Subject: Re: [RFC] kernel/pid.c pid allocation wierdness
Posted by William Lee Irwin III on Wed, 14 Mar 2007 15:03:06 GMT
View Forum Message <> Reply to Message

On Wed, Mar 14, 2007 at 08:12:35AM -0600, Eric W. Biederman wrote:
> If we do dig into this more we need to consider a radix_tree to hold
> the pid values. That could replace both the pid map and the hash
> table, gracefully handle but large and small pid counts, might
> be a smidgin simpler, possibly be more space efficient, and it would
> more easily handle multiple pid namespaces. The downside to using a
> radix tree is that is looks like it will have more cache misses for
> the normal pid map size, and it is yet another change that we would
> need to validate.

Radix trees' space behavior is extremely poor in sparsely-populated
index spaces. There is no way they would save space or even come close
to the current space footprint.

Lock contention here would be a severe functional regression (note
"functional," not "performance;" the lock contention surrounding these
affairs takes down systems vs. mere slowdown nonsense), so it would
necessarily depend on lockless radix tree code for correctness.

The comment block describing the hashtable locking is stale and should
have been updated in tandem with the RCU changes.

-- wli

_____
Containers mailing list
Containers@lists.osdl.org
https://lists.osdl.org/mailman/listinfo/containers