drives
#
# CONFIG_BLK_DEV_IDE_SATA is not set
# CONFIG_BLK_DEV_HD_IDE is not set
CONFIG_BLK_DEV_IDEDISK=y
CONFIG_IDEDISK_MULTI_MODE=y
CONFIG_BLK_DEV_IDECD=y
# CONFIG_BLK_DEV_IDETAPE is not set
# CONFIG_BLK_DEV_IDEFLOPPY is not set
# CONFIG_BLK_DEV_IDESCSI is not set
CONFIG_BLK_DEV_IDEACPI=y
# CONFIG_IDE_TASK_IOCTL is not set
CONFIG_IDE_PROC_FS=y

#
# IDE chipset support/bugfixes
```

```
#
CONFIG_IDE_GENERIC=y
# CONFIG_BLK_DEV_PLATFORM is not set
# CONFIG_BLK_DEV_CMD640 is not set
# CONFIG_BLK_DEV_IDEPNP is not set

#
# PCI IDE chipsets support
#
CONFIG_BLK_DEV_IDEPCI=y
# CONFIG_IDEPCI_SHARE_IRQ is not set
CONFIG_IDEPCI_PCIBUS_ORDER=y
# CONFIG_BLK_DEV_OFFBOARD is not set
# CONFIG_BLK_DEV_GENERIC is not set
# CONFIG_BLK_DEV_OPTI621 is not set
# CONFIG_BLK_DEV_RZ1000 is not set
CONFIG_BLK_DEV_IDEDMA_PCI=y
# CONFIG_BLK_DEV_AEC62XX is not set
# CONFIG_BLK_DEV_ALI15X3 is not set
CONFIG_BLK_DEV_AMD74XX=y
CONFIG_BLK_DEV_ATIIXP=y
# CONFIG_BLK_DEV_CMD64X is not set
# CONFIG_BLK_DEV_TRIFLEX is not set
# CONFIG_BLK_DEV_CY82C693 is not set
# CONFIG_BLK_DEV_CS5520 is not set
# CONFIG_BLK_DEV_CS5530 is not set
# CONFIG_BLK_DEV_HPT34X is not set
# CONFIG_BLK_DEV_HPT366 is not set
# CONFIG_BLK_DEV_JMICRON is not set
# CONFIG_BLK_DEV_SC1200 is not set
CONFIG_BLK_DEV_PIIIX=y
# CONFIG_BLK_DEV_IT8213 is not set
# CONFIG_BLK_DEV_IT821X is not set
# CONFIG_BLK_DEV_NS87415 is not set
# CONFIG_BLK_DEV_PDC202XX_OLD is not set
CONFIG_BLK_DEV_PDC202XX_NEW=y
# CONFIG_BLK_DEV_SVWKS is not set
# CONFIG_BLK_DEV_SIIIMAGE is not set
# CONFIG_BLK_DEV_SIS5513 is not set
# CONFIG_BLK_DEV_SLC90E66 is not set
# CONFIG_BLK_DEV_TRM290 is not set
# CONFIG_BLK_DEV_VIA82CXXX is not set
# CONFIG_BLK_DEV_TC86C001 is not set
# CONFIG_IDE_ARM is not set
CONFIG_BLK_DEV_IDEDMA=y
CONFIG_IDE_ARCH_OBSOLETE_INIT=y
# CONFIG_BLK_DEV_HD is not set
```

```
#
# SCSI device support
#
# CONFIG_RAID_ATTRS is not set
CONFIG_SCSI=y
CONFIG_SCSI_DMA=y
# CONFIG_SCSI_TGT is not set
CONFIG_SCSI_NETLINK=y
# CONFIG_SCSI_PROC_FS is not set

#
# SCSI support type (disk, tape, CD-ROM)
#
CONFIG_BLK_DEV_SD=y
# CONFIG_CHR_DEV_ST is not set
# CONFIG_CHR_DEV_OSST is not set
CONFIG_BLK_DEV_SR=y
# CONFIG_BLK_DEV_SR_VENDOR is not set
CONFIG_CHR_DEV_SG=y
# CONFIG_CHR_DEV_SCH is not set

#
# Some SCSI devices (e.g. CD jukebox) support multiple LUNs
#
# CONFIG_SCSI_MULTI_LUN is not set
CONFIG_SCSI_CONSTANTS=y
# CONFIG_SCSI_LOGGING is not set
# CONFIG_SCSI_SCAN_ASYNC is not set
CONFIG_SCSI_WAIT_SCAN=m

#
# SCSI Transports
#
CONFIG_SCSI_SPI_ATTRS=y
CONFIG_SCSI_FC_ATTRS=y
# CONFIG_SCSI_ISCSI_ATTRS is not set
# CONFIG_SCSI_SAS_LIBSAS is not set
# CONFIG_SCSI_SRP_ATTRS is not set
CONFIG_SCSI_LOWLEVEL=y
# CONFIG_ISCSI_TCP is not set
# CONFIG_BLK_DEV_3W_XXXX_RAID is not set
# CONFIG_SCSI_3W_9XXX is not set
# CONFIG_SCSI_ACARD is not set
# CONFIG_SCSI_AACRAID is not set
# CONFIG_SCSI_AIC7XXX is not set
# CONFIG_SCSI_AIC7XXX_OLD is not set
CONFIG_SCSI_AIC79XX=y
CONFIG_AIC79XX_CMDS_PER_DEVICE=32
```

```
CONFIG_AIC79XX_RESET_DELAY_MS=4000
# CONFIG_AIC79XX_DEBUG_ENABLE is not set
CONFIG_AIC79XX_DEBUG_MASK=0
# CONFIG_AIC79XX_REG_PRETTY_PRINT is not set
# CONFIG_SCSI_AIC94XX is not set
# CONFIG_SCSI_ADVANSYS is not set
# CONFIG_SCSI_ARCMSR is not set
# CONFIG_MEGARAID_NEWGEN is not set
# CONFIG_MEGARAID_LEGACY is not set
# CONFIG_MEGARAID_SAS is not set
# CONFIG_SCSI_HPTIOP is not set
# CONFIG_SCSI_BUSLOGIC is not set
# CONFIG_SCSI_DMX3191D is not set
# CONFIG_SCSI_EATA is not set
# CONFIG_SCSI_FUTURE_DOMAIN is not set
# CONFIG_SCSI_GDTH is not set
# CONFIG_SCSI_IPS is not set
# CONFIG_SCSI_INITIO is not set
# CONFIG_SCSI_INIA100 is not set
# CONFIG_SCSI_STEX is not set
# CONFIG_SCSI_SYM53C8XX_2 is not set
# CONFIG_SCSI_IPR is not set
# CONFIG_SCSI_QLOGIC_1280 is not set
# CONFIG_SCSI_QLA_FC is not set
# CONFIG_SCSI_QLA_ISCSI is not set
# CONFIG_SCSI_LPFC is not set
# CONFIG_SCSI_DC395x is not set
# CONFIG_SCSI_DC390T is not set
# CONFIG_SCSI_DEBUG is not set
# CONFIG_SCSI_SRP is not set
CONFIG_ATA=y
# CONFIG_ATA_NONSTANDARD is not set
CONFIG_ATA_ACPI=y
CONFIG_SATA_AHCI=y
CONFIG_SATA_SVW=y
CONFIG_ATA_PIIX=y
# CONFIG_SATA_MV is not set
CONFIG_SATA_NV=y
# CONFIG_PDC_ADMA is not set
# CONFIG_SATA_QSTOR is not set
# CONFIG_SATA_PROMISE is not set
# CONFIG_SATA_SX4 is not set
CONFIG_SATA_SIL=y
# CONFIG_SATA_SIL24 is not set
# CONFIG_SATA_SIS is not set
# CONFIG_SATA_ULI is not set
CONFIG_SATA_VIA=y
# CONFIG_SATA_VITESSE is not set
```

```
# CONFIG_SATA_INIC162X is not set
# CONFIG_PATA_ACPI is not set
# CONFIG_PATA_ALI is not set
# CONFIG_PATA_AMD is not set
# CONFIG_PATA_ARTOP is not set
# CONFIG_PATA_ATIIXP is not set
# CONFIG_PATA_CMD640_PCI is not set
# CONFIG_PATA_CMD64X is not set
# CONFIG_PATA_CS5520 is not set
# CONFIG_PATA_CS5530 is not set
# CONFIG_PATA_CYPRESS is not set
# CONFIG_PATA_EFAR is not set
# CONFIG_ATA_GENERIC is not set
# CONFIG_PATA_HPT366 is not set
# CONFIG_PATA_HPT37X is not set
# CONFIG_PATA_HPT3X2N is not set
# CONFIG_PATA_HPT3X3 is not set
# CONFIG_PATA_IT821X is not set
# CONFIG_PATA_IT8213 is not set
# CONFIG_PATA_JMICRON is not set
# CONFIG_PATA_TRIFLEX is not set
# CONFIG_PATA_MARVELL is not set
# CONFIG_PATA_MPIIX is not set
# CONFIG_PATA_OLDPIIX is not set
# CONFIG_PATA_NETCELL is not set
# CONFIG_PATA_NS87410 is not set
# CONFIG_PATA_NS87415 is not set
# CONFIG_PATA_OPTI is not set
# CONFIG_PATA_OPTIDMA is not set
# CONFIG_PATA_PDC_OLD is not set
# CONFIG_PATA_RADISYS is not set
# CONFIG_PATA_RZ1000 is not set
# CONFIG_PATA_SC1200 is not set
# CONFIG_PATA_SERVERWORKS is not set
# CONFIG_PATA_PDC2027X is not set
# CONFIG_PATA_SIL680 is not set
# CONFIG_PATA_SIS is not set
# CONFIG_PATA_VIA is not set
# CONFIG_PATA_WINBOND is not set
CONFIG_MD=y
# CONFIG_BLK_DEV_MD is not set
CONFIG_BLK_DEV_DM=y
# CONFIG_DM_DEBUG is not set
# CONFIG_DM_CRYPT is not set
# CONFIG_DM_SNAPSHOT is not set
# CONFIG_DM_MIRROR is not set
# CONFIG_DM_ZERO is not set
# CONFIG_DM_MULTIPATH is not set
```

```
# CONFIG_DM_DELAY is not set
# CONFIG_DM_UEVENT is not set
CONFIG_FUSION=y
CONFIG_FUSION_SPI=y
# CONFIG_FUSION_FC is not set
# CONFIG_FUSION_SAS is not set
CONFIG_FUSION_MAX_SGE=128
# CONFIG_FUSION_CTL is not set
# CONFIG_FUSION_LOGGING is not set

#
# IEEE 1394 (FireWire) support
#
# CONFIG_FIREWIRE is not set
CONFIG_IEEE1394=y

#
# Subsystem Options
#
# CONFIG_IEEE1394_VERBOSEDEBUG is not set

#
# Controllers
#

#
# Texas Instruments PCILynx requires I2C
#
CONFIG_IEEE1394_OHCI1394=y

#
# Protocols
#
# CONFIG_IEEE1394_VIDEO1394 is not set
# CONFIG_IEEE1394_SBP2 is not set
# CONFIG_IEEE1394_ETH1394_ROM_ENTRY is not set
# CONFIG_IEEE1394_ETH1394 is not set
# CONFIG_IEEE1394_DV1394 is not set
CONFIG_IEEE1394_RAWIO=y
# CONFIG_I2O is not set
CONFIG_MACINTOSH_DRIVERS=y
# CONFIG_MAC_EMUMOUSEBTN is not set
CONFIG_NETDEVICES=y
CONFIG_NETDEVICES_MULTIQUEUE=y
# CONFIG_DUMMY is not set
# CONFIG_BONDING is not set
# CONFIG_MACVLAN is not set
# CONFIG_EQUALIZER is not set
```

```
CONFIG_TUN=y
# CONFIG_VETH is not set
# CONFIG_NET_SB1000 is not set
# CONFIG_IP1000 is not set
# CONFIG_ARCNET is not set
# CONFIG_PHYLIB is not set
CONFIG_NET_ETHERNET=y
CONFIG_MII=y
# CONFIG_HAPPYMEAL is not set
# CONFIG_SUNGEM is not set
# CONFIG_CASSINI is not set
CONFIG_NET_VENDOR_3COM=y
CONFIG_VORTEX=y
# CONFIG_TYphoon is not set
CONFIG_NET_TULIP=y
# CONFIG_DE2104X is not set
CONFIG_TULIP=y
# CONFIG_TULIP_MWI is not set
# CONFIG_TULIP_MMIO is not set
# CONFIG_TULIP_NAPI is not set
# CONFIG_DE4X5 is not set
# CONFIG_WINBOND_840 is not set
# CONFIG_DM9102 is not set
# CONFIG_ULI526X is not set
# CONFIG_HP100 is not set
# CONFIG_IBM_NEW_EMAC_ZMII is not set
# CONFIG_IBM_NEW_EMAC_RGMII is not set
# CONFIG_IBM_NEW_EMAC_TAH is not set
# CONFIG_IBM_NEW_EMAC_EMAC4 is not set
CONFIG_NET_PCI=y
# CONFIG_PCNET32 is not set
CONFIG_AMD8111_ETH=y
# CONFIG_AMD8111E_NAPI is not set
# CONFIG_ADAPTEC_STARFIRE is not set
CONFIG_B44=y
CONFIG_B44_PCI_AUTOSELECT=y
CONFIG_B44_PCICORE_AUTOSELECT=y
CONFIG_B44_PCI=y
CONFIG_FORCEDETH=y
# CONFIG_FORCEDETH_NAPI is not set
# CONFIG_EEPRO100 is not set
CONFIG_E100=y
# CONFIG_FEALNX is not set
# CONFIG_NATSEMI is not set
# CONFIG_NE2K_PCI is not set
CONFIG_8139CP=y
CONFIG_8139TOO=y
# CONFIG_8139TOO_PIO is not set
```

```
# CONFIG_8139TOO_TUNE_TWISTER is not set
# CONFIG_8139TOO_8129 is not set
# CONFIG_8139_OLD_RX_RESET is not set
# CONFIG_SIS900 is not set
# CONFIG_EPIC100 is not set
# CONFIG_SUNDANCE is not set
# CONFIG_VIA_RHINE is not set
# CONFIG_SC92031 is not set
CONFIG_NETDEV_1000=y
# CONFIG_ACENIC is not set
# CONFIG_DL2K is not set
CONFIG_E1000=y
# CONFIG_E1000_NAPI is not set
# CONFIG_E1000_DISABLE_PACKET_SPLIT is not set
# CONFIG_E1000E is not set
# CONFIG_NS83820 is not set
# CONFIG_HAMACHI is not set
# CONFIG_YELLOWFIN is not set
# CONFIG_R8169 is not set
# CONFIG_SIS190 is not set
# CONFIG_SKGE is not set
# CONFIG_SKY2 is not set
# CONFIG_SK98LIN is not set
# CONFIG_VIA_VELOCITY is not set
CONFIG_TIGON3=y
CONFIG_BNX2=y
# CONFIG_QLA3XXX is not set
# CONFIG_ATL1 is not set
CONFIG_NETDEV_10000=y
# CONFIG_CHELSIO_T1 is not set
# CONFIG_CHELSIO_T3 is not set
# CONFIG_IXGBE is not set
# CONFIG_IXGB is not set
CONFIG_S2IO=m
# CONFIG_S2IO_NAPI is not set
# CONFIG_MYRI10GE is not set
# CONFIG_NETXEN_NIC is not set
# CONFIG_NIU is not set
# CONFIG_MLX4_CORE is not set
# CONFIG_TEHUTI is not set
# CONFIG_TR is not set

#
# Wireless LAN
#
# CONFIG_WLAN_PRE80211 is not set
# CONFIG_WLAN_80211 is not set
```

```
#
# USB Network Adapters
#
# CONFIG_USB_CATC is not set
# CONFIG_USB_KAWETH is not set
# CONFIG_USB_PEGASUS is not set
# CONFIG_USB_RTL8150 is not set
# CONFIG_USB_USBNET_MII is not set
# CONFIG_USB_USBNET is not set
# CONFIG_WAN is not set
# CONFIG_FDDI is not set
# CONFIG_HIPPI is not set
# CONFIG_PPP is not set
# CONFIG_SLIP is not set
# CONFIG_NET_FC is not set
# CONFIG_SHAPER is not set
CONFIG_NETCONSOLE=y
# CONFIG_NETCONSOLE_DYNAMIC is not set
CONFIG_NETPOLL=y
# CONFIG_NETPOLL_TRAP is not set
CONFIG_NET_POLL_CONTROLLER=y
# CONFIG_ISDN is not set
# CONFIG_PHONE is not set

#
# Input device support
#
CONFIG_INPUT=y
# CONFIG_INPUT_FF_MEMLESS is not set
# CONFIG_INPUT_POLLDEV is not set

#
# Userland interfaces
#
CONFIG_INPUT_MOUSEDEV=y
CONFIG_INPUT_MOUSEDEV_PSAUX=y
CONFIG_INPUT_MOUSEDEV_SCREEN_X=1024
CONFIG_INPUT_MOUSEDEV_SCREEN_Y=768
# CONFIG_INPUT_JOYDEV is not set
CONFIG_INPUT_EVDEV=y
# CONFIG_INPUT_EVBUG is not set

#
# Input Device Drivers
#
CONFIG_INPUT_KEYBOARD=y
CONFIG_KEYBOARD_ATKBD=y
# CONFIG_KEYBOARD_SUNKBD is not set
```

```
# CONFIG_KEYBOARD_LKKBD is not set
# CONFIG_KEYBOARD_XTKBD is not set
# CONFIG_KEYBOARD_NEWTON is not set
# CONFIG_KEYBOARD_STOWAWAY is not set
CONFIG_INPUT_MOUSE=y
CONFIG_MOUSE_PS2=y
CONFIG_MOUSE_PS2_ALPS=y
CONFIG_MOUSE_PS2_LOGIPS2PP=y
CONFIG_MOUSE_PS2_SYNAPTICS=y
CONFIG_MOUSE_PS2_LIFEBOOK=y
CONFIG_MOUSE_PS2_TRACKPOINT=y
# CONFIG_MOUSE_PS2_TOUCHKIT is not set
# CONFIG_MOUSE_SERIAL is not set
# CONFIG_MOUSE_APPLETOUCH is not set
# CONFIG_MOUSE_VSXXXAA is not set
# CONFIG_INPUT_JOYSTICK is not set
# CONFIG_INPUT_TABLET is not set
# CONFIG_INPUT_TOUCHSCREEN is not set
# CONFIG_INPUT_MISC is not set
```

```
#
# Hardware I/O ports
```

```
#
CONFIG_SERIO=y
CONFIG_SERIO_I8042=y
# CONFIG_SERIO_SERPORT is not set
# CONFIG_SERIO_CT82C710 is not set
# CONFIG_SERIO_PCIPS2 is not set
CONFIG_SERIO_LIBPS2=y
# CONFIG_SERIO_RAW is not set
# CONFIG_GAMEPORT is not set
```

```
#
# Character devices
```

```
#
CONFIG_VT=y
CONFIG_VT_CONSOLE=y
CONFIG_HW_CONSOLE=y
# CONFIG_VT_HW_CONSOLE_BINDING is not set
# CONFIG_SERIAL_NONSTANDARD is not set
```

```
#
# Serial drivers
```

```
#
CONFIG_SERIAL_8250=y
CONFIG_SERIAL_8250_CONSOLE=y
CONFIG_FIX_EARLYCON_MEM=y
CONFIG_SERIAL_8250_PCI=y
```

```
CONFIG_SERIAL_8250_PNP=y
CONFIG_SERIAL_8250_NR_UARTS=4
CONFIG_SERIAL_8250_RUNTIME_UARTS=4
# CONFIG_SERIAL_8250_EXTENDED is not set
```

```
#
# Non-8250 serial port support
#
CONFIG_SERIAL_CORE=y
CONFIG_SERIAL_CORE_CONSOLE=y
# CONFIG_SERIAL_JSM is not set
CONFIG_UNIX98_PTYS=y
CONFIG_LEGACY_PTYS=y
CONFIG_LEGACY_PTY_COUNT=256
# CONFIG_IPMI_HANDLER is not set
CONFIG_HW_RANDOM=y
CONFIG_HW_RANDOM_INTEL=y
CONFIG_HW_RANDOM_AMD=y
# CONFIG_NVRAM is not set
CONFIG_RTC=y
# CONFIG_R3964 is not set
# CONFIG_APPLICOM is not set
# CONFIG_MWAVE is not set
# CONFIG_PC8736x_GPIO is not set
CONFIG_RAW_DRIVER=y
CONFIG_MAX_RAW_DEVS=256
CONFIG_HPET=y
# CONFIG_HPET_RTC_IRQ is not set
CONFIG_HPET_MMAP=y
# CONFIG_HANGCHECK_TIMER is not set
# CONFIG_TCG_TPM is not set
# CONFIG_TELCLOCK is not set
CONFIG_DEVPORT=y
# CONFIG_I2C is not set
```

```
#
# SPI support
#
# CONFIG_SPI is not set
# CONFIG_SPI_MASTER is not set
# CONFIG_W1 is not set
# CONFIG_POWER_SUPPLY is not set
# CONFIG_HWMON is not set
# CONFIG_WATCHDOG is not set
```

```
#
# Sonics Silicon Backplane
#
```

```
CONFIG_SSB_POSSIBLE=y
CONFIG_SSB=y
CONFIG_SSB_PCIHOST_POSSIBLE=y
CONFIG_SSB_PCIHOST=y
# CONFIG_SSB_DEBUG is not set
CONFIG_SSB_DRIVER_PCICORE_POSSIBLE=y
CONFIG_SSB_DRIVER_PCICORE=y

#
# Multifunction device drivers
#
# CONFIG_MFD_SM501 is not set

#
# Multimedia devices
#
# CONFIG_VIDEO_DEV is not set
# CONFIG_DVB_CORE is not set
CONFIG_DAB=y
# CONFIG_USB_DABUSB is not set

#
# Graphics support
#
CONFIG_AGP=y
CONFIG_AGP_AMD64=y
CONFIG_AGP_INTEL=y
# CONFIG_AGP_SIS is not set
# CONFIG_AGP_VIA is not set
# CONFIG_DRM is not set
# CONFIG_VGASTATE is not set
# CONFIG_VIDEO_OUTPUT_CONTROL is not set
# CONFIG_FB is not set
# CONFIG_BACKLIGHT_LCD_SUPPORT is not set

#
# Display device support
#
# CONFIG_DISPLAY_SUPPORT is not set

#
# Console display driver support
#
CONFIG_VGA_CONSOLE=y
CONFIG_VGACON_SOFT_SCROLLBACK=y
CONFIG_VGACON_SOFT_SCROLLBACK_SIZE=256
CONFIG_VIDEO_SELECT=y
CONFIG_DUMMY_CONSOLE=y
```

```
#
# Sound
#
CONFIG_SOUND=y

#
# Advanced Linux Sound Architecture
#
# CONFIG_SND is not set

#
# Open Sound System
#
CONFIG_SOUND_PRIME=y
# CONFIG_SOUND_TRIDENT is not set
# CONFIG_SOUND_MSNDCLAS is not set
# CONFIG_SOUND_MSNDPIN is not set
# CONFIG_SOUND_OSS is not set
CONFIG_HID_SUPPORT=y
CONFIG_HID=y
# CONFIG_HID_DEBUG is not set
# CONFIG_HIDRAW is not set

#
# USB Input Devices
#
CONFIG_USB_HID=y
# CONFIG_USB_HIDINPUT_POWERBOOK is not set
# CONFIG_HID_FF is not set
# CONFIG_USB_HIDDEV is not set
CONFIG_USB_SUPPORT=y
CONFIG_USB_ARCH_HAS_HCD=y
CONFIG_USB_ARCH_HAS_OHCI=y
CONFIG_USB_ARCH_HAS_EHCI=y
CONFIG_USB=y
# CONFIG_USB_DEBUG is not set

#
# Miscellaneous USB options
#
CONFIG_USB_DEVICEFS=y
# CONFIG_USB_DEVICE_CLASS is not set
# CONFIG_USB_DYNAMIC_MINORS is not set
# CONFIG_USB_SUSPEND is not set
# CONFIG_USB_PERSIST is not set
# CONFIG_USB_OTG is not set
```

```
#
# USB Host Controller Drivers
#
CONFIG_USB_EHCI_HCD=y
# CONFIG_USB_EHCI_SPLIT_ISO is not set
# CONFIG_USB_EHCI_ROOT_HUB_TT is not set
# CONFIG_USB_EHCI_TT_NEWSCHED is not set
# CONFIG_USB_ISP116X_HCD is not set
CONFIG_USB_OHCI_HCD=y
# CONFIG_USB_OHCI_HCD_SSB is not set
# CONFIG_USB_OHCI_BIG_ENDIAN_DESC is not set
# CONFIG_USB_OHCI_BIG_ENDIAN_MMIO is not set
CONFIG_USB_OHCI_LITTLE_ENDIAN=y
CONFIG_USB_UHCI_HCD=y
# CONFIG_USB_SL811_HCD is not set
# CONFIG_USB_R8A66597_HCD is not set

#
# USB Device Class drivers
#
# CONFIG_USB_ACM is not set
CONFIG_USB_PRINTER=y

#
# NOTE: USB_STORAGE enables SCSI, and 'SCSI disk support'
#

#
# may also be needed; see USB_STORAGE Help for more information
#
CONFIG_USB_STORAGE=y
# CONFIG_USB_STORAGE_DEBUG is not set
# CONFIG_USB_STORAGE_DATAFAB is not set
# CONFIG_USB_STORAGE_FREECOM is not set
# CONFIG_USB_STORAGE_ISD200 is not set
# CONFIG_USB_STORAGE_DPCM is not set
# CONFIG_USB_STORAGE_USBAT is not set
# CONFIG_USB_STORAGE_SDDR09 is not set
# CONFIG_USB_STORAGE_SDDR55 is not set
# CONFIG_USB_STORAGE_JUMPSHOT is not set
# CONFIG_USB_STORAGE_ALAUDA is not set
# CONFIG_USB_STORAGE_KARMA is not set
# CONFIG_USB_LIBUSUAL is not set

#
# USB Imaging devices
#
# CONFIG_USB_MDC800 is not set
```

```
# CONFIG_USB_MICROTEK is not set
CONFIG_USB_MON=y

#
# USB port drivers
#

#
# USB Serial Converter support
#
# CONFIG_USB_SERIAL is not set

#
# USB Miscellaneous drivers
#
# CONFIG_USB_EMI62 is not set
# CONFIG_USB_EMI26 is not set
# CONFIG_USB_ADUTUX is not set
# CONFIG_USB_AUERSWALD is not set
# CONFIG_USB_RIO500 is not set
# CONFIG_USB_LEGOTOWER is not set
# CONFIG_USB_LCD is not set
# CONFIG_USB_BERRY_CHARGE is not set
# CONFIG_USB_LED is not set
# CONFIG_USB_CYPRESS_CY7C63 is not set
# CONFIG_USB_CYTHERM is not set
# CONFIG_USB_PHIDGET is not set
# CONFIG_USB_IDMOUSE is not set
# CONFIG_USB_FTDI_ELAN is not set
# CONFIG_USB_APPLEDISPLAY is not set
# CONFIG_USB_SISUSBVGA is not set
# CONFIG_USB_LD is not set
# CONFIG_USB_TRANCEVIBRATOR is not set
# CONFIG_USB_IOWARRIOR is not set
# CONFIG_USB_TEST is not set

#
# USB DSL modem support
#

#
# USB Gadget Support
#
# CONFIG_USB_GADGET is not set
# CONFIG_MMC is not set
# CONFIG_NEW_LEDS is not set
# CONFIG_INFINIBAND is not set
# CONFIG_EDAC is not set
```

```
# CONFIG_RTC_CLASS is not set
# CONFIG_DMADEVICES is not set
CONFIG_VIRTUALIZATION=y
# CONFIG_KVM is not set

#
# Userspace I/O
#
# CONFIG_UIO is not set

#
# Firmware Drivers
#
# CONFIG_EDD is not set
# CONFIG_DELL_RBU is not set
# CONFIG_DCDBAS is not set
CONFIG_DMIID=y

#
# File systems
#
CONFIG_EXT2_FS=y
CONFIG_EXT2_FS_XATTR=y
CONFIG_EXT2_FS_POSIX_ACL=y
# CONFIG_EXT2_FS_SECURITY is not set
# CONFIG_EXT2_FS_XIP is not set
CONFIG_EXT3_FS=y
CONFIG_EXT3_FS_XATTR=y
CONFIG_EXT3_FS_POSIX_ACL=y
# CONFIG_EXT3_FS_SECURITY is not set
# CONFIG_EXT4DEV_FS is not set
CONFIG_JBD=y
# CONFIG_JBD_DEBUG is not set
CONFIG_FS_MBCACHE=y
CONFIG_REISERFS_FS=y
# CONFIG_REISERFS_CHECK is not set
# CONFIG_REISERFS_PROC_INFO is not set
CONFIG_REISERFS_FS_XATTR=y
CONFIG_REISERFS_FS_POSIX_ACL=y
# CONFIG_REISERFS_FS_SECURITY is not set
# CONFIG_JFS_FS is not set
CONFIG_FS_POSIX_ACL=y
# CONFIG_XFS_FS is not set
# CONFIG_GFS2_FS is not set
# CONFIG_OCFS2_FS is not set
# CONFIG_MINIX_FS is not set
# CONFIG_ROMFS_FS is not set
CONFIG_INOTIFY=y
```

```
CONFIG_INOTIFY_USER=y
# CONFIG_QUOTA is not set
CONFIG_DNOTIFY=y
# CONFIG_AUTOFS_FS is not set
CONFIG_AUTOFS4_FS=y
# CONFIG_FUSE_FS is not set
CONFIG_GENERIC_ACL=y

#
# CD-ROM/DVD Filesystems
#
CONFIG_ISO9660_FS=y
CONFIG_JOLIET=y
# CONFIG_ZISOFS is not set
# CONFIG_UDF_FS is not set

#
# DOS/FAT/NT Filesystems
#
CONFIG_FAT_FS=y
CONFIG_MSDOS_FS=y
CONFIG_VFAT_FS=y
CONFIG_FAT_DEFAULT_CODEPAGE=437
CONFIG_FAT_DEFAULT_IOCHARSET="iso8859-1"
# CONFIG_NTFS_FS is not set

#
# Pseudo filesystems
#
CONFIG_PROC_FS=y
CONFIG_PROC_KCORE=y
CONFIG_PROC_SYSCTL=y
CONFIG_SYSFS=y
CONFIG_TMPFS=y
CONFIG_TMPFS_POSIX_ACL=y
CONFIG_HUGETLBFS=y
CONFIG_HUGETLB_PAGE=y
# CONFIG_CONFIGFS_FS is not set

#
# Miscellaneous filesystems
#
# CONFIG_ADFS_FS is not set
# CONFIG_AFFS_FS is not set
# CONFIG_HFS_FS is not set
# CONFIG_HFSPLUS_FS is not set
# CONFIG_BEFS_FS is not set
# CONFIG_BFS_FS is not set
```

```
# CONFIG_EFS_FS is not set
# CONFIG_CRAMFS is not set
# CONFIG_VXFS_FS is not set
# CONFIG_HPFS_FS is not set
# CONFIG_QNX4FS_FS is not set
# CONFIG_SYSV_FS is not set
# CONFIG_UFS_FS is not set
CONFIG_NETWORK_FILESYSTEMS=y
CONFIG_NFS_FS=y
CONFIG_NFS_V3=y
# CONFIG_NFS_V3_ACL is not set
# CONFIG_NFS_V4 is not set
# CONFIG_NFS_DIRECTIO is not set
CONFIG_NFSD=y
CONFIG_NFSD_V3=y
# CONFIG_NFSD_V3_ACL is not set
# CONFIG_NFSD_V4 is not set
CONFIG_NFSD_TCP=y
CONFIG_ROOT_NFS=y
CONFIG_LOCKD=y
CONFIG_LOCKD_V4=y
CONFIG_EXPORTFS=y
CONFIG_NFS_COMMON=y
CONFIG_SUNRPC=y
# CONFIG_SUNRPC_BIND34 is not set
# CONFIG_RPCSEC_GSS_KRB5 is not set
# CONFIG_RPCSEC_GSS_SPKM3 is not set
# CONFIG_SMB_FS is not set
# CONFIG_CIFS is not set
# CONFIG_NCP_FS is not set
# CONFIG_CODA_FS is not set
# CONFIG_AFS_FS is not set
```

```
#
# Partition Types
#
CONFIG_PARTITION_ADVANCED=y
# CONFIG_ACORN_PARTITION is not set
# CONFIG_OSF_PARTITION is not set
# CONFIG_AMIGA_PARTITION is not set
# CONFIG_ATARI_PARTITION is not set
# CONFIG_MAC_PARTITION is not set
CONFIG_MSDOS_PARTITION=y
# CONFIG_BSD_DISKLABEL is not set
# CONFIG_MINIX_SUBPARTITION is not set
# CONFIG_SOLARIS_X86_PARTITION is not set
# CONFIG_UNIXWARE_DISKLABEL is not set
# CONFIG_LDM_PARTITION is not set
```

```
# CONFIG_SGI_PARTITION is not set
# CONFIG_ULTRIX_PARTITION is not set
# CONFIG_SUN_PARTITION is not set
# CONFIG_KARMA_PARTITION is not set
CONFIG_EFI_PARTITION=y
# CONFIG_SYSV68_PARTITION is not set
CONFIG_NLS=y
CONFIG_NLS_DEFAULT="iso8859-1"
CONFIG_NLS_CODEPAGE_437=y
# CONFIG_NLS_CODEPAGE_737 is not set
# CONFIG_NLS_CODEPAGE_775 is not set
# CONFIG_NLS_CODEPAGE_850 is not set
# CONFIG_NLS_CODEPAGE_852 is not set
# CONFIG_NLS_CODEPAGE_855 is not set
# CONFIG_NLS_CODEPAGE_857 is not set
# CONFIG_NLS_CODEPAGE_860 is not set
# CONFIG_NLS_CODEPAGE_861 is not set
# CONFIG_NLS_CODEPAGE_862 is not set
# CONFIG_NLS_CODEPAGE_863 is not set
# CONFIG_NLS_CODEPAGE_864 is not set
# CONFIG_NLS_CODEPAGE_865 is not set
# CONFIG_NLS_CODEPAGE_866 is not set
# CONFIG_NLS_CODEPAGE_869 is not set
# CONFIG_NLS_CODEPAGE_936 is not set
# CONFIG_NLS_CODEPAGE_950 is not set
# CONFIG_NLS_CODEPAGE_932 is not set
# CONFIG_NLS_CODEPAGE_949 is not set
# CONFIG_NLS_CODEPAGE_874 is not set
# CONFIG_NLS_ISO8859_8 is not set
# CONFIG_NLS_CODEPAGE_1250 is not set
# CONFIG_NLS_CODEPAGE_1251 is not set
CONFIG_NLS_ASCII=y
CONFIG_NLS_ISO8859_1=y
# CONFIG_NLS_ISO8859_2 is not set
# CONFIG_NLS_ISO8859_3 is not set
# CONFIG_NLS_ISO8859_4 is not set
# CONFIG_NLS_ISO8859_5 is not set
# CONFIG_NLS_ISO8859_6 is not set
# CONFIG_NLS_ISO8859_7 is not set
# CONFIG_NLS_ISO8859_9 is not set
# CONFIG_NLS_ISO8859_13 is not set
# CONFIG_NLS_ISO8859_14 is not set
CONFIG_NLS_ISO8859_15=y
# CONFIG_NLS_KOI8_R is not set
# CONFIG_NLS_KOI8_U is not set
CONFIG_NLS_UTF8=y
# CONFIG_DLM is not set
CONFIG_INSTRUMENTATION=y
```

```
CONFIG_PROFILING=y
CONFIG_OPROFILE=y
CONFIG_KPROBES=y
# CONFIG_MARKERS is not set

#
# Kernel hacking
#
CONFIG_TRACE_IRQFLAGS_SUPPORT=y
# CONFIG_PRINTK_TIME is not set
# CONFIG_ENABLE_MUST_CHECK is not set
CONFIG_MAGIC_SYSRQ=y
CONFIG_UNUSED_SYMBOLS=y
CONFIG_DEBUG_FS=y
# CONFIG_HEADERS_CHECK is not set
CONFIG_DEBUG_KERNEL=y
# CONFIG_DEBUG_SHIRQ is not set
CONFIG_DETECT_SOFTLOCKUP=y
# CONFIG_SCHED_DEBUG is not set
# CONFIG_SCHEDSTATS is not set
CONFIG_TIMER_STATS=y
# CONFIG_DEBUG_SLAB is not set
# CONFIG_DEBUG_RT_MUTEXES is not set
# CONFIG_RT_MUTEX_TESTER is not set
# CONFIG_DEBUG_SPINLOCK is not set
# CONFIG_DEBUG_MUTEXES is not set
# CONFIG_DEBUG_LOCK_ALLOC is not set
# CONFIG_PROVE_LOCKING is not set
# CONFIG_LOCK_STAT is not set
# CONFIG_DEBUG_SPINLOCK_SLEEP is not set
# CONFIG_DEBUG_LOCKING_API_SELFTESTS is not set
# CONFIG_DEBUG_KOBJECT is not set
CONFIG_DEBUG_BUGVERBOSE=y
# CONFIG_DEBUG_INFO is not set
# CONFIG_DEBUG_VM is not set
# CONFIG_DEBUG_LIST is not set
# CONFIG_DEBUG_SG is not set
# CONFIG_FRAME_POINTER is not set
# CONFIG_FORCED_INLINING is not set
# CONFIG_BOOT_PRINTK_DELAY is not set
# CONFIG_RCU_TORTURE_TEST is not set
# CONFIG_LKDTM is not set
# CONFIG_FAULT_INJECTION is not set
# CONFIG_SAMPLES is not set
# CONFIG_DEBUG_RODATA is not set
# CONFIG_IOMMU_DEBUG is not set
CONFIG_DEBUG_STACKOVERFLOW=y
# CONFIG_DEBUG_STACK_USAGE is not set
```

```
#
# Security options
#
# CONFIG_KEYS is not set
# CONFIG_SECURITY is not set
# CONFIG_SECURITY_FILE_CAPABILITIES is not set
# CONFIG_CRYPTO is not set
```

```
#
# Library routines
#
CONFIG_BITREVERSE=y
# CONFIG_CRC_CCITT is not set
# CONFIG_CRC16 is not set
# CONFIG_CRC_ITU_T is not set
CONFIG_CRC32=y
# CONFIG_CRC7 is not set
# CONFIG_LIBCRC32C is not set
CONFIG_ZLIB_INFLATE=y
CONFIG_PLIST=y
CONFIG_HAS_IOMEM=y
CONFIG_HAS_IOPORT=y
CONFIG_HAS_DMA=y
```

```
#include <stdio.h>
#include <stdlib.h>
#include <sched.h>
#include <sys/syscall.h>
#include <unistd.h>
#include <signal.h>
#include <string.h>
#include <errno.h>
#include <libgen.h>
#include <fcntl.h>
#include <sys/types.h>
#include <sys/wait.h>
```

```
#include "clone.h"
```

```
extern pid_t getpgid(pid_t pid);
extern pid_t getsid(pid_t pid);
```

```
static const char* procname;
```

```
static void usage(const char *name)
{
    printf("usage: %s [-h] [-c] [-muipq]"
```

```

"[command [arg ..]]\n", name);
printf("\n");
printf(" -h this message\n");
printf("\n");
printf(" -c use 'clone' rather than 'unshare' system call\n");
printf(" -m mount namespace\n");
printf(" -u utsname namespace\n");
printf(" -i ipc namespace\n");
printf(" -p pid namespace\n");
printf(" -q mqueue namespace\n");
printf("\n");
printf("(C) Copyright IBM Corp. 2006\n");
printf("\n");
exit(1);
}

static void print_my_info(const char *procname, char *ttyname)
{
printf("procname %s, ttyname %s, pid %d, ppid %d, pgid %d, sid %d\n",
procname, ttyname, getpid(), getppid(), getpgid(0),
getsid(0));
}

/*
 * Copied following opentty() from Fedora's util-linux rpm
 * I just changed the "FATAL" message below from syslog()
 * to printf
 */
static void
opentty(const char * tty) {
    int i, fd, flags;

    fd = open(tty, O_RDWR | O_NONBLOCK);
    if (fd == -1) {
printf("FATAL: can't reopen tty: %s", strerror(errno));
        sleep(1);
        exit(1);
    }

    flags = fcntl(fd, F_GETFL);
    flags &= ~O_NONBLOCK;
    fcntl(fd, F_SETFL, flags);

    for (i = 0; i < fd; i++)
        close(i);
    for (i = 0; i < 3; i++)
        if (fd != i)
            dup2(fd, i);
}

```

```

        if (fd >= 3)
            close(fd);
    }
// Code copy end

int do_child(void *vargv)
{
    char **argv = (char **)vargv;
    execve(argv[0], argv, __environ);
    perror("execve");
    return 1;
}

int main(int argc, char *argv[])
{
    int c;
    unsigned long flags = 0;
    char ttyname[256];
    int status;
    int ret, use_clone = 0;
    int pid;

    procname = basename(argv[0]);

    memset(ttyname, '\0', sizeof(ttyname));
    readlink("/proc/self/fd/0", ttyname, sizeof(ttyname));

    while ((c = getopt(argc, argv, "+muiphqc")) != EOF) {
        switch (c) {
            case 'm': flags |= CLONE_NEWNS; break;
            case 'c': use_clone = 1; break;
            case 'u': flags |= CLONE_NEWUTS; break;
            case 'i': flags |= CLONE_NEWIPC; break;
            case 'p': flags |= CLONE_NEWPID|CLONE_NEWNS; break;
            case 'q': flags |= CLONE_NEWMQ; break;
            case 'h':
            default:
                usage(procname);
        }
    };

    argv = &argv[optind];
    argc = argc - optind;

    if (use_clone) {
        void *childstack, *stack = malloc(getpagesize());

        if (!stack) {

```

```

perror("malloc");
return -1;
}
childstack = stack + getpagesize());

printf("about to clone with %lx\n", flags);
pid = clone(do_child, childstack, flags, (void *)argv);
if (pid == -1) {
perror("clone");
return -1;
}
} else {
if ((pid = fork()) == 0) {
// Child.
print_my_info(procname, ttyname);

opentty(ttyname);

printf("about to unshare with %lx\n", flags);
ret = unshare(flags);
if (ret < 0) {
perror("unshare");
return 1;
}

return do_child((void*)argv);
}

}
printf("Parent waiting for pid %d\n", pid);
if ((ret = waitpid(pid, &status, __WALL)) < 0)
printf("waitpid() returns %d, errno %d\n", ret, errno);

exit(0);
}

```

---

Containers mailing list  
Containers@lists.linux-foundation.org  
<https://lists.linux-foundation.org/mailman/listinfo/containers>

### File Attachments

- 
- 1) [cgrouop-sched-config](#), downloaded 372 times
  - 2) [ns\\_exec.c](#), downloaded 348 times
- 

---

Subject: Re: [BUG]: Crash with CONFIG\_FAIR\_CGROUP\_SCHED=y

Posted by [Srivatsa Vaddagiri](#) on Fri, 09 Nov 2007 06:52:40 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

On Thu, Nov 08, 2007 at 03:48:05PM -0800, sukadev@us.ibm.com wrote:  
> With CONFIG\_FAIR\_CGROUP\_SCHED=y, following commands on 2.6.24-rc1 crash  
> the system.

Thanks for reporting the problem. It was caused because of the fact that current task isn't kept in its runqueue in case of sched\_fair class tasks.

With the patch below, I could run ns\_exec w/o any crash. Can you pls verify it works for you as well?

Ingo,

Once Suka verifies that the patch fixes his crash, I would request you to include the same in your tree and route it to Linus.

--

current task is not present in its runqueue in case of sched\_fair class tasks. Take care of this fact in rt\_mutex\_setprio(), sched\_setscheduler() and sched\_move\_task() routines.

Signed-off-by : Srivatsa Vaddagiri <vatsa@linux.vnet.ibm.com>

---

```
kernel/sched.c | 45 ++++++-----  
1 files changed, 25 insertions(+), 20 deletions(-)
```

Index: current/kernel/sched.c

```
=====
```

```
--- current.orig/kernel/sched.c  
+++ current/kernel/sched.c  
@@ -3986,11 +3986,13 @@ void rt_mutex_setprio(struct task_struct  
    oldprio = p->prio;  
    on_rq = p->se.on_rq;  
    running = task_running(rq, p);  
- if (on_rq) {  
+ if (on_rq)  
    dequeue_task(rq, p, 0);  
- if (running)  
- p->sched_class->put_prev_task(rq, p);  
- }  
+ /* current task is not kept in its runqueue in case of sched_fair class.  
+ * Hence we need the 'on_rq?' and 'running?' tests to be separate.  
+ */  
+ if (running)
```

```

+ p->sched_class->put_prev_task(rq, p);

if (rt_prio(prio))
    p->sched_class = &rt_sched_class;
@@ -3999,9 +4001,9 @@ void rt_mutex_setprio(struct task_struct

    p->prio = prio;

+ if (running)
+ p->sched_class->set_curr_task(rq);
    if (on_rq) {
- if (running)
- p->sched_class->set_curr_task(rq);
    enqueue_task(rq, p, 0);
    inc_load(rq, p);
    /*
@@ -4298,18 +4300,20 @@ recheck:
    update_rq_clock(rq);
    on_rq = p->se.on_rq;
    running = task_running(rq, p);
- if (on_rq) {
+ if (on_rq)
    deactivate_task(rq, p, 0);
- if (running)
- p->sched_class->put_prev_task(rq, p);
- }
+ /* current task is not kept in its runqueue in case of sched_fair class.
+ * Hence we need the 'on_rq?' and 'running?' tests to be separate.
+ */
+ if (running)
+ p->sched_class->put_prev_task(rq, p);

    oldprio = p->prio;
    __setscheduler(rq, p, policy, param->sched_priority);

+ if (running)
+ p->sched_class->set_curr_task(rq);
    if (on_rq) {
- if (running)
- p->sched_class->set_curr_task(rq);
    activate_task(rq, p, 0);
    /*
    * Reschedule if we are currently running on this runqueue and
@@ -7036,19 +7040,20 @@ void sched_move_task(struct task_struct
    running = task_running(rq, tsk);
    on_rq = tsk->se.on_rq;

- if (on_rq) {

```

```

+ if (on_rq)
    dequeue_task(rq, tsk, 0);
- if (unlikely(running))
-   tsk->sched_class->put_prev_task(rq, tsk);
- }
+ /* current task is not kept in its runqueue in case of sched_fair class.
+  * Hence we need the 'on_rq?' and 'running?' tests to be separate.
+  */
+ if (unlikely(running))
+   tsk->sched_class->put_prev_task(rq, tsk);

    set_task_cfs_rq(tsk);

- if (on_rq) {
-   if (unlikely(running))
-     tsk->sched_class->set_curr_task(rq);
+ if (unlikely(running))
+   tsk->sched_class->set_curr_task(rq);
+ if (on_rq)
    enqueue_task(rq, tsk, 0);
- }

done:
    task_rq_unlock(rq, &flags);

```

--  
 Regards,  
 vatsa

---

Containers mailing list  
 Containers@lists.linux-foundation.org  
<https://lists.linux-foundation.org/mailman/listinfo/containers>

---

Subject: Re: [BUG]: Crash with CONFIG\_FAIR\_CGROUP\_SCHED=y  
 Posted by [Dmitry Adamushko](#) on Fri, 09 Nov 2007 08:45:21 GMT  
[View Forum Message](#) <> [Reply to Message](#)

Hi Srivatsa,

```

> [ ... ]
> --
>
> current task is not present in its runqueue in case of sched_fair class
> tasks. Take care of this fact in rt_mutex_setprio(),
> sched_setscheduler() and sched_move_task() routines.
>

```

```

> Signed-off-by : Srivatsa Vaddagiri <vatsa@linux.vnet.ibm.com>
>
>
> ---
> kernel/sched.c | 45 ++++++-----
> 1 files changed, 25 insertions(+), 20 deletions(-)
>
> Index: current/kernel/sched.c
> =====
> --- current.orig/kernel/sched.c
> +++ current/kernel/sched.c
> @@ -3986,11 +3986,13 @@ void rt_mutex_setprio(struct task_struct
>     oldprio = p->prio;
>     on_rq = p->se.on_rq;
>     running = task_running(rq, p);
> -   if (on_rq) {
> +   if (on_rq)
>         dequeue_task(rq, p, 0);
> -   if (running)
> -       p->sched_class->put_prev_task(rq, p);
> -   }
> +   /* current task is not kept in its runqueue in case of sched_fair class.
> +    * Hence we need the 'on_rq?' and 'running?' tests to be separate.

```

Humm... the 'current' is not kept within the tree but  
current->se.on\_rq is supposed to be '1',  
so the old code looks ok to me (at least for the 'leaf' elements).

Maybe you were able to get more useful oops on your site?

```

> --
> Regards,
> vatsa
>

```

```

--
Best regards,
Dmitry Adamushko

```

---

Containers mailing list  
Containers@lists.linux-foundation.org  
<https://lists.linux-foundation.org/mailman/listinfo/containers>

---

Subject: Re: [BUG]: Crash with CONFIG\_FAIR\_CGROUP\_SCHED=y  
Posted by [Srivatsa Vaddagiri](#) on Fri, 09 Nov 2007 10:04:38 GMT

On Fri, Nov 09, 2007 at 09:45:21AM +0100, Dmitry Adamushko wrote:  
> Humm... the 'current' is not kept within the tree but  
> current->se.on\_rq is supposed to be '1' ,  
> so the old code looks ok to me (at least for the 'leaf' elements).

You are damned right! Sorry my mistake with the previous analysis and (as I now find out) testing :(

There are couple of problems discovered by Suka's test:

- The test requires the cgroup filesystem to be mounted with atleast the cpu and ns options (i.e both namespace and cpu controllers are active in the same hierarchy).

```
# mkdir /dev/cpuctl
# mount -t cgroup -ocpu,ns none cpuctl
(or simply)
# mount -t cgroup none cpuctl -> Will activate all controllers
in same hierarchy.
```

- The test invokes clone() with CLONE\_NEWNS set. This causes a a new child to be created, also a new group (do\_fork->copy\_namespaces->ns\_cgroup\_clone->cgroup\_clone) and the child is attached to the new group (cgroup\_clone->attach\_task->sched\_move\_task). At this point in time, the child's scheduler related fields are uninitialized (including its on\_rq field, which it has inherited from parent). As a result sched\_move\_task thinks its on runqueue, when it isn't.

As a solution to this problem, I moved sched\_fork() call, which initializes scheduler related fields on a new task, before copy\_namespaces(). I am not sure though whether moving up will cause other side-effects. Do you see any issue?

- The second problem exposed by this test is that task\_new\_fair() assumes that parent and child will be part of the same group (which needn't be as this test shows). As a result, cfs\_rq->curr can be NULL for the child.

The solution is to test for curr pointer being NULL in task\_new\_fair().

With the patch below, I could run ns\_exec() fine w/o a crash.

Suka, can you verify whether this patch fixes your problem?

--

Fix copy\_namespace() <-> sched\_fork() dependency in do\_fork, by moving up sched\_fork().

Also introduce a NULL pointer check for 'curr' in task\_new\_fair().

Signed-off-by : Srivatsa Vaddagiri <vatsa@linux.vnet.ibm.com>

---

kernel/fork.c | 6 +++---  
kernel/sched\_fair.c | 2 +-  
2 files changed, 4 insertions(+), 4 deletions(-)

Index: current/kernel/fork.c

=====

--- current.orig/kernel/fork.c  
+++ current/kernel/fork.c  
@@ -1121,6 +1121,9 @@ static struct task\_struct \*copy\_process(  
p->blocked\_on = NULL; /\* not blocked yet \*/  
#endif

+ /\* Perform scheduler related setup. Assign this task to a CPU. \*/  
+ sched\_fork(p, clone\_flags);

+  
+ if ((retval = security\_task\_alloc(p))  
+ goto bad\_fork\_cleanup\_policy;  
+ if ((retval = audit\_alloc(p))  
@@ -1210,9 +1213,6 @@ static struct task\_struct \*copy\_process(  
INIT\_LIST\_HEAD(&p->ptrace\_children);  
INIT\_LIST\_HEAD(&p->ptrace\_list);

- /\* Perform scheduler related setup. Assign this task to a CPU. \*/  
- sched\_fork(p, clone\_flags);

-  
+ /\* Now that the task is set up, run cgroup callbacks if  
+ \* necessary. We need to run them before the task is visible  
+ \* on the tasklist. \*/

Index: current/kernel/sched\_fair.c

=====

--- current.orig/kernel/sched\_fair.c  
+++ current/kernel/sched\_fair.c  
@@ -1023,7 +1023,7 @@ static void task\_new\_fair(struct rq \*rq,  
place\_entity(cfs\_rq, se, 1);

+ if (sysctl\_sched\_child\_runs\_first && this\_cpu == task\_cpu(p) &&  
- curr->vruntime < se->vruntime) {  
+ curr && curr->vruntime < se->vruntime) {

/\*  
\* Upon rescheduling, sched\_class::put\_prev\_task() will place  
\* 'current' within the tree based on its new key value.

---

Containers mailing list  
Containers@lists.linux-foundation.org  
<https://lists.linux-foundation.org/mailman/listinfo/containers>

---

---

Subject: Re: [BUG]: Crash with CONFIG\_FAIR\_CGROUP\_SCHED=y  
Posted by [Ingo Molnar](#) on Fri, 09 Nov 2007 10:25:23 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

\* Srivatsa Vaddagiri <[vatsa@linux.vnet.ibm.com](mailto:vatsa@linux.vnet.ibm.com)> wrote:

> With the patch below, I could run ns\_exec() fine w/o a crash.  
>  
> Suka, can you verify whether this patch fixes your problem?

thanks, applied. I'll wait for confirmation from Suka before sending it to Linus.

Ingo

---

Containers mailing list  
Containers@lists.linux-foundation.org  
<https://lists.linux-foundation.org/mailman/listinfo/containers>

---

---

Subject: Re: [BUG]: Crash with CONFIG\_FAIR\_CGROUP\_SCHED=y  
Posted by [Dmitry Adamushko](#) on Fri, 09 Nov 2007 10:59:15 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

On 09/11/2007, Srivatsa Vaddagiri <[vatsa@linux.vnet.ibm.com](mailto:vatsa@linux.vnet.ibm.com)> wrote:

> [ ... ]  
>  
> As a solution to this problem, I moved sched\_fork() call, which  
> initializes scheduler related fields on a new task, before  
> copy\_namespaces(). I am not sure though whether moving up will  
> cause other side-effects. Do you see any issue?

Should be ok (IMHO and at first glance :-)

> - The second problem exposed by this test is that task\_new\_fair()  
> assumes that parent and child will be part of the same group (which  
> needn't be as this test shows). As a result, cfs\_rq->curr can be NULL  
> for the child.

Would it be better, logically-wise, to use `is_same_group()` instead?  
Although, we can't have 2 groups with `cfs_rq->curr != NULL` on the same CPU... so if the child belongs to another group, it's `cfs_rq->curr` is automatically NULL indeed.

--  
Best regards,  
Dmitry Adamushko

---

Containers mailing list  
Containers@lists.linux-foundation.org  
<https://lists.linux-foundation.org/mailman/listinfo/containers>

---

---

Subject: Re: [BUG]: Crash with CONFIG\_FAIR\_CGROUP\_SCHED=y  
Posted by [Srivatsa Vaddagiri](#) on Fri, 09 Nov 2007 12:01:16 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

On Fri, Nov 09, 2007 at 11:59:15AM +0100, Dmitry Adamushko wrote:  
> > - The second problem exposed by this test is that `task_new_fair()`  
> > assumes that parent and child will be part of the same group (which  
> > needn't be as this test shows). As a result, `cfs_rq->curr` can be NULL  
> > for the child.  
>  
> Would it be better, logically-wise, to use `is_same_group()` instead?  
> Although, we can't have 2 groups with `cfs_rq->curr != NULL` on the same  
> CPU... so if the child belongs to another group, it's `cfs_rq->curr` is  
> automatically NULL indeed.

Yeah ..I feel safe with an explicit `!curr` check, perhaps with a comment like below added to explain when `curr` can be NULL?

---  
kernel/sched\_fair.c | 1 +  
1 files changed, 1 insertion(+)

Index: current/kernel/sched\_fair.c

```
=====
--- current.orig/kernel/sched_fair.c
+++ current/kernel/sched_fair.c
@@ -1022,6 +1022,7 @@ static void task_new_fair(struct rq *rq,
    update_curr(cfs_rq);
    place_entity(cfs_rq, se, 1);
```

```
+ /* 'curr' will be NULL if the child belongs to a different group */
  if (sysctl_sched_child_runs_first && this_cpu == task_cpu(p) &&
      curr && curr->vruntime < se->vruntime) {
  /*
```

--  
Regards,  
vatsa

---

Containers mailing list  
Containers@lists.linux-foundation.org  
<https://lists.linux-foundation.org/mailman/listinfo/containers>

---

---

Subject: Re: [BUG]: Crash with CONFIG\_FAIR\_CGROUP\_SCHED=y  
Posted by [serue](#) on Fri, 09 Nov 2007 16:05:41 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Quoting Srivatsa Vaddagiri (vatsa@linux.vnet.ibm.com):  
> On Fri, Nov 09, 2007 at 09:45:21AM +0100, Dmitry Adamushko wrote:  
> > Humm... the 'current' is not kept within the tree but  
> > current->se.on\_rq is supposed to be '1' ,  
> > so the old code looks ok to me (at least for the 'leaf' elements).  
>  
> You are damned right! Sorry my mistake with the previous analysis and  
> (as I now find out) testing :(  
>  
> There are couple of problems discovered by Suka's test:  
>  
> - The test requires the cgroup filesystem to be mounted with  
> atleast the cpu and ns options (i.e both namespace and cpu  
> controllers are active in the same hierarchy).  
>  
> # mkdir /dev/cpuctl  
> # mount -t cgroup -ocpu,ns none cpuctl  
> (or simply)  
> # mount -t cgroup none cpuctl -> Will activate all controllers  
> in same hierarchy.  
>  
> - The test invokes clone() with CLONE\_NEWNS set. This causes a a new child  
> to be created, also a new group (do\_fork->copy\_namespaces->ns\_cgroup\_clone->  
> cgroup\_clone) and the child is attached to the new group (cgroup\_clone->  
> attach\_task->sched\_move\_task). At this point in time, the child's scheduler  
> related fields are uninitialized (including its on\_rq field, which it has  
> inherited from parent). As a result sched\_move\_task thinks its on  
> runqueue, when it isn't.

>  
 > As a solution to this problem, I moved sched\_fork() call, which  
 > initializes scheduler related fields on a new task, before  
 > copy\_namespaces(). I am not sure though whether moving up will  
 > cause other side-effects. Do you see any issue?  
 >  
 > - The second problem exposed by this test is that task\_new\_fair()  
 > assumes that parent and child will be part of the same group (which  
 > needn't be as this test shows). As a result, cfs\_rq->curr can be NULL  
 > for the child.  
 >  
 > The solution is to test for curr pointer being NULL in  
 > task\_new\_fair().  
 >  
 >  
 > With the patch below, I could run ns\_exec() fine w/o a crash.  
 >  
 > Suka, can you verify whether this patch fixes your problem?

Works on my machine. Thanks!

> --  
 >  
 > Fix copy\_namespace() <-> sched\_fork() dependency in do\_fork, by moving  
 > up sched\_fork().  
 >  
 > Also introduce a NULL pointer check for 'curr' in task\_new\_fair().  
 >  
 > Signed-off-by : Srivatsa Vaddagiri <vatsa@linux.vnet.ibm.com>

Tested-by: Serge Hallyn <serue@us.ibm.com>

>  
 > ---  
 > kernel/fork.c | 6 +++---  
 > kernel/sched\_fair.c | 2 +-  
 > 2 files changed, 4 insertions(+), 4 deletions(-)  
 >  
 > Index: current/kernel/fork.c  
 > =====  
 > --- current.orig/kernel/fork.c  
 > +++ current/kernel/fork.c  
 > @@ -1121,6 +1121,9 @@ static struct task\_struct \*copy\_process(  
 > p->blocked\_on = NULL; /\* not blocked yet \*/  
 > #endif  
 >  
 > + /\* Perform scheduler related setup. Assign this task to a CPU. \*/  
 > + sched\_fork(p, clone\_flags);

```

> +
> if ((retval = security_task_alloc(p)))
> goto bad_fork_cleanup_policy;
> if ((retval = audit_alloc(p)))
> @@ -1210,9 +1213,6 @@ static struct task_struct *copy_process(
> INIT_LIST_HEAD(&p->ptrace_children);
> INIT_LIST_HEAD(&p->ptrace_list);
>
> - /* Perform scheduler related setup. Assign this task to a CPU. */
> - sched_fork(p, clone_flags);
> -
> /* Now that the task is set up, run cgroup callbacks if
> * necessary. We need to run them before the task is visible
> * on the tasklist. */
> Index: current/kernel/sched_fair.c
> =====
> --- current.orig/kernel/sched_fair.c
> +++ current/kernel/sched_fair.c
> @@ -1023,7 +1023,7 @@ static void task_new_fair(struct rq *rq,
> place_entity(cfs_rq, se, 1);
>
> if (sysctl_sched_child_runs_first && this_cpu == task_cpu(p) &&
> - curr->vruntime < se->vruntime) {
> + curr && curr->vruntime < se->vruntime) {
> /*
> * Upon rescheduling, sched_class::put_prev_task() will place
> * 'current' within the tree based on its new key value.
>
> -----
> Containers mailing list
> Containers@lists.linux-foundation.org
> https://lists.linux-foundation.org/mailman/listinfo/containers

```

---

Containers mailing list  
Containers@lists.linux-foundation.org  
<https://lists.linux-foundation.org/mailman/listinfo/containers>

---

Subject: Re: [BUG]: Crash with CONFIG\_FAIR\_CGROUP\_SCHED=y  
Posted by [Sukadev Bhattiprolu](#) on Sat, 10 Nov 2007 23:13:47 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Serge E. Hallyn [serue@us.ibm.com] wrote:  
| Quoting Srivatsa Vaddagiri (vatsa@linux.vnet.ibm.com):  
| > On Fri, Nov 09, 2007 at 09:45:21AM +0100, Dmitry Adamushko wrote:  
| > > Humm... the 'current' is not kept within the tree but  
| > > current->se.on\_rq is supposed to be '1',  
| > > so the old code looks ok to me (at least for the 'leaf' elements).  
| >

| > You are damned right! Sorry my mistake with the previous analysis and  
| > (as I now find out) testing :(  
| >  
| > There are couple of problems discovered by Suka's test:  
| >  
| > - The test requires the cgroup filesystem to be mounted with  
| > atleast the cpu and ns options (i.e both namespace and cpu  
| > controllers are active in the same hierarchy).  
| >  
| > # mkdir /dev/cpuctl  
| > # mount -t cgroup -ocpu,ns none cpuctl  
| > (or simply)  
| > # mount -t cgroup none cpuctl -> Will activate all controllers  
| > in same hierarchy.  
| >  
| > - The test invokes clone() with CLONE\_NEWNS set. This causes a a new child  
| > to be created, also a new group (do\_fork->copy\_namespaces->ns\_cgroup\_clone->  
| > cgroup\_clone) and the child is attached to the new group (cgroup\_clone->  
| > attach\_task->sched\_move\_task). At this point in time, the child's scheduler  
| > related fields are uninitialized (including its on\_rq field, which it has  
| > inherited from parent). As a result sched\_move\_task thinks its on  
| > runqueue, when it isn't.  
| >  
| > As a solution to this problem, I moved sched\_fork() call, which  
| > initializes scheduler related fields on a new task, before  
| > copy\_namespaces(). I am not sure though whether moving up will  
| > cause other side-effects. Do you see any issue?  
| >  
| > - The second problem exposed by this test is that task\_new\_fair()  
| > assumes that parent and child will be part of the same group (which  
| > needn't be as this test shows). As a result, cfs\_rq->curr can be NULL  
| > for the child.  
| >  
| > The solution is to test for curr pointer being NULL in  
| > task\_new\_fair().  
| >  
| >  
| > With the patch below, I could run ns\_exec() fine w/o a crash.  
| >  
| > Suka, can you verify whether this patch fixes your problem?  
|  
| Works on my machine. Thanks!

And mine too. Thanks,

|  
| > --

| >  
| > Fix copy\_namespace() <-> sched\_fork() dependency in do\_fork, by moving  
| > up sched\_fork().  
| >  
| > Also introduce a NULL pointer check for 'curr' in task\_new\_fair().  
| >  
| > Signed-off-by : Srivatsa Vaddagiri <vatsa@linux.vnet.ibm.com>  
|  
| Tested-by: Serge Hallyn <serue@us.ibm.com>  
Tested-by: Sukadev Bhattiprolu <sukadev@us.ibm.com>

---

Containers mailing list  
Containers@lists.linux-foundation.org  
<https://lists.linux-foundation.org/mailman/listinfo/containers>

---