### SCHED\_DEADLINE desiderata

... and slightly crazy ideas

Juri Lelli Principal Software Engineer Daniel Bristot de Oliveira Principal Software Engineer



Actually... How many of you know what SCHED\_DEADLINE is? :-)



SCHED\_DEADLINE desiderata

### Why SCHED\_DEADLINE?

Introduction to the EDF Scheduling.



### SCHED\_DEADLINE

#### In a nutshell

```
chrt -d --sched-runtime 5000000 \
--sched-period 16666666 0 \setminus
video processing tool
```

Video processing task must process 60 frames per second, so 1 frame every ~16ms (and an estimated upper bound of 5ms to process it).

Earliest Deadline First (EDF) : dynamic deadlines Constant Bandwidth Server (CBS) : temporal isolation

Wikipedia article - https://en.wikipedia.org/wiki/SCHED\_DEADLINE

LWN articles - https://lwn.net/Articles/743740/

https://lwn.net/Articles/743946/









Consider a system with the following tasks:

Task	WCET	Period = Deadline	U
t1	1	4	0.250
t2	2	6	0.330
t3	3	8	0.375
<u>Σ</u> (U)			0.958 ( < 1)



### Fixed priority scheduling:





### EDF scheduling:





SCHED\_DEADLINE desiderata

8

## What we'll discuss today

Proxy Execution Capacity awareness Cgroups support

RT Throttling

Schedulability



### **Proxy Execution**

Better Priority Inheritance for SCHED\_DEADLINE tasks (among DEADLINE entities and entities from other scheduling classes)



### What's the problem?

- Current Priority Inheritance mechanism is not safe for !root
  - Deadline inheritance ( ... also slightly incorrect)
  - Priority boosted tasks are outside runtime enforcement

- We would need to inherit donors' bandwidth (runtime/period)
- And **keep runtime enforcement on** while doing that
- Basically let the mutex owner execute using the scheduling context of a (several) donor(s)









SCHED\_DEADLINE desiderata





### Cgroups support

Cgroups based bandwidth management and Hierarchical scheduling



#### **Cgroups support**

#### Cgroups based bandwidth management

- System administrator could reserve a fraction of total bandwidth to users
- Not hierarchical DEADLINE entities still scheduled by "root scheduler"
- DEADLINE\_GROUP\_SCHED (requires RT\_GROUP\_SCHED)
  - Is RT\_GROUP\_SCHED actually used/useful?
- DEADLINE and RT share bandwidth
  - cpu.rt\_{runtime,period}\_us
  - cpu.dl\_bw maximum available bandwidth
  - cpu.dl\_total\_bw currently allocated bandwidth



Hierarchical scheduling - Hierarchical Constant Bandwidth Server (H-CBS)

- Nest SCHED\_{FIFO,RR,CFS} entities within SCHED\_DEADLINE
- Good, for example, for pipelines of tasks (end-to-end deadline)





SCHED\_DEADLINE desiderata

### **RT** Throttling

Re-working RT Throttling to use DL servers



#### The RT Throttling mechanism avoids starvation

But leaves the CPU idle if no "NORMAL" tasks are ready run





#### Use a DL server instead

But things are not that simple 







## Schedulability improvements

Adaptive and semi-partitioned scheduler



#### The current scheduler is Global

- One runqueue for all CPUs
- But there are some corner cases that... hum...





#### Other possible Deadline Scheduler







#### 6/9 17 18 Partitioned tasks 6/9 17 18 3/8/9 17 18 Semi-partitioned tasks

### The partitioned and semi-partitioned scheduler





Final scheduling:





SCHED\_DEADLINE desiderata

### Capacity awareness

Scheduling on asymmetric CPU capacity architectures (e.g., big.LITTLE)



Issues with asymmetric CPU capacity architectures (e.g., big.LITTLE)

- Admission control may overprovision (all cores "look the same")
- Load balancing doesn't consider cores' speed/capacity

An RFC<sup>1</sup> proposed to fix them by using cpu\_capacity<sup>2</sup>:

- Admission control checks over sum of cpu capacities
- Load balancer checks cpu capacities against tasks bandwidth requirement

1 - https://lore.kernel.org/lkml/20190506044836.2914-1-luca.abeni@santannapisa.it/

2 - https://elixir.bootlin.com/linux/latest/source/Documentation/ABI/testing/sysfs-devices-system-cpu#L475



Power Management and Scheduling in the Linux Kernel (OSPM) Summit IV edition

Pisa (Italy) - May 11-13, 2020 Submit a topic! (until Feb 10th) http://retis.sssup.it/ospm-summit/





# Thank you

Red Hat is the world's leading provider of enterprise open source software solutions. Award-winning support, training, and consulting services make Red Hat a trusted adviser to the Fortune 500.

- in linkedin.com/company/red-hat
- youtube.com/user/RedHatVideos
  - facebook.com/redhatinc
  - twitter.com/RedHat

