
Subject: [PATCH 4/4] pid: Limit cap_set_all to the current pid namespace

Posted by [ebiederm](#) on Wed, 12 Dec 2007 13:56:14 GMT

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

Use find_ge_pid in cap_set_all we only consider tasks in the current pid namespace.

This doesn't walk the task_list so on a good day in a pid namespace it should be more scalable, more importantly this is the same idiom used in proc and kill_pid_ns_info so it should be reasonably maintainable.

Signed-off-by: Eric W. Biederman <ebiederm@xmission.com>

kernel/capability.c | 15 ++++++++-----

1 files changed, 11 insertions(+), 4 deletions(-)

diff --git a/kernel/capability.c b/kernel/capability.c

index 652a2c5..a11bb24 100644

--- a/kernel/capability.c

+++ b/kernel/capability.c

@@ -198,12 +198,19 @@ static inline int cap_set_all(kernel_cap_t *effective,
kernel_cap_t *inheritable,
kernel_cap_t *permitted)

{

- struct task_struct *g, *target;

+ struct pid_namespace *ns = task_active_pid_ns(current);

+ struct task_struct *target;

int ret = -EPERM;

int found = 0;

+ struct pid *pid;

+ int nr;

- do_each_thread(g, target) {

- if (target == current || is_container_init(target->group_leader))

+ nr = 0;

+ while ((pid = find_ge_pid(nr + 1, ns))) {

+ nr = pid_nr_ns(pid, ns);

+ target = pid_task(pid, PIDTYPE_PID);

+

+ if (target == current || task_tgid(target) == ns->child_reaper)

continue;

found = 1;

if (security_capset_check(target, effective, inheritable,

@@ -211,7 +218,7 @@ static inline int cap_set_all(kernel_cap_t *effective,
continue;

ret = 0;

security_capset_set(target, effective, inheritable, permitted);

```
- } while_each_thread(g, target);  
+ }
```

```
    if (!found)  
        ret = 0;
```

```
--
```

1.5.3.rc6.17.g1911

Containers mailing list

Containers@lists.linux-foundation.org

<https://lists.linux-foundation.org/mailman/listinfo/containers>
