
Subject: [PATCH -mm] remove node_ prefix_from ns subsystem
Posted by [Cedric Le Goater](#) on Wed, 21 May 2008 15:47:21 GMT
[View Forum Message](#) <> [Reply to Message](#)

From: Cedric Le Goater <clg@fr.ibm.com>

This is a slight change in the namespace cgroup subsystem api.

Signed-off-by: Cedric Le Goater <clg@fr.ibm.com>

kernel/cgroup.c | 2 +-
1 file changed, 1 insertion(+), 1 deletion(-)

Index: 2.6.26-rc2-mm1/kernel/cgroup.c

```
=====
--- 2.6.26-rc2-mm1.orig/kernel/cgroup.c
+++ 2.6.26-rc2-mm1/kernel/cgroup.c
@@ -2903,7 +2903,7 @@ int cgroup_clone(struct task_struct *tsk
     cg = tsk->cgroups;
     parent = task_cgroup(tsk, subsys->subsys_id);

-    snprintf(nodename, MAX_CGROUP_TYPE_NAMELEN, "node_%d", tsk->pid);
+    snprintf(nodename, MAX_CGROUP_TYPE_NAMELEN, "%d", tsk->pid);

    /* Pin the hierarchy */
    atomic_inc(&parent->root->sb->s_active);
```

Containers mailing list
Containers@lists.linux-foundation.org
<https://lists.linux-foundation.org/mailman/listinfo/containers>

Subject: Re: [PATCH -mm] remove node_ prefix_from ns subsystem
Posted by [serue](#) on Wed, 21 May 2008 20:03:47 GMT
[View Forum Message](#) <> [Reply to Message](#)

Quoting Cedric Le Goater (clg@fr.ibm.com):
> From: Cedric Le Goater <clg@fr.ibm.com>
>
> This is a slight change in the namespace cgroup subsystem api.
>
> Signed-off-by: Cedric Le Goater <clg@fr.ibm.com>

I agree that is preferable.

Acked-by: Serge Hallyn <serue@us.ibm.com>

> ---

```

> kernel/cgroup.c | 2 +-
> 1 file changed, 1 insertion(+), 1 deletion(-)
>
> Index: 2.6.26-rc2-mm1/kernel/cgroup.c
> =====
> --- 2.6.26-rc2-mm1.orig/kernel/cgroup.c
> +++ 2.6.26-rc2-mm1/kernel/cgroup.c
> @@ -2903,7 +2903,7 @@ int cgroup_clone(struct task_struct *tsk
>  cg = tsk->cgroups;
>  parent = task_cgroup(tsk, subsys->subsys_id);
>
> - snprintf(nodename, MAX_CGROUP_TYPE_NAMELEN, "node_%d", tsk->pid);
> + snprintf(nodename, MAX_CGROUP_TYPE_NAMELEN, "%d", tsk->pid);
>
> /* Pin the hierarchy */
> atomic_inc(&parent->root->sb->s_active);

```

Containers mailing list
Containers@lists.linux-foundation.org
<https://lists.linux-foundation.org/mailman/listinfo/containers>

Subject: Re: [PATCH -mm] remove node_ prefix_ from ns subsystem
Posted by [akpm](#) on Thu, 22 May 2008 05:06:57 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Wed, 21 May 2008 17:47:21 +0200 Cedric Le Goater <clg@fr.ibm.com> wrote:

```

> From: Cedric Le Goater <clg@fr.ibm.com>
>
> This is a slight change in the namespace cgroup subsystem api.
>

```

What are the reasons for making this change?

```

> ---
> kernel/cgroup.c | 2 +-
> 1 file changed, 1 insertion(+), 1 deletion(-)
>
> Index: 2.6.26-rc2-mm1/kernel/cgroup.c
> =====
> --- 2.6.26-rc2-mm1.orig/kernel/cgroup.c
> +++ 2.6.26-rc2-mm1/kernel/cgroup.c
> @@ -2903,7 +2903,7 @@ int cgroup_clone(struct task_struct *tsk
>  cg = tsk->cgroups;
>  parent = task_cgroup(tsk, subsys->subsys_id);
>
> - snprintf(nodename, MAX_CGROUP_TYPE_NAMELEN, "node_%d", tsk->pid);
>

```

```
> + snprintf(nodename, MAX_CGROUP_TYPE_NAMELEN, "%d", tsk->pid);
>
> /* Pin the hierarchy */
> atomic_inc(&parent->root->sb->s_active);
```

This is user-visible, isn't it? Been there since October. How can we safely change it now?

Containers mailing list
Containers@lists.linux-foundation.org
<https://lists.linux-foundation.org/mailman/listinfo/containers>

Subject: Re: [PATCH -mm] remove node_prefix_from ns subsystem
Posted by [Paul Menage](#) on Thu, 22 May 2008 07:11:56 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Wed, May 21, 2008 at 10:06 PM, Andrew Morton
<akpm@linux-foundation.org> wrote:

```
>> --- 2.6.26-rc2-mm1.orig/kernel/cgroup.c
>> +++ 2.6.26-rc2-mm1/kernel/cgroup.c
>> @@ -2903,7 +2903,7 @@ int cgroup_clone(struct task_struct *tsk
>>     cg = tsk->cgroups;
>>     parent = task_cgroup(tsk, subsys->subsys_id);
>>
>> -   snprintf(nodename, MAX_CGROUP_TYPE_NAMELEN, "node_%d", tsk->pid);
>> +   snprintf(nodename, MAX_CGROUP_TYPE_NAMELEN, "%d", tsk->pid);
>>
>>     /* Pin the hierarchy */
>>     atomic_inc(&parent->root->sb->s_active);
>>
>
> This is user-visible, isn't it? Been there since October. How can we
> safely change it now?
```

It is user-visible, but I'd be extremely surprised if anyone was depending on it or even using it at this point.

Paul

Containers mailing list
Containers@lists.linux-foundation.org
<https://lists.linux-foundation.org/mailman/listinfo/containers>

Subject: Re: [PATCH -mm] remove node_prefix_from ns subsystem
Posted by [akpm](#) on Thu, 22 May 2008 07:32:13 GMT

On Thu, 22 May 2008 00:11:56 -0700 "Paul Menage" <menage@google.com> wrote:

```
> On Wed, May 21, 2008 at 10:06 PM, Andrew Morton
> <akpm@linux-foundation.org> wrote:
> > --- 2.6.26-rc2-mm1.orig/kernel/cgroup.c
> > +++ 2.6.26-rc2-mm1/kernel/cgroup.c
> > @@ -2903,7 +2903,7 @@ int cgroup_clone(struct task_struct *tsk
> >     cg = tsk->cgroups;
> >     parent = task_cgroup(tsk, subsys->subsys_id);
> >
> > -    snprintf(nodename, MAX_CGROUP_TYPE_NAMELEN, "node_%d", tsk->pid);
> > +    snprintf(nodename, MAX_CGROUP_TYPE_NAMELEN, "%d", tsk->pid);
> >
> >     /* Pin the hierarchy */
> >     atomic_inc(&parent->root->sb->s_active);
> >
> > This is user-visible, isn't it? Been there since October. How can we
> > safely change it now?
>
> It is user-visible, but I'd be extremely surprised if anyone was
> depending on it or even using it at this point.
```

After six months?

Oh well, I guess the number is small, and they'll know how to fix it.

What is the reason for making this change?

Containers mailing list
Containers@lists.linux-foundation.org
<https://lists.linux-foundation.org/mailman/listinfo/containers>

Subject: Re: [PATCH -mm] remove node_ prefix_ from ns subsystem
Posted by [Cedric Le Goater](#) on Thu, 22 May 2008 07:36:23 GMT
[View Forum Message](#) <> [Reply to Message](#)

Andrew Morton wrote:

```
> On Thu, 22 May 2008 00:11:56 -0700 "Paul Menage" <menage@google.com> wrote:
>
>> On Wed, May 21, 2008 at 10:06 PM, Andrew Morton
>> <akpm@linux-foundation.org> wrote:
>>>> --- 2.6.26-rc2-mm1.orig/kernel/cgroup.c
>>>> +++ 2.6.26-rc2-mm1/kernel/cgroup.c
>>>> @@ -2903,7 +2903,7 @@ int cgroup_clone(struct task_struct *tsk
>>>>     cg = tsk->cgroups;
```

```

>>>> parent = task_cgroup(tsk, subsys->subsys_id);
>>>>
>>>> -   snprintf(nodename, MAX_CGROUP_TYPE_NAMELEN, "node_%d", tsk->pid);
>>>> +   snprintf(nodename, MAX_CGROUP_TYPE_NAMELEN, "%d", tsk->pid);
>>>>
>>>> /* Pin the hierarchy */
>>>> atomic_inc(&parent->root->sb->s_active);
>>> This is user-visible, isn't it? Been there since October. How can we
>>> safely change it now?
>> It is user-visible, but I'd be extremely surprised if anyone was
>> depending on it or even using it at this point.
>
> After six months?

```

yes that's what we believe also.

```

> Oh well, I guess the number is small, and they'll know how to fix it.
>
> What is the reason for making this change?

```

the 'node_' prefix really is superfluous and misleading. what's a node ?

C.

Containers mailing list
Containers@lists.linux-foundation.org
<https://lists.linux-foundation.org/mailman/listinfo/containers>

Subject: Re: [PATCH -mm] remove node_ prefix_ from ns subsystem
Posted by [akpm](#) on Thu, 22 May 2008 08:10:53 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Thu, 22 May 2008 09:36:23 +0200 Cedric Le Goater <clg@fr.ibm.com> wrote:

```

> > What is the reason for making this change?
>
> the 'node_' prefix really is superfluous and misleading. what's a node ?

```

For gawd's sake guys. I've asked three times and now I'm getting rhetorical questions for answers. It's like drawing teeth.

Will someone please write a halfway decent description of this change?
One which (at least) covers all the questions I've been asking?

What the change is, why it is being made, what the user-visible presentation is, what the impact upon users is, why we think it won't be a problem, etc? The stuff which should have been right there from

day one, before the code change was even made?

Thanks.

Containers mailing list
Containers@lists.linux-foundation.org
<https://lists.linux-foundation.org/mailman/listinfo/containers>

Subject: Re: [PATCH -mm] remove node_prefix_from ns subsystem
Posted by [Paul Menage](#) on Thu, 22 May 2008 08:23:35 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi Andrew,

On Thu, May 22, 2008 at 1:10 AM, Andrew Morton
<akpm@linux-foundation.org> wrote:

>
> What the change is, why it is being made, what the user-visible
> presentation is, what the impact upon users is, why we think it won't
> be a problem, etc? The stuff which should have been right there from
> day one, before the code change was even made?

The change is that previously when `cgroup_clone()` was called (currently only from the `unshare` path in `ns_proxy` cgroup, you'd get a new group named "`node_$pid`" whereas now you'll get a group named after just your pid.)

The only users who would notice it are those who are using the `ns_proxy` cgroup subsystem to auto-create cgroups when namespaces are unshared - something of an experimental feature, which I think really needs more complete container/namespace support in order to be useful. I suspect the only users are Cedric and Serge, or maybe a few others on containers@lists.linux-foundation.org. And in fact it would only be noticed by the users who make the assumption about how the name is generated, rather than getting it from the `/proc/<pid>/cgroups` file for the process in question.

Whether the change is actually needed or not I'm fairly agnostic on, but I guess it is more elegant to just use the pid as the new group name rather than adding a fairly arbitrary "`node_`" prefix on the front.

Paul

Containers mailing list
Containers@lists.linux-foundation.org
<https://lists.linux-foundation.org/mailman/listinfo/containers>

Subject: Re: [PATCH -mm] remove node_prefix_from ns subsystem

Posted by [akpm](#) on Thu, 22 May 2008 08:36:12 GMT

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

On Thu, 22 May 2008 01:23:35 -0700 "Paul Menage" <menage@google.com> wrote:

> Hi Andrew,

>

> On Thu, May 22, 2008 at 1:10 AM, Andrew Morton

> <akpm@linux-foundation.org> wrote:

> >

> > What the change is, why it is being made, what the user-visible

> > presentation is, what the impact upon users is, why we think it won't

> > be a problem, etc? The stuff which should have been right there from

> > day one, before the code change was even made?

>

> The change is that previously when cgroup_clone() was called

> (currently only from the unshare path in ns_proxy cgroup, you'd get a

> new group named "node_\$pid" whereas now you'll get a group named after

> just your pid.)

>

> The only users who would notice it are those who are using the

> ns_proxy cgroup subsystem to auto-create cgroups when namespaces are

> unshared - something of an experimental feature, which I think really

> needs more complete container/namespace support in order to be useful.

> I suspect the only users are Cedric and Serge, or maybe a few others

> on containers@lists.linux-foundation.org. And in fact it would only be

> noticed by the users who make the assumption about how the name is

> generated, rather than getting it from the /proc/<pid>/cgroups file

> for the process in question.

>

> Whether the change is actually needed or not I'm fairly agnostic on,

> but I guess it is more elegant to just use the pid as the new group

> name rather than adding a fairly arbitrary "node_" prefix on the

> front.

>

Well I suppose that as a non-back-compatible change we should feed it into 2.6.25.x as well. It's a bit unusual, but so doing will reduce the number of hey-where-did-my-file-go discoveries.

Containers mailing list

Containers@lists.linux-foundation.org

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

Subject: Re: [PATCH -mm] remove node_prefix_from ns subsystem

Posted by [Cedric Le Goater](#) on Thu, 22 May 2008 09:21:21 GMT

Andrew Morton wrote:

> On Thu, 22 May 2008 01:23:35 -0700 "Paul Menage" <menage@google.com> wrote:

>

>> Hi Andrew,

>>

>> On Thu, May 22, 2008 at 1:10 AM, Andrew Morton

>> <akpm@linux-foundation.org> wrote:

>>> What the change is, why it is being made, what the user-visible

>>> presentation is, what the impact upon users is, why we think it won't

>>> be a problem, etc? The stuff which should have been right there from

>>> day one, before the code change was even made?

>> The change is that previously when `cgroup_clone()` was called

>> (currently only from the `unshare` path in `ns_proxy` cgroup, you'd get a

>> new group named "`node_$pid`" whereas now you'll get a group named after

>> just your pid.)

>>

>> The only users who would notice it are those who are using the

>> `ns_proxy` cgroup subsystem to auto-create cgroups when namespaces are

>> unshared - something of an experimental feature, which I think really

>> needs more complete container/namespace support in order to be useful.

>> I suspect the only users are Cedric and Serge, or maybe a few others

>> on `containers@lists.linux-foundation.org`. And in fact it would only be

>> noticed by the users who make the assumption about how the name is

>> generated, rather than getting it from the `/proc/<pid>/cgroups` file

>> for the process in question.

>>

>> Whether the change is actually needed or not I'm fairly agnostic on,

>> but I guess it is more elegant to just use the pid as the new group

>> name rather than adding a fairly arbitrary "`node_`" prefix on the

>> front.

Thanks Paul.

> Well I suppose that as a non-back-compatible change we should feed it

> into 2.6.25.x as well. It's a bit unusual, but so doing will reduce the

> number of hey-where-did-my-file-go discoveries.

Thanks andrew. I have been keeping that patch for so long that I didn't realize that the `ns` cgroup subsystem had been already merged. sorry for the noise.

C.

Containers mailing list

`Containers@lists.linux-foundation.org`

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