Subject: [PATCH 4/9] allow killing tasks in your own or child userns
Posted by serge on Thu, 17 Feb 2011 15:03:25 GMT
View Forum Message <> Reply to Message

Changelog:
 Dec 8: Fixed bug in my check_kill_permission pointed out by
         Eric Biederman.
 Dec 13: Apply Eric's suggestion to pass target task into kill_ok_by_cred()
         for clarity
 Dec 31: address comment by Eric Biederman:
  don't need cred/tcred in check_kill_permission.
 Jan 1: use const cred struct.
 Jan 11: Per Bastian Blank's advice, clean up kill_ok_by_cred().
 Feb 16: kill_ok_by_cred: fix bad parentheses

Signed-off-by: Serge E. Hallyn <serge.hallyn@canonical.com>
---
 kernel/signal.c | 30 ++++++++++++++++++++++--------
 1 files changed, 22 insertions(+), 8 deletions(-)

diff --git a/kernel/signal.c b/kernel/signal.c
index 4e3cff1..ffe4bdf 100644
--- a/kernel/signal.c
+++ b/kernel/signal.c
@@ -636,13 +636,33 @@ static inline bool si_fromuser(const struct siginfo *info)
 }

 /*
+ * called with RCU read lock from check_kill_permission()
+ */
+static inline int kill_ok_by_cred(struct task_struct *t)
+{
+ const struct cred *cred = current_cred();
+ const struct cred *tcred = __task_cred(t);
+
+ if (cred->user->user_ns == tcred->user->user_ns &&
+   (cred->euid == tcred->suid ||
+   cred->euid == tcred->uid ||
+   cred->uid == tcred->suid ||
+   cred->uid == tcred->uid))
+ return 1;
+
+ if (ns_capable(tcred->user->user_ns, CAP_KILL))
+ return 1;
+
+ return 0;
+}
+

```
+/*
  * Bad permissions for sending the signal
  * - the caller must hold the RCU read lock
  */
 static int check_kill_permission(int sig, struct siginfo *info,
      struct task_struct *t)
 {
- const struct cred *cred, *tcred;
  struct pid *sid;
  int error;

@@ -656,14 +676,8 @@ static int check_kill_permission(int sig, struct siginfo *info,
  if (error)
   return error;

- cred = current_cred();
- tcred = __task_cred(t);
  if (!same_thread_group(current, t) &&
-     (cred->euid ^ tcred->suid) &&
-     (cred->euid ^ tcred->uid) &&
-     (cred->uid  ^ tcred->suid) &&
-     (cred->uid  ^ tcred->uid) &&
-     !capable(CAP_KILL)) {
+     !kill_ok_by_cred(t)) {
   switch (sig) {
   case SIGCONT:
    sid = task_session(t);
--
1.7.0.4
```