

# soft lockup和hard lockup的检测原理

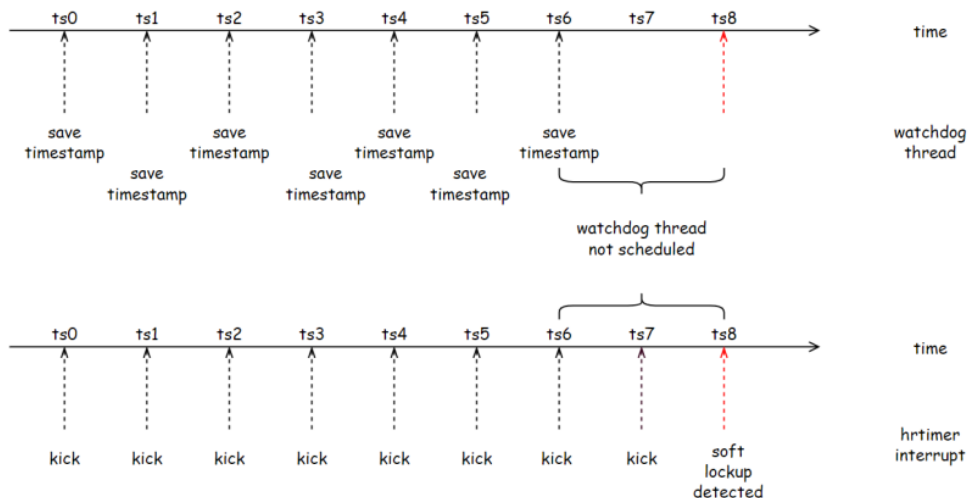
云计算基础架构 2022-10-14 23:29 发表于江苏

对于lockup问题处理，首先需要明确什么是lockup？它在什么情况下会被触发？本文就来聊聊这两种lockup是怎么检测的。

soft lockup 内核在每个CPU上都启动了一个watchdog线程，该线程被定期唤醒并记录per cpu的时间戳，同时启动的还有per-cpu的hrtimer，当hrtimer中断到来时会触发中断处理，在中断处理函数中会读取当前时间戳并与watchdog线程记录的时间做比较，如果两者相差超过一定范围（可以配置的watchdog\_thresh）就会触发soft lockup异常。

这种情况就所检测的情况为：watchdog线程已经有超过watchdog\_thresh的时间未被调度执行了，说明系统调度出现了问题，有其他进程长时间的霸占CPU不让出。

下面是我画的原理图：

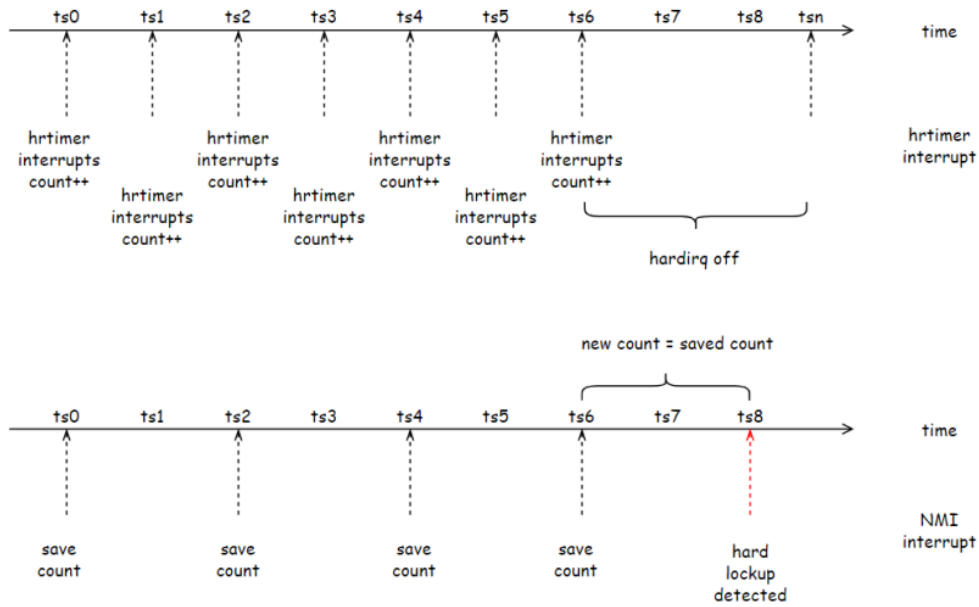


<https://mp.weixin.qq.com/s/193c3f3ng-coch-nz8f0k9ey9se>

内核为每个CPU创建了watchdog线程，此线程主要做的动作就是负责保存当前的时间戳，执行完成后就会进入sleep状态等待下一次被唤醒执行，实际上这个时间戳记录的就是watchdog运行的时间戳。同时内核创建了per-cpu的hrtimer定时器负责定期唤醒watchdog线程以及做soft lockup的检测。hrtimer定时器在处理函数中会获取watchdog上一次被唤醒执行的时间，计算出到当前的时间差，如果检测到watchdog已经超过一定时间未被调度到了，那么就认为触发了soft lockup。

hard lockup 对于hard lockup，它所检测的情况为：系统中的中断是否被长时间关闭，当某个CPU上的中断被关闭超过一定的时间，那么也可以认为系统运行已经存在了问题，正常情况下中断是不允许被长时间关闭的。

下面是我画的原理图：



- per-cpu hrtimer interrupt:
  1. update per-cpu interrupt count
- NMI interrupt:
  1. save hrtimer interrupt count
  2. detect hard lockup ( new count = saved count)

<https://blog.csdn.net/nikeyou>

对于hard lockup来说，主要为了检测中断被异常关闭的场景，那么这里就不得不提一下NMI中断，NMI中断是不可屏蔽中断，也就是说内核中关闭的中断只是普通的中断，对于NMI中断来说，它是无法被关闭的。有了这个大前提，我们才能有办法检测这种情况。

它的检测原理是，使用per-cpu hrtimer中断来记录中断次数，同时配置perf event来触发NMI中断，每次NMI中断到来时都会读取hrtimer的中断次数，并且做一下保存，当NMI中断处理中检测到hrtimer中断次数与上一次保存的值相等时，说明在两次NMI中断期间，hrtimer中断都没有到来，可以认为普通中断被关闭了。

喜欢此内容的人还喜欢

15种windows权限维持的方法

安全初心

