Subject: mem leak
Posted by kapper on Thu, 24 Apr 2008 00:55:38 GMT
View Forum Message <> Reply to Message

hi
I suspect three openvz-servers administered by us do suffer similiar problems like
http://forum.openvz.org/index.php?t=tree&goto=26664

first - all RHEL4 based openvz-servers do not have this problem here.

only RHEL5 based openvz-systems are affected.

roughly 45 days after reboot - those machines - running not a very high load - have eaten up 8 GB swap and 4 GB RAM.

stopping VEs does only free several 100 Megs. stopping openvz alltogether also doesn't help - it seems the kernel is eating up all of this.

only a reboot fixes this.

latest verified kernel with this problem:
2.6.18-53.1.6.el5.028stab053.6

releases of kernels before are unknown but I remember having to reboot those boxes at least for two more times, so probably the problem was introduced during Summer 2007.

we're now running 2.6.18-53.1.13.el5.028stab053.10 - but this only for a few hours - so any suggestions are very welcome.

kindest regards
harald kapper

Subject: Re: mem leak
Posted by vaverin on Thu, 24 Apr 2008 06:40:55 GMT
View Forum Message <> Reply to Message

Hi Harald,

i would note that I do not see any kernel-related problems here:
linux kernel by design frees the memory only when it really required -- due performance optimisation reasons.

By this way kernel without any significat load will use all RAM, for example for disk cache.
however when somebody will request new memory -- kernel will drop oldest cache entries (it is quite fast operations) and give this memory to application.

Thereofre from kernel point of view the situation when kernel and running userspace programs eats toghether all memory -- is normal.

However from userspace point of view huge swap memory usage can point to memory leak in some aplication:
Huge swap usage points to huge userspace memory consumption (kernel do not uses swap for kernel structures, swap contains only rarely used programm's data). It may be normal, or may be caused by some memory leaks in applications.

thank you,
Vasily Averin

Subject: Re: mem leak
Posted by kapper on Thu, 24 Apr 2008 07:32:17 GMT
View Forum Message <> Reply to Message

Dear Vaverin,
I would absolutely agree - but....

thins on the hardware-node look bad after roughly 45 days uptime:

```
free -m
     total       used       free     shared    buffers     cached
Mem:  3946       3846        100          0         34        585
-/+ buffers/cache:415        515
Swap: 8191       8191          0
```

which similar 2.6.9 kernel based host-nodes do not share as a common resource-problem.

this view doesn't change when killing all VEs - swap is still at 8 GB USED - which is for an idle host - definitely nothing normal.

the same box after the reboot yesterday and several hours uptime:

```
free -m
    total       used       free     shared    buffers     cached
Mem: 3946       3782        163          0        201       2430
-/+ buffers/cache:1150     2796
Swap:8191          0       8191
```

also three different hardware-nodes do experience the same problem running the EL5 kernel, EL4 kernel based boxes do not suffer this very problem.

finally there are some error-messages that might enlighten anyone on this problem:

WARNING: Kernel Errors Present

[<ffffffff8005ee39>] error_exit+0x0/0x84 ...: 14 Time(s)

and on the other box

WARNING: Kernel Errors Present
[<ffffffff8005ee39>] error_exit+0x0/0x84 ...: 30 Time(s)

and the serverlog shows the outofmemory-killing going on as we invoke yum to upgrade the kernel...

Apr 23 23:49:22 hostname kernel: yum invoked oom-killer: gfp_mask=0x201d2, order=0, oomkilladj=0
Apr 23 23:49:22 hostname kernel:
Apr 23 23:49:22 hostname kernel: Call Trace:
Apr 23 23:49:22 hostname kernel: [<ffffffff800c6672>] out_of_memory+0x9f/0x25e
Apr 23 23:49:22 hostname kernel: [<ffffffff8000eda3>] __alloc_pages+0x242/0x332
Apr 23 23:49:22 hostname kernel: [<ffffffff800124c4>] __do_page_cache_readahead+0x95/0x1d9
Apr 23 23:49:22 hostname kernel: [<ffffffff8002999f>] sync_page+0x0/0x42
Apr 23 23:49:22 hostname kernel: [<ffffffff8005b8d4>] getnstimeofday+0x10/0x28
Apr 23 23:49:22 hostname kernel: [<ffffffff8009d3c7>] ktime_get_ts+0x1a/0x4d
Apr 23 23:49:22 hostname kernel: [<ffffffff800bf752>] delayacct_end+0x5d/0x86
Apr 23 23:49:22 hostname kernel: [<ffffffff80032e9c>] blockable_page_cache_readahead+0x53/0xb2
Apr 23 23:49:22 hostname kernel: [<ffffffff8002fdef>] make_ahead_window+0x82/0x9e
Apr 23 23:49:22 hostname kernel: [<ffffffff80013763>] page_cache_readahead+0x17f/0x1af
Apr 23 23:49:22 hostname kernel: [<ffffffff8000b992>] do_generic_mapping_read+0x126/0x3f8
Apr 23 23:49:22 hostname kernel: [<ffffffff8000c917>] file_read_actor+0x0/0x13c
Apr 23 23:49:22 hostname kernel: [<ffffffff8000bdb0>] __generic_file_aio_read+0x14c/0x190
Apr 23 23:49:22 hostname kernel: [<ffffffff80016800>] generic_file_aio_read+0x34/0x39
Apr 23 23:49:22 hostname kernel: [<ffffffff8000c633>] do_sync_read+0xc7/0x104
Apr 23 23:49:22 hostname kernel: [<ffffffff80066051>] do_page_fault+0x4e9/0x7ef
Apr 23 23:49:22 hostname kernel: [<ffffffff8009b432>] autoremove_wake_function+0x0/0x2e
Apr 23 23:49:22 hostname kernel: [<ffffffff80061597>] __sched_text_start+0x177/0xee2
Apr 23 23:49:22 hostname kernel: [<ffffffff8000af5c>] vfs_read+0xaa/0x150
Apr 23 23:49:22 hostname kernel: [<ffffffff8000b125>] fget_light+0x18/0x7c
Apr 23 23:49:22 hostname kernel: [<ffffffff80012c4a>] sys_pread64+0x54/0xc4
Apr 23 23:49:22 hostname kernel: [<ffffffff8005e166>] system_call+0x7e/0x83
Apr 23 23:49:22 hostname kernel:
Apr 23 23:49:22 hostname kernel: Mem-info:
Apr 23 23:49:22 hostname kernel: Node 0 DMA per-cpu:
Apr 23 23:49:22 hostname kernel: cpu 0 hot: high 0, batch 1 used:0
Apr 23 23:49:22 hostname kernel: cpu 0 cold: high 0, batch 1 used:0
Apr 23 23:49:22 hostname kernel: cpu 1 hot: high 0, batch 1 used:0
Apr 23 23:49:22 hostname kernel: cpu 1 cold: high 0, batch 1 used:0
Apr 23 23:49:22 hostname kernel: Node 0 DMA32 per-cpu:
Apr 23 23:49:22 hostname kernel: cpu 0 hot: high 186, batch 31 used:125
Apr 23 23:49:22 hostname kernel: cpu 0 cold: high 62, batch 15 used:50

Apr 23 23:49:22 hostname kernel: cpu 1 hot: high 186, batch 31 used:15
Apr 23 23:49:22 hostname kernel: cpu 1 cold: high 62, batch 15 used:14
Apr 23 23:49:22 hostname kernel: Node 0 Normal per-cpu:
Apr 23 23:49:22 hostname kernel: cpu 0 hot: high 186, batch 31 used:57
Apr 23 23:49:22 hostname kernel: cpu 0 cold: high 62, batch 15 used:57
Apr 23 23:49:22 hostname kernel: cpu 1 hot: high 186, batch 31 used:36
Apr 23 23:49:22 hostname kernel: cpu 1 cold: high 62, batch 15 used:46
Apr 23 23:49:22 hostname kernel: Node 0 HighMem per-cpu: empty
Apr 23 23:49:22 hostname kernel: Free pages: 21556kB (0kB HighMem)
Apr 23 23:49:22 hostname kernel: Active:482971 inactive:471530 dirty:500 writeback:0 unstable:0 free:5389 slab:26825 mapped-file:9603 mapped-anon:844193 pagetables:10871
Apr 23 23:49:22 hostname kernel: Node 0 DMA free:10980kB min:20kB low:24kB high:28kB active:0kB inactive:0kB present:10532kB pages_scanned:0 all_unreclaimable? yes
Apr 23 23:49:22 hostname kernel: lowmem_reserve[]: 0 3502 4006 4006
Apr 23 23:49:22 hostname kernel: Node 0 DMA32 free:9196kB min:7072kB low:8840kB high:10608kB active:1727524kB inactive:1716528kB present:3586732kB pages_scanned:3724 all_unreclaimable? no
Apr 23 23:49:22 hostname kernel: lowmem_reserve[]: 0 0 504 504
Apr 23 23:49:22 hostname kernel: Node 0 Normal free:1380kB min:1016kB low:1268kB high:1524kB active:204360kB inactive:169592kB present:516096kB pages_scanned:4 all_unreclaimable? no
Apr 23 23:49:22 hostname kernel: lowmem_reserve[]: 0 0 0 0
Apr 23 23:49:22 hostname kernel: Node 0 HighMem free:0kB min:128kB low:128kB high:128kB active:0kB inactive:0kB present:0kB pages_scanned:0 all_unreclaimable? no
Apr 23 23:49:22 hostname kernel: lowmem_reserve[]: 0 0 0 0
Apr 23 23:49:22 hostname kernel: Node 0 DMA: 3*4kB 5*8kB 3*16kB 4*32kB 4*64kB 2*128kB 2*256kB 1*512kB 1*1024kB 0*2048kB 2*4096kB = 10980kB
Apr 23 23:49:22 hostname kernel: Node 0 DMA32: 31*4kB 2*8kB 18*16kB 32*32kB 7*64kB 1*128kB 0*256kB 0*512kB 1*1024kB 1*2048kB 1*4096kB = 9196kB
Apr 23 23:49:22 hostname kernel: Node 0 Normal: 105*4kB 8*8kB 10*16kB 1*32kB 3*64kB 0*128kB 0*256kB 1*512kB 0*1024kB 0*2048kB 0*4096kB = 1380kB
Apr 23 23:49:22 hostname kernel: Node 0 HighMem: empty
Apr 23 23:49:22 hostname kernel: Swap cache: add 2338321, delete 2336791, find 32746461/32778627, race 0+92+75
Apr 23 23:49:22 hostname kernel: Free swap = 0kB
Apr 23 23:49:22 hostname kernel: Total swap = 8388600kB
Apr 23 23:49:22 hostname kernel: Free swap: 0kB
Apr 23 23:49:22 hostname kernel: 1179648 pages of RAM
Apr 23 23:49:22 hostname kernel: 169401 reserved pages
Apr 23 23:49:22 hostname kernel: 73976 pages shared
Apr 23 23:49:22 hostname kernel: 1531 pages swap cached
Apr 23 23:49:22 hostname kernel: Top 10 caches:
Apr 23 23:49:22 hostname kernel: filp   : size   1617920 objsize   256
Apr 23 23:49:22 hostname kernel: page_beancounter   : size   61296640 objsize   64
Apr 23 23:49:22 hostname kernel: buffer_head   : size   1445888 objsize   96
Apr 23 23:49:22 hostname kernel: radix_tree_node   : size   3624960 objsize   536
Apr 23 23:49:22 hostname kernel: size-2048   : size   1688960 objsize   2048
Apr 23 23:49:22 hostname kernel: size-128   : size   1941504 objsize   128

Apr 23 23:49:22 hostname kernel: vm_area_struct : size 5017600 objsize 176
Apr 23 23:49:22 hostname kernel: ext3_inode_cache : size 8697216 objsize 808
Apr 23 23:49:22 hostname kernel: task_struct : size 1425408 objsize 2192
Apr 23 23:49:22 hostname kernel: dentry_cache : size 5943296 objsize 248
Apr 23 23:49:22 hostname kernel: Out of memory: Killed process 12866 (apache2).
Apr 23 23:49:22 hostname kernel: Mem-info:
Apr 23 23:49:22 hostname kernel: Node 0 DMA per-cpu:
Apr 23 23:49:22 hostname kernel: cpu 0 hot: high 0, batch 1 used:0
Apr 23 23:49:22 hostname kernel: cpu 0 cold: high 0, batch 1 used:0
Apr 23 23:49:22 hostname kernel: cpu 1 hot: high 0, batch 1 used:0
Apr 23 23:49:22 hostname kernel: cpu 1 cold: high 0, batch 1 used:0
Apr 23 23:49:22 hostname kernel: Node 0 DMA32 per-cpu:
Apr 23 23:49:22 hostname kernel: cpu 0 hot: high 186, batch 31 used:125
Apr 23 23:49:22 hostname kernel: cpu 0 cold: high 62, batch 15 used:49
Apr 23 23:49:22 hostname kernel: cpu 1 hot: high 186, batch 31 used:15
Apr 23 23:49:22 hostname kernel: cpu 1 cold: high 62, batch 15 used:14
Apr 23 23:49:22 hostname kernel: Node 0 Normal per-cpu:
Apr 23 23:49:22 hostname kernel: cpu 0 hot: high 186, batch 31 used:22
Apr 23 23:49:22 hostname kernel: cpu 0 cold: high 62, batch 15 used:61
Apr 23 23:49:22 hostname kernel: cpu 1 hot: high 186, batch 31 used:36
Apr 23 23:49:22 hostname kernel: cpu 1 cold: high 62, batch 15 used:46
Apr 23 23:49:22 hostname kernel: Node 0 HighMem per-cpu: empty
Apr 23 23:49:22 hostname kernel: Free pages: 24136kB (0kB HighMem)
Apr 23 23:49:22 hostname kernel: Active:482668 inactive:471214 dirty:128 writeback:0 unstable:0 free:6034 slab:26851 mapped-file:9603 mapped-anon:844193 pagetables:10871
Apr 23 23:49:22 hostname kernel: Node 0 DMA free:10980kB min:20kB low:24kB high:28kB active:0kB inactive:0kB present:10532kB pages_scanned:0 all_unreclaimable? yes
Apr 23 23:49:22 hostname kernel: lowmem_reserve[]: 0 3502 4006 4006
Apr 23 23:49:22 hostname kernel: Node 0 DMA32 free:11536kB min:7072kB low:8840kB high:10608kB active:1726548kB inactive:1715316kB present:3586732kB pages_scanned:364 all_unreclaimable? no
Apr 23 23:49:22 hostname kernel: lowmem_reserve[]: 0 0 504 504
Apr 23 23:49:22 hostname kernel: Node 0 Normal free:1620kB min:1016kB low:1268kB high:1524kB active:204124kB inactive:169540kB present:516096kB pages_scanned:6 all_unreclaimable? no
Apr 23 23:49:22 hostname kernel: lowmem_reserve[]: 0 0 0 0
Apr 23 23:49:22 hostname kernel: Node 0 HighMem free:0kB min:128kB low:128kB high:128kB active:0kB inactive:0kB present:0kB pages_scanned:0 all_unreclaimable? no
Apr 23 23:49:22 hostname kernel: lowmem_reserve[]: 0 0 0 0
Apr 23 23:49:22 hostname kernel: Node 0 DMA: 3*4kB 5*8kB 3*16kB 4*32kB 4*64kB 2*128kB 2*256kB 1*512kB 1*1024kB 0*2048kB 2*4096kB = 10980kB
Apr 23 23:49:22 hostname kernel: Node 0 DMA32: 584*4kB 18*8kB 18*16kB 32*32kB 7*64kB 1*128kB 0*256kB 0*512kB 1*1024kB 1*2048kB 1*4096kB = 11536kB
Apr 23 23:49:22 hostname kernel: Node 0 Normal: 165*4kB 8*8kB 10*16kB 1*32kB 3*64kB 0*128kB 0*256kB 1*512kB 0*1024kB 0*2048kB 0*4096kB = 1620kB
Apr 23 23:49:22 hostname kernel: Node 0 HighMem: empty
Apr 23 23:49:22 hostname kernel: Swap cache: add 2338321, delete 2336791, find 32746461/32778627, race 0+92+75

Apr 23 23:49:22 hostname kernel: Free swap = 0kB
Apr 23 23:49:22 hostname kernel: Total swap = 8388600kB
Apr 23 23:49:22 hostname kernel: Free swap: 0kB
Apr 23 23:49:22 hostname kernel: 1179648 pages of RAM
Apr 23 23:49:22 hostname kernel: 169401 reserved pages
Apr 23 23:49:22 hostname kernel: 74036 pages shared
Apr 23 23:49:22 hostname kernel: 1531 pages swap cached
Apr 23 23:49:22 hostname kernel: OOM killed process apache2 (pid=12893, ve=419) exited, free=6096 gen=147.
Apr 23 23:54:21 hostname shutdown[14605]: shutting down for system reboot
Apr 23 23:54:21 hostname init: Switching to runlevel: 6
Apr 23 23:54:23 hostname smartd[22658]: smartd received signal 15: Terminated
Apr 23 23:54:23 hostname smartd[22658]: smartd is exiting (exit status 0)
Apr 23 23:54:33 hostname ntpd[6324]: ntpd exiting on signal 15
Apr 23 23:54:35 hostname kernel: Kernel logging (proc) stopped.
Apr 23 23:54:35 hostname kernel: Kernel log daemon terminating.
Apr 23 23:54:36 hostname exiting on signal 15
Apr 24 00:43:11 hostname syslogd 1.4.1: restart.

any advise is greatly appreciated
hk

Subject: Re: mem leak
Posted by vaverin on Thu, 24 Apr 2008 07:49:54 GMT
View Forum Message <> Reply to Message

Dear Harald,

kernel errors messages are very important. Any kernel error can lead to unpredictible results. It may be crashes or hangs, but memory leakage is also possible.

To found the real cause of the trouble we need to look at the first
error message (because the following errors can be caused by previous).

Could you please check your logs about the oops messages?
http://wiki.openvz.org/When_you_have_an_oops

Also I can do it for You if you give me an access to your node (of course via Private Messaging)

Thank you,
Vasily Averin

Subject: Re: mem leak
Posted by kapper on Thu, 24 Apr 2008 08:22:40 GMT

hi
the kernel was:
Linux 2.6.18-53.1.13.el5.028stab053.6 compiled by openvz for RHEL5.

now running
Linux hwnode 2.6.18-53.1.13.el5.028stab053.10 #1 SMP Tue Apr 1 14:45:45 MSD 2008 x86_64 x86_64 x86_64 GNU/Linux

---

the first error happened on April 21st 10.55 localtime.

should I post the full log or simply private-message it so someone?

the log starts (very first error):

Apr 21 10:55:02 hwnode kernel: apache invoked oom-killer: gfp_mask=0x200d2, order=0, oomkilladj=0
Apr 21 10:55:02 hwnode kernel:
Apr 21 10:55:02 hwnode kernel: Call Trace:
Apr 21 10:55:05 hwnode kernel: [<ffffffff800c6672>] out_of_memory+0x9f/0x25e
Apr 21 10:55:05 hwnode kernel: [<ffffffff8000eda3>] __alloc_pages+0x242/0x332
Apr 21 10:55:05 hwnode kernel: [<ffffffff80010900>] do_wp_page+0x315/0x6b5
Apr 21 10:55:05 hwnode kernel: [<ffffffff80008e8f>] __handle_mm_fault+0xf6c/0x1041
Apr 21 10:55:05 hwnode kernel: [<ffffffff80098ed5>] attach_pid+0x8c/0xb9
Apr 21 10:55:06 hwnode kernel: [<ffffffff8003cbe6>] remove_wait_queue+0x1c/0x2c
Apr 21 10:55:08 hwnode kernel: [<ffffffff80029214>] do_wait+0xad8/0xb74
Apr 21 10:55:08 hwnode kernel: [<ffffffff8006601e>] do_page_fault+0x4b6/0x7ef
Apr 21 10:55:09 hwnode kernel: [<ffffffff800f72ba>] compat_core_sys_select+0x1bf/0x1d0
Apr 21 10:55:10 hwnode kernel: [<ffffffff8005ee39>] error_exit+0x0/0x84
Apr 21 10:55:10 hwnode kernel:
Apr 21 10:55:10 hwnode kernel: Mem-info:
Apr 21 10:55:10 hwnode kernel: Node 0 DMA per-cpu:
Apr 21 10:55:10 hwnode kernel: cpu 0 hot: high 0, batch 1 used:0
Apr 21 10:55:10 hwnode kernel: cpu 0 cold: high 0, batch 1 used:0
Apr 21 10:55:10 hwnode kernel: cpu 1 hot: high 0, batch 1 used:0
Apr 21 10:55:10 hwnode kernel: cpu 1 cold: high 0, batch 1 used:0
Apr 21 10:55:10 hwnode kernel: Node 0 DMA32 per-cpu:
Apr 21 10:55:10 hwnode kernel: cpu 0 hot: high 186, batch 31 used:39
Apr 21 10:55:10 hwnode kernel: cpu 0 cold: high 62, batch 15 used:58
Apr 21 10:55:10 hwnode kernel: cpu 1 hot: high 186, batch 31 used:33
Apr 21 10:55:10 hwnode kernel: cpu 1 cold: high 62, batch 15 used:59
Apr 21 10:55:10 hwnode kernel: Node 0 Normal per-cpu:
Apr 21 10:55:10 hwnode kernel: cpu 0 hot: high 186, batch 31 used:18
Apr 21 10:55:10 hwnode kernel: cpu 0 cold: high 62, batch 15 used:50
Apr 21 10:55:10 hwnode kernel: cpu 1 hot: high 186, batch 31 used:28

Apr 21 10:55:10 hwnode kernel: cpu 1 cold: high 62, batch 15 used:56
Apr 21 10:55:10 hwnode kernel: Node 0 HighMem per-cpu: empty
Apr 21 10:55:10 hwnode kernel: Free pages: 20384kB (0kB HighMem)
Apr 21 10:55:10 hwnode kernel: Active:796844 inactive:123293 dirty:260 writeback:12112 unstable:0 free:5096 slab:60543 mapped-file:21985 mapped-anon:734226 pagetables:11267
Apr 21 10:55:10 hwnode kernel: Node 0 DMA free:10980kB min:20kB low:24kB high:28kB active:0kB inactive:0kB present:10532kB pages_scanned:0 all_unreclaimable? yes
Apr 21 10:55:10 hwnode kernel: lowmem_reserve[]: 0 3502 4006 4006
Apr 21 10:55:10 hwnode kernel: Node 0 DMA32 free:8452kB min:7072kB low:8840kB high:10608kB active:2861528kB inactive:454940kB present:3586732kB pages_scanned:8220 all_unreclaimable? no
Apr 21 10:55:10 hwnode kernel: lowmem_reserve[]: 0 0 504 504
Apr 21 10:55:10 hwnode kernel: Node 0 Normal free:952kB min:1016kB low:1268kB high:1524kB active:325976kB inactive:38232kB present:516096kB pages_scanned:96 all_unreclaimable? no
Apr 21 10:55:10 hwnode kernel: lowmem_reserve[]: 0 0 0 0
Apr 21 10:55:10 hwnode kernel: Node 0 HighMem free:0kB min:128kB low:128kB high:128kB active:0kB inactive:0kB present:0kB pages_scanned:0 all_unreclaimable? no
Apr 21 10:55:10 hwnode kernel: lowmem_reserve[]: 0 0 0 0
Apr 21 10:55:10 hwnode kernel: Node 0 DMA: 3*4kB 5*8kB 3*16kB 4*32kB 4*64kB 2*128kB 2*256kB 1*512kB 1*1024kB 0*2048kB 2*4096kB = 10980kB
Apr 21 10:55:10 hwnode kernel: Node 0 DMA32: 313*4kB 0*8kB 0*16kB 1*32kB 0*64kB 0*128kB 0*256kB 0*512kB 1*1024kB 1*2048kB 1*4096kB = 8452kB
Apr 21 10:55:10 hwnode kernel: Node 0 Normal: 96*4kB 1*8kB 1*16kB 1*32kB 0*64kB 0*128kB 0*256kB 1*512kB 0*1024kB 0*2048kB 0*4096kB = 952kB
Apr 21 10:55:10 hwnode kernel: Node 0 HighMem: empty
Apr 21 10:55:10 hwnode kernel: Swap cache: add 2307383, delete 2290105, find 31844745/31873186, race 0+89+67
Apr 21 10:55:10 hwnode kernel: Free swap = 0kB
Apr 21 10:55:10 hwnode kernel: Total swap = 8388600kB
Apr 21 10:55:10 hwnode kernel: Free swap: 0kB
Apr 21 10:55:10 hwnode kernel: 1179648 pages of RAM
Apr 21 10:55:10 hwnode kernel: 169401 reserved pages
Apr 21 10:55:10 hwnode kernel: 184758 pages shared
Apr 21 10:55:10 hwnode kernel: 17228 pages swap cached

thank you in advance
hk

Subject: Re: mem leak
Posted by kir on Thu, 24 Apr 2008 08:40:27 GMT
View Forum Message <> Reply to Message

Can you show us what 'top' shows on a "loaded" box?

Run top, then press 'M' (capital M, i.e. usually Shift+M), then copy-paste what you see on the screen to here.

Subject: Re: mem leak
Posted by kapper on Thu, 24 Apr 2008 08:45:38 GMT
View Forum Message <> Reply to Message

hi
absolutely - though I probably have to wait for 40 days now after rebooting all three boxes in order to get what you want.

now it shows:

top - 10:44:09 up 10:01, 1 user, load average: 0.06, 0.09, 0.18
Tasks: 318 total, 1 running, 317 sleeping, 0 stopped, 0 zombie
Cpu(s): 0.3%us, 0.3%sy, 0.0%ni, 99.2%id, 0.0%wa, 0.0%hi, 0.2%si, 0.0%st
Mem: 4040988k total, 3945596k used, 95392k free, 213668k buffers
Swap: 8388600k total, 0k used, 8388600k free, 2513336k cached

PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
20537	root	15	0	339m	112m	3760	S	0	2.9	0:13.96	yum-updatesd
10730	root	15	0	121m	105m	1912	S	0	2.7	4:17.38	pdns_recursor
8521	100	15	0	134m	50m	4504	S	0	1.3	8:49.31	mysqld
12793	103	15	0	160m	42m	5640	S	0	1.1	0:55.73	mysqld
20653	root	18	0	164m	27m	1684	S	0	0.7	0:10.86	python
21154	33	15	0	41660	24m	3560	S	0	0.6	0:13.76	apache
20866	33	15	0	41552	24m	3564	S	0	0.6	0:10.36	apache
20860	33	15	0	41432	23m	3252	S	0	0.6	0:09.10	apache
21312	33	15	0	41424	23m	3248	S	0	0.6	0:13.96	apache
20699	33	15	0	40996	23m	3588	S	0	0.6	0:16.37	apache
16445	101	15	0	161m	23m	5428	S	0	0.6	0:01.33	mysqld
21460	33	15	0	40992	23m	3248	S	0	0.6	0:08.47	apache
21341	33	15	0	40156	22m	3556	S	0	0.6	0:07.32	apache
22947	33	15	0	40316	22m	3220	S	0	0.6	0:00.94	apache
22672	33	15	0	40140	22m	3216	S	0	0.6	0:03.60	apache
14785	101	15	0	158m	22m	5492	S	0	0.6	0:00.93	mysqld

16420 101 15 0 158m 22m 5192 S 0 0.6 0:01.37 mysqld

20506 101 15 0 158m 21m 5360 S 0 0.6 0:00.76 mysqld

10995 101 15 0 158m 21m 5392 S 0 0.6 0:01.63 mysqld

17028 33 15 0 115m 16m 4176 S 0 0.4 0:00.17 apache2

16504 33 15 0 115m 16m 4036 S 0 0.4 0:03.14 apache2

21437 33 15 0 115m 16m 4020 S 0 0.4 0:02.39 apache2

16505 33 15 0 115m 16m 3996 S 0 0.4 0:02.69 apache2

21433 33 15 0 115m 16m 3996 S 0 0.4 0:02.73 apache2

16502 33 15 0 115m 16m 3960 S 0 0.4 0:02.76 apache2

16499 33 15 0 115m 16m 3952 S 0 0.4 0:02.56 apache2

16503 33 18 0 115m 16m 3952 S 0 0.4 0:02.03 apache2

18992 root 18 0 122m 10m 5764 S 0 0.3 0:00.06 apache2

19133 33 18 0 122m 9.8m 4744 S 0 0.2 0:10.95 apache2

19130 33 15 0 122m 9976 4728 S 0 0.2 0:11.36 apache2

21345 33 16 0 123m 9972 4712 S 0 0.2 0:02.85 apache2

21336 33 15 0 123m 9968 4704 S 0 0.2 0:07.21 apache2

19132 33 15 0 123m 9960 4712 S 0 0.2 0:11.40 apache2

19131 33 15 0 122m 9956 4720 S 0 0.2 0:15.42 apache2

19129 33 15 0 122m 9952 4708 S 0 0.2 0:06.72 apache2

21346 33 15 0 122m 9928 4664 S 0 0.2 0:13.24 apache2

21062 33 15 0 122m 9916 4660 S 0 0.2 0:17.51 apache2

21337 33 15 0 123m 9868 4620 S 0 0.2 0:03.11 apache2

19178 33 15 0 97348 9192 4544 S 0 0.2 0:00.40 apache2

19179 33 18 0 97580 9156 4536 S 0 0.2 0:00.30 apache2

19176	33	15	0	97604	9124	4512	S	0	0.2	0:00.41	apache2
21141	33	15	0	97600	8792	4148	S	0	0.2	0:00.29	apache2
19177	33	15	0	97320	8788	4176	S	0	0.2	0:00.33	apache2
19175	33	18	0	97568	8760	4156	S	0	0.2	0:00.35	apache2
19080	root	18	0	97032	8692	4508	S	0	0.2	0:00.06	apache2
20933	33	15	0	82368	8692	4248	S	0	0.2	0:00.28	apache2
21434	33	15	0	97564	8684	4116	S	0	0.2	0:00.20	apache2
15006	root	18	0	103m	8656	5276	S	0	0.2	0:00.06	apache2
20929	33	15	0	82360	8536	4128	S	0	0.2	0:00.51	apache2
20931	33	15	0	82364	8528	4140	S	0	0.2	0:00.69	apache2
18676	33	15	0	81952	8412	4044	S	0	0.2	0:00.42	apache2
18677	33	15	0	81876	8228	3940	S	0	0.2	0:00.32	apache2

Subject: Re: mem leak
Posted by vaverin on Thu, 24 Apr 2008 08:50:44 GMT
View Forum Message <> Reply to Message

Hi Harald,

oom-killer messages are not an errors,
oom-killer kills the processes to free the memory in case when kernel cannot free it by any another ways

messages like "WARNING: Kernel Errors Present" -- is more interesting for me.

Could you please give me an access to your node? If not -- I would like to look at the /var/log/dmesg file (to understand configuration of your node) , /var/log/messages* files (to find all kernel error messages) and to dmesg (to look at the current errors). Please sent it to me via PM or via email to vvs@parallels.com

thank you,
Vasily Averin

Subject: Re: mem leak
Posted by kir on Thu, 24 Apr 2008 08:53:58 GMT
View Forum Message <> Reply to Message

Yup, so far the situation looks pretty decent -- although yum-updatesd eats a bit too much memory, it's still OK.

Please re-check that in a week or two, and post your results here.

Subject: Re: mem leak
Posted by vaverin on Thu, 24 Apr 2008 10:49:07 GMT
View Forum Message <> Reply to Message

Hi Harald,

thank you for the logs. I've found the follwing messages:

1) Machine check exception message:
Mar 30 12:38:04 k4 kernel: Machine check events logged
it points to some hardware troubles. It may not be related to our current issue, but anyway -- I would recommend You to found its cause.
Some motherboards saves MCE events to the motherboard BIOS. Sometimes userspace daemon writes it to the /var/log/mce.log file.

2) disk-related issue:
Apr 10 13:50:05 k4 kernel:
ata2.00: exception Emask 0x0 SAct 0x70000 SErr 0x0 action 0x2 frozen
...
ata2: hard resetting port
ata2: port is slow to respond, please be patient (Status 0x80)
ata2: COMRESET failed (errno=-16)
ata2: hard resetting port
ata2: SATA link up 3.0 Gbps (SStatus 123 SControl 300)
ata2.00: configured for UDMA/133
ata2: EH complete
SCSI device sdb: 488397168 512-byte hdwr sectors (250059 MB)
sdb: Write Protect is off

scsi subsystem has detected some issue, and started Error Handler to resolve it. Ideally it should not lead to any troubles, however if this operation was not handled properly then it can be the cause of the following troubles on your node.

3) unfrtunaely I've not found any "WARNING: Kernel Errors Present" messages in the logs.

4) I've investigate OOM killer messages and they looks straange for me, I'll consult with my colleguaes and tell You about result a bit later.

Thank you,
Vasily Averin

Subject: Re: mem leak
Posted by vaverin on Thu, 24 Apr 2008 10:50:34 GMT
View Forum Message <> Reply to Message

btw, if you have similair issue on the some other nodes -- please send me its logs too, it would be useful.

thank you,
Vasily Averin