Subject: Re: Re: Hang with fair cgroup scheduler (reproducer is attached.)
Posted by Dhaval Giani on Fri, 14 Dec 2007 15:38:23 GMT
View Forum Message <> Reply to Message

On Fri, Dec 14, 2007 at 09:06:07PM +0530, Dhaval Giani wrote:
> On Fri, Dec 14, 2007 at 11:24:28PM +0900, kamezawa.hiroyu@jp.fujitsu.com wrote:
> > >> just to be sure SMP does matter here (most likely yes, I guess).
> > >>
> > >
> > >NUMA? I am not able to reproduce it here locally on an x86 8 CPU box.
> > >
> > yes. I used NUMA. 2 Nodes/4CPU x 2
> >
>
> OK, I got hold of an IA64 box, non numa and have managed to reproduce
> it.
>

Actually no, its another bug. Thanks for the program!

reg[3330]: NaT consumption 2216203124768 [1]
Modules linked in: ipv6 button binfmt_misc nls_iso8859_1 loop dm_mod tg3
ext3 jbd fan thermal processor sg mptspi mptscsih mptbase
scsi_transport_spi via82cxxx sd_mod scsi_mod ide_disk ide_core

Pid: 3330, CPU 3, comm: reg
psr : 00001210085a2010 ifs : 8000000000000308 ip : [<a0000001002e0481>]
Not tainted
ip is at rb_erase+0x301/0x7e0
unat: 0000000000000000 pfs : 0000000000000308 rsc : 0000000000000003
rnat: 0000000000000000 bsps: 0000000000000000 pr : a5565666a9556959
ldrs: 0000000000000000 ccv : 0000000000000000 fpsr: 0009804c0270033f
csd : 0000000000000000 ssd : 0000000000000000
b0 : a000000100076290 b6 : a000000100086b20 b7 : a000000100076360
f6 : 1003e0000000000000d34 f7 : 1003e000000000000000a
f8 : 1003e0000000000000000 f9 : 1003e0000000000000152
f10 : 1003e0000000000000004 f11 : 0fff2fffffffff0000000
r1 : a000000100c92030 r2 : e000000244bd0068 r3 : e000000245882000
r8 : e000000245882000 r9 : e000000241e6eda0 r10 : 0000000000000001
r11 : e0000002458f0070 r12 : e0000002458a7d80 r13 : e0000002458a0000
r14 : e000000244bd0060 r15 : e000000244bd0058 r16 : 0000000000000000
r17 : e000000245920d34 r18 : 0000000000000000 r19 : 0000000000000000
r20 : e000000245920c90 r21 : 0000000000000001 r22 : a000000100076360
r23 : a000000100a7f2f8 r24 : a000000100a7f2b0 r25 : e0000002458c0058
r26 : e000000004e05b10 r27 : 0000000000000001 r28 : 0000000000000000
r29 : a000000100a7f2e0 r30 : a000000100a7f2b0 r31 : e000000245920098

Call Trace:

```
 [<a000000100014a80>] show_stack+0x40/0xa0
                        sp=e0000002458a77d0 bsp=e0000002458a1310
 [<a000000100015380>] show_regs+0x840/0x880
                        sp=e0000002458a79a0 bsp=e0000002458a12b8
 [<a0000001000384a0>] die+0x1a0/0x2a0
                        sp=e0000002458a79a0 bsp=e0000002458a1270
 [<a0000001000385f0>] die_if_kernel+0x50/0x80
                        sp=e0000002458a79a0 bsp=e0000002458a1240
 [<a0000001005b1a80>] ia64_fault+0x1180/0x12a0
                        sp=e0000002458a79a0 bsp=e0000002458a11e0
 [<a000000100000b2a0>] ia64_leave_kernel+0x0/0x270
                        sp=e0000002458a7bb0 bsp=e0000002458a11e0
 [<a0000001002e0480>] rb_erase+0x300/0x7e0
                        sp=e0000002458a7d80 bsp=e0000002458a11a0
 [<a000000100076290>] __dequeue_entity+0x70/0xa0
                        sp=e0000002458a7d80 bsp=e0000002458a1170
 [<a000000100076300>] set_next_entity+0x40/0xa0
                        sp=e0000002458a7d80 bsp=e0000002458a1148
 [<a0000001000763a0>] set_curr_task_fair+0x40/0xa0
                        sp=e0000002458a7d80 bsp=e0000002458a1128
 [<a000000100078d90>] sched_move_task+0x2d0/0x340
                        sp=e0000002458a7d80 bsp=e0000002458a10e8
 [<a000000100078e20>] cpu_cgroup_attach+0x20/0x40
                        sp=e0000002458a7d90 bsp=e0000002458a10b0
 [<a0000001000e9370>] attach_task+0x9b0/0xac0
                        sp=e0000002458a7d90 bsp=e0000002458a1058
 [<a0000001000ed4e0>] cgroup_common_file_write+0x340/0x520
                        sp=e0000002458a7dc0 bsp=e0000002458a1010
 [<a0000001000eccd0>] cgroup_file_write+0xf0/0x300
                        sp=e0000002458a7dd0 bsp=e0000002458a0fc0
 [<a000000100017bbd0>] vfs_write+0x1d0/0x320
                        sp=e0000002458a7e20 bsp=e0000002458a0f70
 [<a000000100017c7f0>] sys_write+0x70/0xe0
                        sp=e0000002458a7e20 bsp=e0000002458a0ef8
 [<a000000100000b100>] ia64_ret_from_syscall+0x0/0x20
                        sp=e0000002458a7e30 bsp=e0000002458a0ef8
 [<a000000000010720>] __kernel_syscall_via_break+0x0/0x20
                        sp=e0000002458a8000 bsp=e0000002458a0ef8
```

--
regards,
Dhaval