Subject: [RFC][PATCH 03/14] use pid_nr in procfs
Posted by Sukadev Bhattiprolu on Wed, 21 Mar 2007 03:19:00 GMT
View Forum Message <> Reply to Message

From: Sukadev Bhattiprolu <sukadev@us.ibm.com>
Subject: [RFC][PATCH 03/14] use pid_nr in procfs

With containers, a process can have different pid_t values in different
pid namespaces. To ensure we get the correct pid_t value in any context,
we should use pid_nr() function rather than directly accessing either
task->pid or pid->nr.

Signed-off-by: Sukadev Bhattiprolu <sukadev@us.ibm.com>
---
 fs/proc/array.c | 16 ++++++++++------
 fs/proc/base.c | 4 ++--
 2 files changed, 12 insertions(+), 8 deletions(-)

Index: 2.6.21-rc3-mm2/fs/proc/array.c
===
--- 2.6.21-rc3-mm2.orig/fs/proc/array.c
+++ 2.6.21-rc3-mm2/fs/proc/array.c
@@ -164,11 +164,15 @@ static inline char * task_state(struct t
  struct group_info *group_info;
  int g;
  struct fdtable *fdt = NULL;
+ pid_t ppid = 0;

  rcu_read_lock();
  tracer = tracehook_tracer_task(p);
  tracer_pid = tracer == NULL ? 0 : tracer->pid;

+ if (pid_alive(p))
+ ppid = pid_nr(task_tgid(rcu_dereference(p->parent)));
+
  buffer += sprintf(buffer,
   "State:\t%s\n"
   "SleepAVG:\t%lu%%\n"
@@ -180,8 +184,8 @@ static inline char * task_state(struct t
   "Gid:\t%d\t%d\t%d\t%d\n",
   get_task_state(p),
   (p->sleep_avg/1024)*100/(1020000000/1024),
-        p->tgid, p->pid,
-        pid_alive(p) ? rcu_dereference(p->parent)->tgid : 0,
+   pid_nr(task_tgid(p)), pid_nr(task_pid(p)),
+   ppid,
   tracer_pid,
   p->uid, p->euid, p->suid, p->fsuid,

```
   p->gid, p->egid, p->sgid, p->fsgid);
@@ -387,9 +391,9 @@ static int do_task_stat(struct task_stru
    stime = cputime_add(stime, sig->stime);
   }

-  sid = signal_session(sig);
-  pgid = process_group(task);
-  ppid = rcu_dereference(task->parent)->tgid;
+  sid = pid_nr(task_session(task));
+  pgid = pid_nr(task_pgrp(task));
+  ppid = pid_nr(task_tgid(rcu_dereference(task->parent)));

   unlock_task_sighand(task, &flags);
  }
@@ -419,7 +423,7 @@ static int do_task_stat(struct task_stru
  res = sprintf(buffer,"%d (%s) %c %d %d %d %d %d %u %lu \
 %lu %lu %lu %lu %lu %ld %ld %ld %ld %d 0 %llu %lu %ld %lu %lu %lu %lu %lu \
 %lu %lu %lu %lu %lu %lu %lu %lu %d %d %u %u %llu\n",
-  task->pid,
+  pid_nr(task_pid(task)),
   tcomm,
   state,
   ppid,
Index: 2.6.21-rc3-mm2/fs/proc/base.c
===================================================================
--- 2.6.21-rc3-mm2.orig/fs/proc/base.c
+++ 2.6.21-rc3-mm2/fs/proc/base.c
@@ -2132,7 +2132,7 @@ retry:
  task = NULL;
  pid = find_ge_pid(tgid);
  if (pid) {
-  tgid = pid->nr + 1;
+  tgid = pid_nr(pid) + 1;
   task = pid_task(pid, PIDTYPE_PID);
   /* What we to know is if the pid we have find is the
    * pid of a thread_group_leader.  Testing for task
@@ -2186,7 +2186,7 @@ int proc_pid_readdir(struct file * filp,
  for (task = next_tgid(tgid);
       task;
       put_task_struct(task), task = next_tgid(tgid + 1)) {
-  tgid = task->pid;
+  tgid = pid_nr(task_pid(task));
   filp->f_pos = tgid + TGID_OFFSET;
   if (proc_pid_fill_cache(filp, dirent, filldir, task, tgid) < 0) {
    put_task_struct(task);

--
```

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

Subject: Re: [RFC][PATCH 03/14] use pid_nr in procfs
Posted by ebiederm on Wed, 21 Mar 2007 07:11:31 GMT
View Forum Message <> Reply to Message

sukadev@us.ibm.com writes:

> From: Sukadev Bhattiprolu <sukadev@us.ibm.com>
> Subject: [RFC][PATCH 03/14] use pid_nr in procfs
>
> With containers, a process can have different pid_t values in different
> pid namespaces. To ensure we get the correct pid_t value in any context,
> we should use pid_nr() function rather than directly accessing either
> task->pid or pid->nr.

I almost agree here.

However pid_nr looks at current to figure out which pid namespace to
find the pid in. In the case of procfs we need to look at the
pid_namespace from the super block.

Eric

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

Subject: Re: [RFC][PATCH 03/14] use pid_nr in procfs
Posted by ebiederm on Wed, 21 Mar 2007 07:46:12 GMT
View Forum Message <> Reply to Message

sukadev@us.ibm.com writes:

> From: Sukadev Bhattiprolu <sukadev@us.ibm.com>
> Subject: [RFC][PATCH 03/14] use pid_nr in procfs
>
> With containers, a process can have different pid_t values in different
> pid namespaces. To ensure we get the correct pid_t value in any context,
> we should use pid_nr() function rather than directly accessing either
> task->pid or pid->nr.

To clarify my previous comment. I believe to get this right
we need a factor of pid_nr:
pid_t __pid_nr(struct pid_namespace *ns, struct pid *pid)

That we can pass the pid_namespace from the super block of the
mount to handle the proc case.

This is a case I have wanted to avoid but in this instance I don't
see any other way to get the code correct.

i.e. Since the pid namespace of /proc gets set at mount time who
you are should not vary the result it gives.

Eric

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

---

Subject: Re: [RFC][PATCH 03/14] use pid_nr in procfs
Posted by serue on Wed, 21 Mar 2007 17:09:32 GMT
View Forum Message <> Reply to Message

Quoting Eric W. Biederman (ebiederm@xmission.com):
> sukadev@us.ibm.com writes:
>
> > From: Sukadev Bhattiprolu <sukadev@us.ibm.com>
> > Subject: [RFC][PATCH 03/14] use pid_nr in procfs
> >
> > With containers, a process can have different pid_t values in different
> > pid namespaces. To ensure we get the correct pid_t value in any context,
> > we should use pid_nr() function rather than directly accessing either
> > task->pid or pid->nr.
>
> To clarify my previous comment. I believe to get this right
> we need a factor of pid_nr:
> pid_t __pid_nr(struct pid_namespace *ns, struct pid *pid)
>
> That we can pass the pid_namespace from the super block of the
> mount to handle the proc case.
>
> This is a case I have wanted to avoid but in this instance I don't
> see any other way to get the code correct.
>
> i.e. Since the pid namespace of /proc gets set at mount time who
> you are should not vary the result it gives.

Seems to make sense.

However since meaningful /proc usage under multiple pid namespaces will
require Dave's patch to accomodate multiple proc_mnt's, and since that
patch seems to keep having to change as this pid_ns patchset changes,
how about we drop this patch for now and have (a corrected version of)
it included in Dave's /proc patchset on top of the rest of this set?

-serge

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

---

Subject: Re: [RFC][PATCH 03/14] use pid_nr in procfs
Posted by Dave Hansen on Wed, 21 Mar 2007 17:55:03 GMT
View Forum Message <> Reply to Message

On Wed, 2007-03-21 at 12:09 -0500, Serge E. Hallyn wrote:
>
> However since meaningful /proc usage under multiple pid namespaces
> will
> require Dave's patch to accomodate multiple proc_mnt's, and since that
> patch seems to keep having to change as this pid_ns patchset changes,
> how about we drop this patch for now and have (a corrected version of)
> it included in Dave's /proc patchset on top of the rest of this set?

Sounds sane to me.

-- Dave

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

---