

AVALIAÇÃO DO DETERMINISMO TEMPORAL NO TRATAMENTO DE INTERRUPÇÕES EM PLATAFORMAS DE TEMPO REAL LINUX

Paul Regnier, George Lima, Luciano Barreto

Laboratório de Sistemas Distribuídos (LaSiD)
Mestrado em Mecatrônica
Departamento de Ciência da Computação - UFBA
Apoio FAPESB - TIC 2630-2006

Workshop de Sistemas Operacionais 2008

DESENVOLVIMENTO DE *DoRiS* *an Ethernet Double Ring Service for RTS*

- Protocolo de comunicação em tempo real baseado em Ethernet
- Especificação e verificação formal em TLA+
- Realização de um protótipo num SO de Tempo Real

➡ OBJETIVO: achar uma plataforma operacional com garantias temporais satisfatória em situação de carga variável

Sistema Operacional de Tempo Real (SOTR) com ambiente misto de programação (controle e supervisão em tempo real e aplicações de melhor esforço)

DESENVOLVIMENTO DE *DoRiS* *an Ethernet Double Ring Service for RTS*

- Protocolo de comunicação em tempo real baseado em Ethernet
- Especificação e verificação formal em TLA+
- Realização de um protótipo num SO de Tempo Real

➡ **OBJETIVO:** achar uma plataforma operacional com garantias temporais satisfatória em situação de carga variável

Sistema Operacional de Tempo Real (SOTR) com ambiente misto de programação (controle e supervisão em tempo real e aplicações de melhor esforço)

VANTAGENS

- SO de código livre e aberto
- Boa aceitação nos ambientes de pesquisa acadêmica
- Grande variedade de aplicações (multimídia, banco de dados, sistemas embarcadas...)

PORÉM, LINUX NÃO FORNECE GARANTIAS TEMPORAIS SUFICIENTES

- Resolução temporal
- Escalonamento (resolução do *tick*)
- Tratamento das interrupções

VANTAGENS

- SO de código livre e aberto
- Boa aceitação nos ambientes de pesquisa acadêmica
- Grande variedade de aplicações (multimídia, banco de dados, sistemas embarcadas...)

PORÉM, LINUX NÃO FORNECE GARANTIAS TEMPORAIS SUFICIENTES

- Resolução temporal
- Escalonamento (resolução do *tick*)
- Tratamento das interrupções

LINUX^{Prt}

- Temporizadores de alta resolução (μs)
- Kernel totalmente **preemptível**
- Compartilhamento de recursos \Rightarrow **herança de prioridade**
- 1 *IRQ-line* \Rightarrow 1 **thread de interrupção**
- Opção **IRQF_NODELAY**

LINUX^{Xen}

- **Domínio**: ambiente isolados de execução
- **I-Pipe**: canal hierárquico de interrupções
- Técnica de **indireção das interrupções** - Adeos (*nanokernel*)
- Pilha de rede determinista RTnet
- API de tempo real (RTDM) - modo kernel e/ou usuário

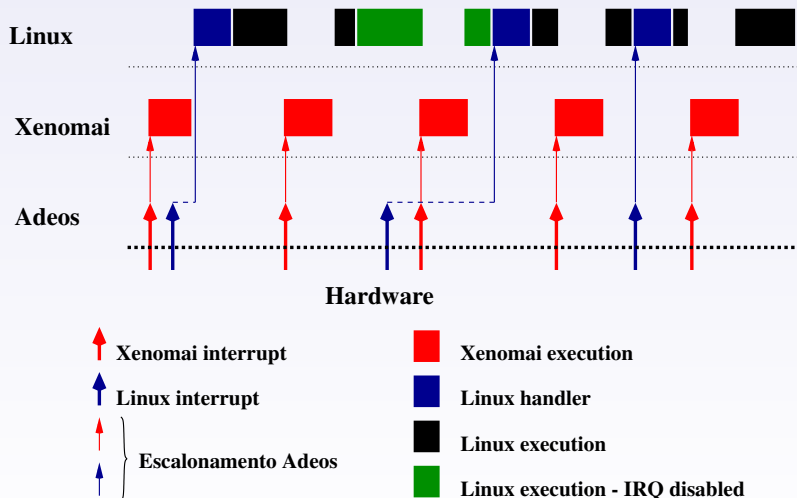
LINUX^{Prt}

- Temporizadores de alta resolução (μs)
- Kernel totalmente **preemptível**
- Compartilhamento de recursos \Rightarrow **herança de prioridade**
- 1 *IRQ-line* \Rightarrow 1 **thread de interrupção**
- Opção **IRQF_NODELAY**

LINUX^{Xen}

- **Domínio**: ambiente isolados de execução
- **I-Pipe**: canal hierárquico de interrupções
- Técnica de **indireção das interrupções** - Adeos (*nanokernel*)
- Pilha de rede determinista RTnet
- API de tempo real (RTDM) - modo kernel e/ou usuário

GERENCIAMENTO DAS INTERRUPÇÕES NO XENOMAI



MÉTRICAS: Lat_{irq} E Lat_{ativ}

LATÊNCIA DE INTERRUPTÃO: Lat_{irq}

O intervalo de tempo entre o instante no qual uma interrupção é requisitada por um dispositivo de hardware e o instante no qual o respectivo tratador começa a executar

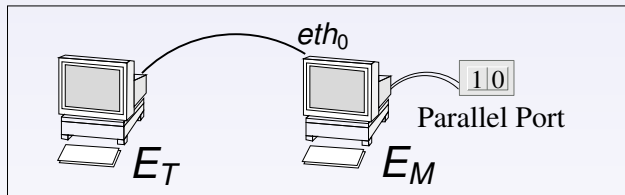
LATÊNCIA DE ATIVAÇÃO: Lat_{ativ}

O intervalo de tempo entre o instante no qual uma tarefa é acordada e o instante no qual esta tarefa começa a executar

TIME STAMP COUNTER (TSC)

Medidas de tempo realizadas pela leitura dos 64 bits do TSC com precisão de $30ns$

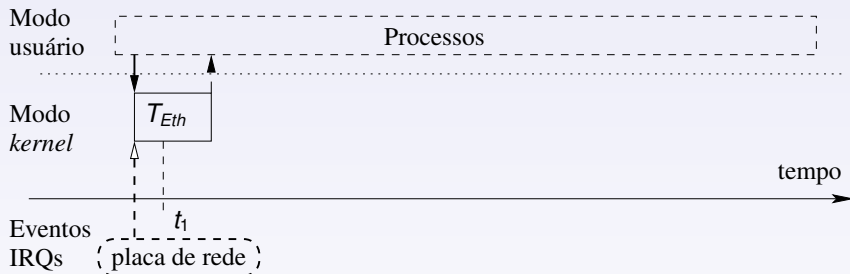
DISPOSITIVO EXPERIMENTAL



MEDIDAS NA ESTAÇÃO E_M

Realizadas no contexto da recepção, em eth_0 , de quadros Ethernet enviados por E_T com uma frequência de 20 Hz

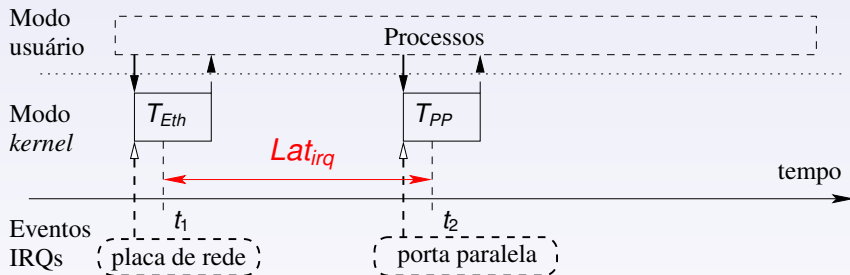
METODOLOGIA EXPERIMENTAL



t_1 em E_M

1. Pacote enviado por E_D chegue em E_M
2. Requisição de uma interrupção na PP

METODOLOGIA EXPERIMENTAL



t_1 em E_M

1. Pacote enviado por E_D chegue em E_M
2. Requisição de uma interrupção na PP

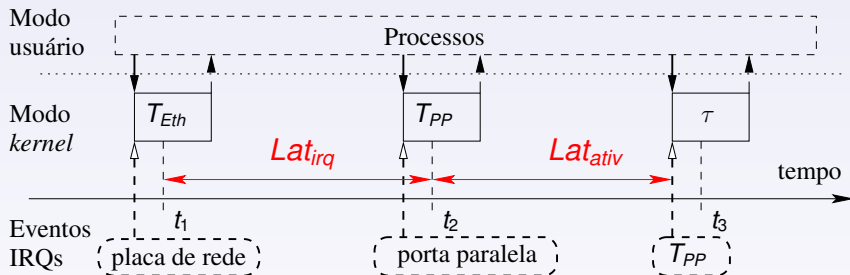
t_2 em E_M

1. Execução do tratador de interrupção da PP
2. Sinal de despertar da tarefa τ

t_3 em E_M

1. A tarefa τ acorda

METODOLOGIA EXPERIMENTAL



t_1 em E_M

1. Pacote enviado por E_D chegue em E_M
2. Requisição de uma interrupção na PP

t_2 em E_M

1. Execução do tratador de interrupção da PP
2. Sinal de despertar da tarefa τ

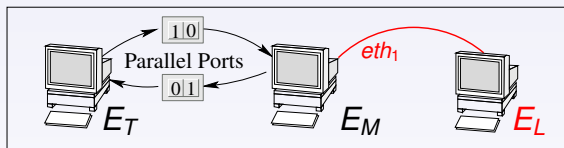
t_3 em E_M

1. A tarefa τ acorda

CARGA DE INTERRUPÇÕES, DE PROCESSAMENTO E E/S

CARGA DE INTERRUPÇÕES

Fluxo de pacotes de 64 *bytes* na frequência de 200kHz entre E_M (servidor) e E_L (load) (100.000 interrupções por segundos)



CARGA DE PROCESSAMENTO E E/S (SHELL-SCRIPT)

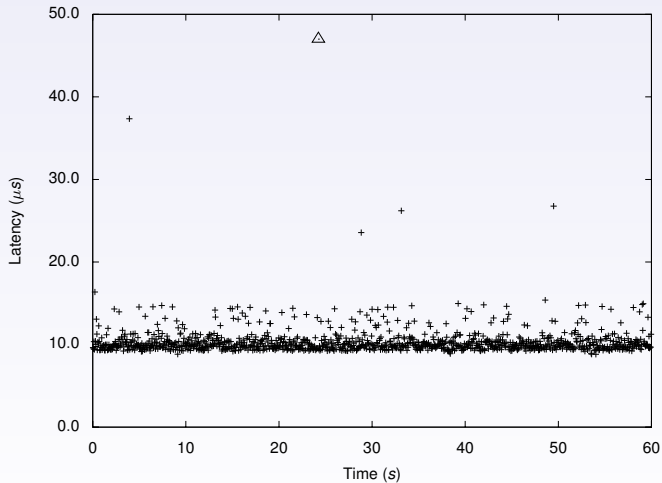
```
while "true"; do
    dd if=/dev/hda2 of=/dev/null bs=1M count=1000
    find / -name '*.c' | xargs egrep include
    tar -cjf /tmp/root.tbz2 /usr/src/linux-xenomai
    cd /usr/src/linux-preempt; make clean; make
done
```

Linux^{Std}	Linux padrão - kernel 2.6.23.9 (opção <i>low-latency</i>)
Linux^{Prt}	Linux, <i>patch</i> Preempt-RT rt12, kernel 2.6.23.9
Linux^{PrtND}	Linux ^{Prt} , opção IRQF_NODELAY para a linha PP-IRQ
Linux^{Xen}	Linux, <i>patch</i> Xenomai 2.4-rc5, kernel 2.6.19.7

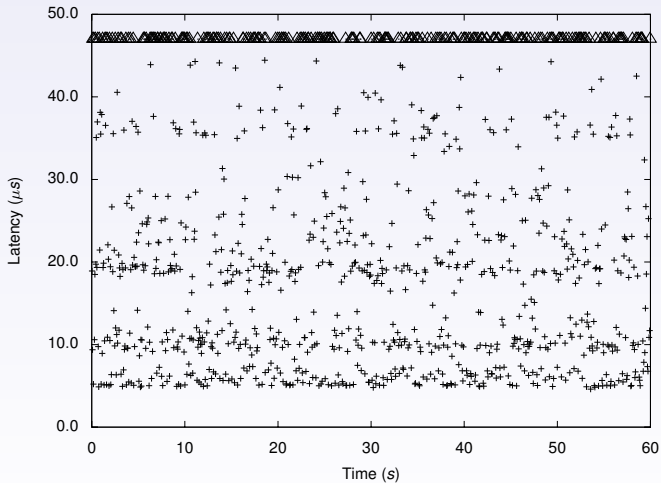
EXPERIMENTOS

- Duração dos experimentos de 600 s (12.000 medidas)
- As figuras correspondem a amostras de 60 s (1.200 medidas)
- Estatísticas correspondem a duração total de 600 s

LINUX^{Std} - Lat_{irq} COM CARGA

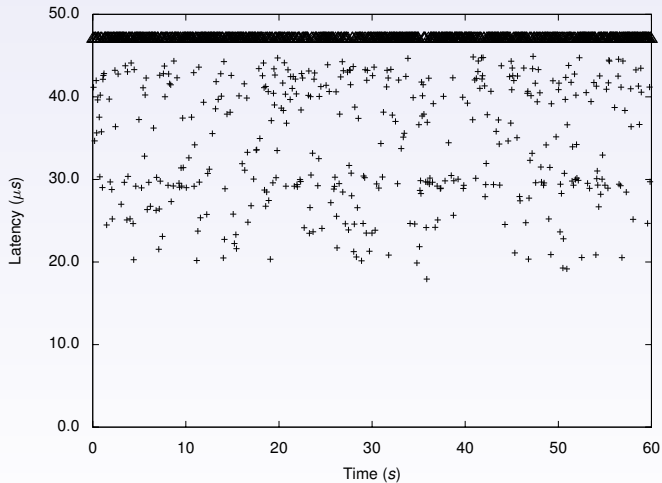


M: 10.4, SD: 1.9, Mn: 8.8, Mx: 67.7



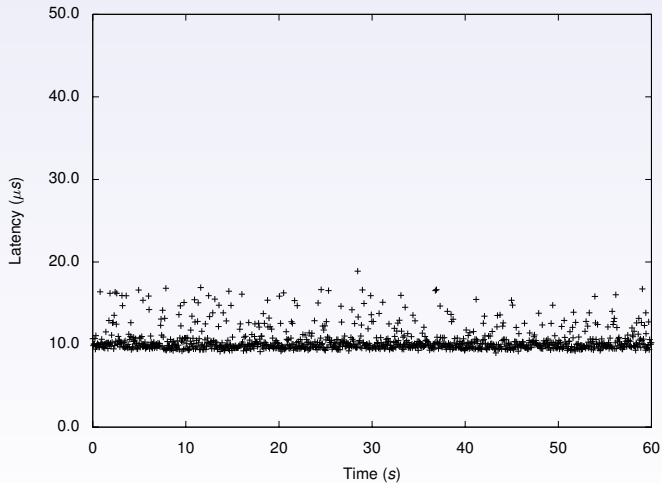
M: 37.3, SD: 48.2, Mn: 4.6, Mx: 617.5

LINUX^{Prt} - Lat_{irq} COM CARGA

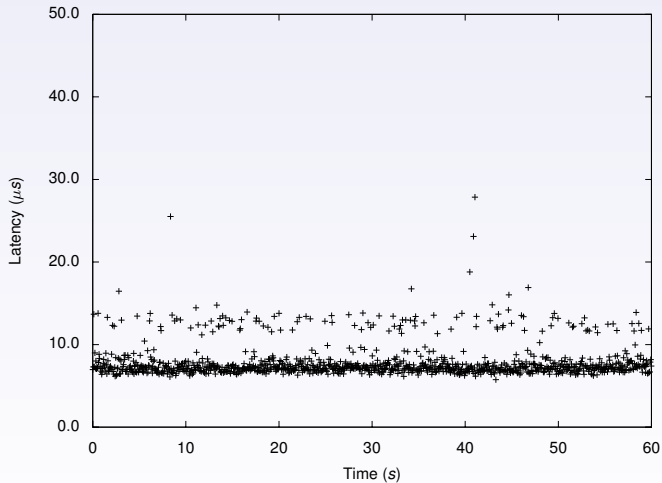


M: 58.5, SD: 26.4, Mn: 17.2, Mx: 245.9

LINUX^{PrtND} - Lat_{irq} COM CARGA

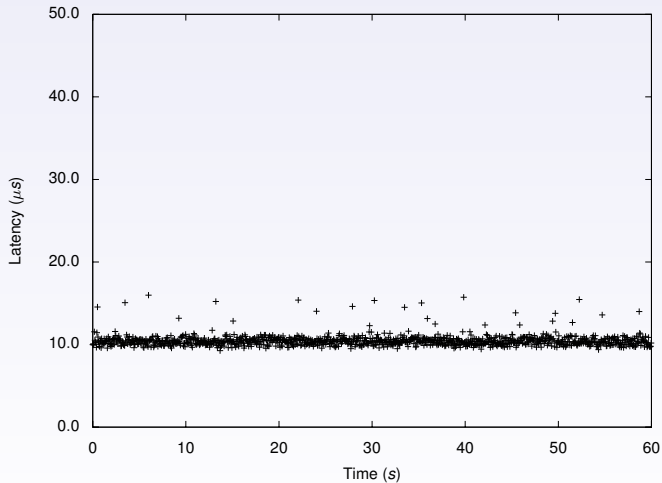


M: 10.6, SD: 1.6, Mn: 8.9, Mx: 35.8

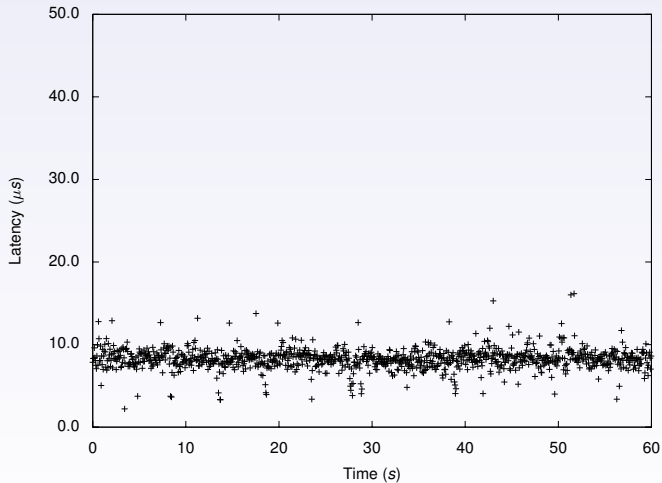


M: 8.0, SD: 2.0, Mn: 5.2, Mx: 31.0

LINUX^{Xen} - Lat_{irq} COM CARGA

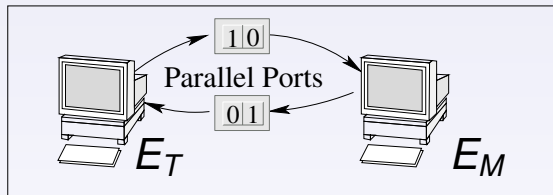


M: 10.2, SD: 0.1, Mn: 8.8, Mx: 20.8



M: 8.7, SD: 0.3, Mn: 1.8, Mx: 18.7

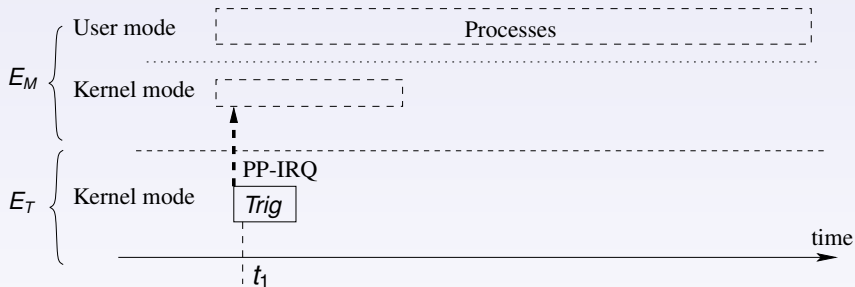
DISPOSITIVO EXPERIMENTAL 2 (DE2)



MEDIDAS NA ESTAÇÃO E_T

- E_T utiliza Linux^{xen}, em modo *single* $\Rightarrow Lat_{irq} \approx \delta$
- E_T requisita interrupções na PP de E_M , com frequência de 20 Hz
- Medidas realizadas em E_T no tratamento das interrupções requisitadas por E_M

METODOLOGIA EXPERIMENTAL - DE2

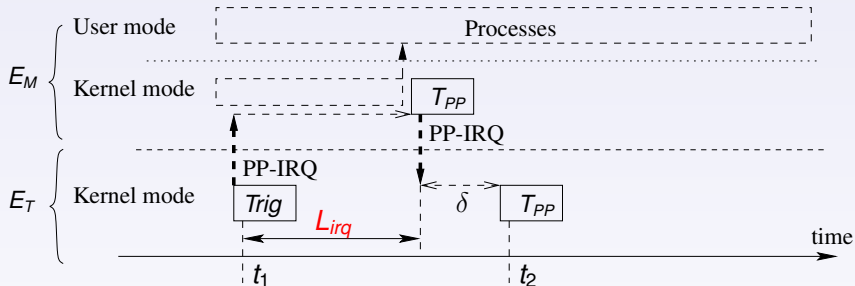


t_1 em E_T IRQ na PP de E_M e gravação de t_1

$t_2 - \delta$ em E_M IRQ na PP de E_T e sinal de despertar de τ
 t_2 em E_T Gravação de t_2

$t_3 - \delta$ em E_M A tarefa τ acorda e IRQ na PP de E_T
 t_3 em E_T Gravação de t_3

METODOLOGIA EXPERIMENTAL - DE2



t_1 em E_T IRQ na PP de E_M e gravação de t_1

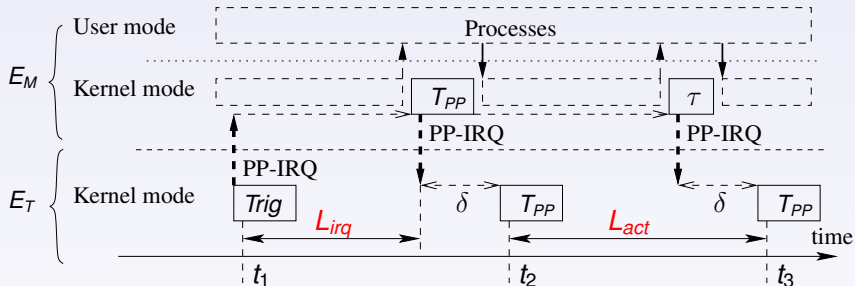
$t_2 - \delta$ em E_M IRQ na PP de E_T e sinal de despertar de τ

t_2 em E_T Gravação de t_2

$t_3 - \delta$ em E_M A tarefa τ acorda e IRQ na PP de E_T

t_3 em E_T Gravação de t_3

METODOLOGIA EXPERIMENTAL - DE2

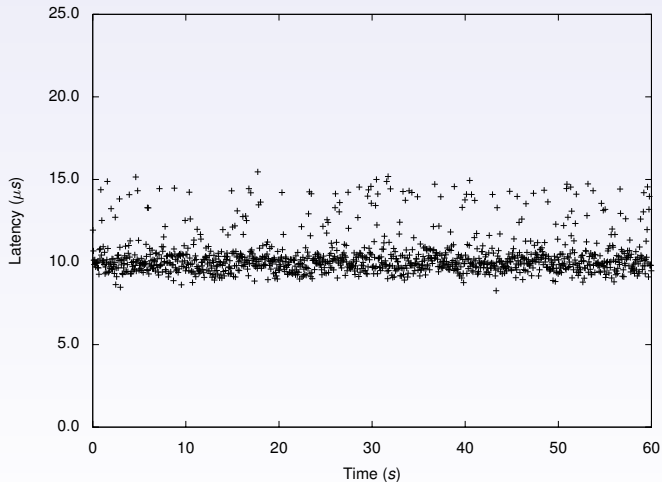


t_1 em E_T IRQ na PP de E_M e gravação de t_1

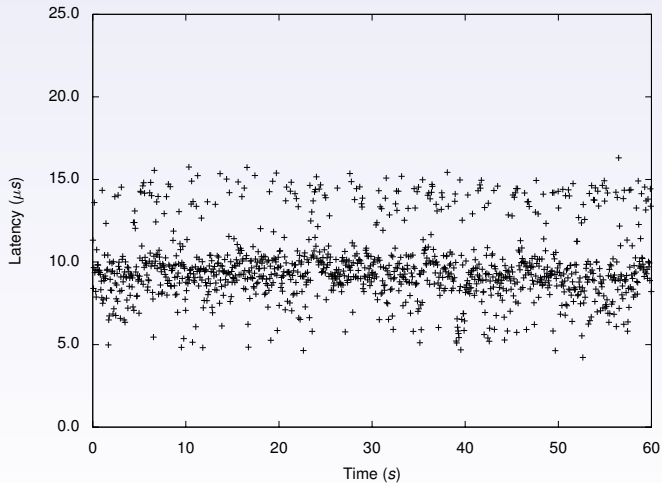
$t_2 - \delta$ em E_M IRQ na PP de E_T e sinal de despertar de τ
 t_2 em E_T Gravação de t_2

$t_3 - \delta$ em E_M A tarefa τ acorda e IRQ na PP de E_T
 t_3 em E_T Gravação de t_3

LINUX^{Xen} - Lat_{irq} COM CARGA -DE2



M: 11.3, SD: 1.2, Mn: 9.0, Mx: 19.7



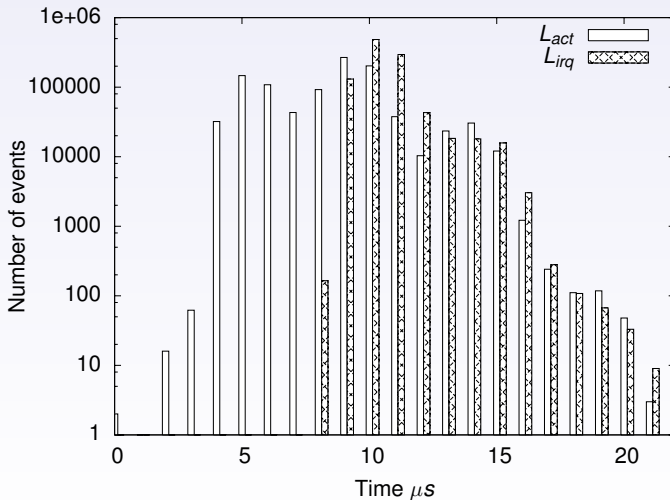
M: 9.8, SD: 2.0, Mn: 2.7, Mx: 20.8

RESULTADOS EM μS PARA AS 4 PLATAFORMAS

Exp. 1	Load	Linux ^{Std}				Linux ^{Prt}				Linux ^{PrtND}				Linux ^{Xen}			
		no		yes		no		yes		no		yes		no		yes	
		L _{irq}	L _{act}	L _{irq}	L _{act}	L _{irq}	L _{act}	L _{irq}	L _{act}	L _{irq}	L _{act}	L _{irq}	L _{act}	L _{irq}	L _{act}	L _{irq}	L _{act}
Mean		8.9	4.6	10.4	37.3	21.5	2.1	58.5	3.8	8.9	5.3	10.6	8.0	9.0	2.1	10.2	8.7
SD		0.3	0.4	1.9	48.2	1.7	0.2	26.4	2.8	0.2	0.3	1.6	2.0	0.1	0.5	0.1	0.3
Min		8.7	4.4	8.8	4.6	20.3	1.2	17.2	1.1	8.8	5.0	8.9	5.2	8.8	1.8	8.8	1.8
Max		18.4	16.2	67.7	617.5	45.1	9.4	245.9	27.4	16.7	13.1	35.8	31.0	11.1	8.4	20.8	18.7
Exp. 2	Mean	9.0	3.6	12.5	19.9	10.2	3.7	31.2	7.2	9.2	4.6	11.8	14.9	9.1	4.0	11.3	9.8
	SD	0.4	0.6	3.2	17.4	0.5	0.4	19.0	3.1	0.4	0.5	2.3	5.6	0.3	0.3	1.2	2.0
	Min	8.8	-1.3	9.0	2.3	10.0	0.8	10.4	2.2	8.9	-0.3	9.1	4.5	8.8	0.3	9.0	2.7
	Max	18.4	19.0	75.0	428.4	30.8	12.7	203.9	21.2	14.9	14.2	49.2	85.0	13.4	9.6	19.7	11.8

LINUX^{Xen} - 20 HORAS - DE2

HISTOGRAMA COM 1.000.000 EVENTOS (ESCALA LOGARÍTMICA)



ROSTEDT AND HART [3] - LINUX^{Prt}

- Latência de interrupção e escalonamento (tarefa periódica)
- Sem carga do processador

SIRO AND AL. [4] - LINUX^{Prt}, RT-LINUX, RTAI

- Desvios de uma tarefa periódica
- *Imbench*, sem carga de interrupção

BENOIT AND YAGHMOUR [7] - LINUX^{Prt}, LINUX^{Xen}

- Latência de interrupção (metodologia 2)
- *Imbench*, com carga de interrupção
- Não publicado, sem latência de ativação

CONCLUSÃO

As duas metodologias experimentais forneceram resultados similares.

LINUX^{Std}

- Não provê garantias temporais
- Piores casos acima de $100\mu s$ e desvios padrões altos

LINUX^{Prt}

- Piores casos abaixo de $50\mu s$ sem *thread* de interrupção
- Degradação sensível de Lat_{irq} com *thread* de interrupção

LINUX^{Xen}

- **Plataforma a mais determinística**
- Piores casos observados de $20\mu s$, desvios padrões de $3\mu s$



D. P. Bovet, M. Cesati

Understanding the Linux Kernel

O'Reilly (3rd), 2005



J. Corbet, A. Rubini and G. Kroah-Hartman

Linux Device Drivers

O'Reilly (3rd), 2005



S. Rostedt and D. V. Hart

Internals of the RT Patch

Proceedings of the Linux Symposium, 2007



A. Siro, C. Emde and N. McGuire

Assessment of the Realtime Preemption Patches (RT-Preempt) and their impact on the general purpose performance of the system

Proceedings of the 9th Real-Time Linux Workshop, 2007

- ▶ P. Gerum, K. Yaghmour et al.
Xenomai
<http://www.xenomai.org>, acessado em julho 08
- ▶ I. MOLNAR et al.
Preempt-RT
<http://rt.wiki.kernel.org> - acessado em julho 08
- ▶ K. Benoit and K. Yaghmour
Preempt-RT and I-pipe: the numbers
<http://marc.info/?l=linux-kernel&m=112086443319815&w=2>,
acessado em julho 2008