Subject: Re: [RFC][PATCH 2/7] VPIDs: pid/vpid conversions
Posted by Herbert Poetzl on Tue, 21 Feb 2006 23:17:29 GMT
View Forum Message <> Reply to Message

On Tue, Feb 21, 2006 at 07:19:01PM +0300, Kirill Korotaev wrote:
>>>The only correct thing you noticed is get_xpid on alpha. But this is
>>>in fact a simple bug and half a year before we didn't care much for
>>>archs others than i386/x86-64/ia64. That's it.
>>sidenote on that, maybe the various archs could
>>switch to C implementations of those 'special'
>>get_xpid() and friends, as I do not think they
>>are a) done that often (might be wrong there)
>>and b) recent gcc should get that right now anyway
>I also wonder why it was required and can't be done in normal way...
>Maybe worth trying to switch to C, really.
definitely
>>>For example, networking is coupled with sysctl, which in turn are
>>>coupled with proc filesystem. And sysfs! You even added a piece of code
>>>in net/core/net-sysfs.c in your patch, which is a dirty hack.
>>>Another example, mqueues and other subsystems which use netlinks and
>>>also depend on network context.
>>>shmem/IPC is dependand on file system context and so on.
>>>So it won't work when one have networking from one container and proc
>>>from another.
>>the question should be: which part of proc should be part
>>of the pid space and which not, definitely the network
>>stuff would _not_ be part of the pid space ...
>Ok, just one simple question:
>how do you propose to handle network sysctls and network
>statistics/information in proc?
well, procfs is called procfs because it is/was?
supposed to contain process information, otherwise
it would have been called netfs or statfs or even
junkfs :)
>_how_ can you imagine this namespaces should work?
>I see no elegant solution for this, do you?
>If there is any, I will be happy with namespaces again.
junkfs parts need to be properly virtualized, the
procfs parts do not.
>>>So I really see no much reasons to have separate namespaces,
>>>but it is ok for me if someone really wants it this way.
>>the reasons are, as I explained several times, that folks
>>use 'virtualization' or 'isolation' for many different
>>things, just because SWsoft only uses it for VPS doesn't
>>meant that it cannot be used for other things
>Out of curiosity, do you have any _working_ examples of other usages?
>I see only theoretical examples from you, but would like to hear from
>anyone who _uses_/_knows_ how to use it.

seems we are going in circles here, I already gave
a detailed list of _actual_ uses which are different
from the VPS approach
>>just consider isolating/virtualizing the network stack,
>>but leaving the processes in the same pid space, how to
>>do that in a sane way with a single reference?
>I see... Any idea why this can be required?
>(without proc? :))
>BTW, if you have virtualized networking, but not isolated fs namespace
>in this case, how are you going to handle unix sockets? Or maybe it's
>another separate namespace?
two httpd servers could easily bind to a subset of
the host IP addresses while sharing the pid space
(and other spaces). guess what, that actually works
and is in use ...
>>>1. ask Linus about the preffered approach. I prepared an email for him
>>>with a description of approaches.
>>why do you propose, if you already did? :)
>because, the question was quite simple, isn't it?
no comment
>>>2. start from networking/netfilters/IPC which are essentially the same
>>>in both projects and help each other.
>>no problem with that, once Eric got there ...
>Kirill
best,
Herbert
PS: as one can see, I gave up on fixing your unreadable
    quoting, so don't expect readability ...