

## Re: [PATCH] Adding Solaris-style "owner of records" to rwlocks

---

*Source:* <http://unix.derkeiler.com/Mailing-Lists/FreeBSD/arch/2006-08/msg00011.html>

---

- *From:* "Attilio Rao" <[attilio@xxxxxxxxxxxx](mailto:attilio@xxxxxxxxxxxx)>
  - *Date:* Tue, 8 Aug 2006 18:08:59 +0200
- 

2006/8/8, Suleiman Souhlal <[ssouhlal@xxxxxxxxxxxx](mailto:ssouhlal@xxxxxxxxxxxx)>:

Attilio Rao wrote:

- > This is a first implementation of the owner of records concept in rwlocks.
- > It allows to avoid the priority inversion problem in the current
- > rwlocks implementation (for readers).
- >
- > The main idea (that John and I discussed) is to have as owner of
- > records the first rlock'er for a "class contention".
- > The implementation consists in adding two flags (RW\_LOCK\_OWNED and
- > RW\_LOCK\_EXEMPTED) which are used in order to not penalize the easy
- > case, and synchronizing the operation of acquiring and dropping the
- > owner of records with the turnstile spin-lock.
- > The main scheme might work in this way:
- >
- > thread1::rlock() -> sets the owner of records
- > thread2::rlock() -> checks for RW\_LOCK\_OWNED bit and, if it is set, go
- > in the easy case
- > thread3::rlock() -> checks for RW\_LOCK\_OWNED...
- > thread4::wlock() -> blocks and land its priority to thread1
- > thread1::runlock() -> disable the owner of records (disowning the
- > associated turnstile) and sets the RW\_LOCK\_EXEMPTED flag. In this way
- > other threads will treat as an easy case.
- > ...

Aren't you missing the hard part: transferring ownership from one reader to another? If you don't, you'll still have priority inversions as soon as the initial reader unlocks..

Exactly, but having a complete owner switching would be:

- 1) too hard to achieve in terms of resource taken
- 2) will imply too many races and we might get a too hard function

With this implementation, only the first rlock (for every class contention) will be penalized while the other are treated as the easy/hard case.

It doesn't completely solve the priority inversion problem, but it's

Re: [PATCH] Adding Solaris-style "owner of records" to rwlocks

the better compromise between performances/correctnes.

Attilio

—

Peace can only be achieved by understanding – A. Einstein

---

freebsd-arch@xxxxxxxxxxx mailing list

<http://lists.freebsd.org/mailman/listinfo/freebsd-arch>

To unsubscribe, send any mail to "freebsd-arch-unsubscribe@xxxxxxxxxxx"