Subject: Re: restarting tests/sleep
Posted by serue on Mon, 09 Jun 2008 16:23:09 GMT
View Forum Message <> Reply to Message

Quoting Nadia Derbey (Nadia.Derbey@bull.net):
> Serge E. Hallyn wrote:
>> Can you check your /var/log/messages for a 'protection fault' message?
>> Then, if there is one for 'sleep', run gdb on sleep, take the IP in
>> the protection fault msg, and do 'x/10 0x8005cd0' or whatever the
>> address is?
>
> No, I'm not getting a protection fault, and I've got nothing generated
> in the /var/log/messages. But when I reran the test to check what you're
> asking me to, I didn't reach the same point: I'm not reaching the [vdso]
> restoring. I'm getting an EFAULT. Here are the last lines from the
> output:
>
> =====
> DEBUG (cr.c::923) next memseg_t is: start bfd99000 end bfdae000 prot 3
> flag 50 offset 3221139456 fnam [stack]
> DEBUG (cr.c::969) Delete segment bfc96000 - bfcab000
> DEBUG (cr.c::971) Restore segment bfd99000 - bfdae000
> [1 cr.c:972 restore_mem()] inj_func(ic, inj_readbuf, 3, remfd,
> memseg->start, memseg->end - memseg->start) : -14
> ERROR (cr.c:926) unexpected item name: '' (size=0)
> =====

Yeah that's insurmountable - notice the stack in the process which was
fork()ed to be the restarted process topped at bfcab000, while the
checkpointed stack topped at bfdae000. You're not allowed to write
above the stack. So the only things to do are

1. keep trying the restart in the hopes you get a task with
   stack topping at or above bfdae000
2. if the checkpointed stack is too high to be likely to be
   restartable, generate a new checkpoint image and you
   should get a lower stack top.

(Dave, maybe you had other ideas I haven't considered)

Good thing cryo is just a proof of concept to exploit the kernel code :)

thanks,
-serge