Subject: Re: [PATCH] Send quota messages via netlink
Posted by serge on Thu, 30 Aug 2007 19:18:09 GMT
View Forum Message <> Reply to Message

Quoting Eric W. Biederman (ebiederm@xmission.com):
> Jan Kara <jack@suse.cz> writes:
> >   There can be arbitrary number of listeners (potentially from different
> > namespaces if I understand it correctly) listening to broadcasts. So I
> > think we should pass some universal identifier rather than try to find out
> > who is listening etc. I think such identifiers would be useful for other
> > things too, won't they?
>
> So internal to the kernel we have such a universal identifier.
> struct user.
>
> There are to practical questions.
> 1) How do we present that information to user space?
> 2) How does user space want to process this information?
>
> If we only want user space to be able to look up a user and send
> him a message. It probably makes sense to do the struct user to
> uid conversion in the proper context in the kernel because we have
> that information.
>
> If this is a general feature that happens to allows us to look up
> the user given the filesystems view of what is going on would be
> easier in the kernel, and not require translation. But it means
> that we can't support 9p and nfs for now. But since we don't support
> quotas on the client end anyway that doesn't sound like a big deal.
>
> The problem with the filesystem view is that there will be occasions
> where we simply can not map a user into it, because the filesystem
> won't have a concept of that particular user.
>
> So we could run into the situation where alice owns the file. Bob
> writes to the file and pushes it over quota. But the filesystem
> has no concept of who bob is. So we won't be able to report that
> it was bob that pushed things over the edge.
>
> > BTW: Do you have some idea, when would be the infrastructure clearer?
>
> So the plan is to get to the point where are uid comparisons in the
> kernel are (user namespace, uid) comparisons. Or possibly struct

Just fyi Eric,

Note that given the amount of churn going on due to pid and network
namespaces, I was seeing completion of user namespaces as something to

be done sometime next year. In the meantime I was only going to do
something with capabilities to restrict root in user namespaces (which I
think will take the form of per-process non-expandable cap_bsets, which
I plan to start basically right now).

But I'll gladly do the userns enhancements earlier if it's actually
wanted. They promise to be great fun :)

-serge

> user comparisons (depending on the context. And struct mount will
> contain the user namespace of whoever mounted the filesystem.
>
> Adding infrastructure to netlink to allow us to do conversions
> as the packets are enqueued for a specific user is something I
> would rather avoid, but that is a path we can go down if we have
> to.
>
> > Whether it makes sence to currently proceed with UIDs and later change it
> > to something generic or whether I should wait before you sort it out :).
>
> A good question. I think things are clear enough that it at least
> makes sense to sketch a solution to the problem even if we don't
> implement it at this point.
>
> I have been hoping Cedric or Serge would jump in because I think those
> are the guys who have been working on the implementation.
>
> Eric
> -
> To unsubscribe from this list: send the line "unsubscribe linux-kernel" in
> the body of a message to majordomo@vger.kernel.org
> More majordomo info at http://vger.kernel.org/majordomo-info.html
> Please read the FAQ at http://www.tux.org/lkml/
_____
Containers mailing list
Containers@lists.linux-foundation.org
https://lists.linux-foundation.org/mailman/listinfo/containers