Subject: [RFC] Containers infrastructure problems
Posted by xemul on Mon, 05 Mar 2007 15:52:34 GMT
View Forum Message <> Reply to Message

Hi.

I'm trying to implement RSS accounting via containers and
I have some difficulties and proposals.

1. Fork

container_fork() is placed before new task obtains its
new mm_struct, files_struct, signal_struct etc. Isn't it
better to move container fork at the place where newly
created task it fully initialized to give controller
possibility to work with new mm, signals etc?

2. Early container usage

Consider the following code:

```
  struct my_container *cnt;

  cnt = my_cnt_from_cont(task_container(current, &my_subsys));
```

the problem is that when it is used before I register my
rss subsystem in initcall task_container returns me
dummytop container which is not my_container actually :(

I've workarounded this issue with

```
static int rss_create(struct container_subsys *ss,
                      struct container *cont)
{
        struct rss_container *rss;

        rss = kzalloc(sizeof(struct rss_container), GFP_KERNEL);
        if (rss == NULL)
                return -ENOMEM;


        ...
        cont->subsys[rss_subsys.subsys_id] = &rss->css;
        return 0;
}

static struct rss_container init_rss_container;

static __init int rss_create_early(struct container_subsys *ss,
```

```
                struct container *cont)
{
        struct rss_container *rss;

        rss = &init_rss_container;
        ...
        cont->subsys[rss_subsys.subsys_id] = &rss->css;
        ss->create = rss_create;
        return 0;
}

static struct container_subsys rss_subsys = {
        .name = "rss",
        .create = rss_create_early,
};

void __init container_rss_init_early(void)
{
        container_register_subsys(&rss_subsys);
}
```

and call container_rss_init_early() from container_init_early()
but this is probably not what we want.

I believe that we need some early container initialization
implemented in a generic way. What do you think?

Thanks,
Pavel.

Subject: Re: [RFC] Containers infrastructure problems
Posted by Paul Menage on Wed, 07 Mar 2007 01:46:41 GMT
View Forum Message <> Reply to Message

On 3/5/07, Pavel Emelianov <xemul@sw.ru> wrote:
> Hi.
>
> I'm trying to implement RSS accounting via containers and
> I have some difficulties and proposals.
>
> 1. Fork
>
> container_fork() is placed before new task obtains its
> new mm_struct, files_struct, signal_struct etc. Isn't it
> better to move container fork at the place where newly
> created task it fully initialized to give controller
> possibility to work with new mm, signals etc?

Yes, that seems reasonable.

>
> 2. Early container usage
>
> Consider the following code:
>
> struct my_container *cnt;
>
> cnt = my_cnt_from_cont(task_container(current, &my_subsys));
>
> the problem is that when it is used before I register my
> rss subsystem in initcall task_container returns me
> dummytop container which is not my_container actually :(

By definition all tasks are in dummytop (the top container in the
dummy hierarchy) since you can't create sub-containers in the dummy
hierarchy.

You're right that before you're registered, the current container
won't have a pointer for your subsystem. But calling
container_register_subsys() from container_rss_init_early(), and
having that called early on from init/main.c should be OK.

Paul

Subject: Re: [RFC] Containers infrastructure problems
Posted by xemul on Wed, 07 Mar 2007 07:05:13 GMT
View Forum Message <> Reply to Message

Paul Menage wrote:
> On 3/5/07, Pavel Emelianov <xemul@sw.ru> wrote:
>> Hi.
>>
>> I'm trying to implement RSS accounting via containers and
>> I have some difficulties and proposals.
>>
>> 1. Fork
>>
>> container_fork() is placed before new task obtains its
>> new mm_struct, files_struct, signal_struct etc. Isn't it
>> better to move container fork at the place where newly
>> created task it fully initialized to give controller
>> possibility to work with new mm, signals etc?
>
> Yes, that seems reasonable.

>
>>
>> 2. Early container usage
>>
>> Consider the following code:
>>
>>   struct my_container *cnt;
>>
>>   cnt = my_cnt_from_cont(task_container(current, &my_subsys));
>>
>> the problem is that when it is used before I register my
>> rss subsystem in initcall task_container returns me
>> dummytop container which is not my_container actually :(
>
> By definition all tasks are in dummytop (the top container in the
> dummy hierarchy) since you can't create sub-containers in the dummy
> hierarchy.
>
> You're right that before you're registered, the current container
> won't have a pointer for your subsystem. But calling
> container_register_subsys() from container_rss_init_early(), and
> having that called early on from init/main.c should be OK.

It is OK, but ->create callback should be aware of
the fact it is called on system boot time and thus
it mustn't call kmalloc() :)

>
> Paul
>