



www.yomocode.com

Linux进程、线程和调度(4)

阅码场™
www.yomocode.com

麦当劳喜欢您来，喜欢您再来



扫描关注
Linux阅码场



第四次课大纲

Linux Deadline调度器（全新内容） FIFO/RR

Linux为什么不是硬实时的

preempt-rt对Linux实时性的改造

多核下负载均衡

中断负载均衡、RPS软中断负载均衡

练习题

用time命令跑1个含有2个死循环线程的进程

用taskset调整多线程依附的CPU

cyclictest

负载均衡

- RT 进程：N个优先级最高的RT分布到N个核
 - ◆ pull_rt_task()
 - ◆ push_rt_task()
- 普通进程
 - ◆ 周期性负载均衡
 - ◆ IDLE时负载均衡
 - ◆ fork和exec时负载均衡

CPU task affinity

■ 设置affinity

```
int pthread_attr_setaffinity_np(pthread_attr_t *, size_t, const cpu_set_t *);  
int pthread_attr_getaffinity_np(pthread_attr_t *, size_t, cpu_set_t *);  
int sched_setaffinity(pid_t pid, unsigned int cpusetsize, cpu_set_t *mask);  
int sched_getaffinity(pid_t pid, unsigned int cpusetsize, cpu_set_t *mask);
```



0x6

taskset

- `taskset -a -p 01 19999`
- `taskset -a -p 02 19999`
- `taskset -a -p 03 19999`

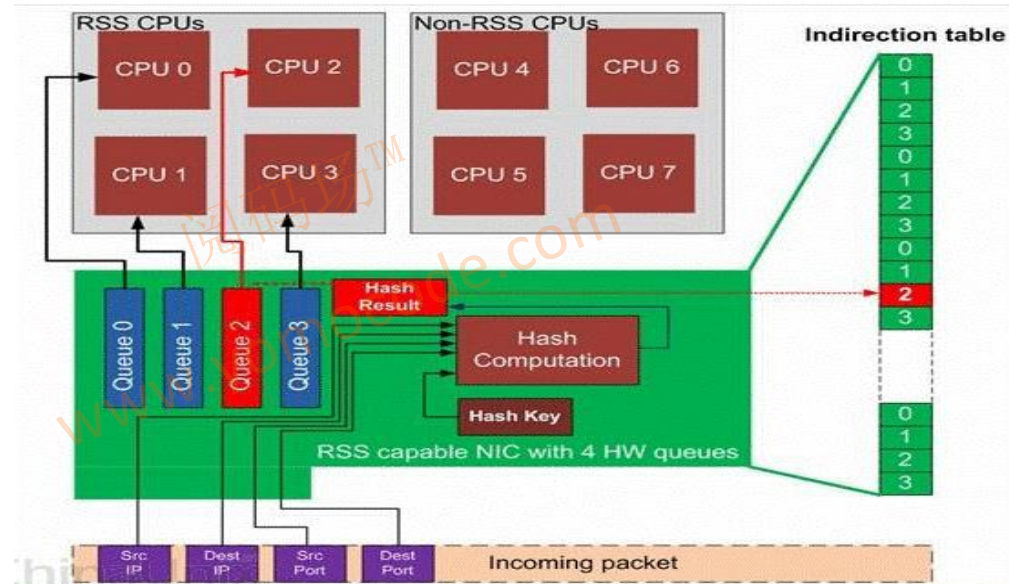
IRQ affinity

■ 分配IRQ到某个CPU

```
[root@boss ~]# echo 01 > /proc/irq/145/smp_affinity  
[root@boss ~]# cat /proc/irq/145/smp_affinity  
00000001
```

■ mq ethernet

```
/proc/irq/74/smp_affinity 000001  
/proc/irq/75/smp_affinity 000002  
/proc/irq/76/smp_affinity 000004  
/proc/irq/77/smp_affinity 000008  
...
```



多核间的softIRQ scaling

- RPS 将包处理负载均衡到多个CPU

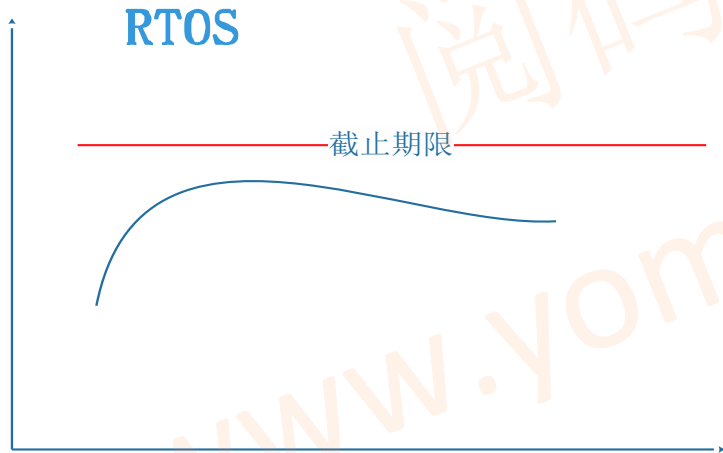
#例如

```
[root@machine1 ~]# echo ffe > /sys/class/net/eth1/queues/rx-0/rps_cpus  
ffe
```

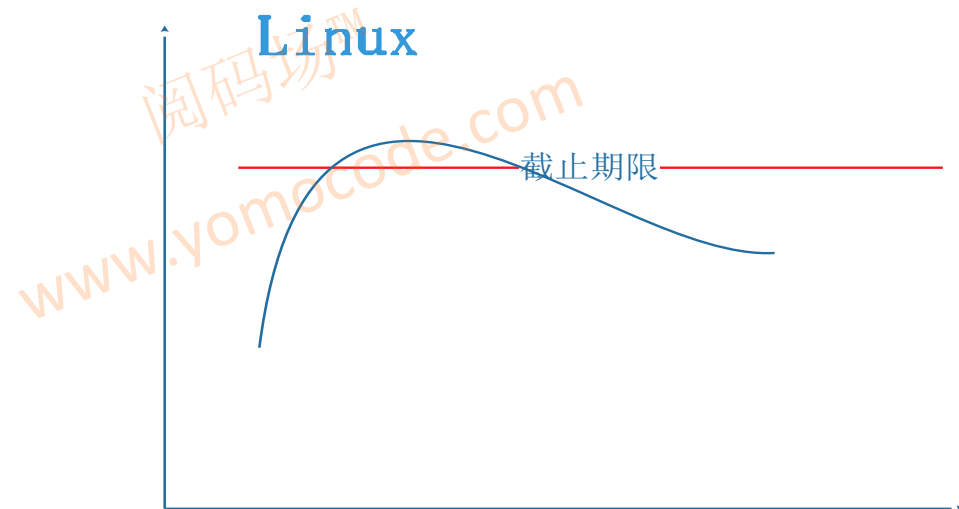
#观察

```
[root@machine1 ~]# watch -d "cat /proc/softirqs | grep NET_RX"
```

Hard realtime - 可预期性

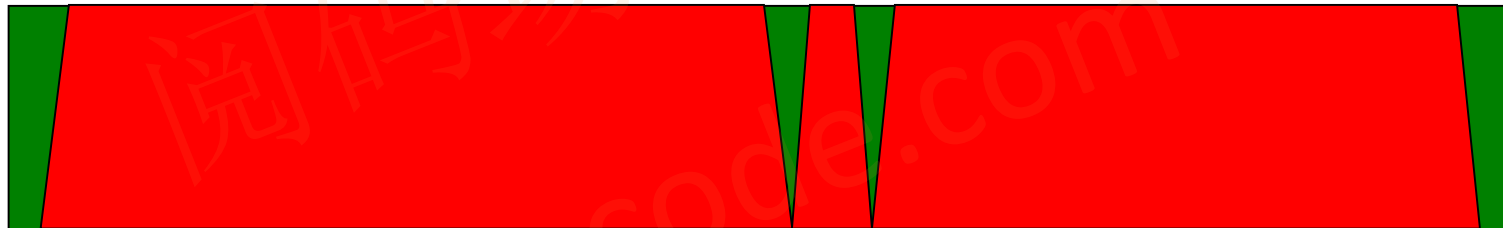


VS.

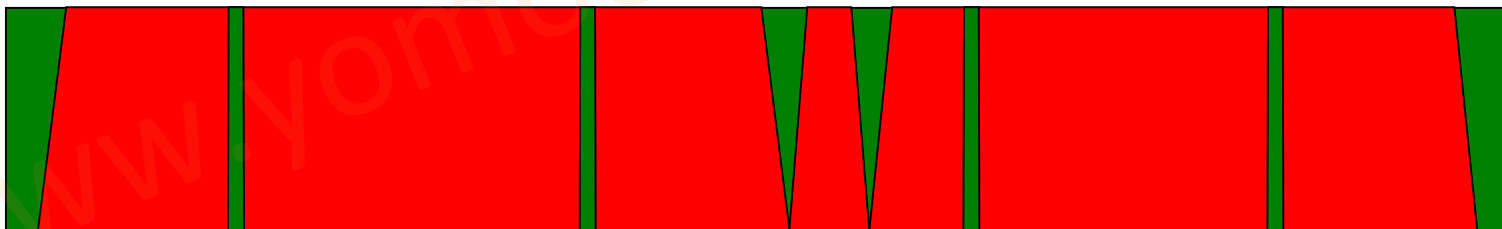


Kernel 越发支持抢占

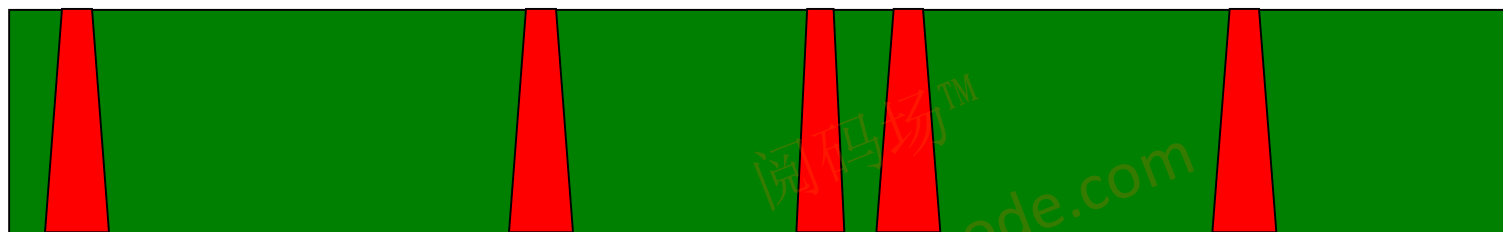
Kernel 2.0



Kernels
2.2-2.4



Preemptible
Kernel 2.4
Kernel 2.6



Real-Time
Kernel 2.6

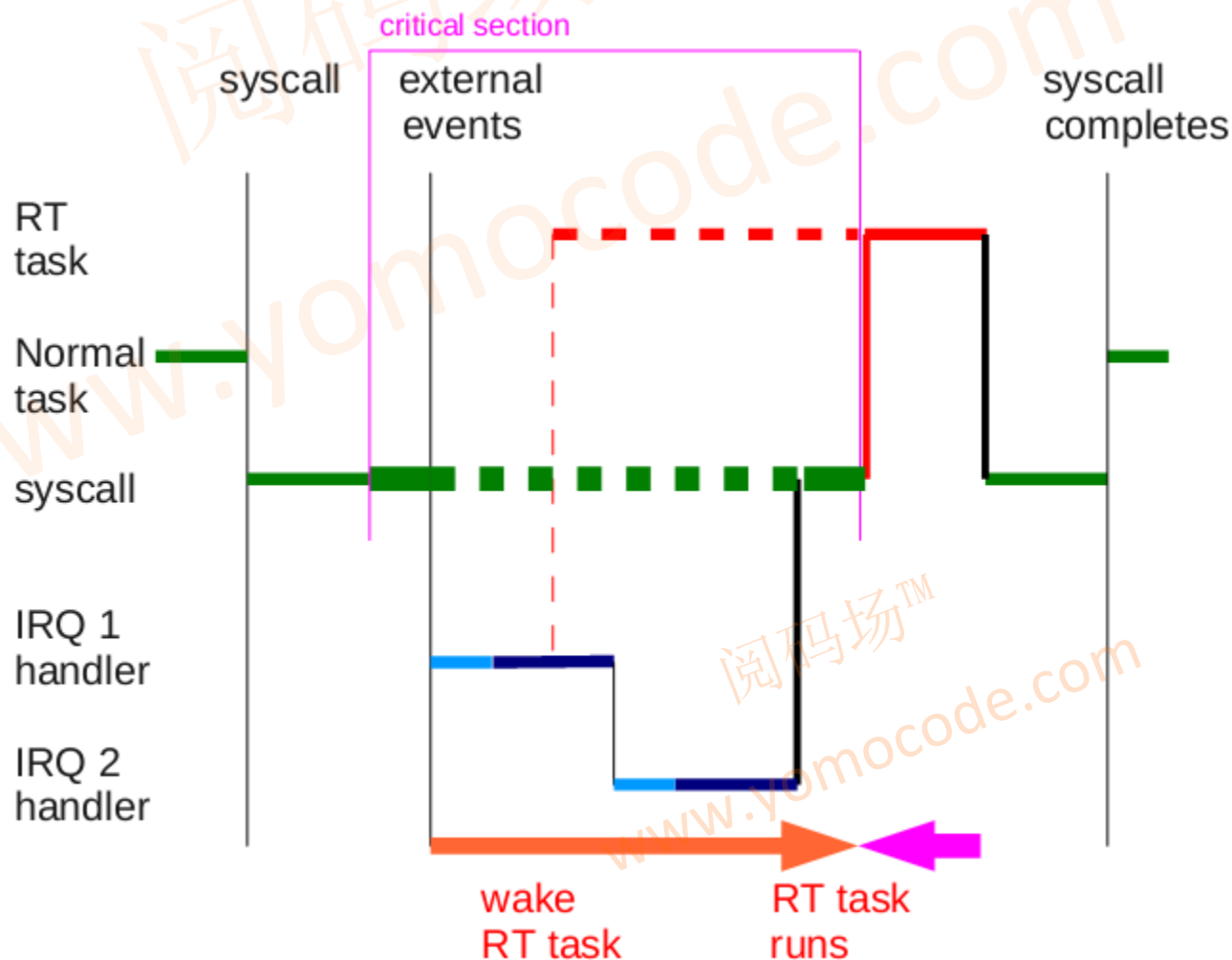


Preemptible



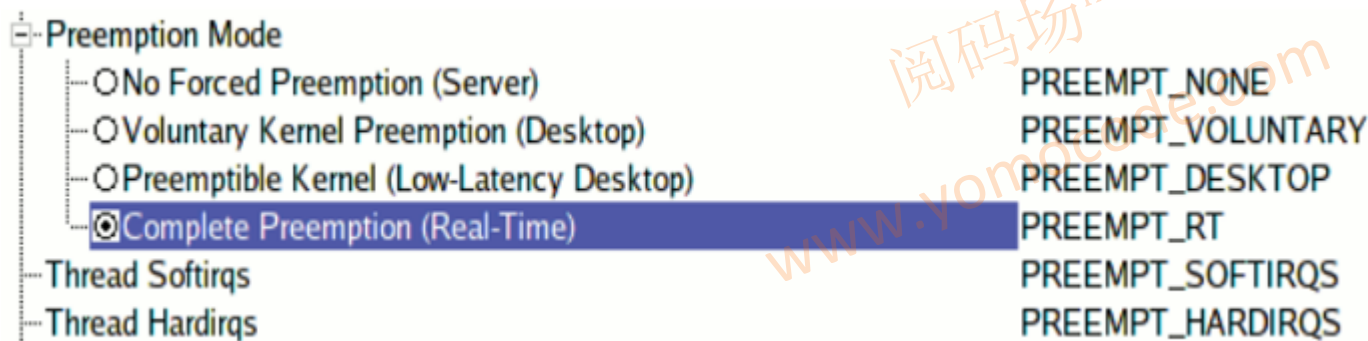
Non-Preemptible

Linux为什么不硬实时



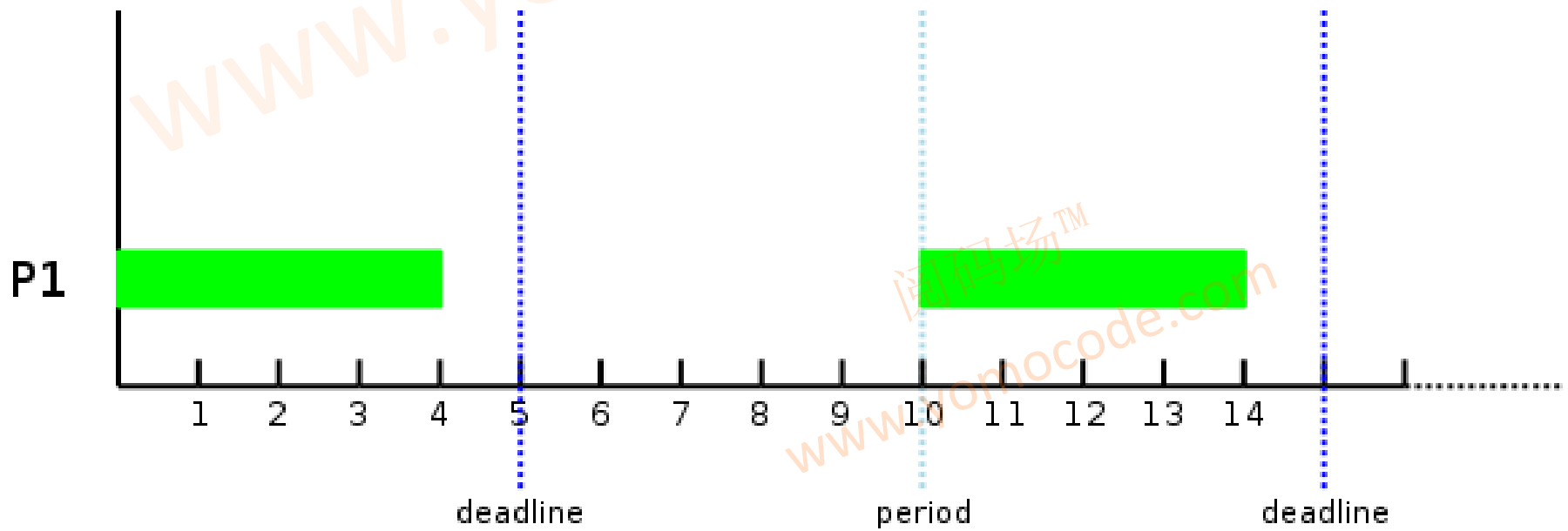
PREEMPT_RT 补丁

- spinlock 迁移为可调度的 mutex, 同时报了 raw_spinlock_t
- 实现优先级继承协议
- 中断线程化
- 软中断线程化



SCHED_DEADLINE

Runtime
Period
Deadline



谢谢!

阅码场™

www.yomocode.com

阅码场™

www.yomocode.com



阅码场出品

www.yomocode.com