

Latency hints for CFS task

Vincent Guittot LPC'22



Content

- Intro
- CFS latency
- Preemption
- User interface

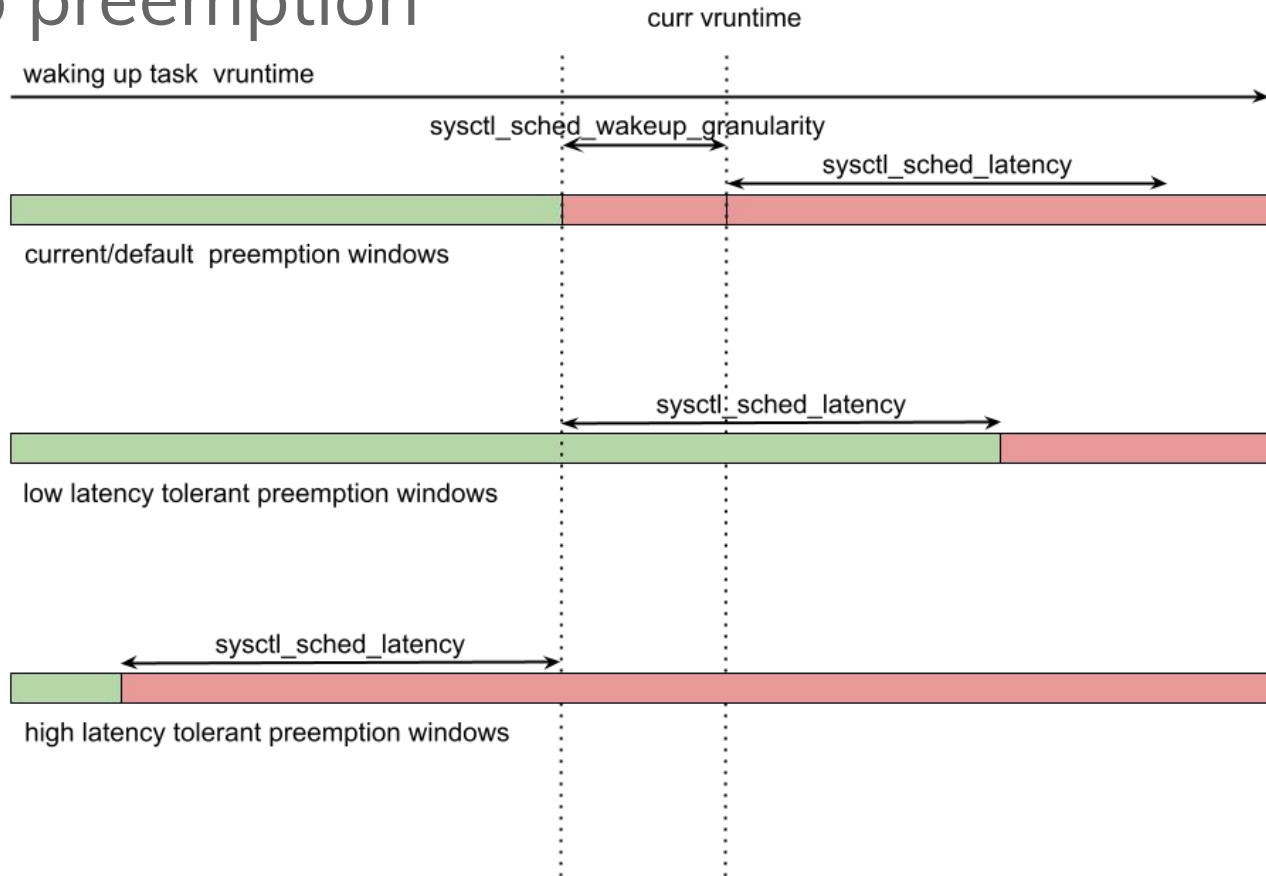
Intro

- CFS (Completely Fair Scheduler) is all about fair share of CPUs
 - Share Compute capacity → Share CPU time
 - Lowest vruntime runs 1st ... almost
- All CFS tasks don't have same goal
 - CPU intensive work
 - Background work
 - Feed a media pipeline
 - Part of user interaction path
- What if 2 CFS tasks have same or close vruntime ?
 - Would be good to schedule the latency sensitive task 1st

CFS latency

- Add a latency hint to CFS entities
 - Describe its sensitivity to the schedule latency
- Try to schedule the latency sensitive task 1st
- Do not preempt current task by non sensitive to latency tasks
- Do not break fairness

Wakeup preemption



Latency hint

- Comparison done at sched_entity level
 - Include sched group
- Add a latency_offset field
 - Set an offset in the sched_latency "virtual" period
- Use the offset at wakeup preempt
 - Preempt task with lower vruntime

Processing capacity figures

```
hackbench -p -1 (2560 / group) -g group
```

```
group
```

1	1.768 (+/- 13%)	0.805 (+/- 2%)	+54%
4	1.634 (+/- 13%)	0.765 (+/- 1%)	+53%
8	1.305 (+/- 4%)	0.745 (+/- 2%)	+43%
16	0.786 (+/- 4%)	0.705 (+/- 2%)	+10%

Scheduling delay figures

```
hackbench -l 10000 -g 16 &  
cyclictest --policy other -D 5 -q -n  
-H 20000 --histfile data.txt  
                latency 0      latency -20  
Min Latencies:    63           63  
Avg Latencies:   1459          214  
Max Latencies:   58274         82358  
50% latencies:    164          87  
75% latencies:    848          91  
85% latencies:   1246          94  
90% latencies:   2149          96  
95% latencies:   8463          99  
99% latencies: >20000        120
```

```
hackbench -l 10000 -g 16 &  
cyclictest --policy other -D 5 -q -n  
-H 20000 --histfile data.txt  
                latency 0      latency -20  
Min Latencies:    60           61  
Avg Latencies:   1077          86  
Max Latencies:   87311         444  
50% latencies:    92           85  
75% latencies:   554           90  
85% latencies:   1019          93  
90% latencies:   1346          96  
95% latencies:   5400         100  
99% latencies:  19044         110
```


User interface

- Define priority between tasks
- Signed value
 - Negative value : Latency sensitive i.e. not nice with others
 - Positive value: Preemption sensitive i.e. nice with others
- Several attempt to find a generic interface
 - Latency nice prio aligned with nice prio
 - Absolute latency offset value in ns ?
- cgroup interface
 - cpu.latency interface
 - set the latency offset in the range [-sched_latency period:sched_latency]

Thank you



