Subject: [PATCH] SUNRPC: check current nsproxy before set of node name on client creation
Posted by Stanislav Kinsbursky on Mon, 13 Aug 2012 11:21:56 GMT
View Forum Message <> Reply to Message

When child reaper exits, it can destroy mount namespace it belong to, and if
there are NFS mounts inside, then it will try to umount them. But in this
point current->nsproxy is set to NULL and all namespaces will be destroyed one
by one. I.e. we can't dereference current->nsproxy to obtain uts namespace.

Signed-off-by: Stanislav Kinsbursky <skinsbursky@parallels.com>
---
 net/sunrpc/clnt.c | 14 ++++++++++--
 1 files changed, 12 insertions(+), 2 deletions(-)

diff --git a/net/sunrpc/clnt.c b/net/sunrpc/clnt.c
index 9a9676e..28ac940 100644
--- a/net/sunrpc/clnt.c
+++ b/net/sunrpc/clnt.c
@@ -279,6 +279,16 @@ void rpc_clients_notifier_unregister(void)

 static void rpc_clnt_set_nodename(struct rpc_clnt *clnt, const char *nodename)
 {
+ const char *nodename;
+
+ /*
+ * We have to protect against dying chilp reaper, which has released
+ * it's nsproxy already and is trying to destroy mount namespace.
+ */
+ if (current->nsproxy == NULL)
+ return;
+
+ nodename = utsname()->nodename;
  clnt->cl_nodelen = strlen(nodename);
  if (clnt->cl_nodelen > UNX_MAXNODENAME)
   clnt->cl_nodelen = UNX_MAXNODENAME;
@@ -365,7 +375,7 @@ static struct rpc_clnt * rpc_new_client(const struct rpc_create_args *args,
stru
  }

  /* save the nodename */
- rpc_clnt_set_nodename(clnt, utsname()->nodename);
+ rpc_clnt_set_nodename(clnt);
  rpc_register_client(clnt);
  return clnt;

@@ -524,7 +534,7 @@ rpc_clone_client(struct rpc_clnt *clnt)
  err = rpc_setup_pipedir(new, clnt->cl_program->pipe_dir_name);

```
  if (err != 0)
   goto out_no_path;
- rpc_clnt_set_nodename(new, utsname()->nodename);
+ rpc_clnt_set_nodename(new);
  if (new->cl_auth)
   atomic_inc(&new->cl_auth->au_count);
  atomic_inc(&clnt->cl_count);
```

Subject: Re: [PATCH] SUNRPC: check current nsproxy before set of node name on client creation
Posted by Jeff Layton on Mon, 13 Aug 2012 11:35:05 GMT
View Forum Message <> Reply to Message

On Mon, 13 Aug 2012 15:21:56 +0400
Stanislav Kinsbursky <skinsbursky@parallels.com> wrote:

> When child reaper exits, it can destroy mount namespace it belong to, and if
> there are NFS mounts inside, then it will try to umount them. But in this
> point current->nsproxy is set to NULL and all namespaces will be destroyed one
> by one. I.e. we can't dereference current->nsproxy to obtain uts namespace.
>
> Signed-off-by: Stanislav Kinsbursky <skinsbursky@parallels.com>
> ---
> net/sunrpc/clnt.c | 14 ++++++++++--
> 1 files changed, 12 insertions(+), 2 deletions(-)
>
> diff --git a/net/sunrpc/clnt.c b/net/sunrpc/clnt.c
> index 9a9676e..28ac940 100644
> --- a/net/sunrpc/clnt.c
> +++ b/net/sunrpc/clnt.c
> @@ -279,6 +279,16 @@ void rpc_clients_notifier_unregister(void)
>
> static void rpc_clnt_set_nodename(struct rpc_clnt *clnt, const char *nodename)

Don't you need to change the prototype of the function here too?

> {
> + const char *nodename;
> +
> + /*
> + * We have to protect against dying chilp reaper, which has released

 nit: "dying child reaper" ?

> + * it's nsproxy already and is trying to destroy mount namespace.

 nit: "its"

> + */
> + if (current->nsproxy == NULL)
> + return;
> +
> + nodename = utsname()->nodename;
>   clnt->cl_nodelen = strlen(nodename);
>   if (clnt->cl_nodelen > UNX_MAXNODENAME)
>   clnt->cl_nodelen = UNX_MAXNODENAME;
> @@ -365,7 +375,7 @@ static struct rpc_clnt * rpc_new_client(const struct rpc_create_args
*args, stru
>   }
>
>   /* save the nodename */
> - rpc_clnt_set_nodename(clnt, utsname()->nodename);
> + rpc_clnt_set_nodename(clnt);
>   rpc_register_client(clnt);
>   return clnt;
>
> @@ -524,7 +534,7 @@ rpc_clone_client(struct rpc_clnt *clnt)
>   err = rpc_setup_pipedir(new, clnt->cl_program->pipe_dir_name);
>   if (err != 0)
>   goto out_no_path;
> - rpc_clnt_set_nodename(new, utsname()->nodename);
> + rpc_clnt_set_nodename(new);
>   if (new->cl_auth)
>   atomic_inc(&new->cl_auth->au_count);
>   atomic_inc(&clnt->cl_count);
>
> --
> To unsubscribe from this list: send the line "unsubscribe linux-nfs" in
> the body of a message to majordomo@vger.kernel.org
> More majordomo info at http://vger.kernel.org/majordomo-info.html

--
Jeff Layton <jlayton@poochiereds.net>

Subject: Re: [PATCH] SUNRPC: check current nsproxy before set of node name on client creation
Posted by Stanislav Kinsbursky on Mon, 13 Aug 2012 11:37:06 GMT
View Forum Message <> Reply to Message

Thanks, Jeff. Will fix.

> On Mon, 13 Aug 2012 15:21:56 +0400

> Stanislav Kinsbursky <skinsbursky@parallels.com> wrote:
>
>> When child reaper exits, it can destroy mount namespace it belong to, and if
>> there are NFS mounts inside, then it will try to umount them. But in this
>> point current->nsproxy is set to NULL and all namespaces will be destroyed one
>> by one. I.e. we can't dereference current->nsproxy to obtain uts namespace.
>>
>> Signed-off-by: Stanislav Kinsbursky <skinsbursky@parallels.com>
>> ---
>> net/sunrpc/clnt.c | 14 ++++++++++--
>> 1 files changed, 12 insertions(+), 2 deletions(-)
>>
>> diff --git a/net/sunrpc/clnt.c b/net/sunrpc/clnt.c
>> index 9a9676e..28ac940 100644
>> --- a/net/sunrpc/clnt.c
>> +++ b/net/sunrpc/clnt.c
>> @@ -279,6 +279,16 @@ void rpc_clients_notifier_unregister(void)
>>
>> static void rpc_clnt_set_nodename(struct rpc_clnt *clnt, const char *nodename)
>
> Don't you need to change the prototype of the function here too?
>
>> {
>> + const char *nodename;
>> +
>> + /*
>> + * We have to protect against dying chilp reaper, which has released
>
> nit: "dying child reaper" ?
>
>> + * it's nsproxy already and is trying to destroy mount namespace.
>
> nit: "its"
>
>> + */
>> + if (current->nsproxy == NULL)
>> + return;
>> +
>> + nodename = utsname()->nodename;
>>   clnt->cl_nodelen = strlen(nodename);
>>   if (clnt->cl_nodelen > UNX_MAXNODENAME)
>>   clnt->cl_nodelen = UNX_MAXNODENAME;
>> @@ -365,7 +375,7 @@ static struct rpc_clnt * rpc_new_client(const struct rpc_create_args
*args, stru
>>   }
>>
>>   /* save the nodename */
>> - rpc_clnt_set_nodename(clnt, utsname()->nodename);

```
>> + rpc_clnt_set_nodename(clnt);
>>    rpc_register_client(clnt);
>>    return clnt;
>>
>> @@ -524,7 +534,7 @@ rpc_clone_client(struct rpc_clnt *clnt)
>>    err = rpc_setup_pipedir(new, clnt->cl_program->pipe_dir_name);
>>    if (err != 0)
>>     goto out_no_path;
>> - rpc_clnt_set_nodename(new, utsname()->nodename);
>> + rpc_clnt_set_nodename(new);
>>    if (new->cl_auth)
>>     atomic_inc(&new->cl_auth->au_count);
>>    atomic_inc(&clnt->cl_count);
>>
>> --
>> To unsubscribe from this list: send the line "unsubscribe linux-nfs" in
>> the body of a message to majordomo@vger.kernel.org
>> More majordomo info at  http://vger.kernel.org/majordomo-info.html
>
>


--
Best regards,
Stanislav Kinsbursky
```