

The Magic Behind PREEMPT_RT

Haris Okanovic
Senior Software Engineer
National Instruments

(Original slides by Julia Cartwright)

Topics

What is a Real-Time Operating System (RTOS)?

What is PREEMPT_RT?

How does PREEMPT_RT make Linux real-time?

RT Systems

A real-time system is a system in which timing requirements are functional.

A real-time operating system (RTOS) manages system resources in a way that satisfies timing requirements of running application(s).

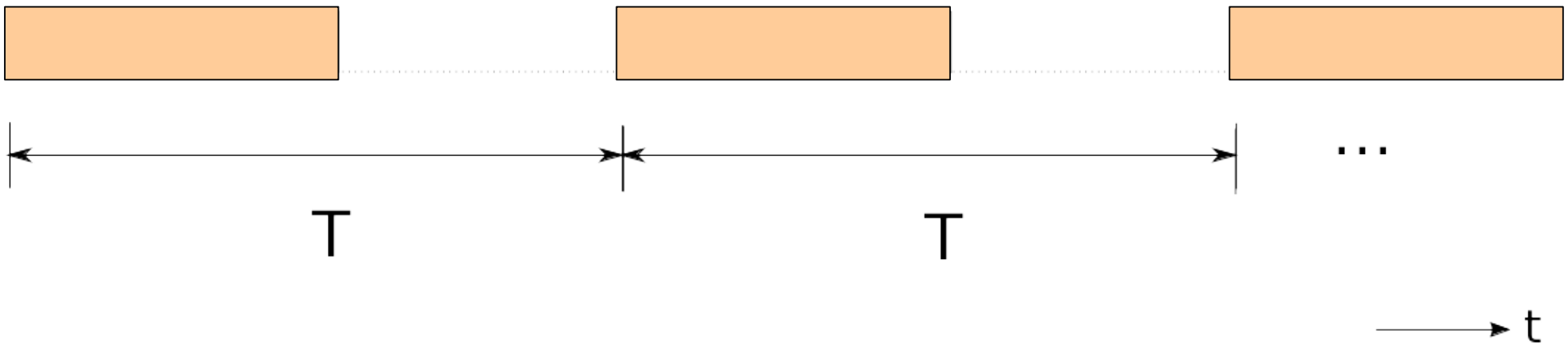
E.g. RT Application

```
for (;;) {  
    t0 = now();  
    feedback = sample();  
    output =  
        calculate(feedback, setpoint);  
    write_new_point(output);  
    sleep_until(t0 + T);  
}
```

E.g. RT Application

```
for (;;) {  
    t0 = now();  
    feedback = sample();  
    output =  
        calculate(feedback, setpoint);  
    write_new_point(output);  
    sleep_until(t0 + T);  
}
```

E.g. RT Application

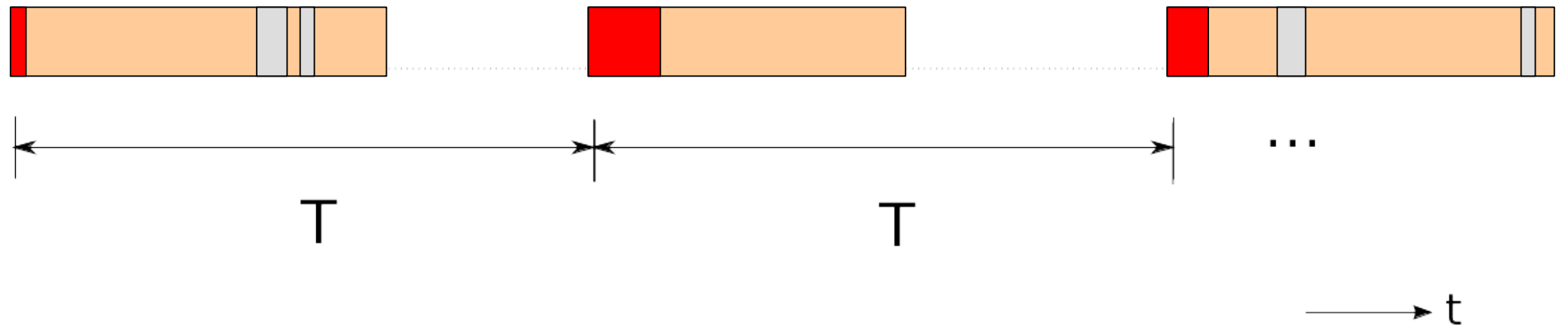


Latency

Wakeup latency (or scheduling latency) is the time between the firing of an external event and the execution of the relevant thread.

Execution latency is the difference between actual and ideal execution times.

E.g. RT Application



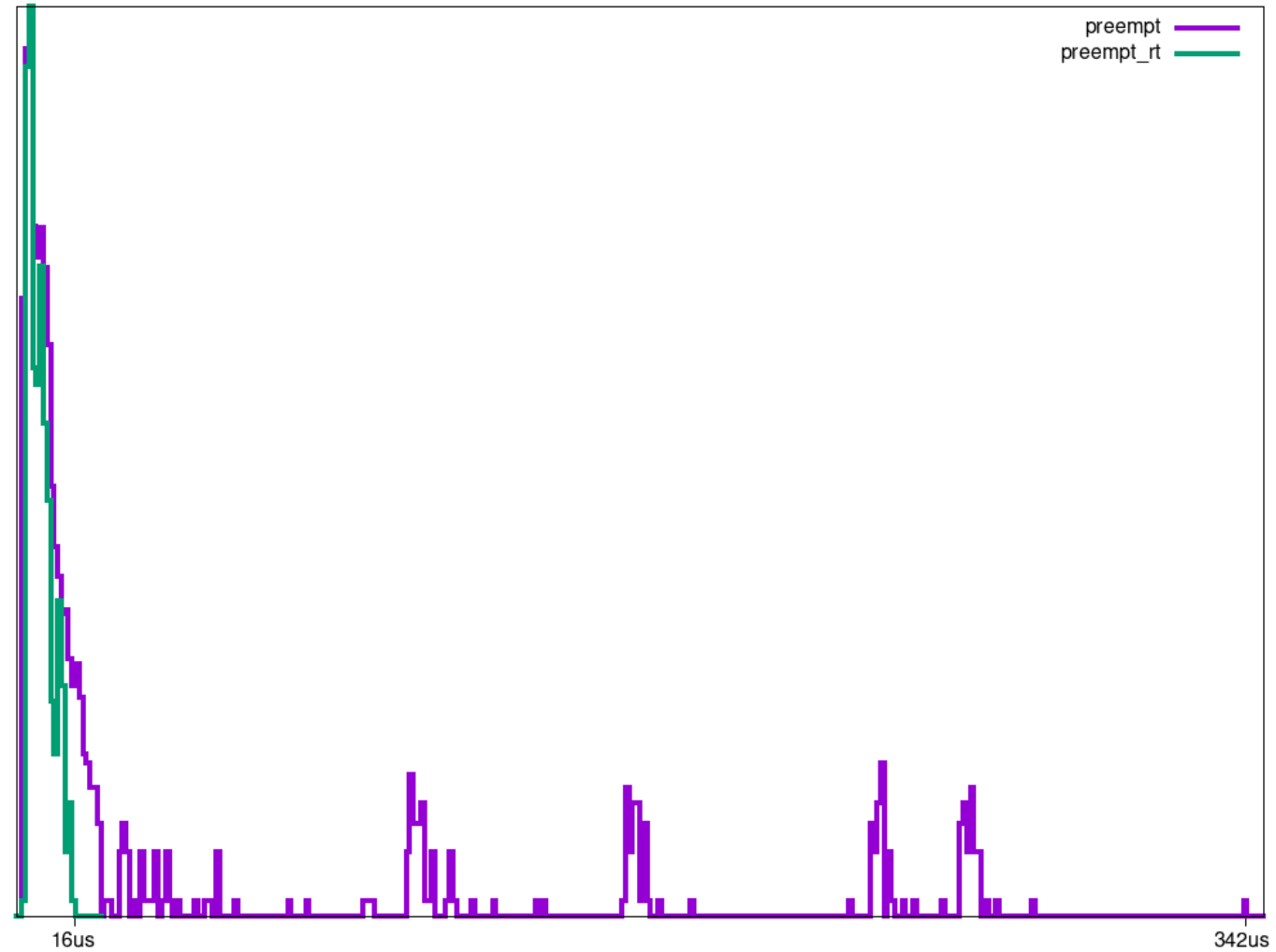
Wakeup Test

```
for (;;) {  
    t0 = now();  
    sleep(expected_T);  
    t1 = now();  
    actual_T = t1 - t0;  
    plot(actual_T - expected_T);  
}
```

Wakeup Test

```
for (;;) {  
    t0 = now();  
    sleep(expected_T);  
    t1 = now();  
    actual_T = t1 - t0;  
    plot(actual_T - expected_T);  
}
```

Wakeup Test



Wakeup Test

Cyclictest:

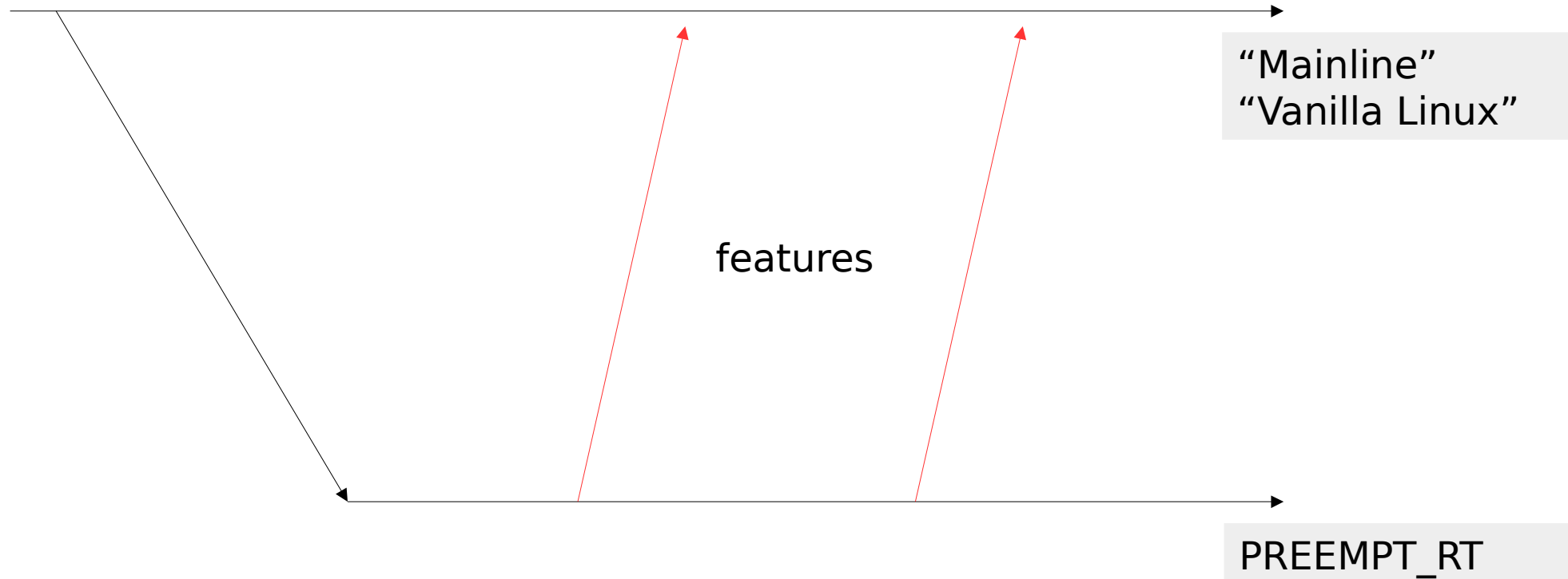
<https://rt.wiki.kernel.org/index.php/Cyclictest>

<https://git.kernel.org/pub/scm/utils/rt-tests/rt-tests.git>

E.g.

```
cyclictest -S -p 98 -m -h 350 --histfile=h.txt
```

PREEMPT_RT Patch



Real-Time Linux Collaborative Project

<https://wiki.linuxfoundation.org/realtime/>

Coalition of numerous organizations and users

Mainline + test PREEMPT_RT patchset

PREEMPT_RT Distributions

<https://www.suse.com/products/realtime>

<https://access.redhat.com/products/red-hat-enterprise-mrg-realtime>

<https://packages.debian.org/linux-image-rt>

<https://www.windriver.com/products/linux>

<https://www.mentor.com/embedded-software/linux>

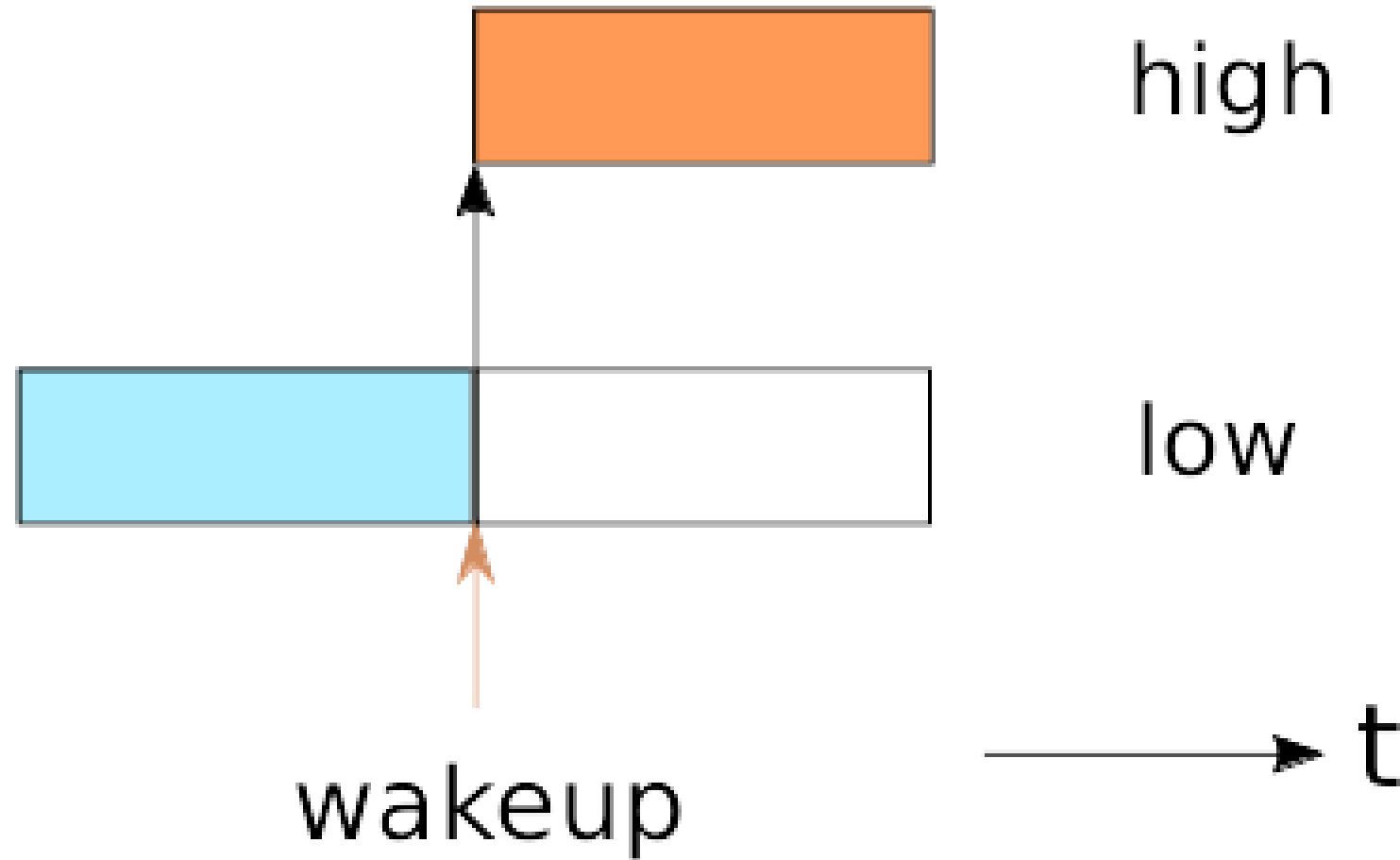
<https://www.ni.com/nilinuxrt>

Preemption

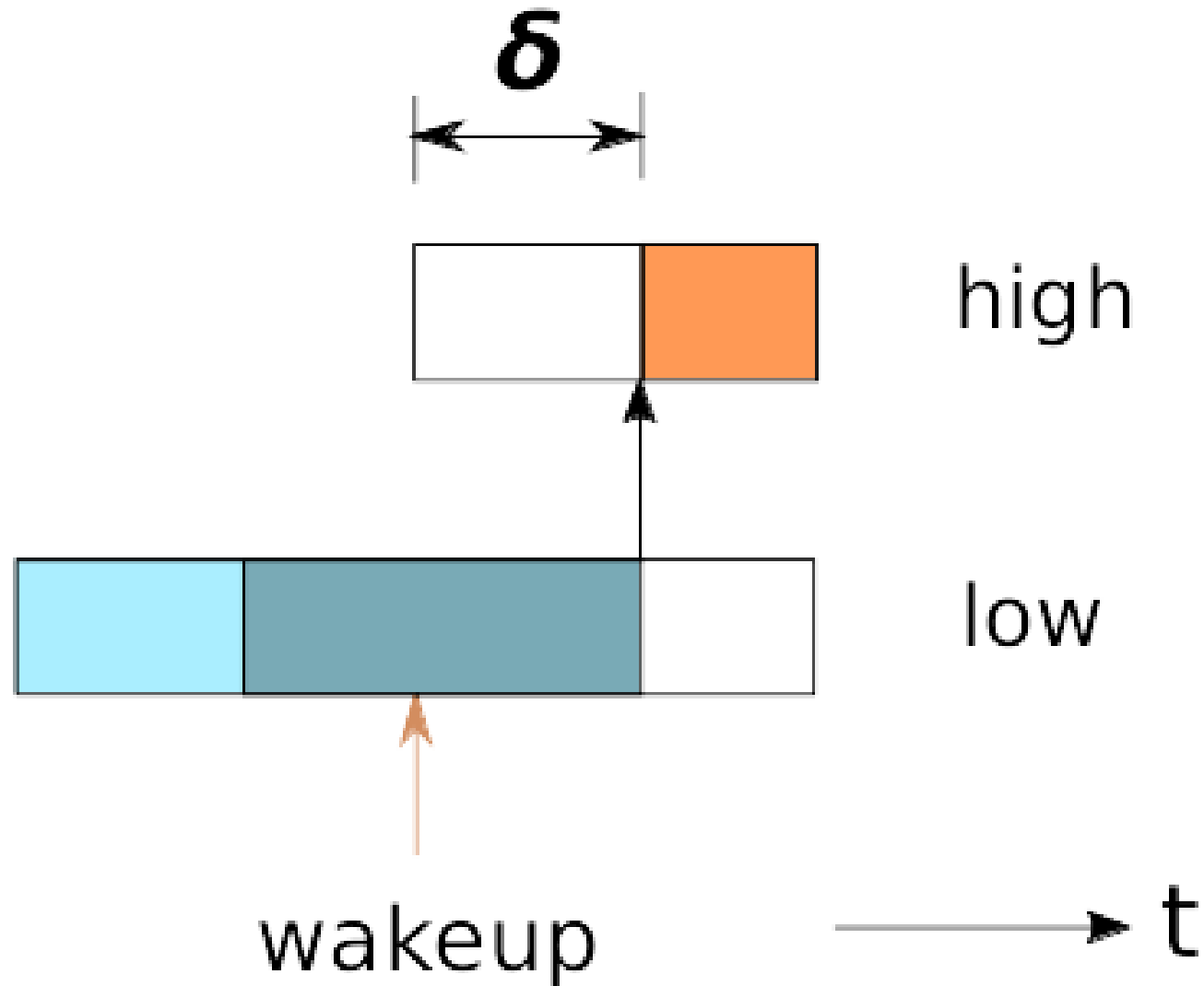
A thread is considered *preempted* when it's execution is suspended due to the wakeup and execution of a higher priority thread.

A thread is considered *preemptible* if the wakeup of a higher priority thread would trigger a preemption event.

Preemption



Preemption



Preemption Model

*A **preemption model** defines, at compile time, which kernel code paths may be preempted.*

Linux Preemption Model

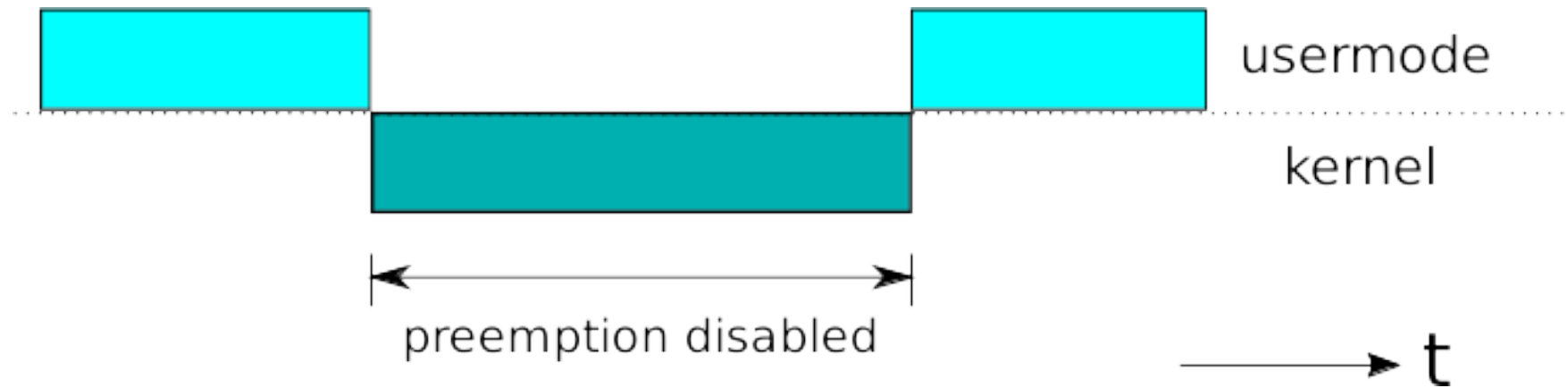
Vanilla

- PREEMPT_NONE
- PREEMPT_VOLUNTARY
- PREEMPT

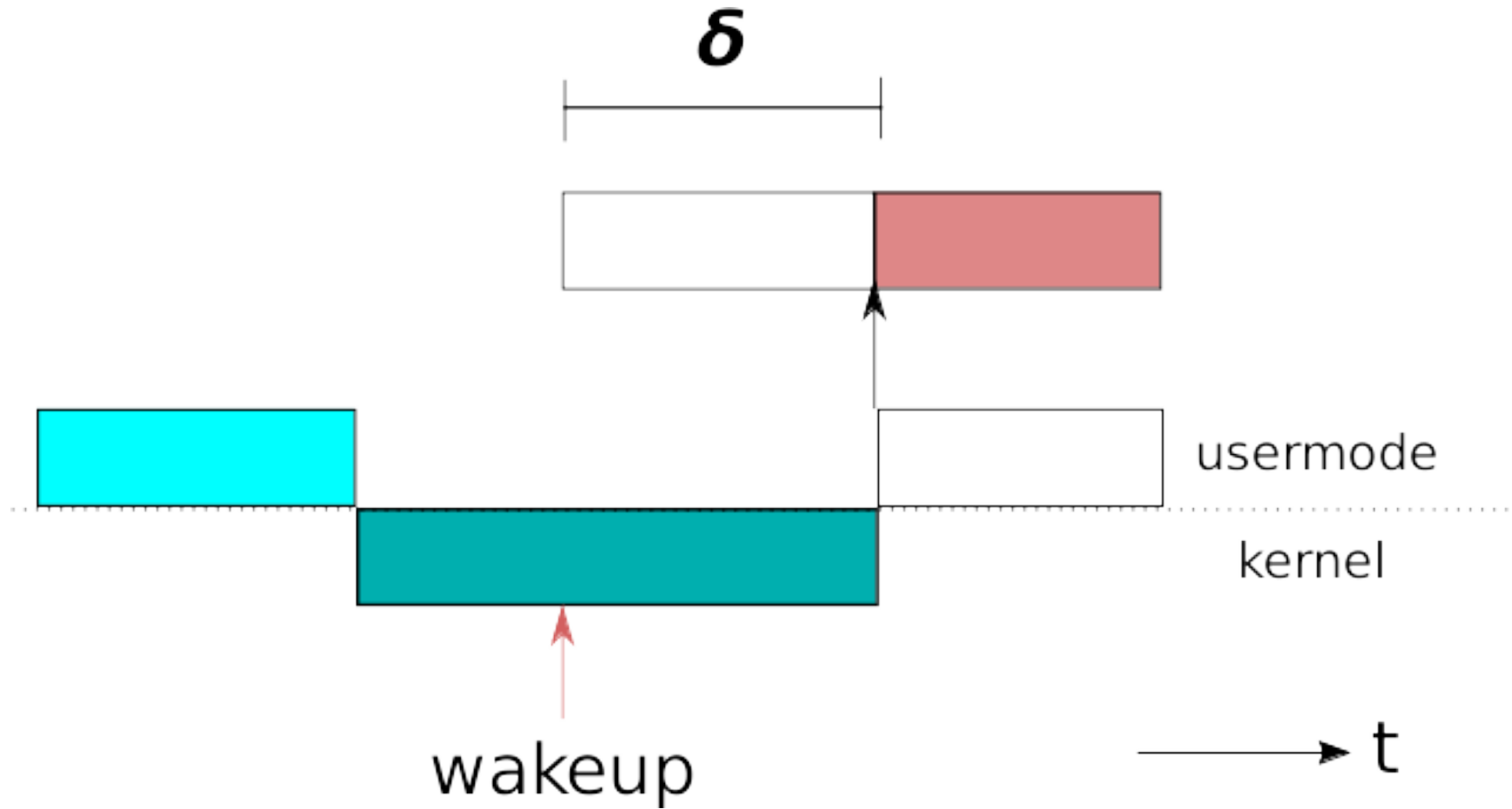
RT Patch

- PREEMPT_RT_FULL

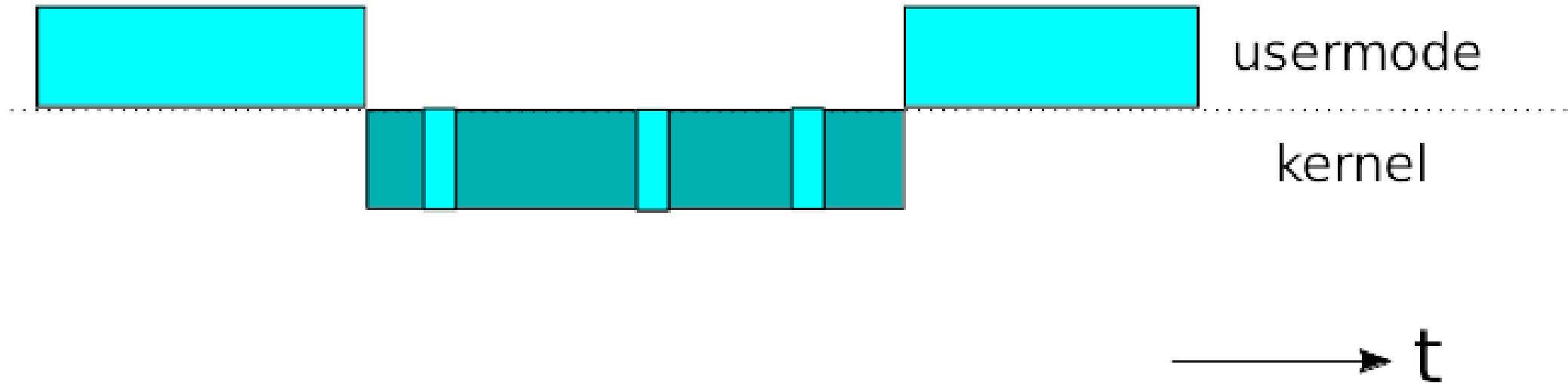
PREEMPT_NONE



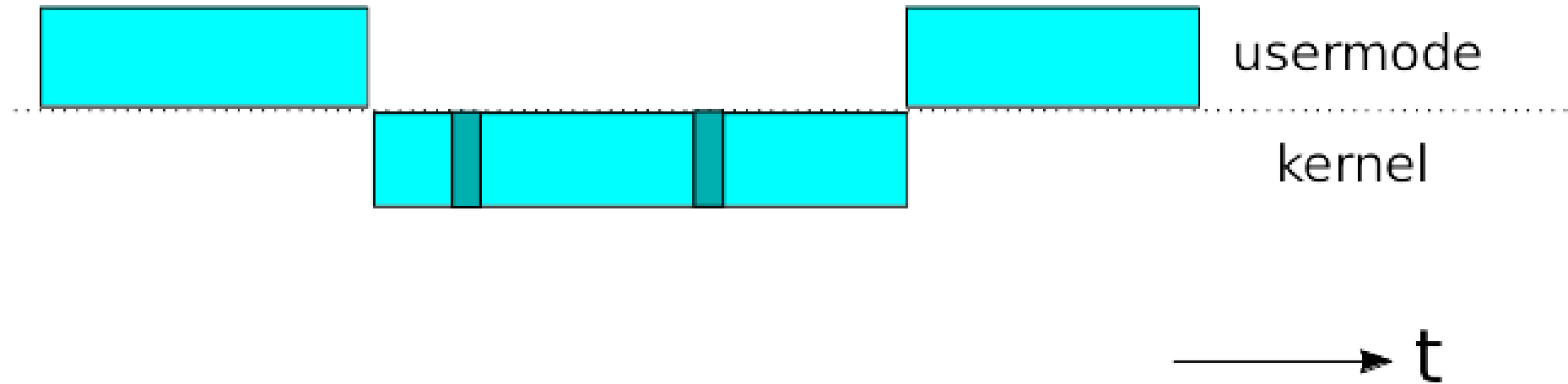
PREEMPT_NONE



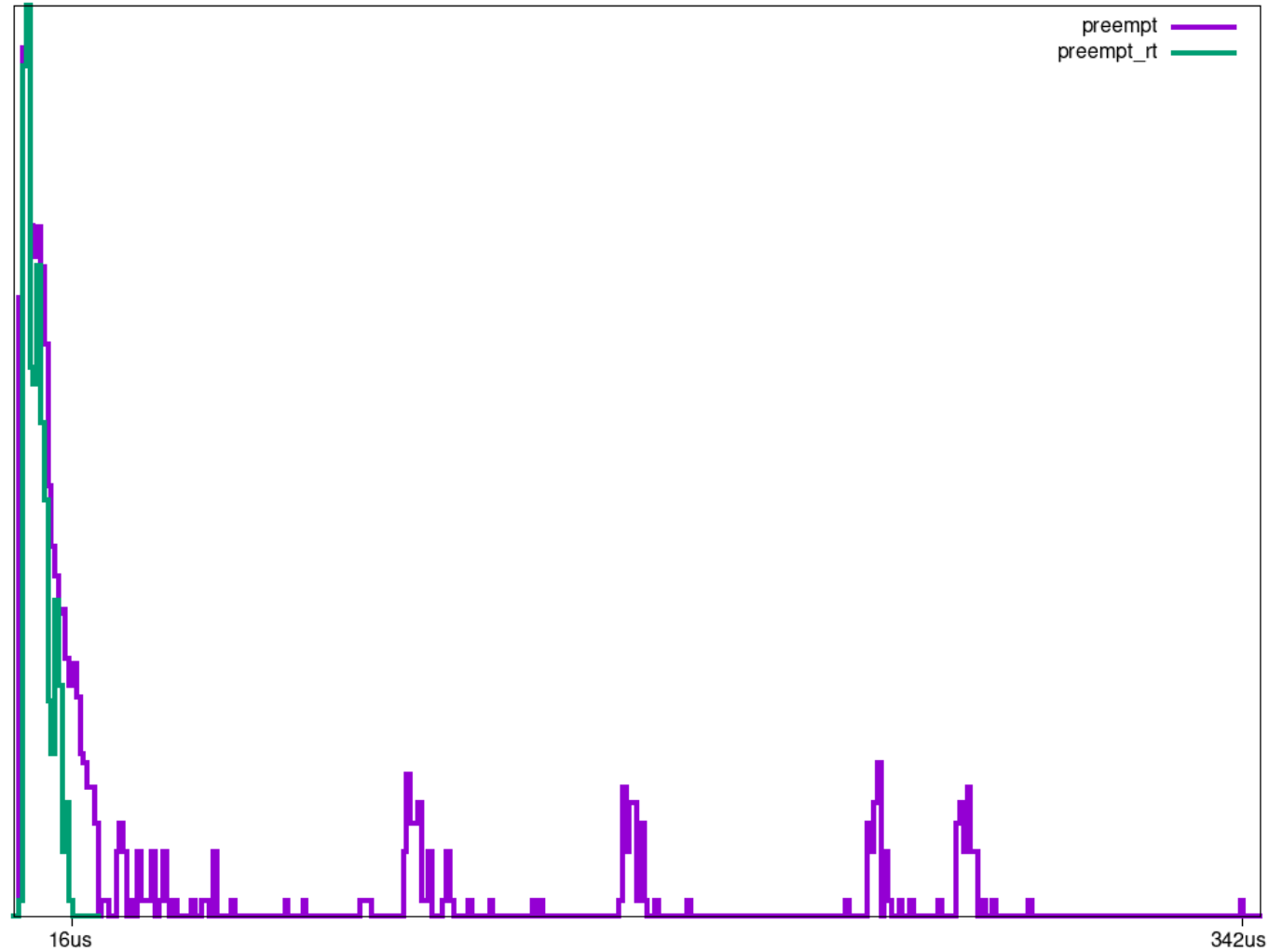
PREEMPT_VOLUNTARY



PREEMPT



PREEMPT_RT Does Better

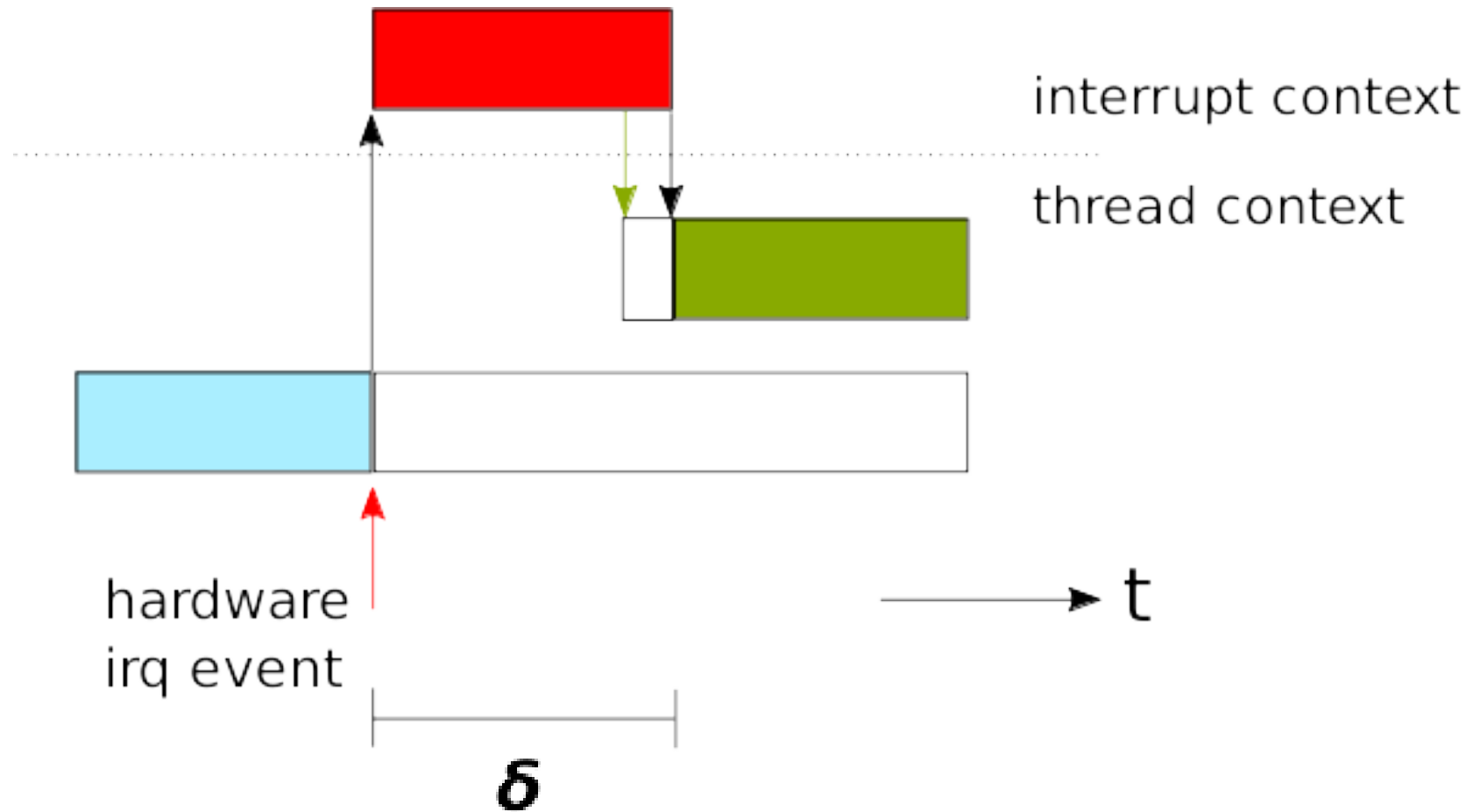


Execution Contexts

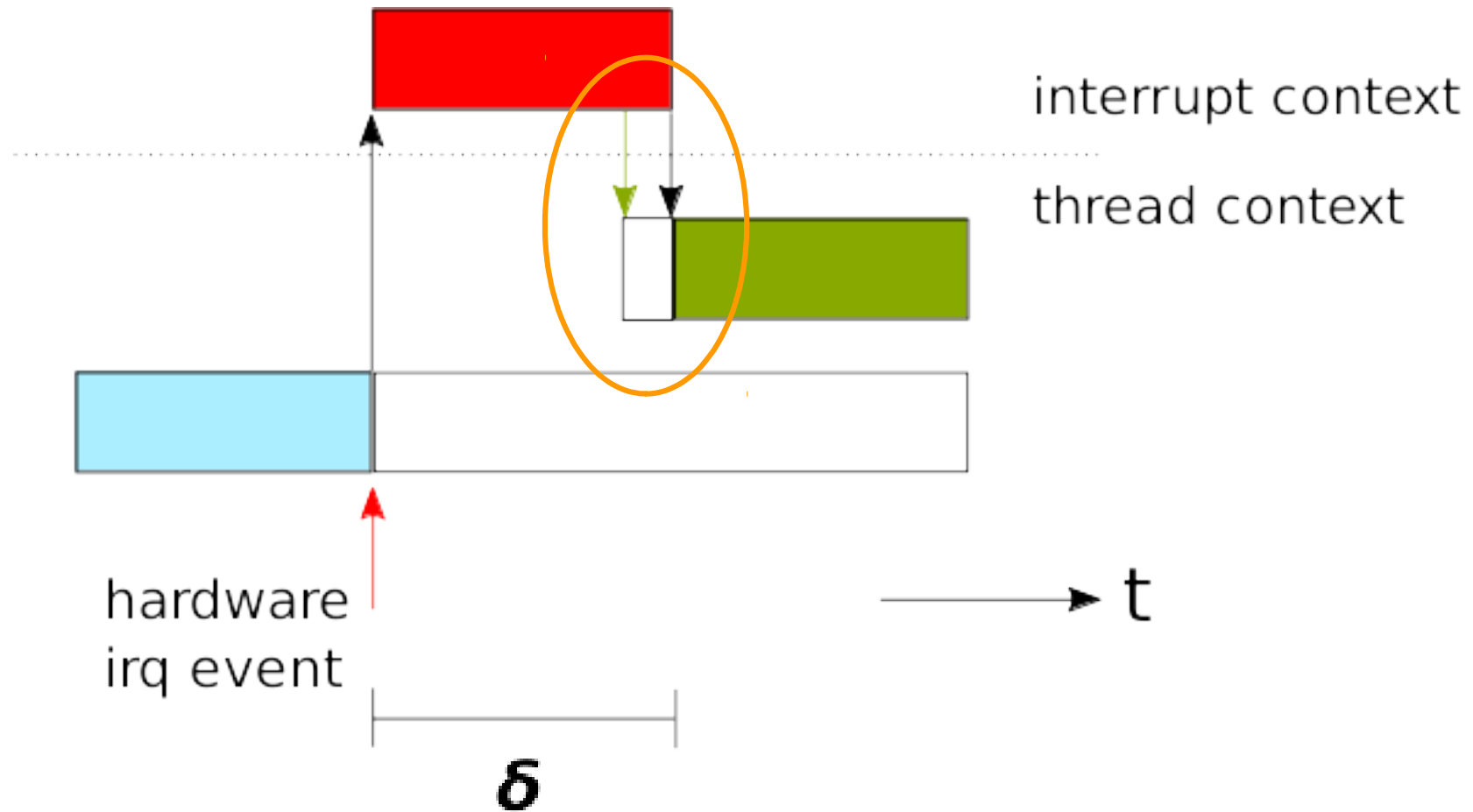
Thread context is any execution flow initiated by software.

Interrupt context is any execution flow initiated by hardware.

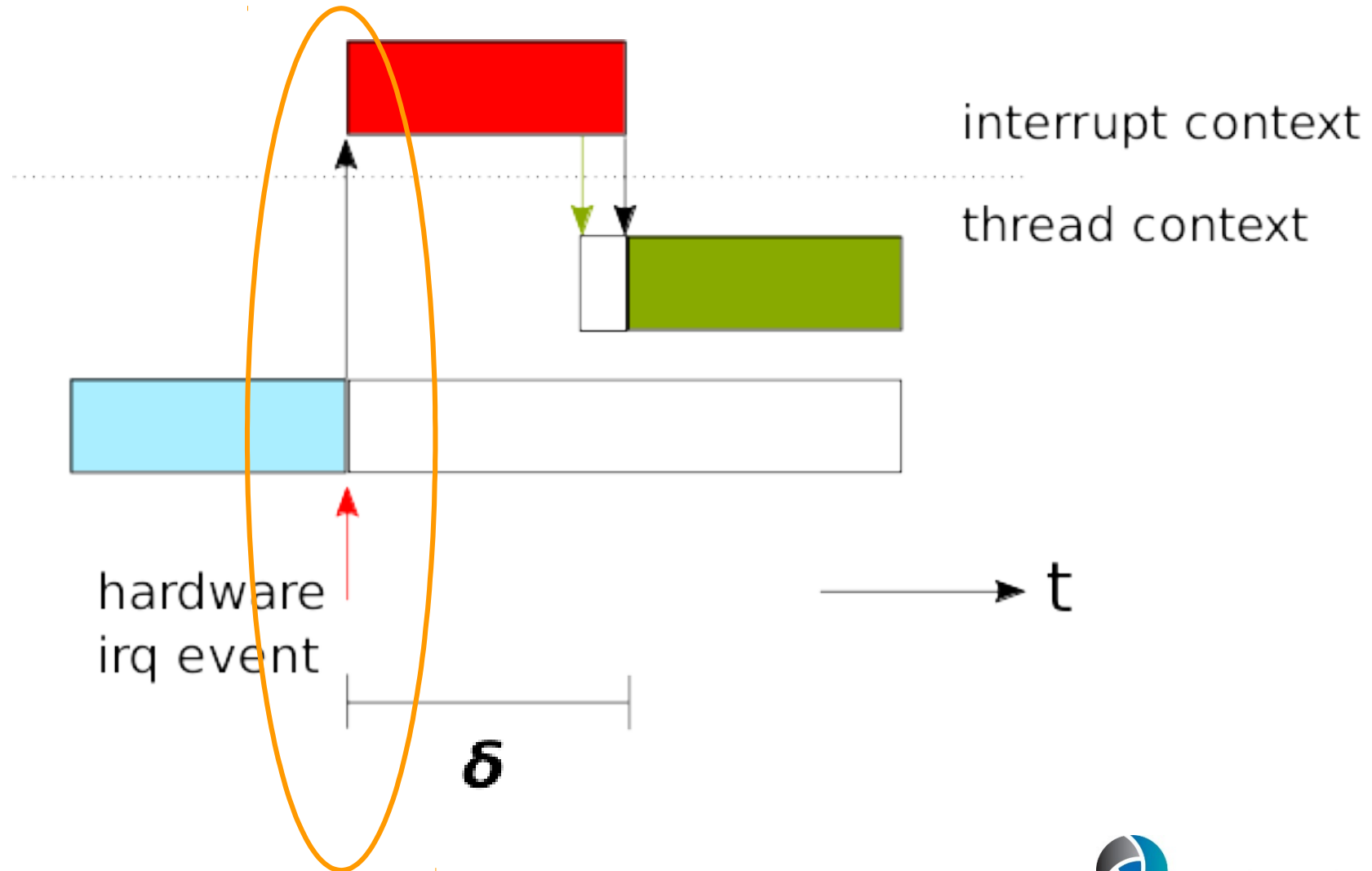
E.g. IRQ Triggered Task Switch



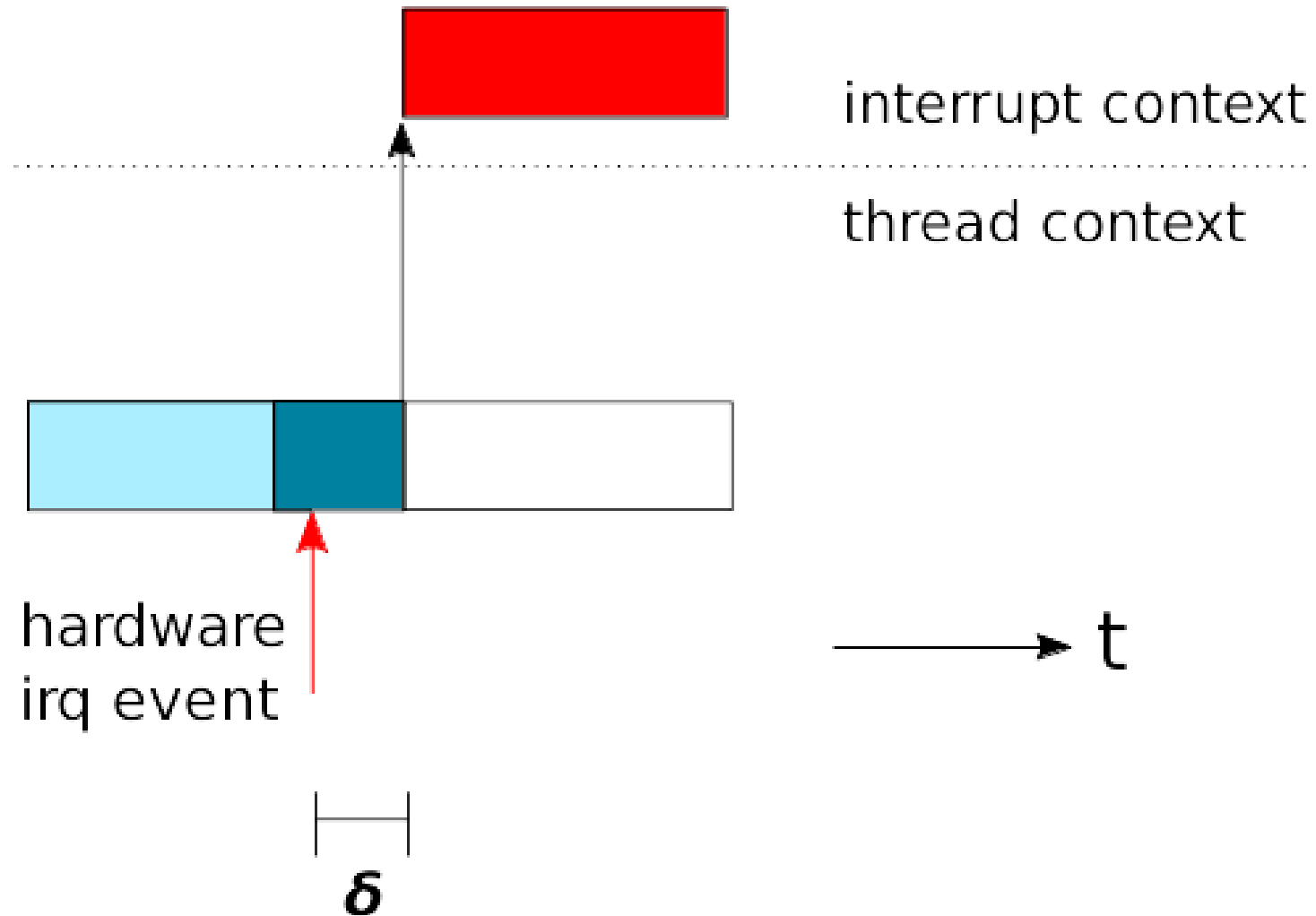
E.g. IRQ Triggered Task Switch



E.g. IRQ Triggered Task Switch



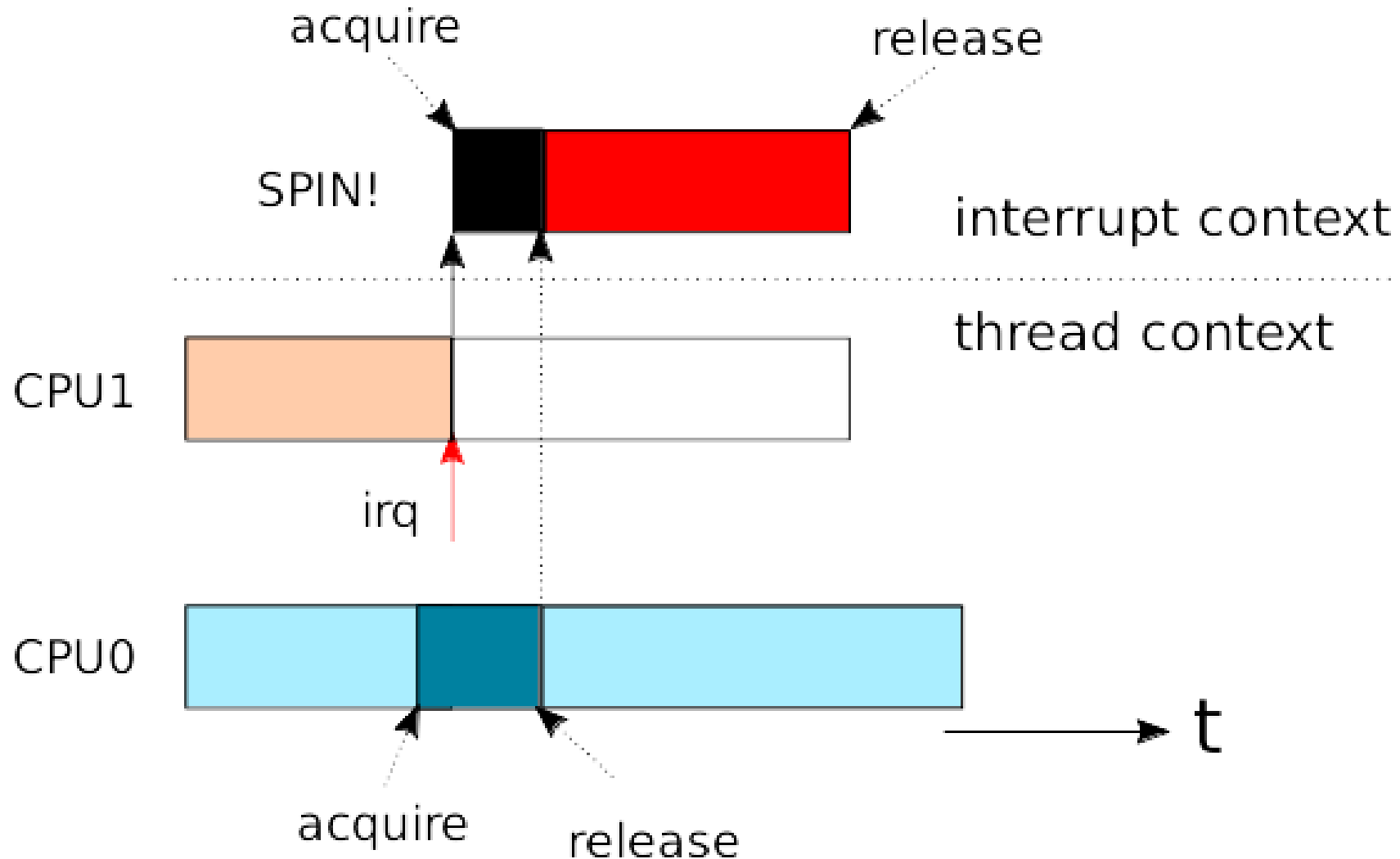
Interrupts Disabled



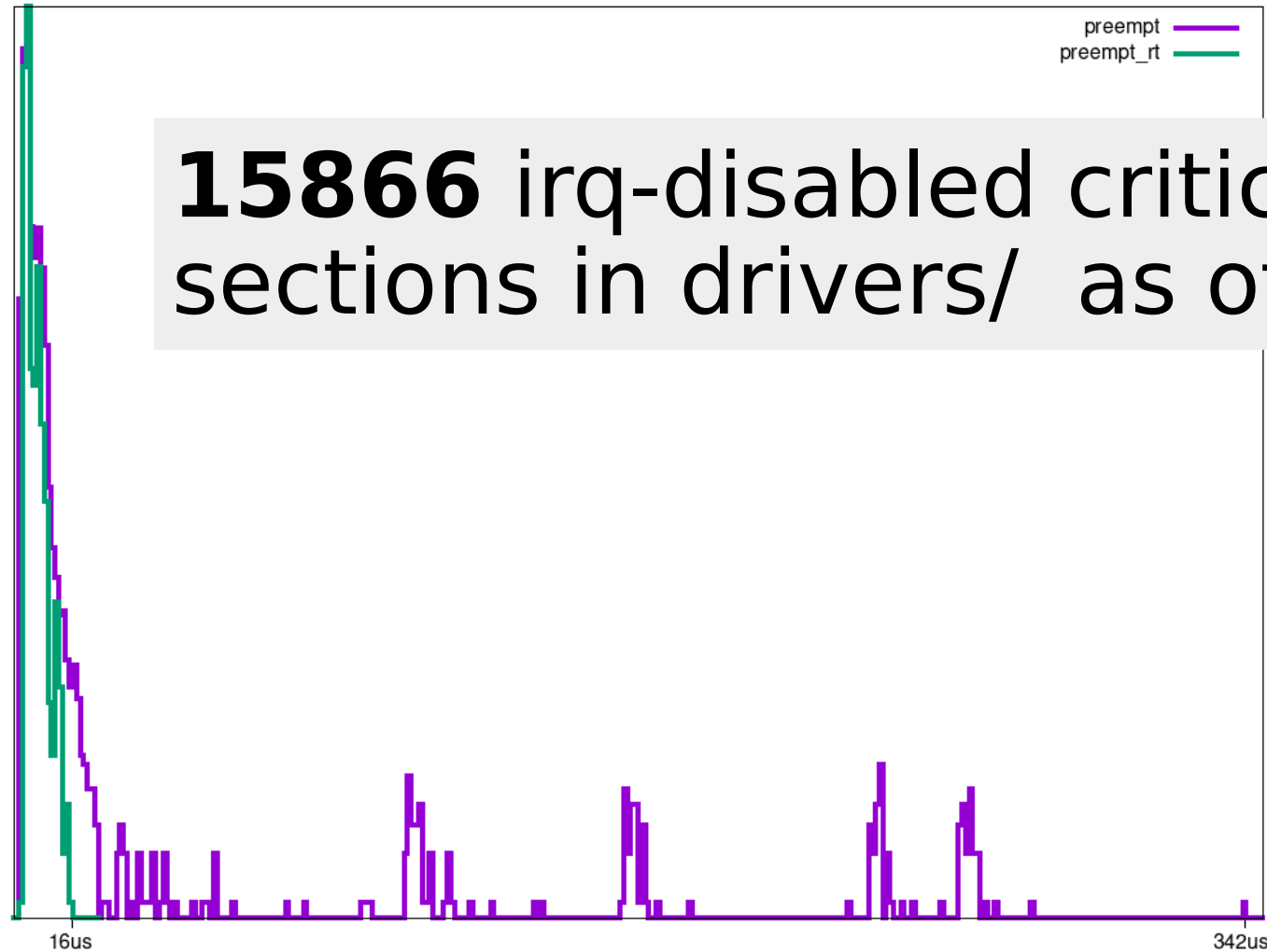
Interrupts Disabled

- Disable/enable interrupts:
 - `local_irq_disable()` / `local_irq_enable()`
- Spin lock/unlock
 - `_spin_lock()` / `_spin_unlock()`
- Spin lock/unlock + disable interrupts:
 - `spin_lock_irq()` / `spin_unlock_irq()`

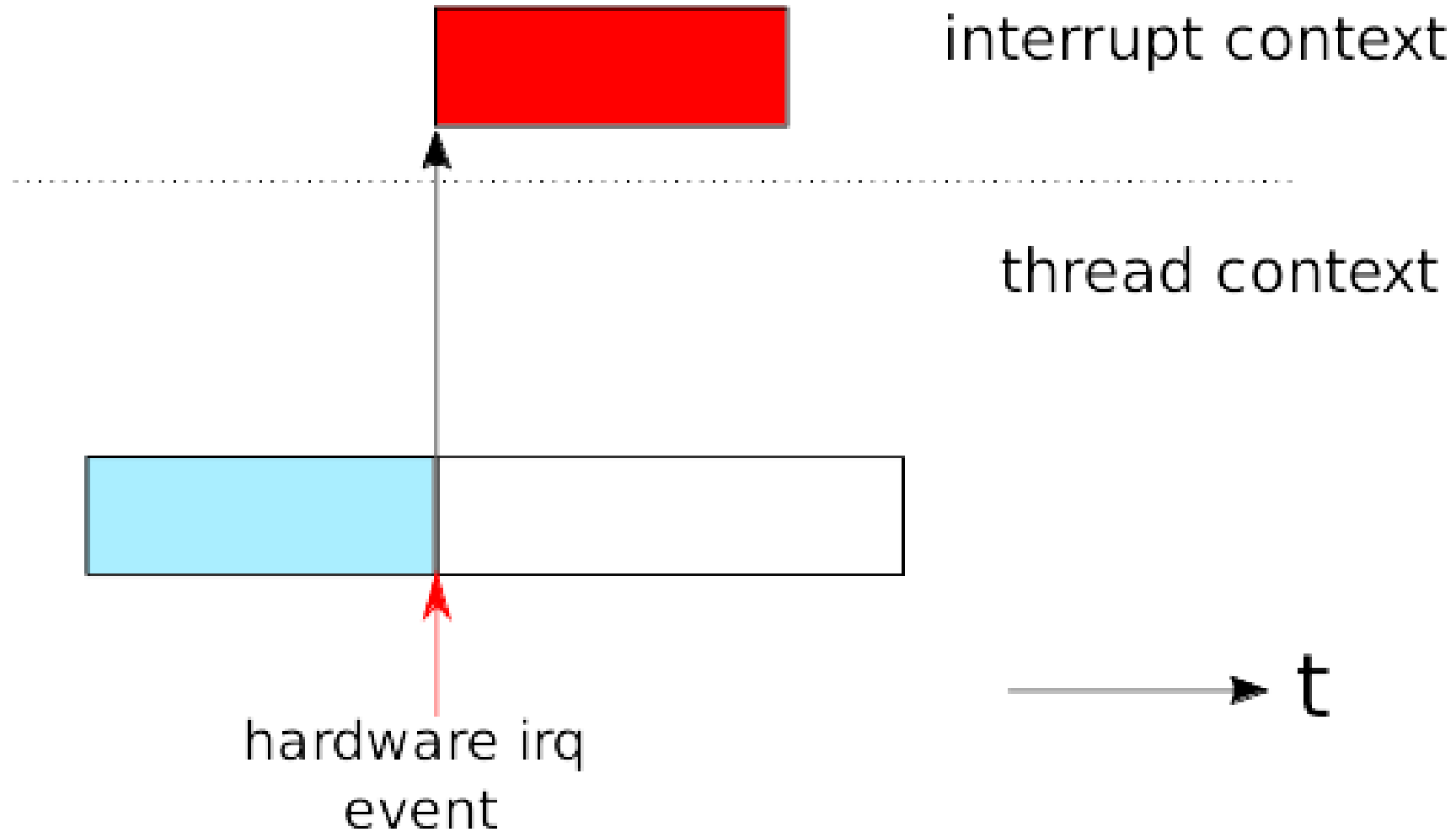
Interrupts Disabled + Spinlock



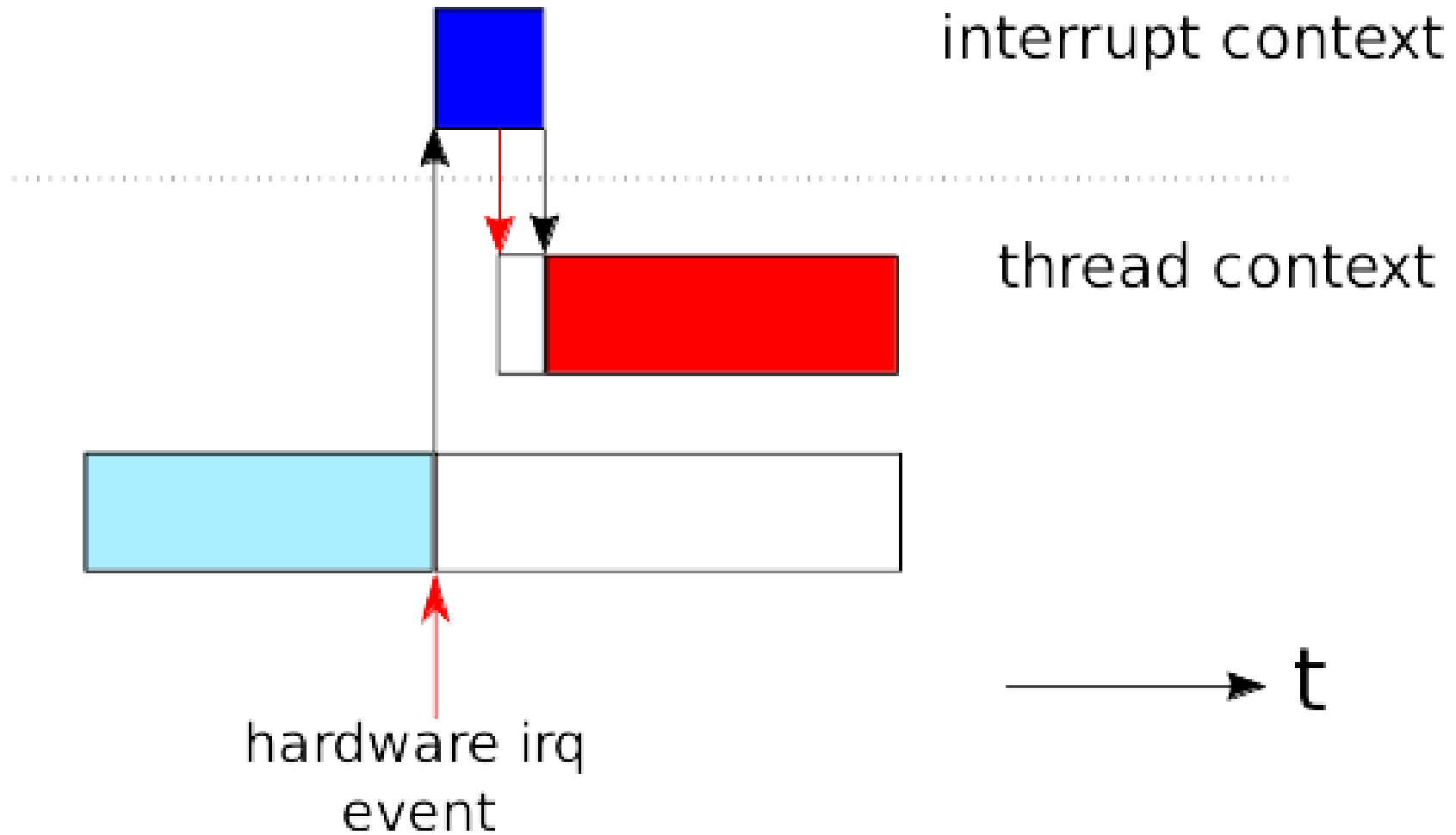
Interrupts Disabled



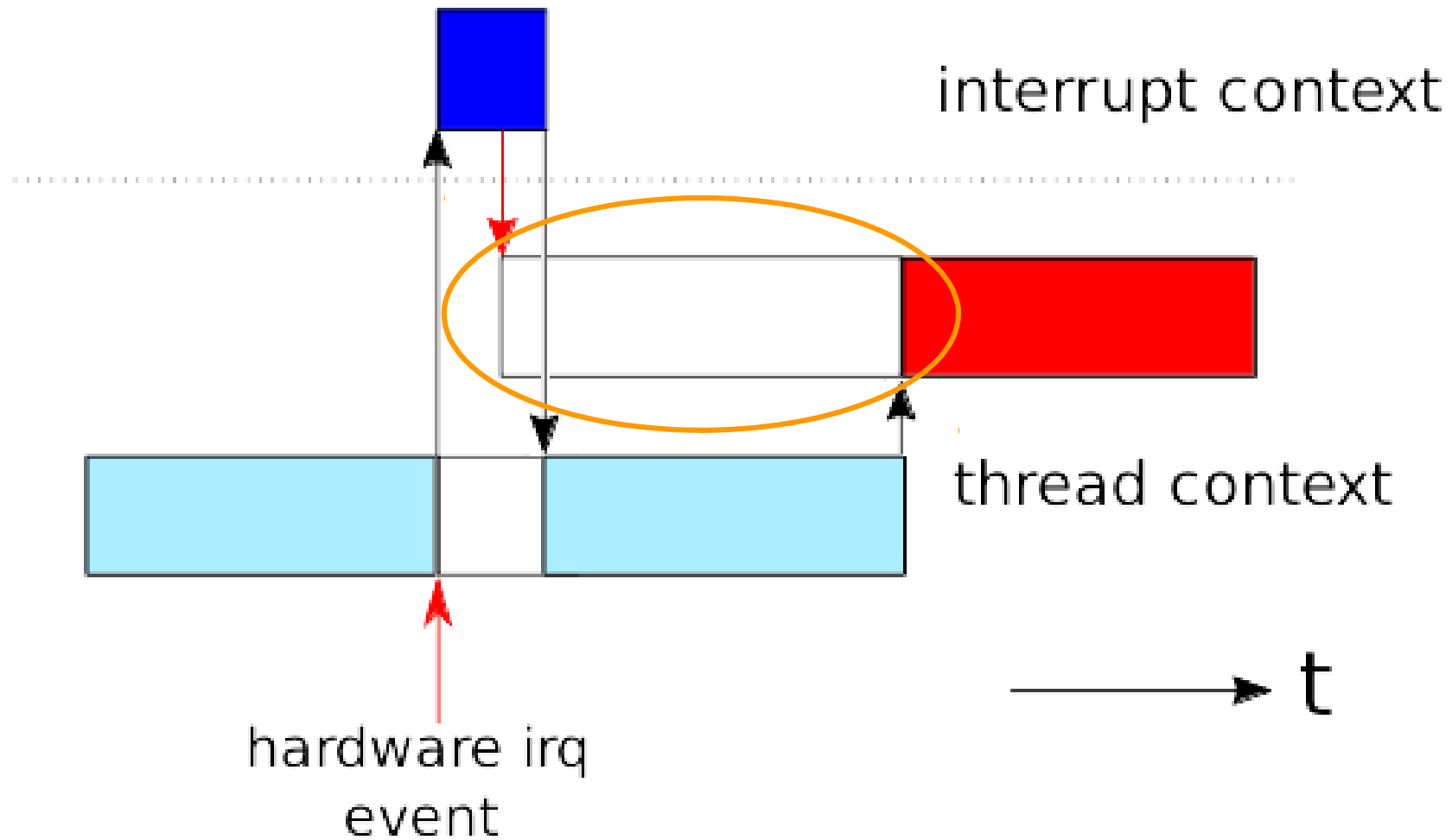
Mainline Interrupt Handling



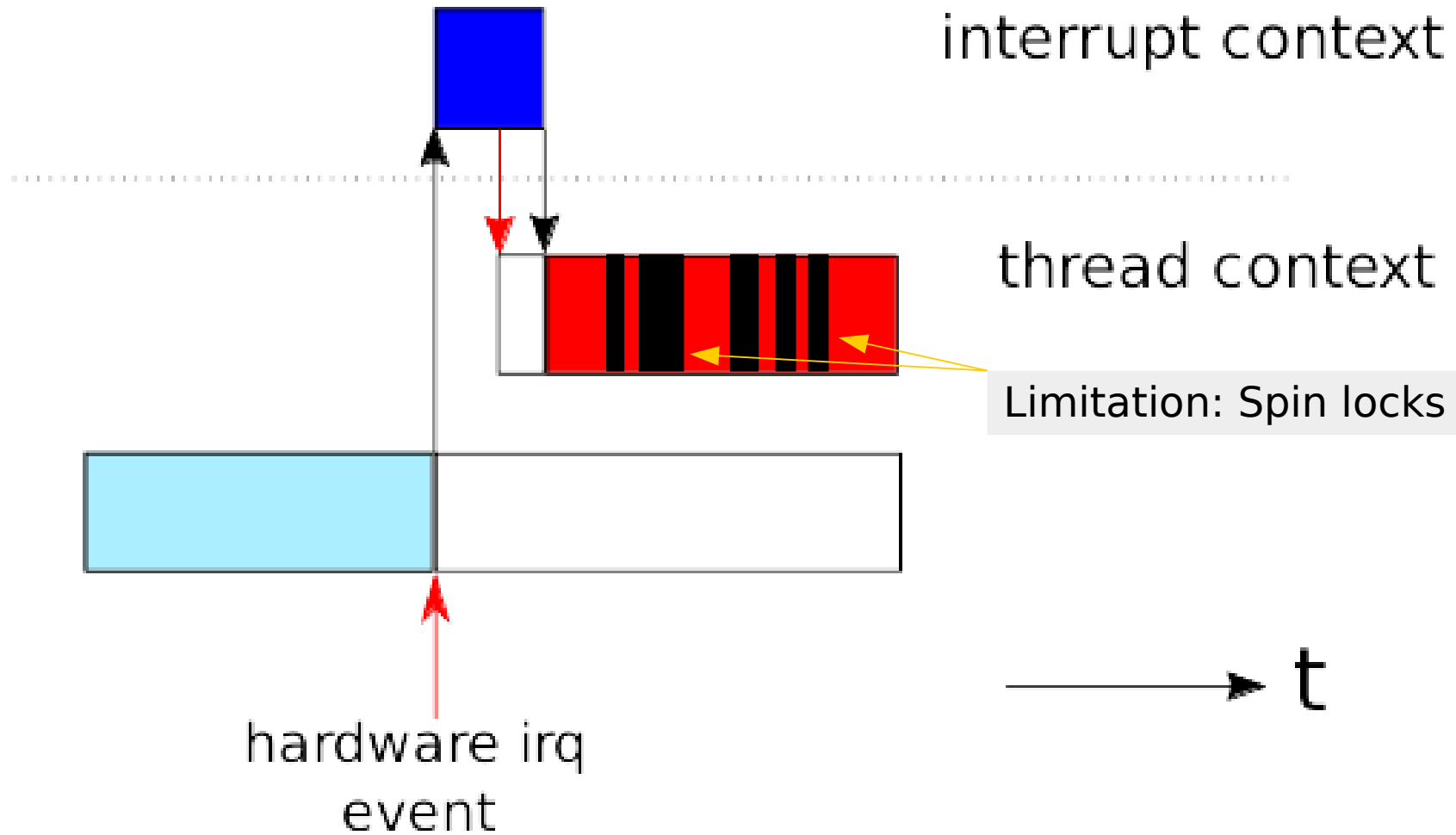
Threadirqs



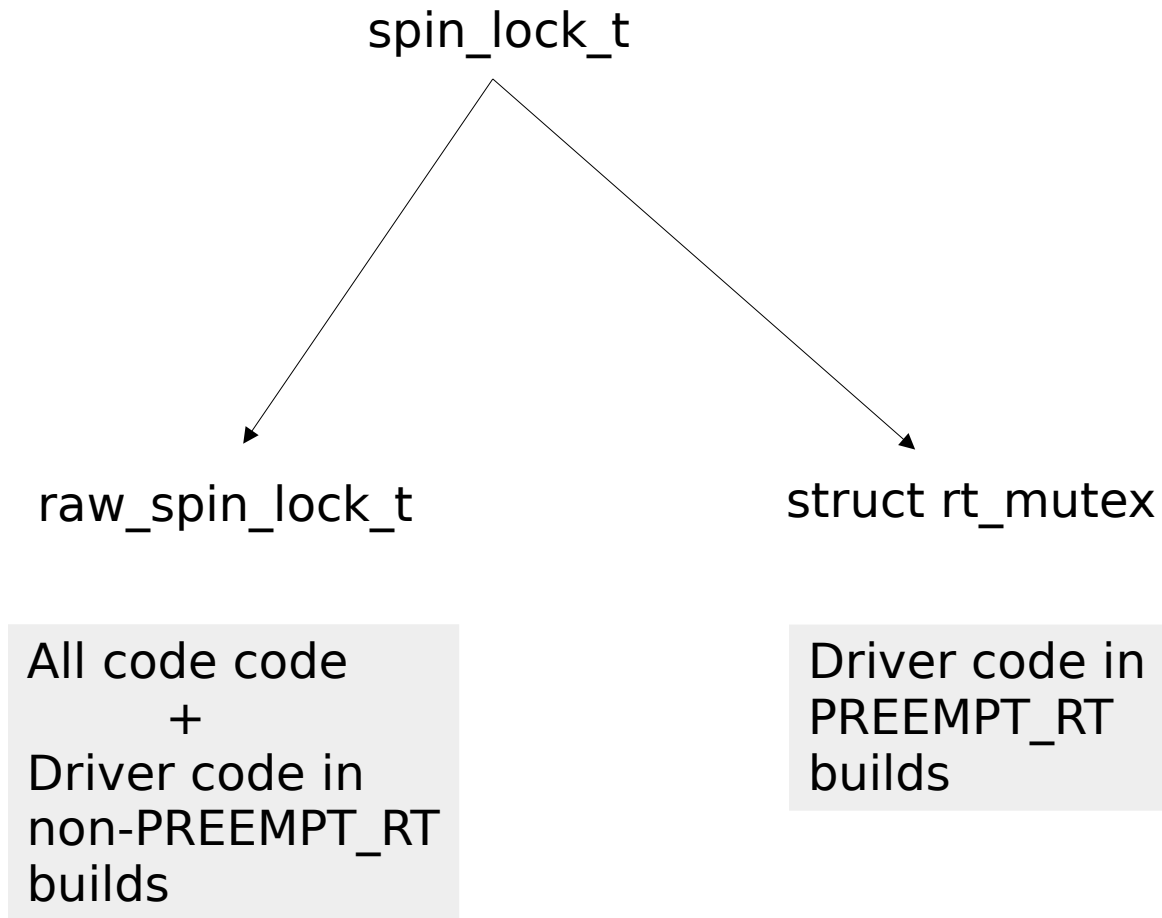
Threadirqs



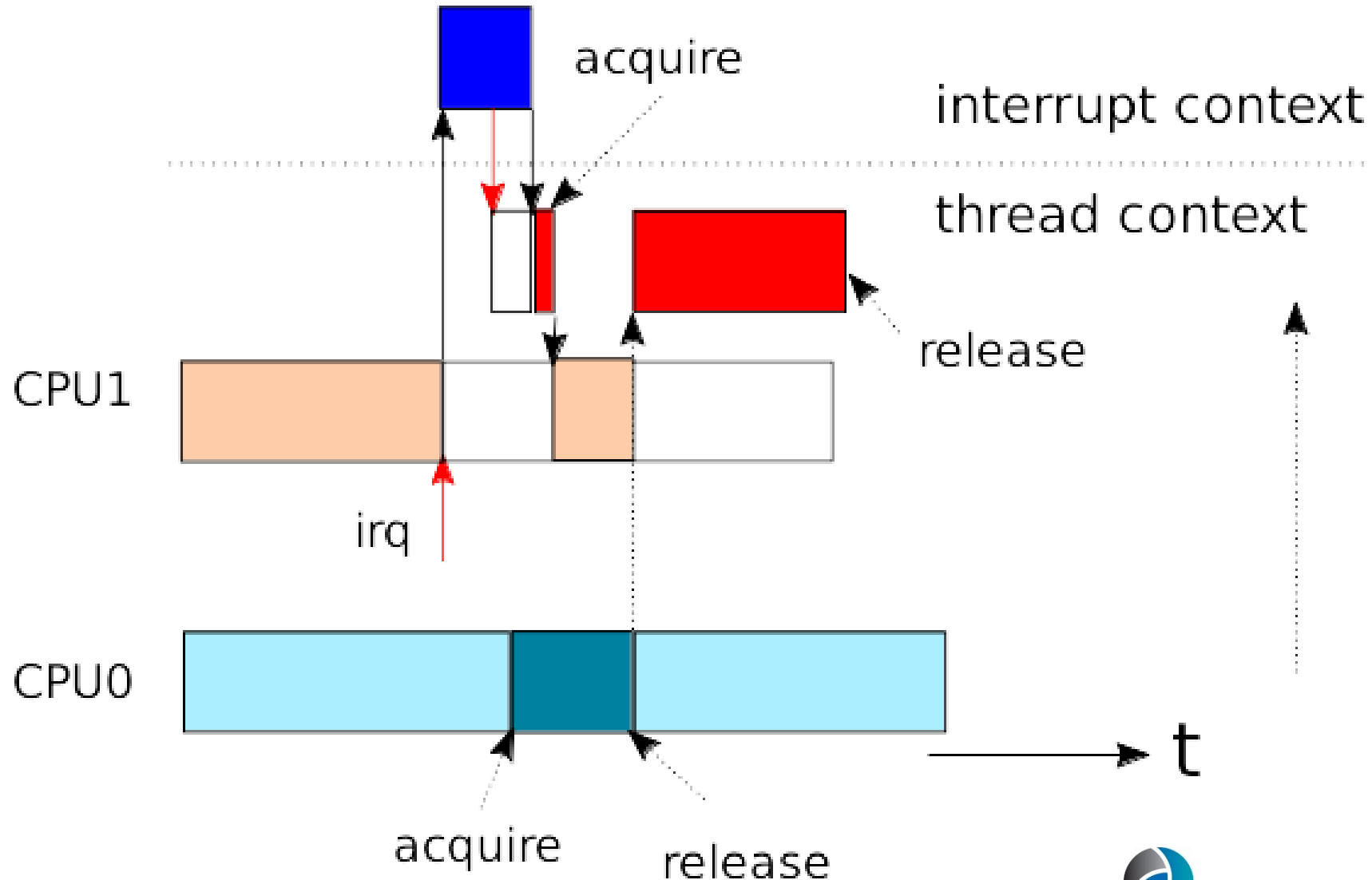
Threadirqs



Sleeping Spinlocks



Sleeping Spinlocks

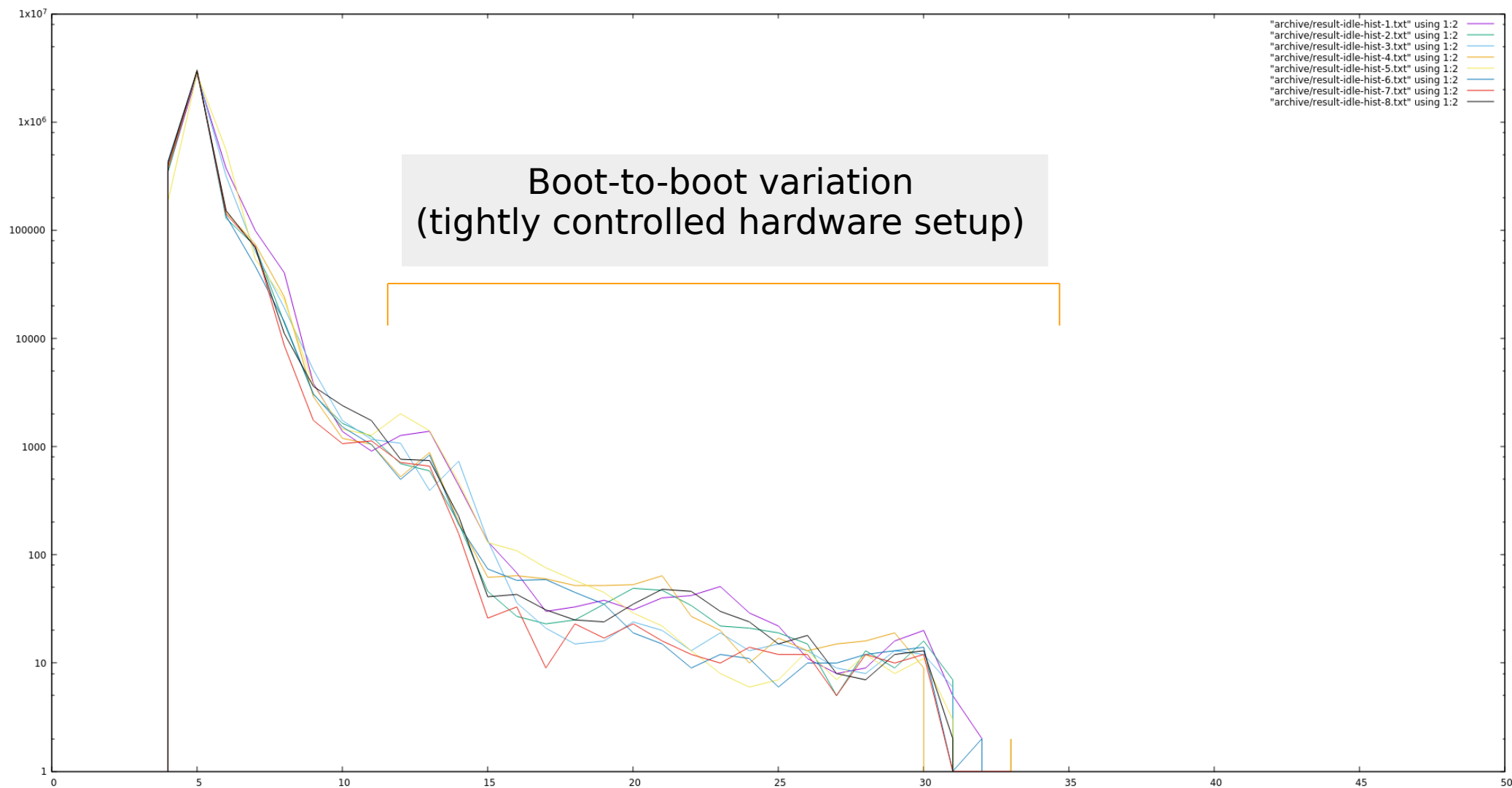


Limitations

Other sources of latency (hardware):

- Cache synchronization delays
- Bus delays
- NUMA: non-uniform memory access
- Unpredictable IO events (Ethernet, WIFI)
- Temperature changes → CPU throttling
- Etc.....

Limitations



Contact Info



Haris Okanovic

Senior Software Engineer

National Instruments

11500 N. Mopac Expy.

Austin, TX 78759

USA

Telephone: 1-512-683-6125

Email: haris.okanovic@ni.com

ni.com/nilinuxrt

Contact Info



Julia Cartwright

Senior Software Engineer

National Instruments

11500 N. Mopac Expy.

Austin, TX 78759

USA

Telephone: 1-512-683-6620

Email: julia.cartwright@ni.com

ni.com/nilinuxrt