
MAC5753 - Sistemas Operacionais

Daniel Macêdo Batista

IME - USP, 26 de Novembro de 2020

Virtualização de E/S: Interfaces de rede

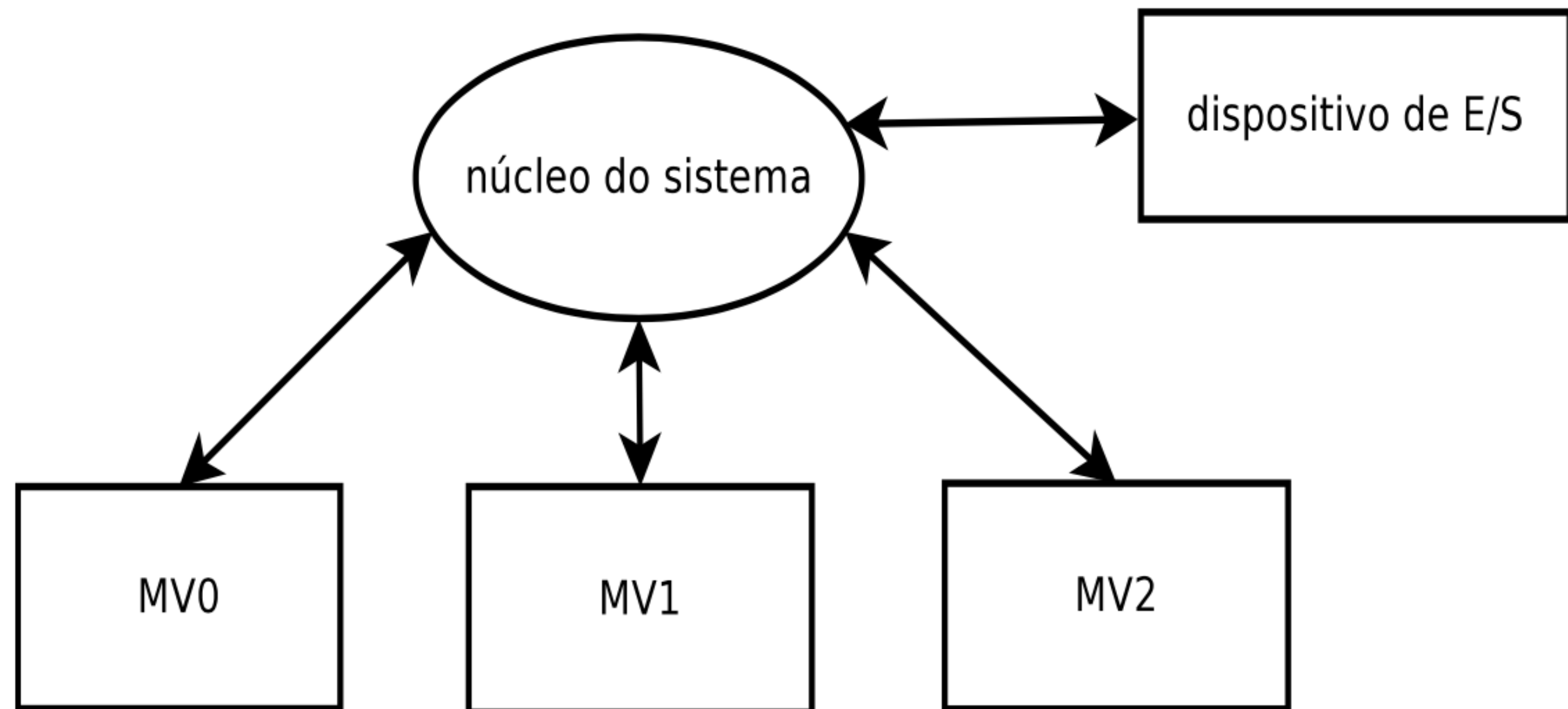
Virtualização de E/S: Interfaces de rede

- Virtualização de E/S privada ou compartilhada
Segurança X Eficiência
- Dispositivo de rede está preparado para ser controlado por uma única máquina
- Como compartilhar o dispositivo físico de rede entre VMs?
- Solução
 - 1) uma máquina gerencia o dispositivo de rede enquanto as outras máquinas acessam o dispositivo por essa máquina (ponte virtual)
 - 2) o dispositivo físico suporta o compartilhamento em uma granularidade menor (Em redes, várias filas, cada uma para uma VM sem passar pela camada de software do hipervisor) – SR-IOV (Single Root Input/Output Virtualization)

Introdução

Virtualização de E/S:
Interfaces de rede

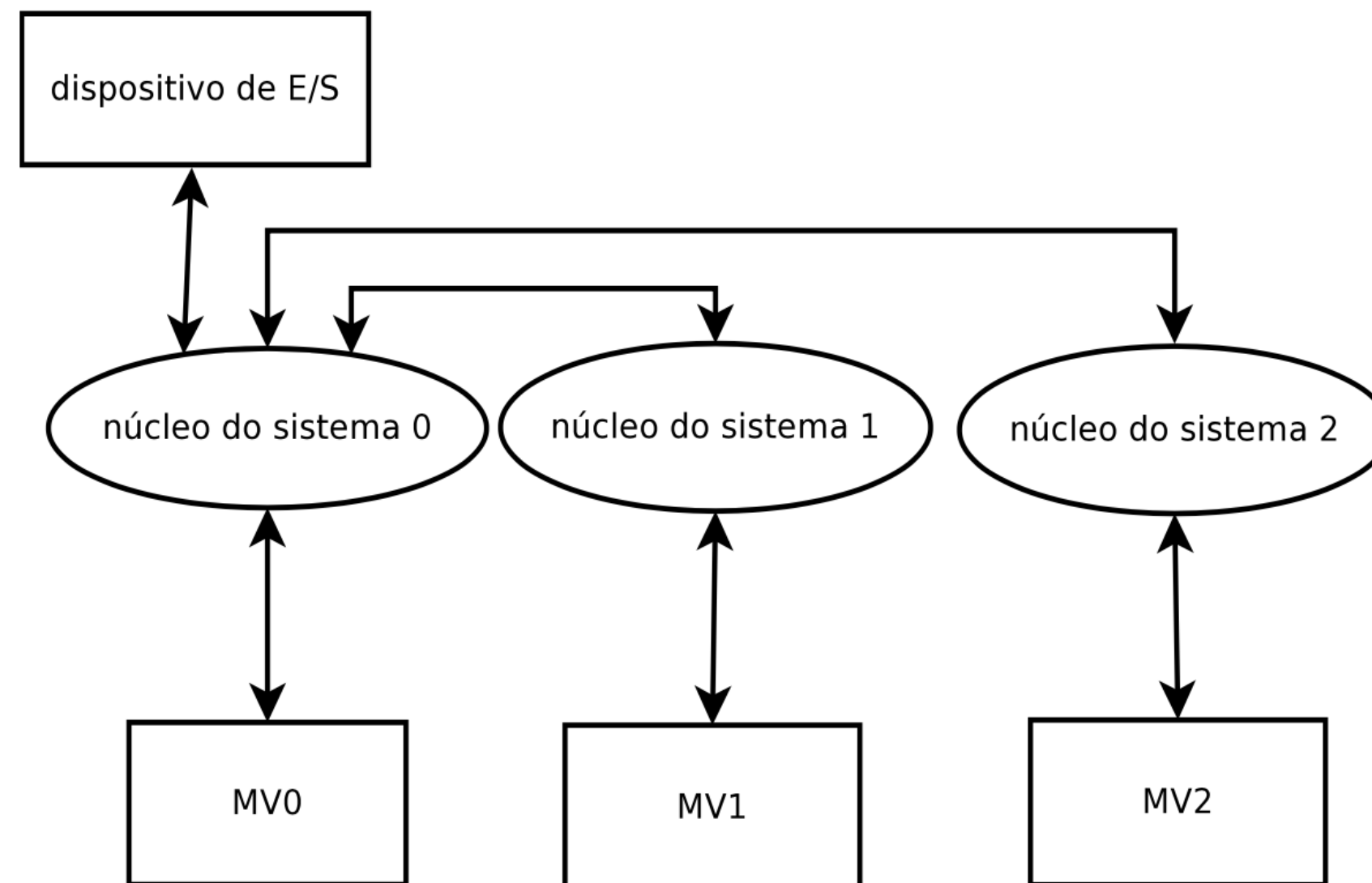
- Com uma máquina gerenciando na virtualização em nível de SO



Introdução

Virtualização de E/S:
Interfaces de rede

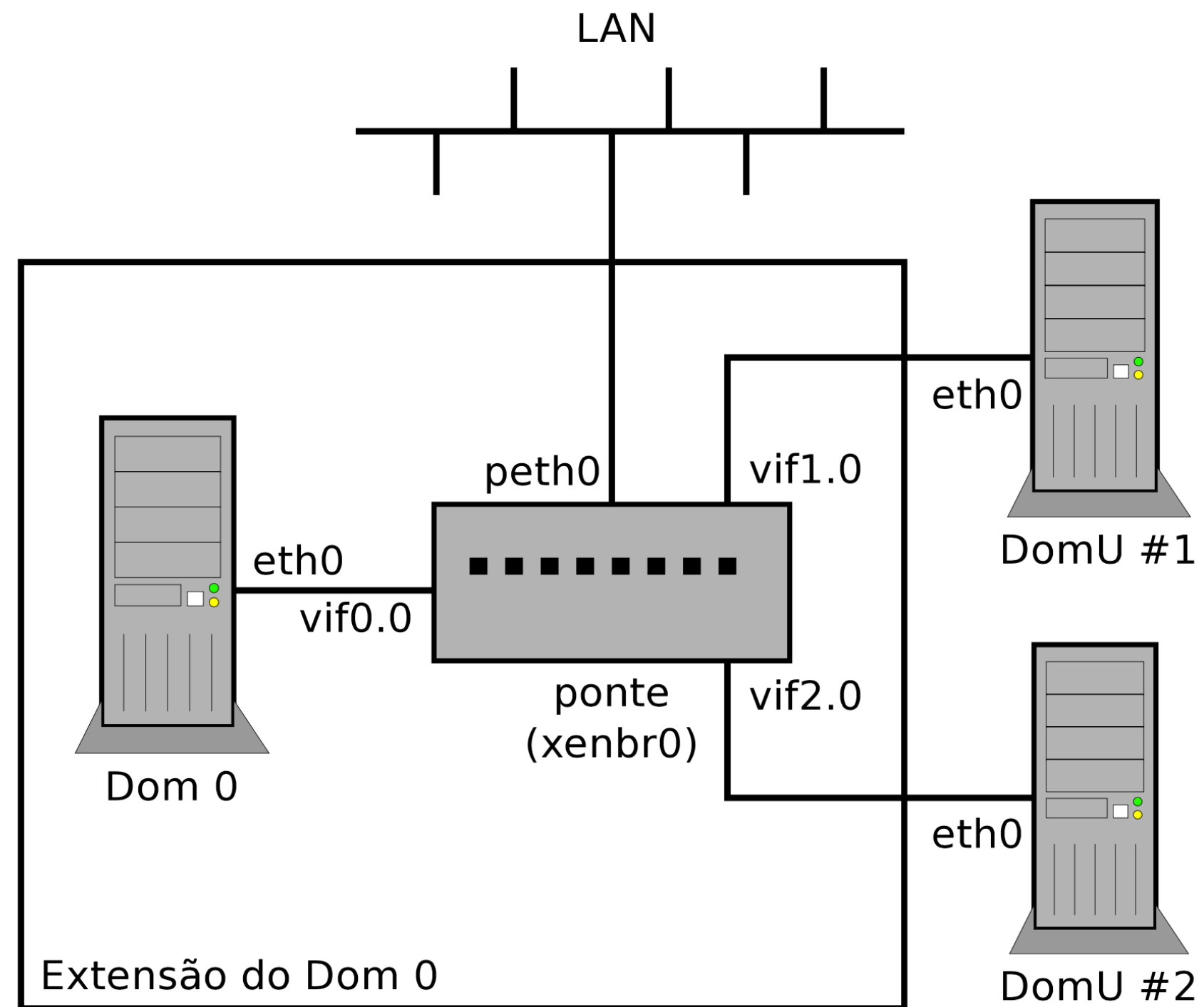
- Com uma máquina gerenciando na virtualização com hipervisor



Introdução

Virtualização de E/S:
Interfaces de rede

□ Exemplo com o Xen



Interrupções de hardware e software

Virtualização de E/S:
Interfaces de rede

- Interrupção é um sinal para o processador avisando que um evento precisa de imediata atenção (Normalmente tarefas de urgência ou de tempo-real)
- Tanto dispositivos de E/S (Hardware) como o núcleo do sistema (Software) podem gerar uma interrupção no sistema operacional (no Linux, o processo do kernel `ksoftirqd` cuida das interrupções de software)
- Hardware
 - não reentrantes (não cede tempo para outros processos)
 - não preemptivas (não pode ser interrompido)
- Software
 - reentrantes
 - não preemptivas

Agregação de interrupções

Virtualização de E/S:
Interfaces de rede

- Em dispositivos de rede, interrupções são geradas a cada pacote recebido
 - pacotes em excesso podem gerar perda de eficiência (e isso pode ser comum de ocorrer com placas em ambientes virtualizados)
- Opção 1: agregação de interrupções por intervalo de tempo
- Opção 2: agregação de interrupções por quantidade de pacotes
- Mas como acertar o valor de tempo ou quantidade de pacotes ideal?
- `ethtool -c <interface>`
 - `tx-frame, rx-frame, tx-usecs, rx-usecs`

- ❑ NAPI (New API) surgiu para melhorar a agregação de interrupções no Linux
- ❑ Conjunto de interfaces oferecido pelo Linux que os drivers dos dispositivos de rede implementam
- ❑ Agrega interrupções se adaptando ao tráfego
 - começa usando interrupções tradicionais e passa para um modo de polling quando o fluxo de pacotes excede um valor
 - enquanto estiver recebendo muitos pacotes, não precisa gerar interrupção de hardware

- Prós
 - quantidade reduzida de interrupções de hardware geradas pelo dispositivo
 - processador fica menos ocupado tratando uma interrupção de hardware
 - adaptação ao tráfego (mais tráfego → menos interrupções)
- Contras
 - existe uma carga adicional com NAPI, já que não é gerada apenas uma interrupção de hardware, como também é gerada uma interrupção de software
 - o driver precisa implementar NAPI
- Como o valor de limite da NAPI influencia no uso de CPU e na largura de banda em dispositivos virtuais?

NAPI - valor limite

Virtualização de E/S:
Interfaces de rede

- ☐ O limite define a quantidade limite de pacotes que a tarefa de recepção poderá coletar por ciclo de varredura
- ☐ Se não alcançar o limite, volta às interrupções de hardware
- ☐ Valores pequenos de limite farão com que menos pacotes sejam coletados por ciclo de varredura, porém, mais vezes a tarefa de recepção é recolocada na fila de varredura
- ☐ Valores pequenos de limite numa rede com alta taxa de pacotes, tende a consumir muita CPU mesmo que reduza as interrupções de hardware

Experimentos

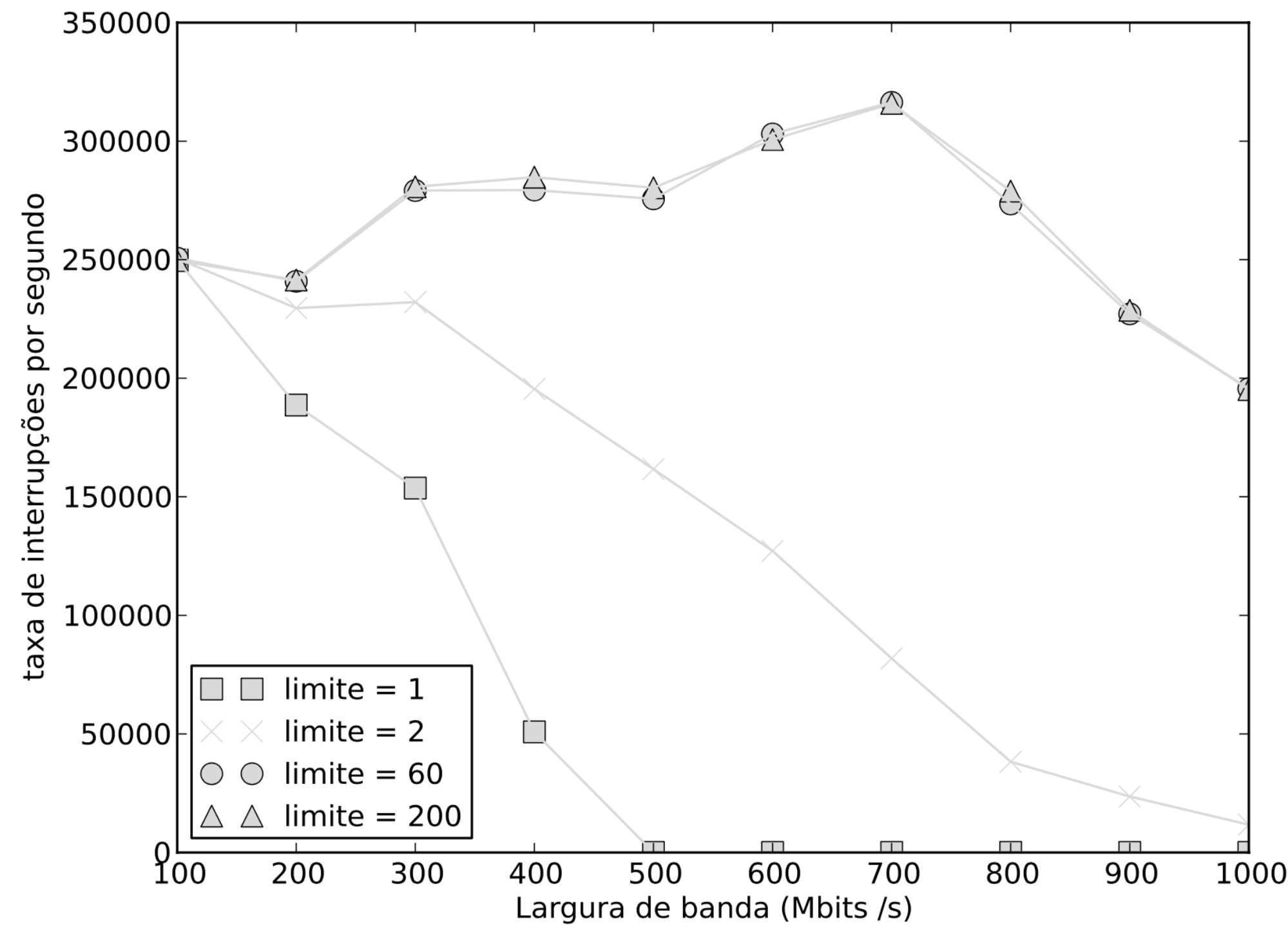
Virtualização de E/S:
Interfaces de rede

- ☐ VirtualBox
- ☐ Rede de 1Gbps
- ☐ Máquina física: 2 núcleos, 4 fluxos de execução
- ☐ VM: 2 fluxos de execução

NAPI - valor limite X interrupções

Virtualização de E/S:
Interfaces de rede

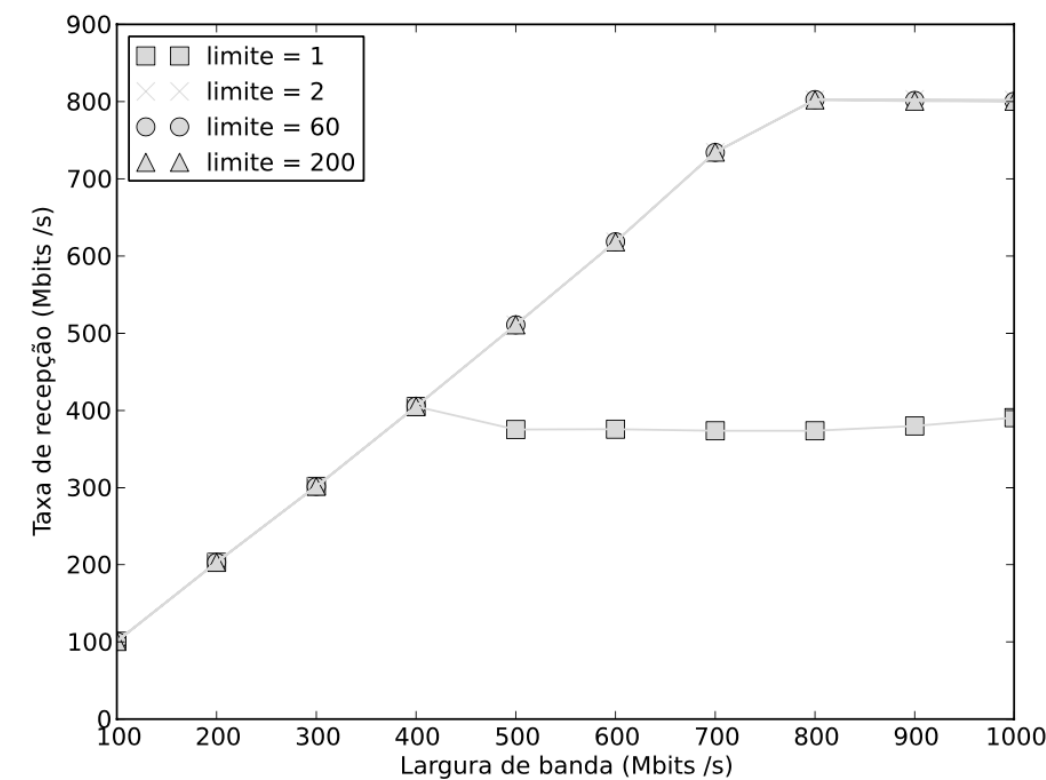
quantidade de interrupções de hardware



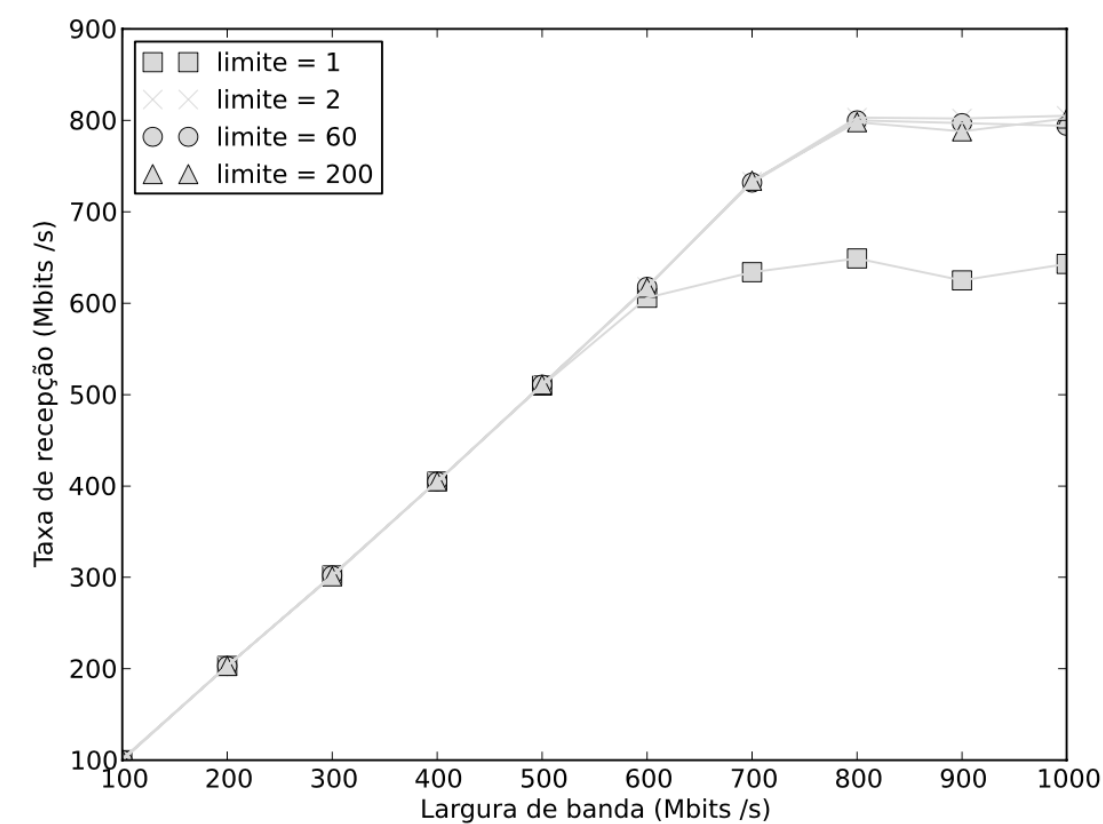
NAPI - valor limite X taxa de recepção X hipervisores

Virtualização de E/S:
Interfaces de rede

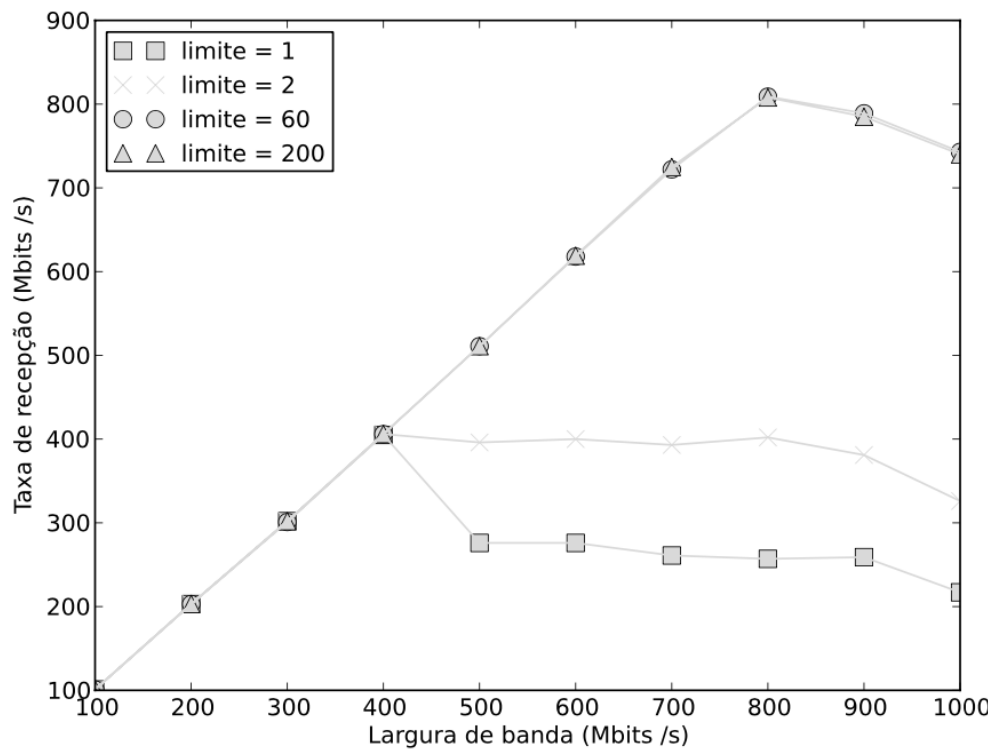
VirtualBox



Xen



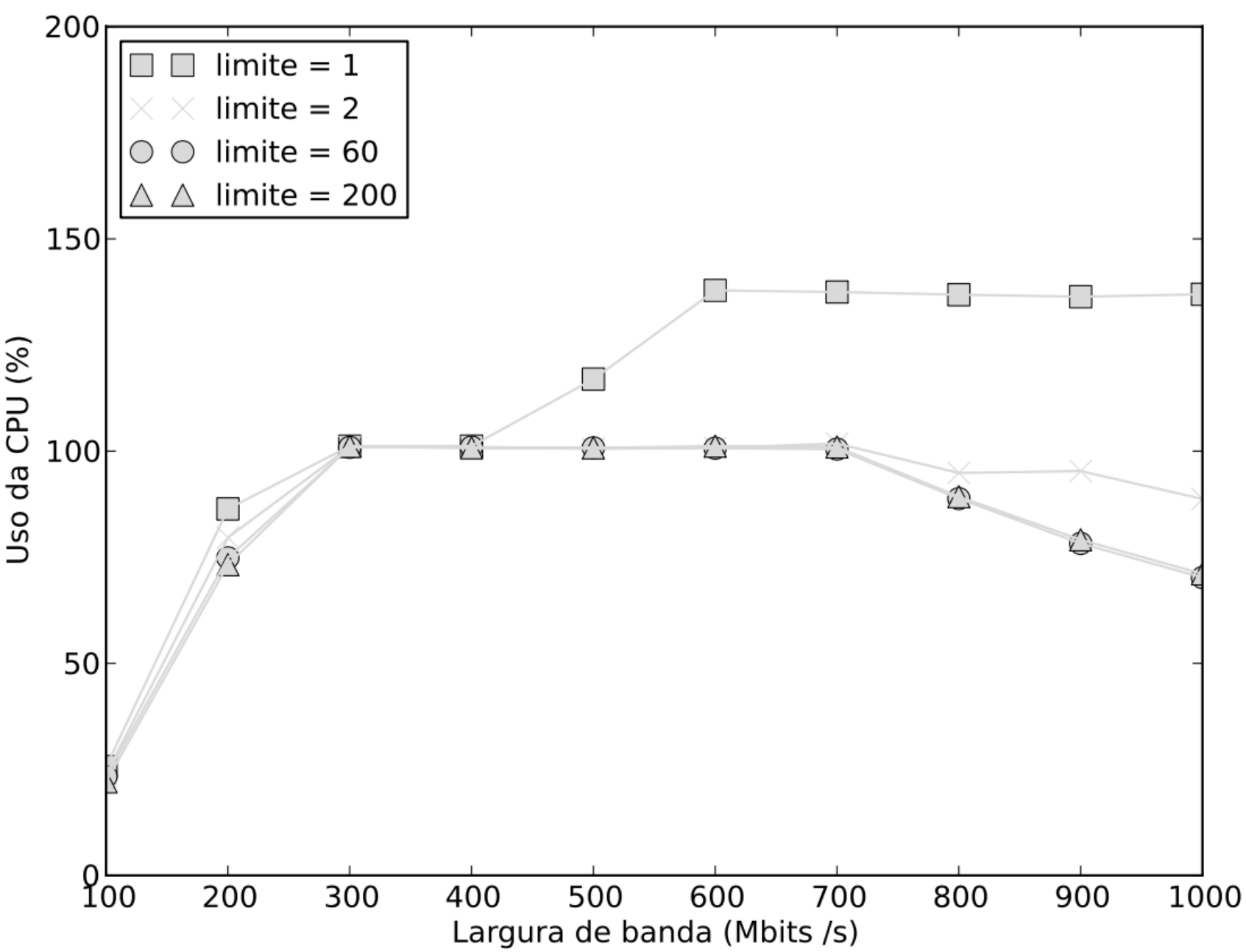
VMware



NAPI - valor limite X uso de CPU

Virtualização de E/S:
Interfaces de rede

uso da CPU na máquina virtual



Conclusões

Virtualização de E/S:
Interfaces de rede

- ☐ Valores altos de limite deram bons resultados
- ☐ O driver e1000, presente em boa parte dos hipervisores, já vem com o limite padrão igual a 64 então não é necessário modificar
- ☐ Separar a CPU usada pela aplicação, da CPU usada pelas interrupções de software pode garantir um desempenho melhor de uso da largura de banda

Eduardo Hideo Kuroda. Análise de Desempenho de Interfaces de Rede Virtualizadas com NAPI. 2013. Dissertação (Mestrado em Ciência da Computação) - Universidade de São Paulo, Orientador: Daniel Macêdo Batista.

dx.doi.org/10.11606/D.45.2016.tde-02022016-155854