

Re: LOR on sleepqueue chain locks, Was: LOR sleepq/scrlock

Source: <http://unix.derkeiler.com/Mailing-Lists/FreeBSD/stable/2008-05/msg00094.html>

- *From:* Aristedes Maniatis <ari@xxxxxxxxxxx>
 - *Date:* Sat, 10 May 2008 12:53:15 +1000
-

On 23/04/2008, at 3:34 AM, John Baldwin wrote:

The
real problem at the bottom of the screen though is a real
issue.
It's a LOR
of two different sleepqueue chain locks. The problem is that
when
setrunnable() encounters a swapped out thread it tries to
wakeup
proc0, but
if proc0 is asleep (which is typical) then its thread lock is a
sleep queue
chain lock, so waking up a swapped out thread from
wakeup() will
usually
trigger this LOR.

I think the best fix is to not have setrunnable() kick proc0
directly.
Perhaps setrunnable() should return an int and return true if
proc0
needs to
be awakened and false otherwise. Then the the sleepq code
(b/c only
sleeping
threads can be swapped out anyway) can return that value
from
sleepq_resume_thread() and can call kick_proc0() directly
once it
has dropped
all of its own locks.

--
John Baldwin

Re: LOR on sleepqueue chain locks, Was: LOR sleepq/scrlock

The way you describe it, it almost sounds like this LOR should be happening for everyone, all the time. To try and eliminate the factors which trigger it for us, we tried the following: removed PAE from kernel, disabled PF. Neither of these things made any difference and the error is fairly quickly reproducible (within a couple of hours running various things to load the machine). The one thing we did not test yet is removing ZFS from the picture. Note also that this box ran for years and years on FreeBSD 4.x without a hiccup (non PAE, ipfw instead of pf and no ZFS of course).

There are two things. 1) Most people who run witness (that I know of) don't run it on spinlocks because of the overhead, so LORs of spin locks are less well-reported than LORs of other locks (mutexes, rwlocks, etc.). 2) You have to have enough load on the box to swap out active processes to get into this situation. Between those I think that is why this is not more widely reported.

Hi John,

Thanks for your efforts so far to track this LOR down. I've been keeping an eye on cvs logs, but haven't seen anything which looks like a patch for this.

- * is this still outstanding?
- * or will it be addressed soon?
- * if not, should I create a PR so that it doesn't get forgotten?
- * in our case, although we can trigger it quickly with some load, the problem occurs (and causes a complete machine lock) even under < 10% load. Not sure if the combination of PAE/ZFS/SCHED ULE exacerbates that in any way compared to a 'standard' build.

Thank you
Ari Maniatis

----->
ish
<http://www.ish.com.au>
Level 1, 30 Wilson Street Newtown 2042 Australia
phone +61 2 9550 5001 fax +61 2 9550 4001
GPG fingerprint CBF8 84B4 738D 4E87 5E5C 5EFA EF6A 7D2E 3E49 102A

freebsd-stable@xxxxxxxxxxxxx mailing list
<http://lists.freebsd.org/mailman/listinfo/freebsd-stable>
To unsubscribe, send any mail to "freebsd-stable-unsubscribe@xxxxxxxxxxxxx"

Re: LOR on sleepqueue chain locks, Was: LOR sleepq/scrlock