Subject: [PATCH 2/3] make clone_children a flag
Posted by Glauber Costa on Sun, 11 Dec 2011 14:45:37 GMT
View Forum Message <> Reply to Message

There is no reason to have a flags field, and then a separate
bool field just to indicate if the clone_children flag is set.
Make it a flag

Signed-off-by: Glauber Costa <glommer@parallels.com>
---
 kernel/cgroup.c | 6 +++---
 1 files changed, 3 insertions(+), 3 deletions(-)

diff --git a/kernel/cgroup.c b/kernel/cgroup.c
index e4b9d3c..fa405ee 100644
--- a/kernel/cgroup.c
+++ b/kernel/cgroup.c
@@ -231,6 +231,7 @@ inline int cgroup_is_removed(const struct cgroup *cgrp)
 /* bits in struct cgroupfs_root flags field */
 enum {
  ROOT_NOPREFIX, /* mounted subsystems have no named prefix */
+ ROOT_CLONE_CHILDREN, /* mounted subsystems starts with clone_children */
 };

 static int cgroup_is_releasable(const struct cgroup *cgrp)
@@ -1062,7 +1063,6 @@ struct cgroup_sb_opts {
  unsigned long subsys_bits;
  unsigned long flags;
  char *release_agent;
- bool clone_children;
  char *name;
  /* User explicitly requested empty subsystem */
  bool none;
@@ -1113,7 +1113,7 @@ static int parse_cgroupfs_options(char *data, struct cgroup_sb_opts
*opts)
    continue;
   }
   if (!strcmp(token, "clone_children")) {
-   opts->clone_children = true;
+   set_bit(ROOT_CLONE_CHILDREN, &opts->flags);
    continue;
   }
   if (!strncmp(token, "release_agent=", 14)) {
@@ -1400,7 +1400,7 @@ static struct cgroupfs_root *cgroup_root_from_opts(struct
cgroup_sb_opts *opts)
  strcpy(root->release_agent_path, opts->release_agent);
 if (opts->name)
  strcpy(root->name, opts->name);

```
- if (opts->clone_children)
+ if (test_bit(ROOT_CLONE_CHILDREN, &opts->flags))
   set_bit(CGRP_CLONE_CHILDREN, &root->top_cgroup.flags);
  return root;
 }
--
1.7.6.4
```

Subject: Re: [PATCH 2/3] make clone_children a flag
Posted by KOSAKI Motohiro on Sun, 11 Dec 2011 18:58:49 GMT
View Forum Message <> Reply to Message

(12/11/11 9:45 AM), Glauber Costa wrote:
> There is no reason to have a flags field, and then a separate
> bool field just to indicate if the clone_children flag is set.
> Make it a flag
>
> Signed-off-by: Glauber Costa<glommer@parallels.com>
> ---
> kernel/cgroup.c | 6 +++---
> 1 files changed, 3 insertions(+), 3 deletions(-)

Reviewed-by: KOSAKI Motohiro <kosaki.motohiro@jp.fujitsu.com>

Subject: Re: [PATCH 2/3] make clone_children a flag
Posted by Tejun Heo on Tue, 13 Dec 2011 15:39:21 GMT
View Forum Message <> Reply to Message

On Sun, Dec 11, 2011 at 03:45:37PM +0100, Glauber Costa wrote:
> There is no reason to have a flags field, and then a separate
> bool field just to indicate if the clone_children flag is set.
> Make it a flag
>
> Signed-off-by: Glauber Costa <glommer@parallels.com>

Doesn't this change how remount conditions are checked?

Thanks.

--
tejun

Subject: Re: [PATCH 2/3] make clone_children a flag

Posted by Li Zefan on Wed, 14 Dec 2011 02:28:17 GMT
View Forum Message <> Reply to Message

Tejun Heo wrote:
> On Sun, Dec 11, 2011 at 03:45:37PM +0100, Glauber Costa wrote:
>> There is no reason to have a flags field, and then a separate
>> bool field just to indicate if the clone_children flag is set.
>> Make it a flag
>>
>> Signed-off-by: Glauber Costa <glommer@parallels.com>
>
> Doesn't this change how remount conditions are checked?
>

Right. Currently we can do this:

```
 # mount -t cgroup xxx /mnt
 # mount -o remount,clone_children /mnt
```

with this patch, the above remount will fail.

But..the current bevaiour of remount is a bit confusing in that remount
with/without "clone_children" has no effect on anything:

```
 # mount -t cgroup -o clone_children xxx /mnt
 # cat /mnt/cgroup.clone_children
 1
 # mount -o remount xxx /mnt
 # mount | grep cgroup
 xxx on /mnt type cgroup (rw,clone_children)
 # cat /mnt/cgroup.clone_children
 1
```

Subject: Re: [PATCH 2/3] make clone_children a flag
Posted by Glauber Costa on Wed, 14 Dec 2011 07:09:14 GMT
View Forum Message <> Reply to Message

On 12/14/2011 06:29 AM, Li Zefan wrote:
> Tejun Heo wrote:
>> On Sun, Dec 11, 2011 at 03:45:37PM +0100, Glauber Costa wrote:
>>> There is no reason to have a flags field, and then a separate
>>> bool field just to indicate if the clone_children flag is set.
>>> Make it a flag
>>>
>>> Signed-off-by: Glauber Costa<glommer@parallels.com>
>>
>> Doesn't this change how remount conditions are checked?

>>

Well, I was thinking it wouldn't, because I patched all callers. But I
forget life is not always that simple: After you mentioned, I checked
and we do test for changes in the flag field explicitly on remount. So I
missed that, indeed.

> Right. Currently we can do this:
>
> # mount -t cgroup xxx /mnt
> # mount -o remount,clone_children /mnt
>
> with this patch, the above remount will fail.
>
> But..the current bevaiour of remount is a bit confusing in that remount
> with/without "clone_children" has no effect on anything:
>
> # mount -t cgroup -o clone_children xxx /mnt
> # cat /mnt/cgroup.clone_children
> 1
> # mount -o remount xxx /mnt
> # mount | grep cgroup
> xxx on /mnt type cgroup (rw,clone_children)
> # cat /mnt/cgroup.clone_children
> 1

That's indeed confusing, and it comes from the fact that we always
inherit clone_children from the parent - which is sane, IMHO. So this
flag only has any value in establishing the initial behaviour of the top
root cgroup. I wonder then if it wouldn't better to just be explicit and
fail in this case ?

Subject: Re: [PATCH 2/3] make clone_children a flag
Posted by Tejun Heo on Wed, 14 Dec 2011 18:18:52 GMT
View Forum Message <> Reply to Message

Hello,

On Wed, Dec 14, 2011 at 11:09:14AM +0400, Glauber Costa wrote:
> That's indeed confusing, and it comes from the fact that we always
> inherit clone_children from the parent - which is sane, IMHO. So
> this flag only has any value in establishing the initial behaviour
> of the top root cgroup. I wonder then if it wouldn't better to just
> be explicit and fail in this case ?

I don't think all current behaviors are sane and if not let's change
them, but those things have to be explicit with proper description and

rationale.

Thanks.

--
tejun

Subject: Re: [PATCH 2/3] make clone_children a flag
Posted by Glauber Costa on Thu, 15 Dec 2011 07:03:38 GMT
View Forum Message <> Reply to Message

On 12/14/2011 10:18 PM, Tejun Heo wrote:
> Hello,
>
> On Wed, Dec 14, 2011 at 11:09:14AM +0400, Glauber Costa wrote:
>> That's indeed confusing, and it comes from the fact that we always
>> inherit clone_children from the parent - which is sane, IMHO. So
>> this flag only has any value in establishing the initial behaviour
>> of the top root cgroup. I wonder then if it wouldn't better to just
>> be explicit and fail in this case ?
>
> I don't think all current behaviors are sane and if not let's change
> them, but those things have to be explicit with proper description and
> rationale.
>

140 % agree to that. As I said, I wrongly believed it to be functionally
equivalent when I sent it, but missed the flags remount check.

If you believe the behavior we now get is saner, I can rewrite the
Changelog and resend it.
