

---

Subject: Re: [PATCH 16/28] [FLAT 1/6] Changes in data structures for flat model  
Posted by [ebiederm](#) on Tue, 19 Jun 2007 19:11:13 GMT

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

---

Pavel Emelianov <xemul@openvz.org> writes:

> sukadev@us.ibm.com wrote:

>> Pavel Emelianov [xemul@openvz.org] wrote:

>> | This patch opens the flat model patches.

>> |

>> | The flat model idea is that struct pid has two numbers. The first one

>> | (pid->nr) is a global one and is unique in the system. The second one

>> | (pid->vnr) is a virtual pid. It is used on the kernel user boundary only.

>>

>> This approach duplicates 5 integers and 2 pointers per process for every

>> process in the system. While this may not be expensive for processes that

>> actually use multiple namespaces, doesn't it waste memory if majority of

>> processes exist only in one namespace ?

>

> task\_struct alignment allows for it. so does the alignment of signal structure.

> and please note that this comes with appropriate ifdefs around. the only problem

> is with struct pid, but we're virtualizing it after all!

>

> moreover - two integers and a pointer to the namespace is the minimal set of

> fields for pid that is visible from two namespaces...

>> | --- ./include/linux/sched.h.flatdata 2007-06-15 15:14:33.000000000 +0400

>> | +++ ./include/linux/sched.h 2007-06-15 15:19:14.000000000 +0400

>> | @@ -482,7 +482,10 @@ struct signal\_struct {

>> | pid\_t session \_\_deprecated;

>> | pid\_t \_\_session;

>> | };

>> | -

>> | +#ifdef CONFIG\_PID\_NS\_FLAT

>> | + pid\_t vpggrp;

>> | + pid\_t vsession;

>> | +#endif

>> | /\* boolean value for session group leader \*/

>> | int leader;

>> |

>> | @@ -944,6 +947,11 @@ struct task\_struct {

>> | unsigned did\_exec:1;

>> | pid\_t pid;

>> | pid\_t tgid;

>> | +#ifdef CONFIG\_PID\_NS\_FLAT

>> | + /\* hash the virtual ids as well \*/

>> | + pid\_t vpid;

>> | + pid\_t vtgid;

>> | +#endif

Adding vpggrp, vsession, vpid, and vtgid is wrong.

A case can probably be made for caching the common case (users view), but we already have fields for that.

For a global view we must use struct pid \*, otherwise we are just asking for trouble.

Eric

---

Containers mailing list

Containers@lists.linux-foundation.org

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

---