

Universidade Federal de Campina Grande
Programa de Pós-Graduação em Ciência da Computação

Avaliação do Impacto de Estratégias de Economia de Energia em Grades Computacionais Entre-Pares

Lesandro Ponciano (Aluno)
Francisco Brasileiro (Orientador)

Sumário

- Introdução
- Trabalhos Relacionados
- Estratégias de Economia de Energia
- Materiais e Métodos de Avaliação
- Apresentação e Análise dos Resultados
- Conclusão e Trabalhos Futuros

Contexto

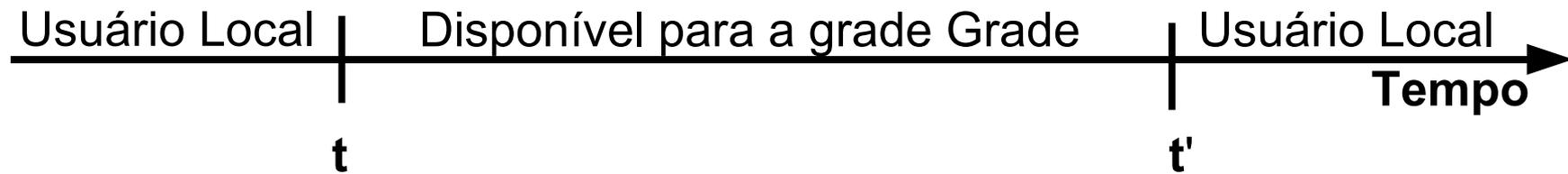
- Sistemas computacionais desenvolvidos visando obter maior poder computacional a qualquer custo
- Aumento no consumo de energia
 - Aumento no custo operacional da infraestrutura de TI nas organizações
 - Impactos ambientais
- Tratamos do consumo de energia no âmbito das Grades Computacionais Entre-Pares

Grades Computacionais Entre-Pares

- Recursos utilizados de forma oportunista
- Cada domínio administrativo pode agregar vários recursos
- Há um mecanismo que incentiva o compartilhamento de recursos entre domínios administrativo

Sessões de Disponibilidade

- Os recursos são disponibilizados à grade em **sessões de disponibilidade**



- O usuário local desabilita algumas configurações de energia definidas pelo Sistema Operacional

Recursos da Grade

- Oferta
- Demanda
- Contenção
 - Demanda $>$ Oferta = Alta Contenção
 - Oferta $>$ Demanda = Baixa Contenção

Descrição do Problema

- Em baixa contenção há recursos ociosos na grade
- Recursos ociosos apresentam entre 49–78% do consumo de energia que apresentariam se estivessem executando alguma tarefa
- Como reduzir o consumo de energia dos recursos mantidos ociosos nos períodos de baixa contenção?

Estados de Dormência

- Configuração Avançada e Interface de Energia (ACPI*)
- Estados de dormência definidos em ACPI*
 - **Sobreaviso** (S3) - suspensão para a RAM
 - **Hibernação** (S4) - suspensão para o disco
- Sobreaviso e Hibernação:
 - Permitem que a máquina seja acordada via *Wake-on-LAN*
 - Mantêm o estado da memória

Sobreaviso e Hibernação

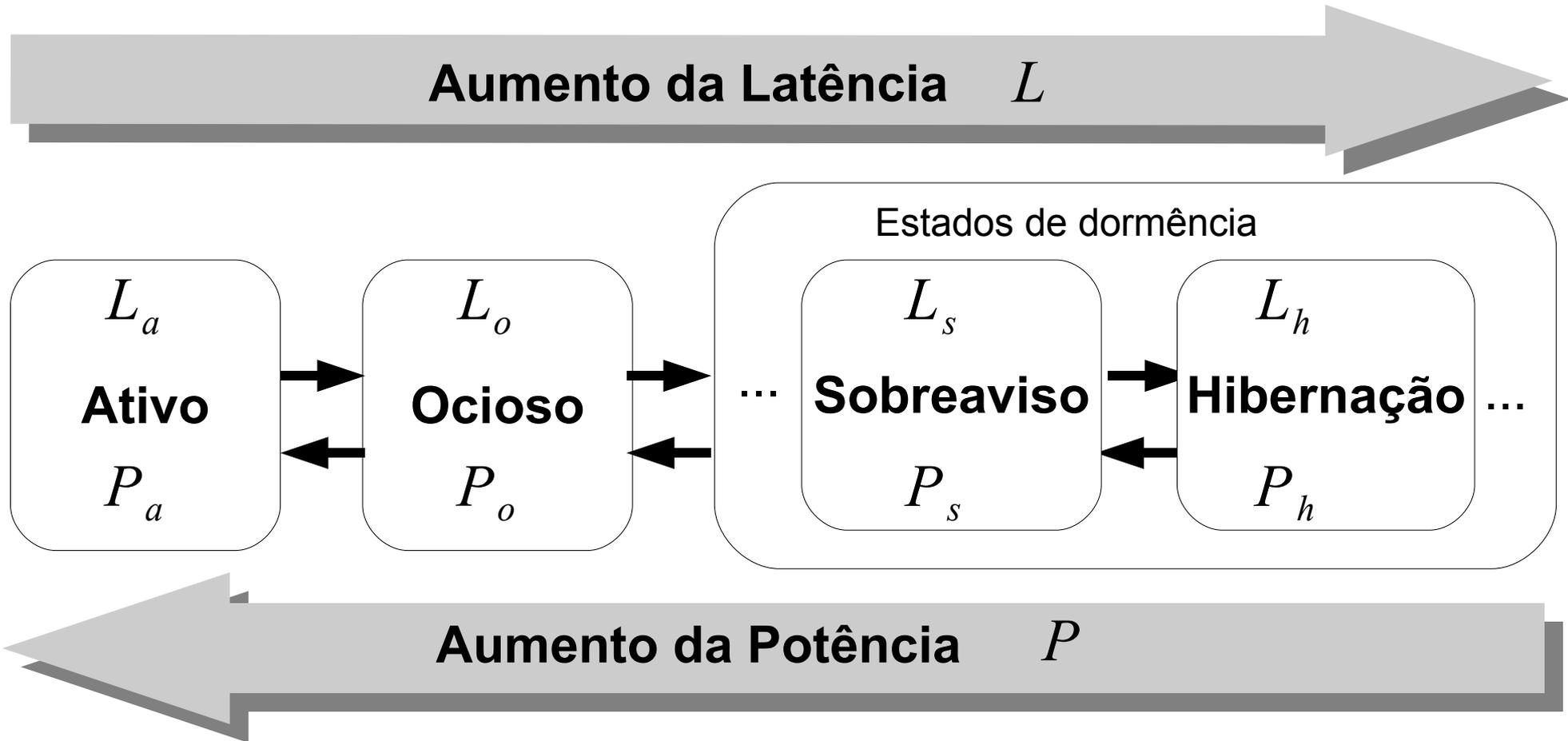
- **Benefício**

- Podem reduzir o consumo de energia dado que apresentam menor potência que o estado ocioso

- **Custos**

- Requerem um tempo para que o recurso seja acordado (latência) e durante esse tempo o recurso opera na potência do estado ativo
- Realizam partidas e paradas das rotações do disco rígido, o que reduz sua vida útil quando realizado em excesso

Sobreaviso vs Hibernação



Trabalhos Relacionados

- **Gerência da Energia**

- **Economia de energia em universidades** [Talebi et al. (2009); Universidade de Indiana (2009); *Energy Star* (2010)]
- ***Proxy-sleeping*** [Reich et al. 2010; Das et al. 2010; Agarwal et al. 2010]
- **Condor 7.2** [Condor 2010]

- **Escalonamento ciente do consumo de energia**

- **Ajuste dinâmico da voltagem e da frequência da CPU (DVFS)** [Sharma e Aggarwal (2009); Lammie et al. (2009)]
- **Ambientes com reserva de recursos e aplicações com restrição de prazo** [Sharma et al.(2009); Garg et al. (2009)]

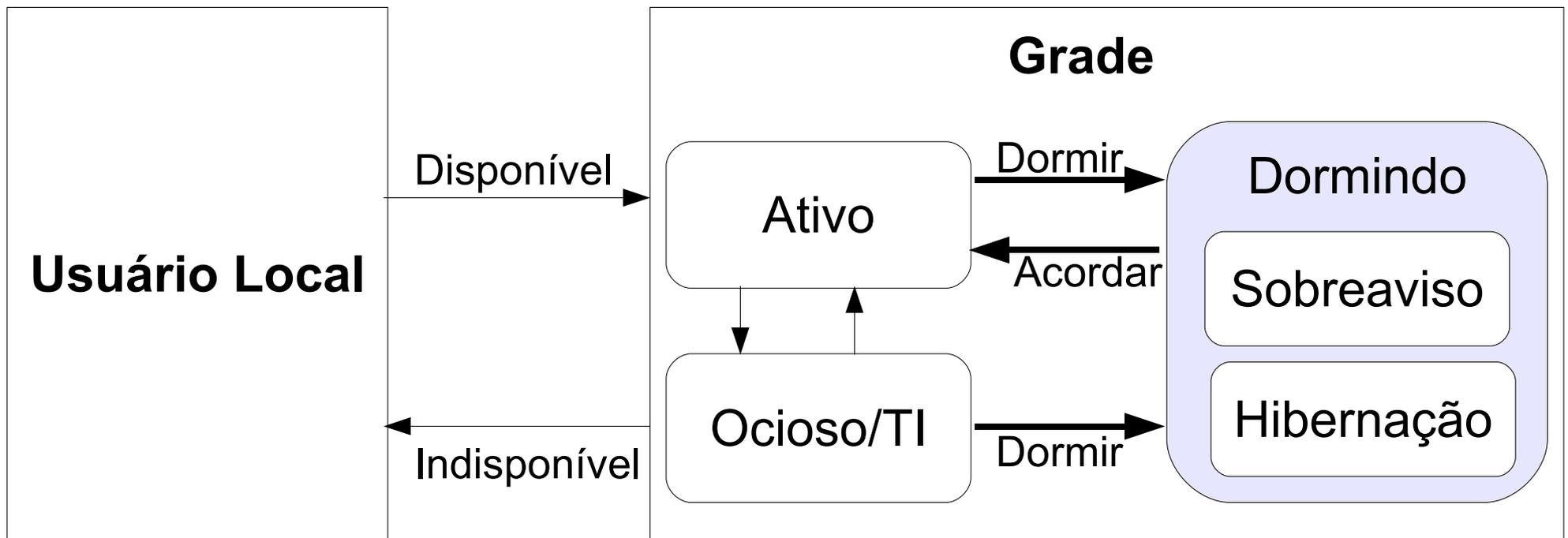
Economia de Energia em Grades Computacionais Entre-Pares



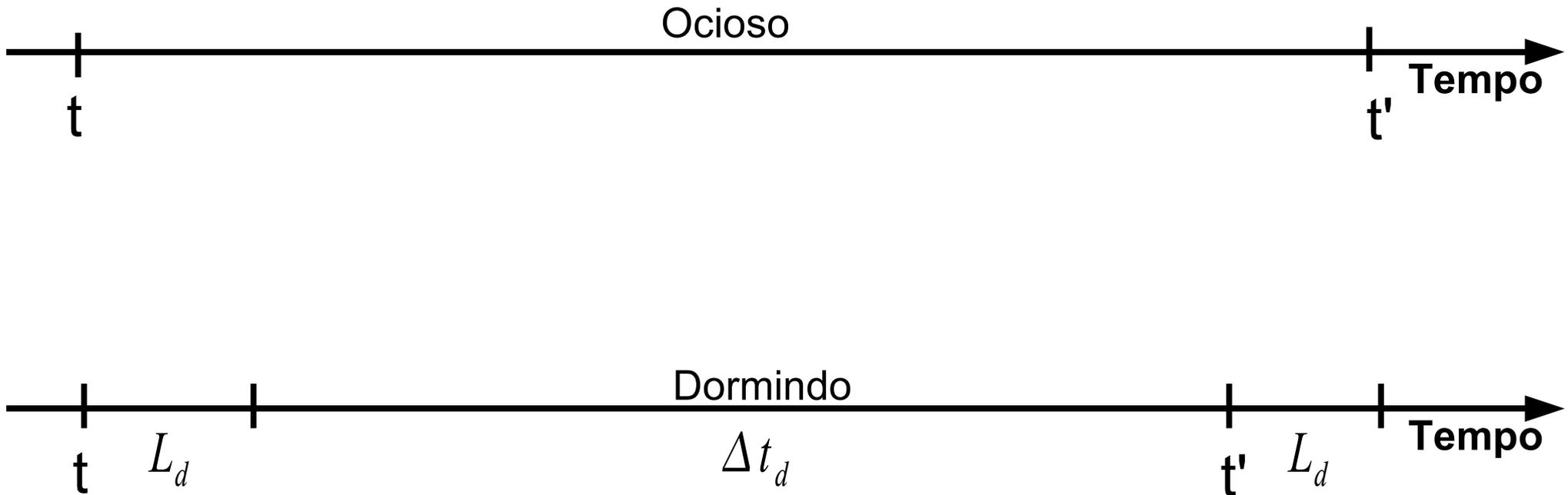
1 - Estados de dormência

Que estado de dormência é mais adequado ao ambiente de grades entres-pares?

Estados de Dormiência

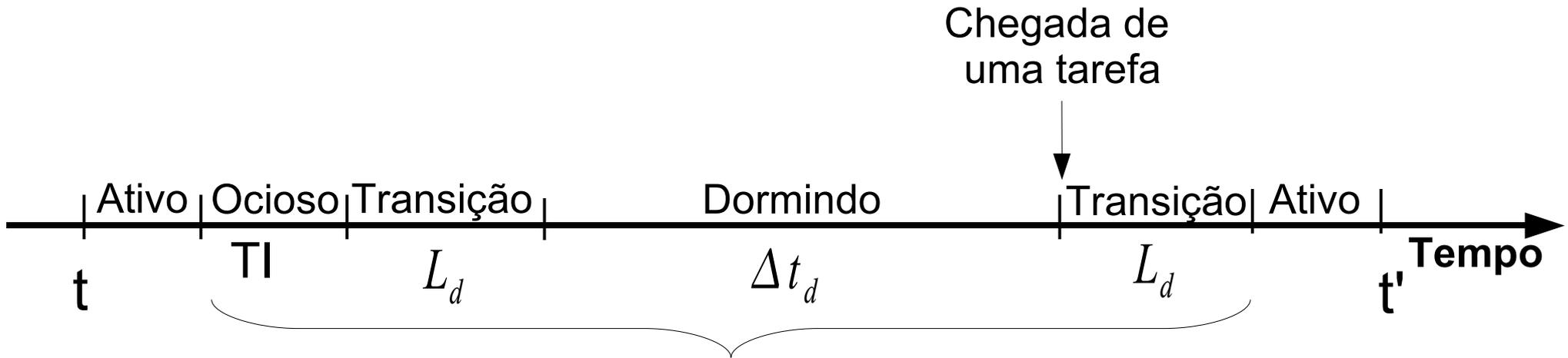


Sessão de dormência



O tamanho de uma **sessão de dormência** é limitado ao tamanho da sessão de disponibilidade

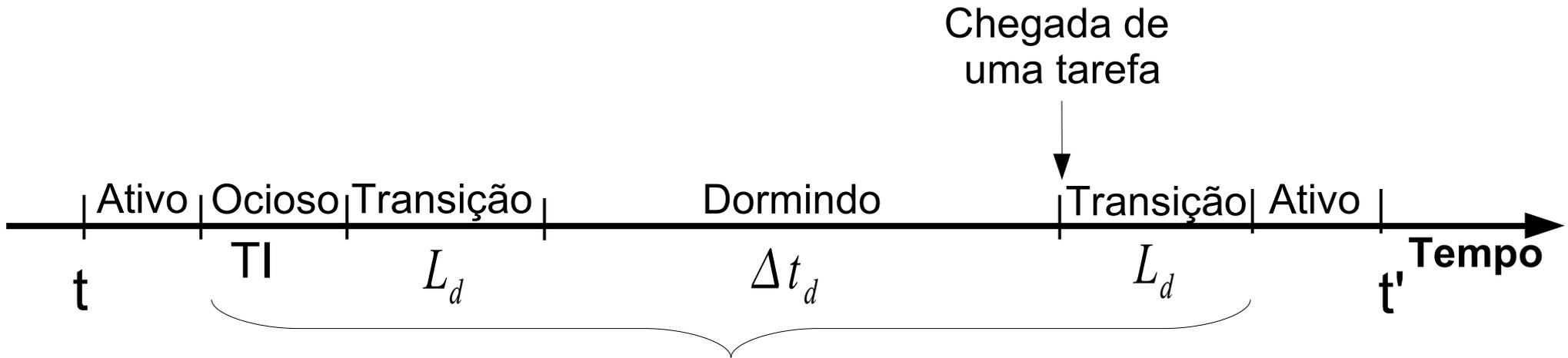
Estado de Dormência vs Estado Ocioso



- Consumo de energia com o uso do Estado de Dormência

$$c_d = 2 \times L_d \times P_a + \Delta t \times P_d + TI \times P_o$$

Estado de Dormência vs Estado Ocioso



- Consumo de energia com o uso do Estado de Dormência

$$c_d = 2 \times L_d \times P_a + \Delta t \times P_d + TI \times P_o$$

- Consumo de energia com o uso do Estado Ocioso

$$c_o = 2 \times L_d \times P_o + \Delta t \times P_o + TI \times P_o$$

Análise dos Estados

- Processador Intel E5200, 2,5 GHz, 2 GB de RAM e 160 GB de HD

$$P_a = 110 \text{ W}$$

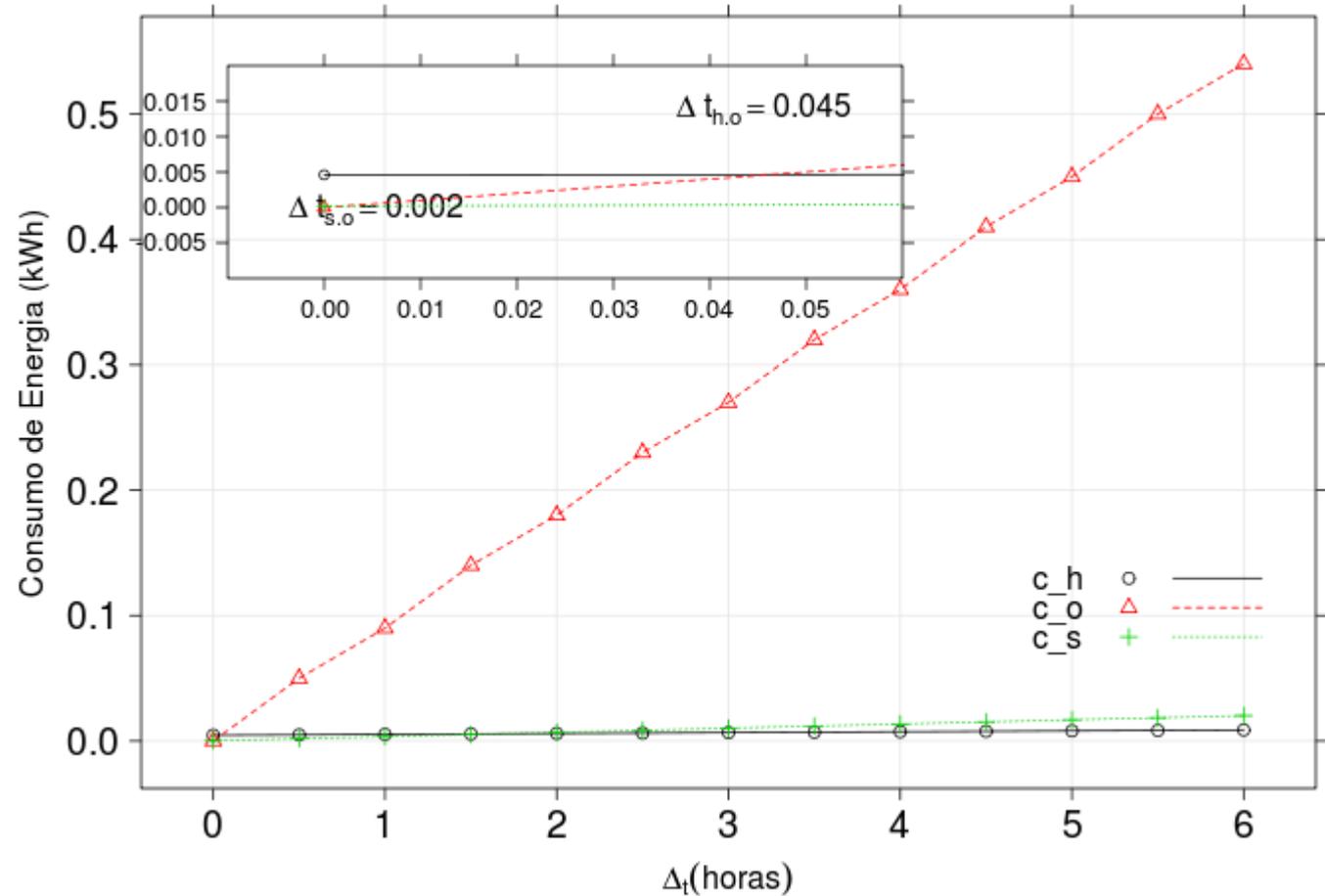
$$P_o = 70 \text{ W}$$

$$P_h = 3,33 \text{ W}$$

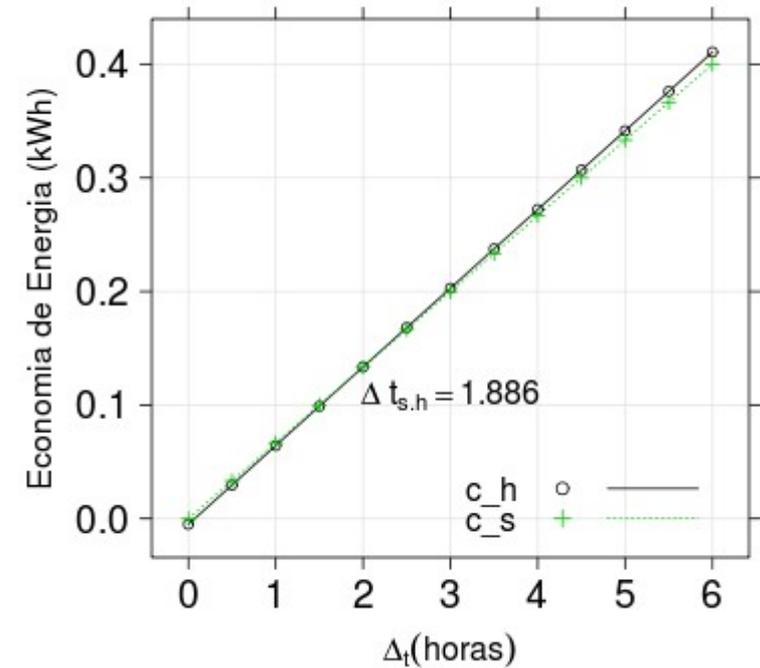
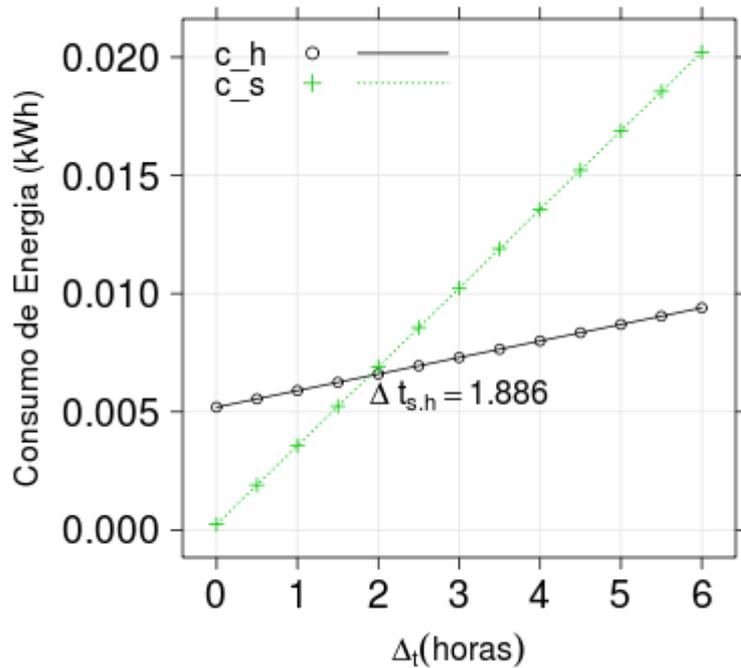
$$P_s = 0,7 \text{ W}$$

$$L_h = 55 \text{ s}$$

$$L_s = 2,5 \text{ s}$$



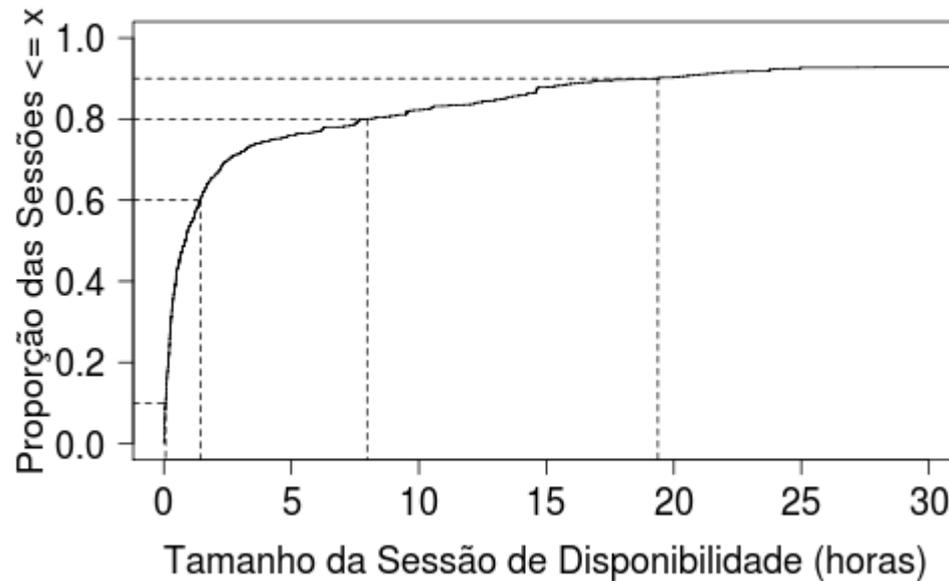
Análise dos Estados



- Economias de energia semelhantes
- Qual o tamanho típico das sessões de disponibilidade?

Sessões de Disponibilidade

- Tamanhos das sessões de disponibilidade de 140 máquinas da grade computacional ao longo de 6 dias

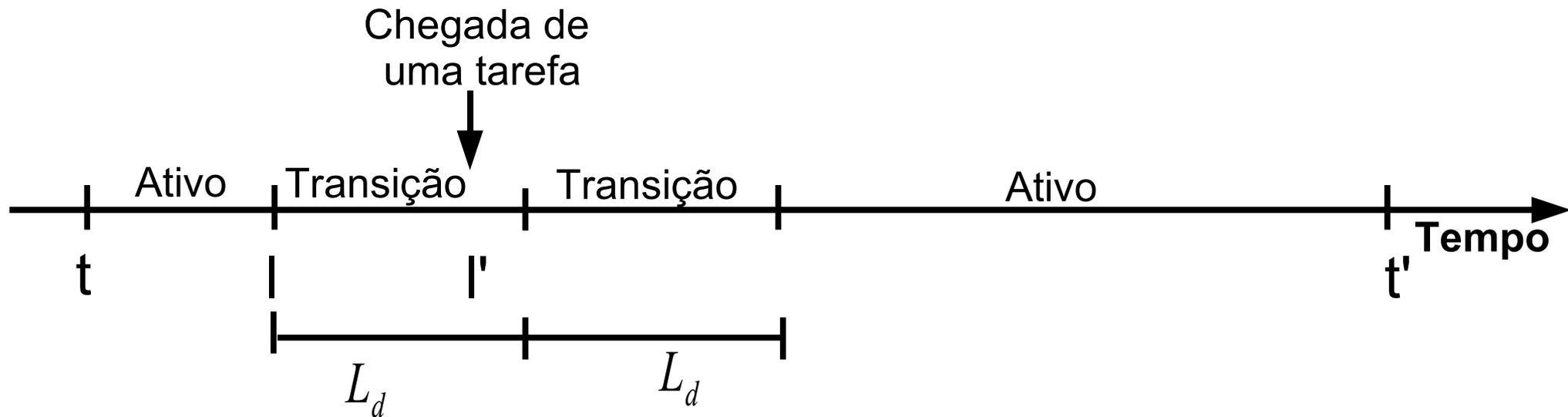


Dados da base de dados da grade Computacional OurGrid
<http://status.ourgrid.org/>

2 - Tempo de Inatividade (TI)

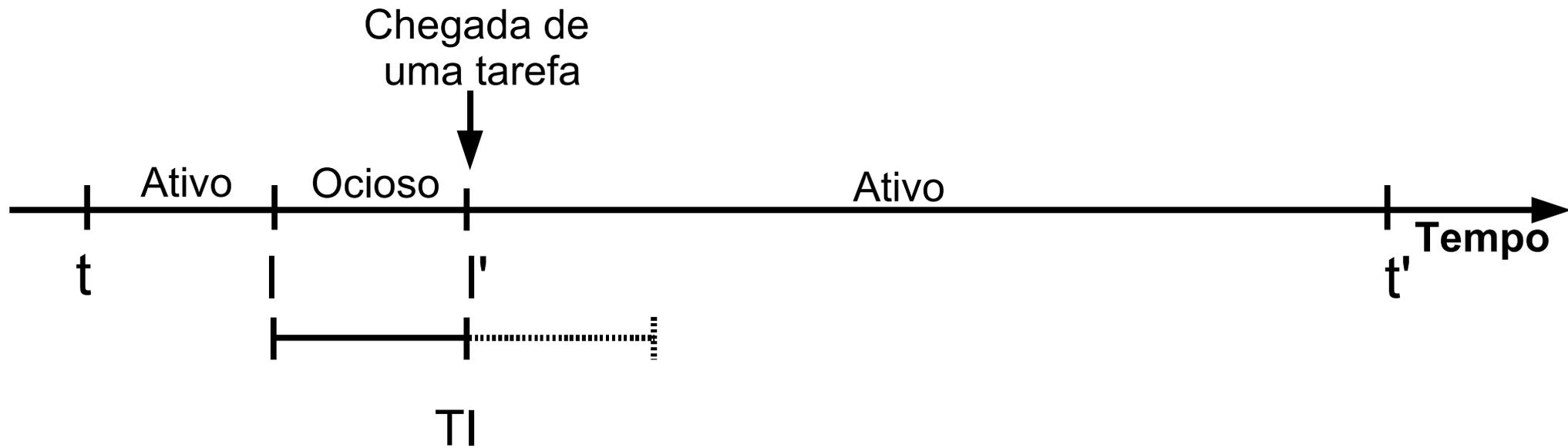
Após quanto tempo de inatividade uma máquina deverá ser adormecida?

Problema do não uso de TI

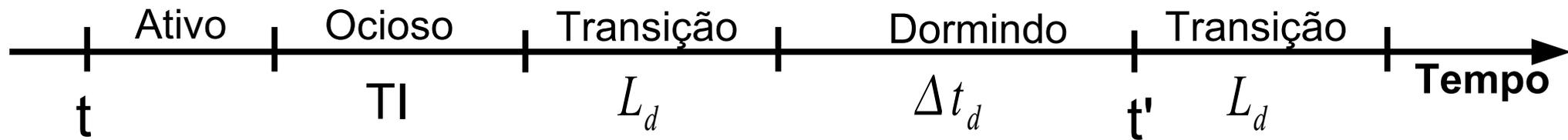


- Aumento:
 - do consumo de energia
 - do tempo de resposta
 - das partidas e paradas do disco rígido

Benefício do uso de TI



Problema do uso de TI



Qual valor de TI utilizar?

- Experiências de outros trabalhos:
 - 300 segundos [Reich et al. 2010]
 - 600 – 900 segundos [Windows 2010, Ubuntu 2010]
 - 1800 segundos [Reich et al. 2010]

3 - Escolha de Recursos

Como escolher quais recursos devem ser acordados quando surgir uma demanda menor que o número de recursos adormecidos?

Quem faz a escolha?

- A gerência da energia é realizada de modo autônomo por cada domínio administrativo da grade computacional entre-pares
- Em cada domínio administrativo dessas grades há um **gerente** que:
 - conhece o estado atual de cada recurso
 - escolhe quais recursos devem ser acordados quando surge uma nova demanda

Estratégias de Escolha de Recursos

- Primeiro os recursos mais recentemente adormecidos (**MRS**)
- Primeiro os recursos menos recentemente adormecidos (**LRS**)
- Primeiro os recursos mais eficientes no aspecto energético (razão entre a frequência da CPU e potência no estado Ativo) - (**EA**)
- Escolha os recursos de modo aleatório - **Aleatório**

Materiais e Métodos de Avaliação

Objetivos da Avaliação

- Estimar a economia de energia e o impacto no tempo de resposta gerados pelos:
 - Estados de dormência
 - Valores de TI
 - Estratégias de escolha

- Avaliar o impacto das estratégias na vida útil dos discos rígidos

Métricas da Avaliação

- Avalia-se **configurações da grade** formadas por um estado de dormência, um valor de TI e uma estratégia de escolha
- De uma configuração A em relação a uma configuração de referência \bar{A} , calcula-se:

- **Economia de Energia:**

$$\xi^A = \frac{E^{\bar{A}} - E^A}{E^{\bar{A}}} \times 100, \text{ onde } E^X \text{ é a energia consumida}$$

pela configuração X

- **Atraso no tempo de resposta:**

$$\beta^A = \frac{m^A - m^{\bar{A}}}{m^{\bar{A}}} \times 100, \text{ onde } m^X \text{ é o tempo de resposta da}$$

aplicação na configuração X

- **Número de transições**

Modelo e Configuração da Simulação

- Modelo de Simulação inspirado no *middleware* OurGrid
 - Cada domínio administrativo tem um gerente chamado *peer*
 - Os *peers* da grade utilizam a NoF como mecanismo de incentivo à colaboração
- Utiliza rastros que descrevem a variação na disponibilidade de recursos e na submissão de aplicações
 - Rastro de disponibilidade da grade OurGrid
 - Modelo de aplicações do tipo Saco-de-Tarefas (Iosup et al., 2008), limitando a tarefas de até 35 minutos

Configuração da Grade

- Potência
 - Ativo: CPU, HD, memória, placa mãe
 - Ocioso: CPU, HD, memória, placa mãe
 - Sobreaviso: 3,33 Watts
 - Hibernação: 0,7 Watts
- Latência
 - Ativo: 0
 - Ocioso: 0
 - Sobreaviso: 2,5 segundos
 - Hibernação: 55 segundos

Cenários Avaliados

- Análise da sensibilidade dos resultados às variações da potência e da latência entre -100% e 200%
 - Variações na Latência com Potência Fixa
 - Variações na Potência com Latência Fixa
 - Latência e Potência variam na mesma proporção
 - Latência e Potência variam em proporção inversa

Apresentação e Análise dos Resultados

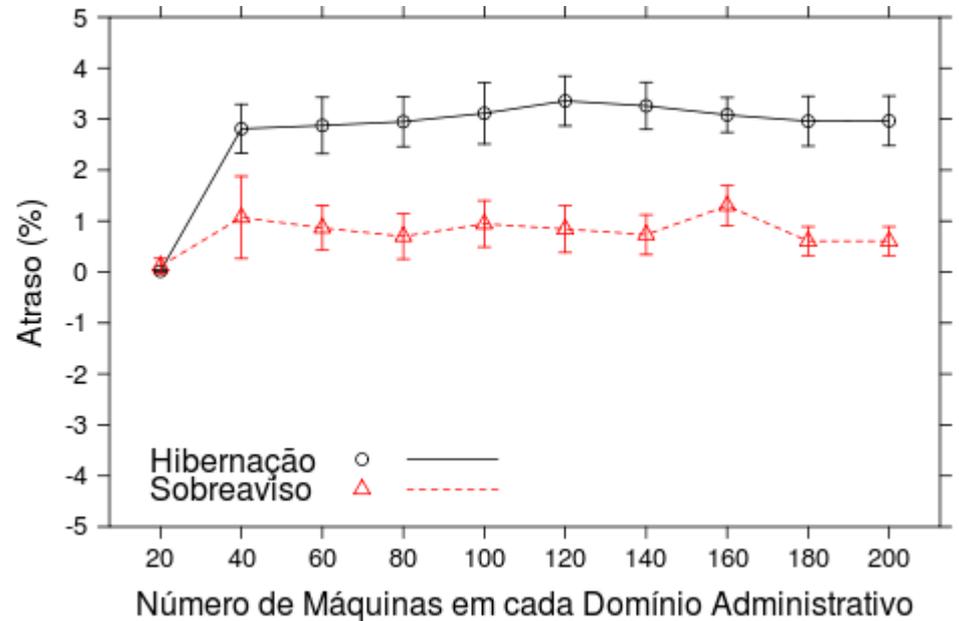
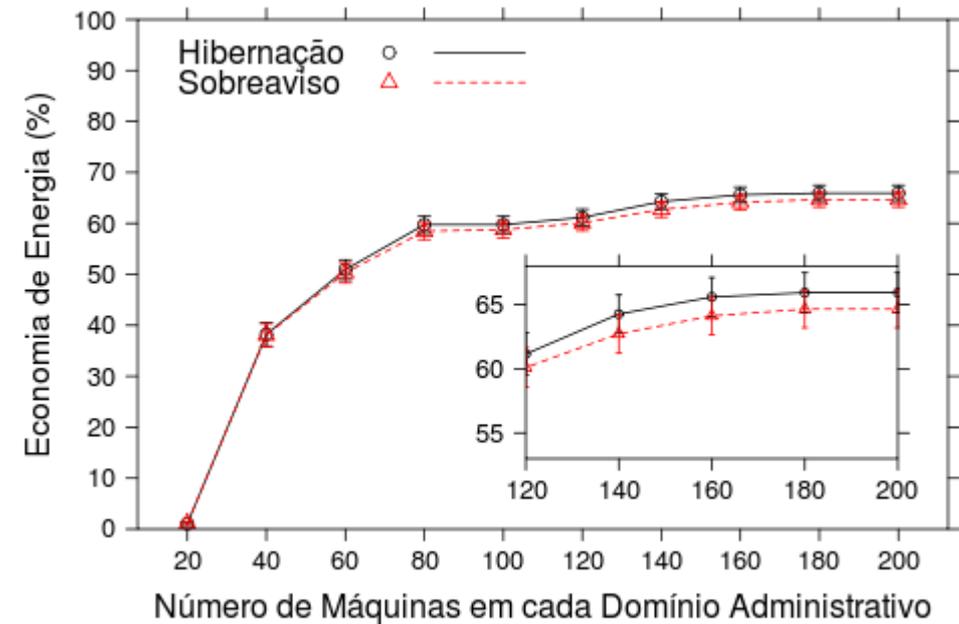
Estados de Dormiência

- **Configuração da grade**

- Escolha: MRS
- TI: 0

- **Resultado**

- Até 65,53% de economia
- Até 3,8% de atraso
- Sobreaviso gera menor atraso



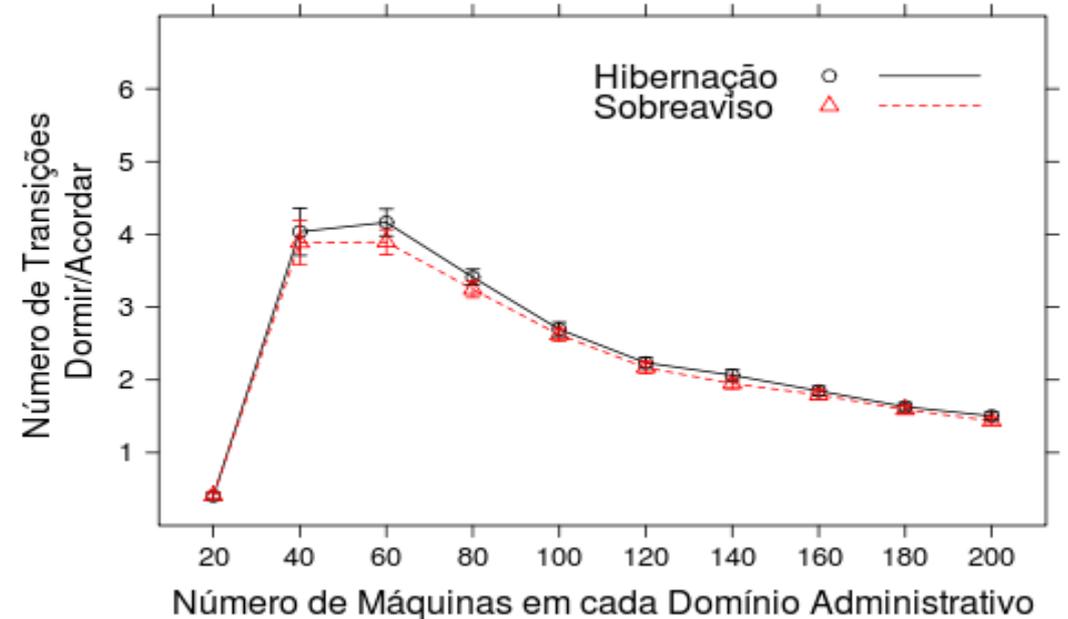
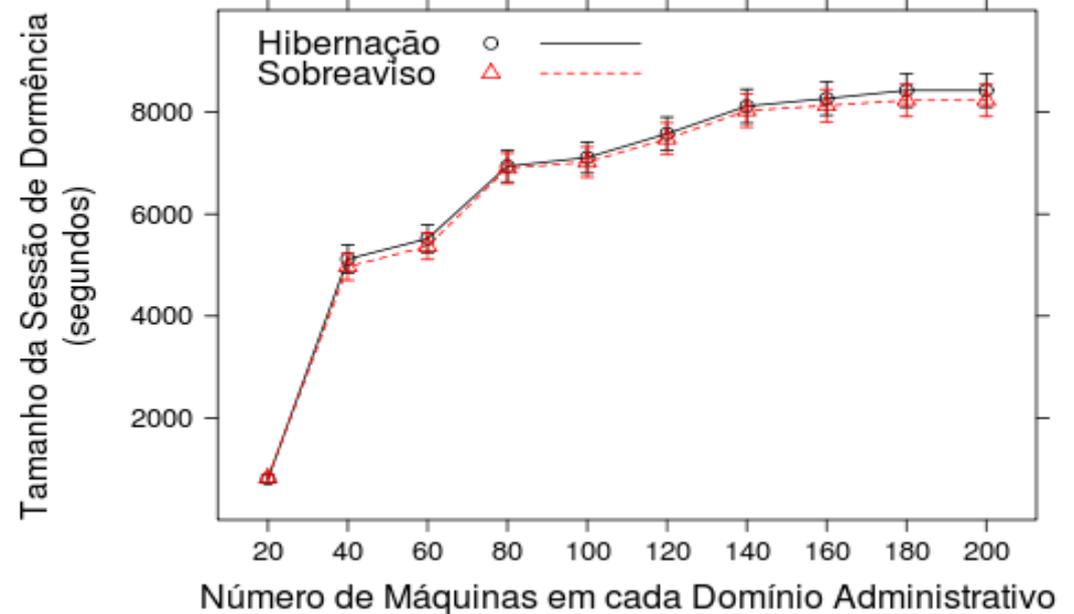
Estados de Dormência

- **Configuração da grade**

- Escolha: MRS
- TI: 0

- **Resultado**

- Sessões de dormência de quase 2 horas
- Mais transições em média contenção



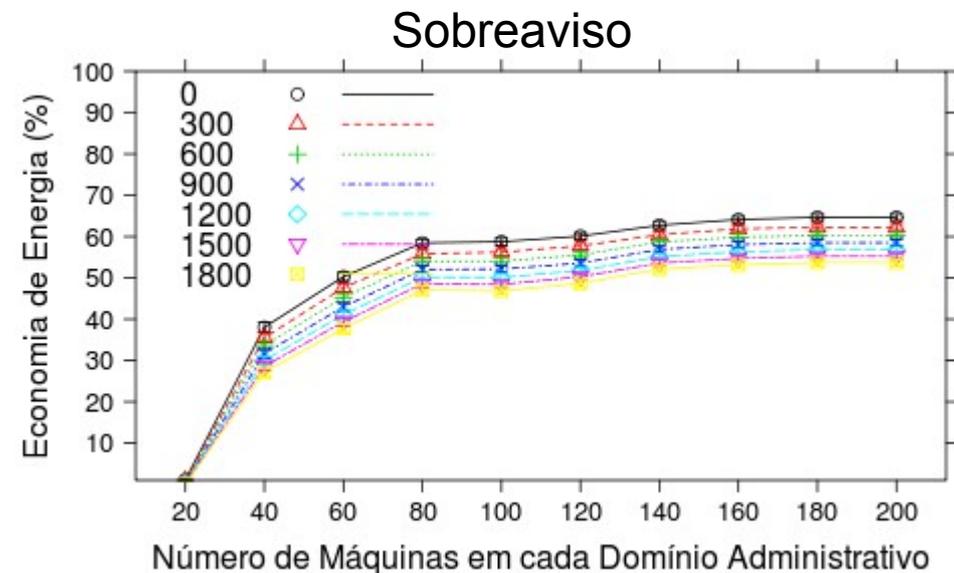
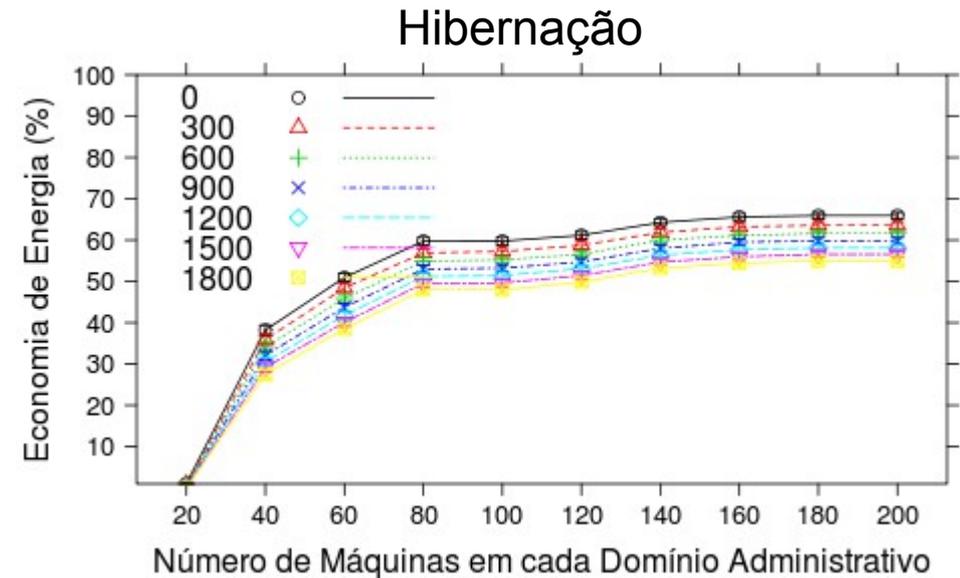
Tempo de Inatividade (TI)

- **Configuração da grade**

- Escolha: MRS

- **Resultado**

- Até 10% de redução da economia de energia com o aumento de TI de 0 para 1.800 segundos
- Até 2% de redução da economia com o aumento de TI de 0 para 300



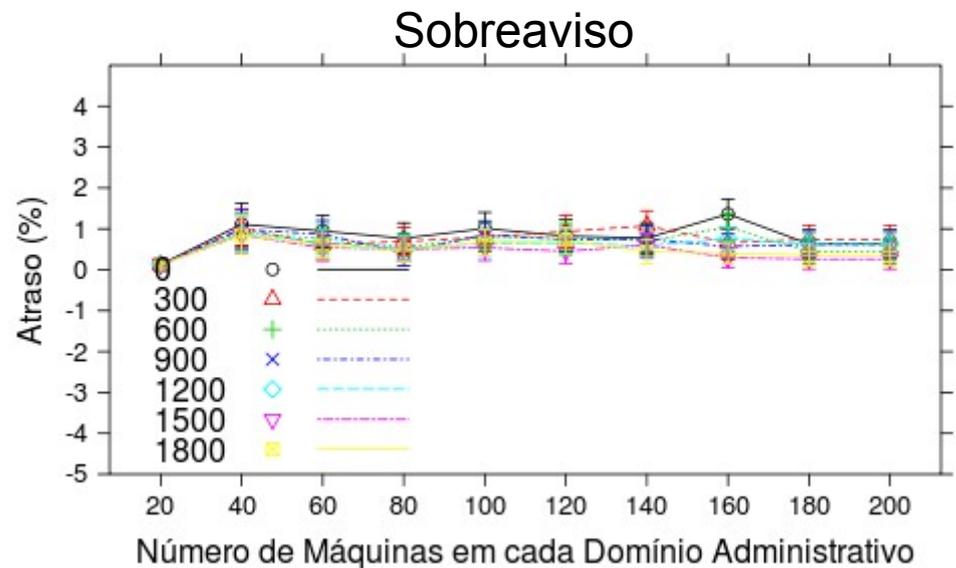
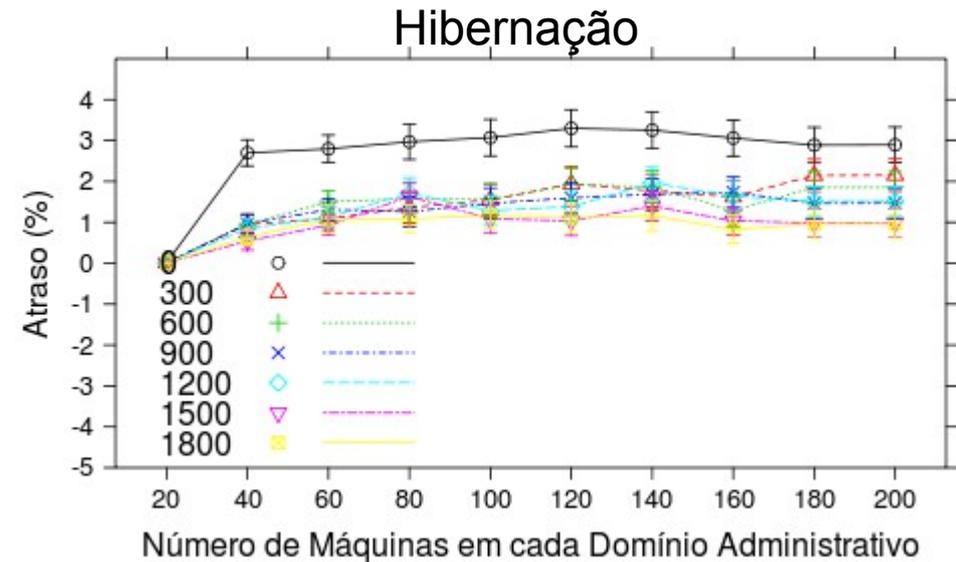
Tempo de Inatividade (TI)

- **Configuração da grade**

- Escolha: MRS

- **Resultado**

- Valores de TI maiores que 0 reduzem o atraso gerado pela estratégia Hibernação
- Variar TI não gera impacto significativo no atraso gerado pela estratégia Sobreaviso



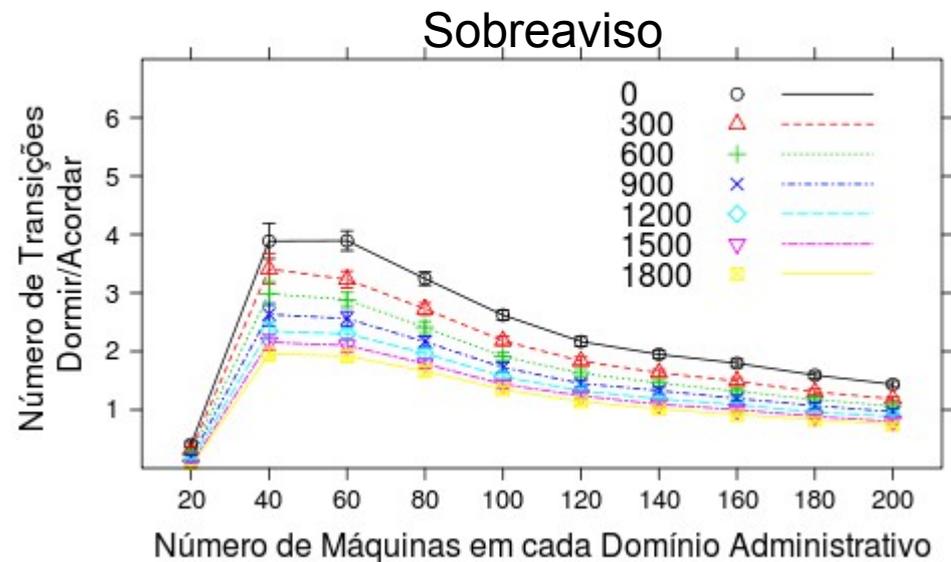
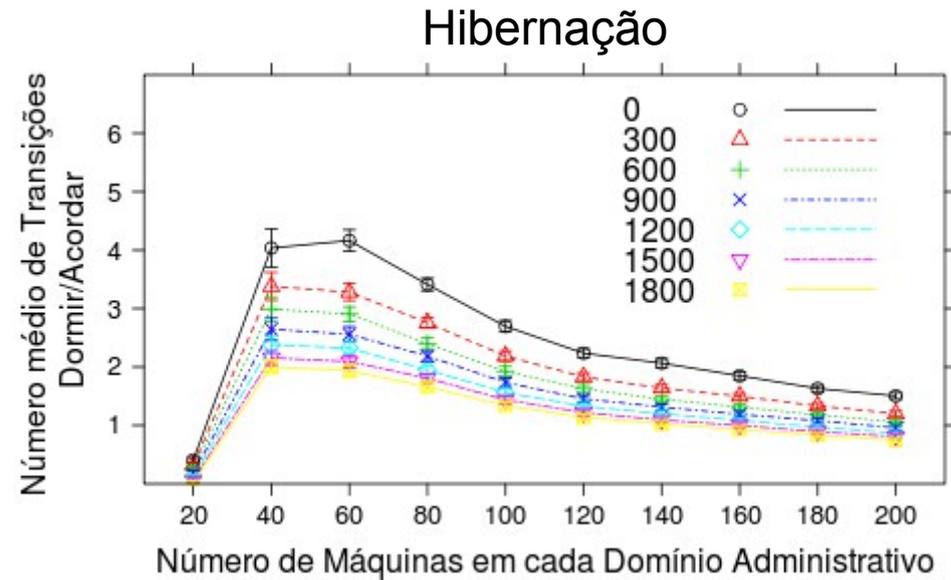
Tempo de Inatividade (TI)

- **Configuração da grade**

- Escolha: MRS

- **Resultado**

- Aumentar o valor de TI permite reduzir o número de transições realizadas



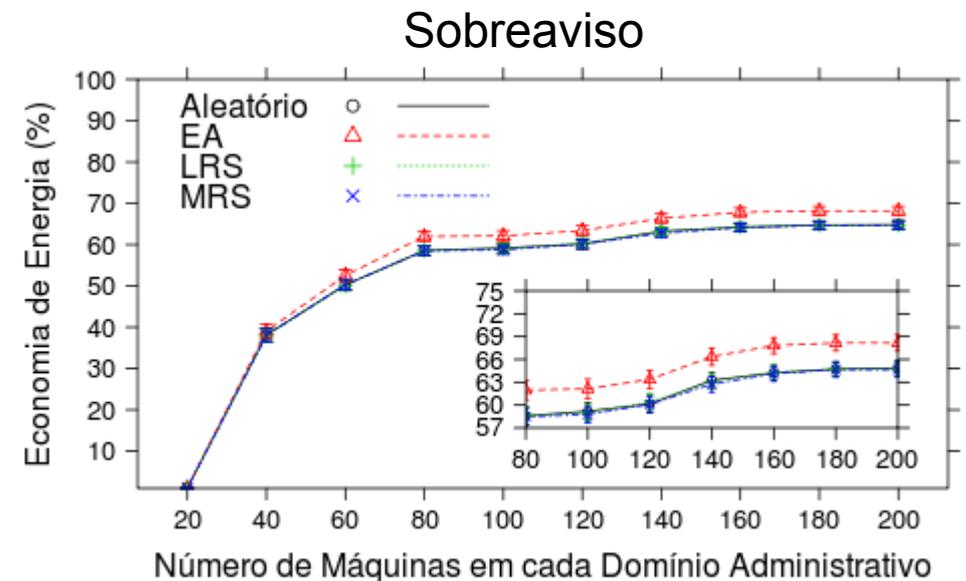
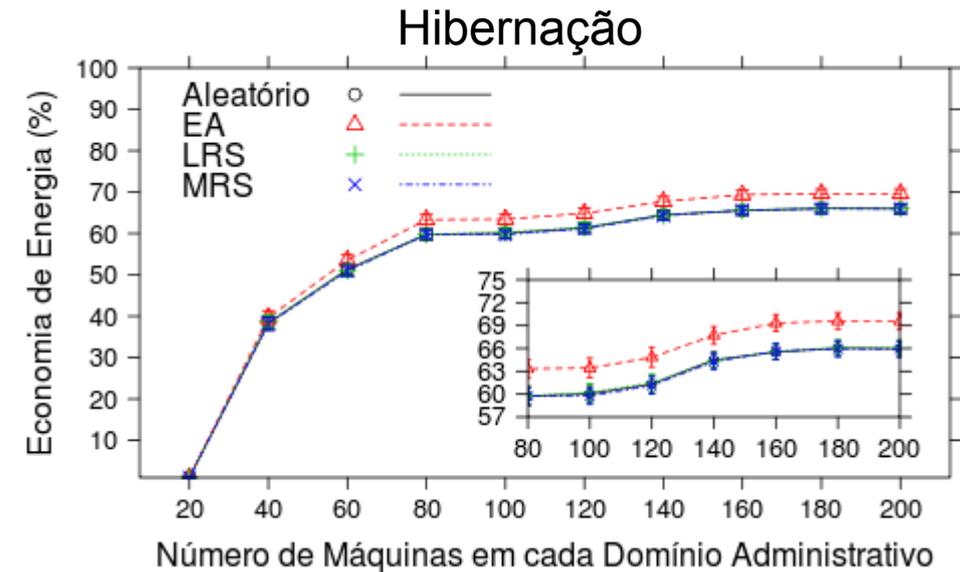
Estratégia de Escolha

- **Configuração da grade**

- TI: 0

- **Resultado**

- AE permite aumentar a economia de energia em até 3%
- Não há diferença significativa entre as demais estratégias



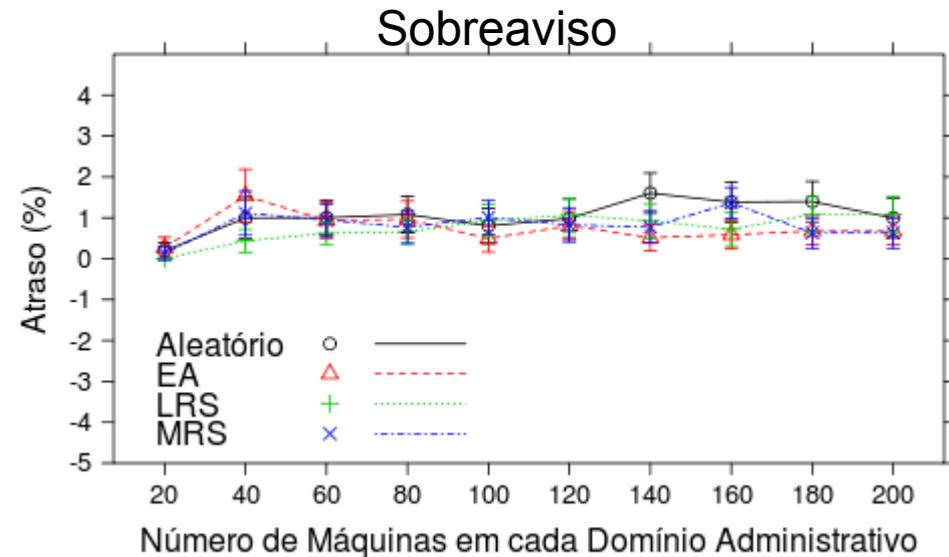
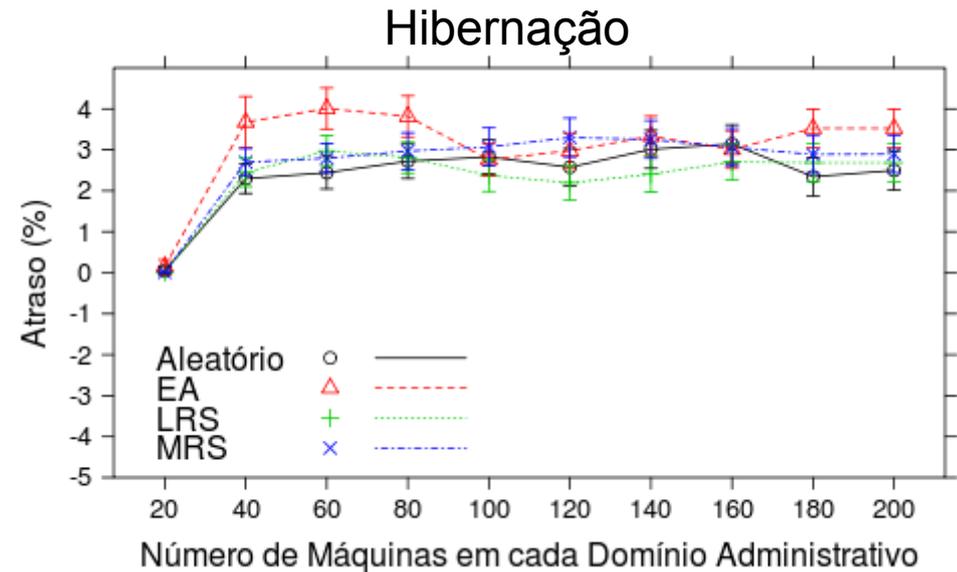
Estratégia de Escolha

- **Configuração da grade**

- TI: 0

- **Resultado**

- A estratégia EA combinada com a estratégia Hibernação gera maior atraso em média contenção



Estratégia de Escolha

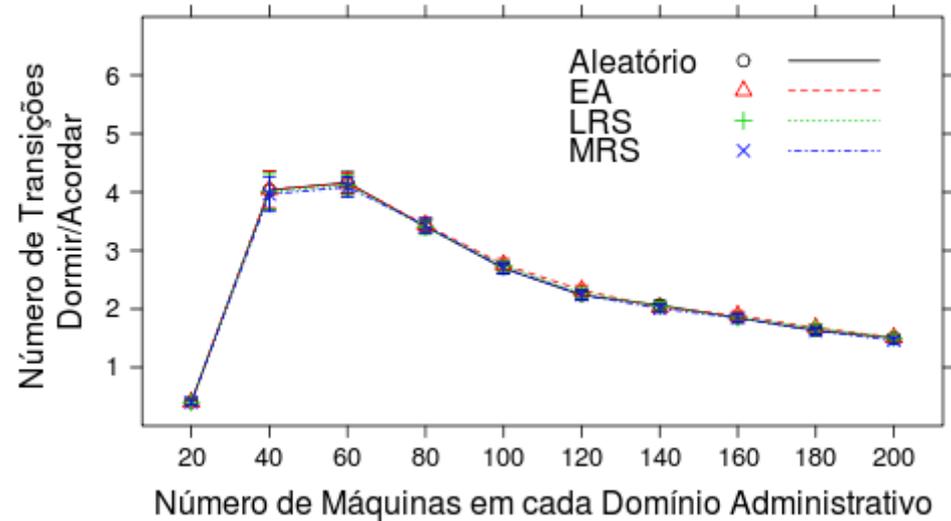
- **Configuração da grade**

- TI: 0

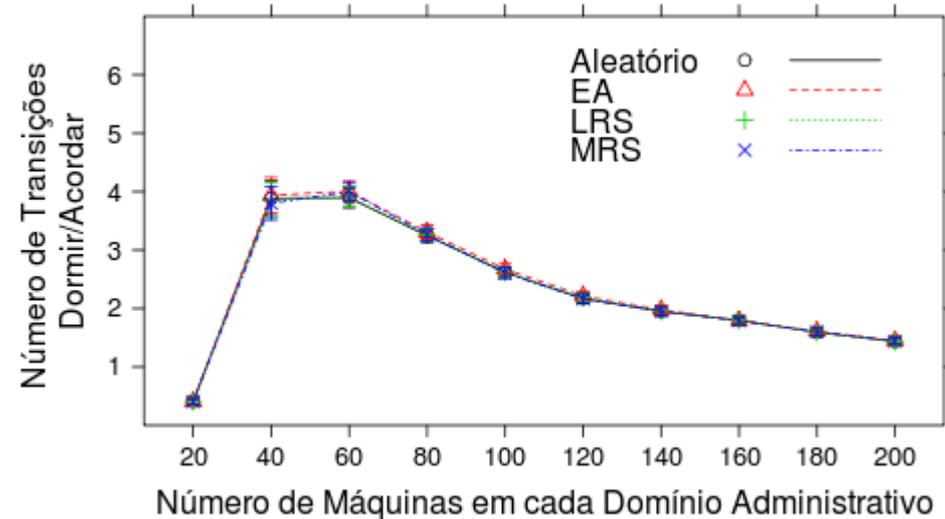
- **Resultado**

- Não há diferença entre o número de transições gerado pelas estratégias
- Mais transições em média contenção

Hibernação

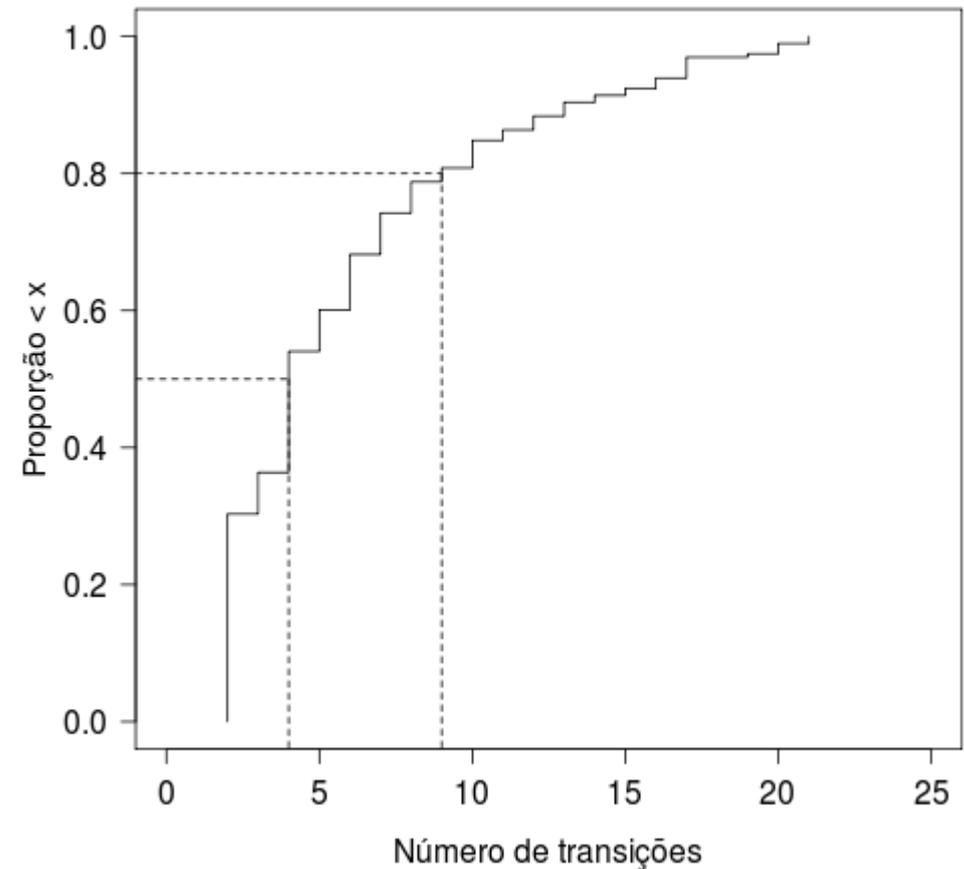


Sobreaviso



Análise do Número de transições

- Usuários que não faziam uso de uma estratégia de dormência
 - A grade aumenta o número de transições
- Usuários que faziam uso de uma estratégia de dormência
 - 4,98 transições por dia
 - Nos cenários avaliados, a grade reduz o número de transições



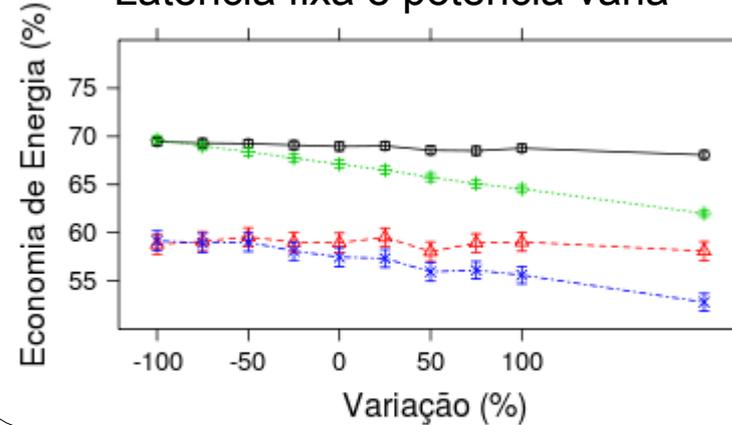
Dados de 60 máquinas da grade Computacional OurGrid

Análise de sensibilidade

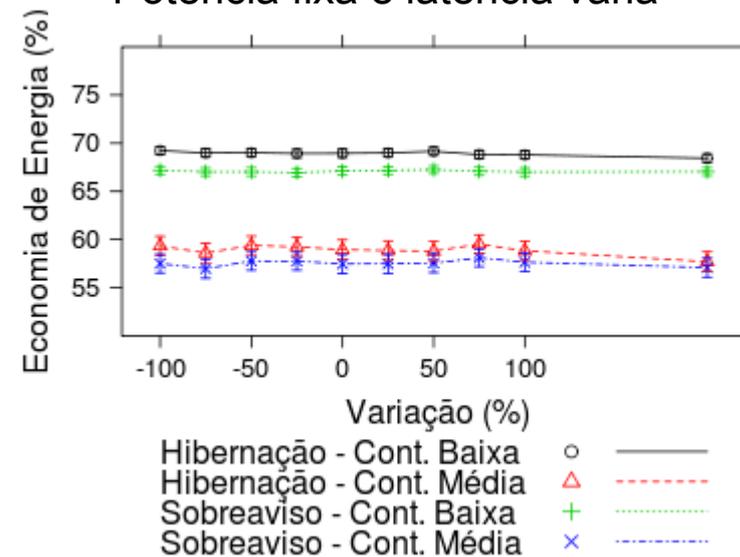
Geram um impacto de até 5%
na economia provida pela
estratégia Sobreaviso

Não impactam a economia de
energia

Latência fixa e potência varia



Potência fixa e latência varia

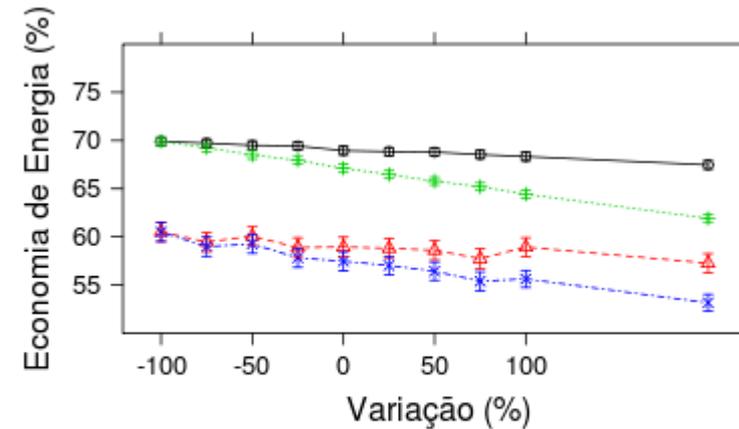


Análise de sensibilidade

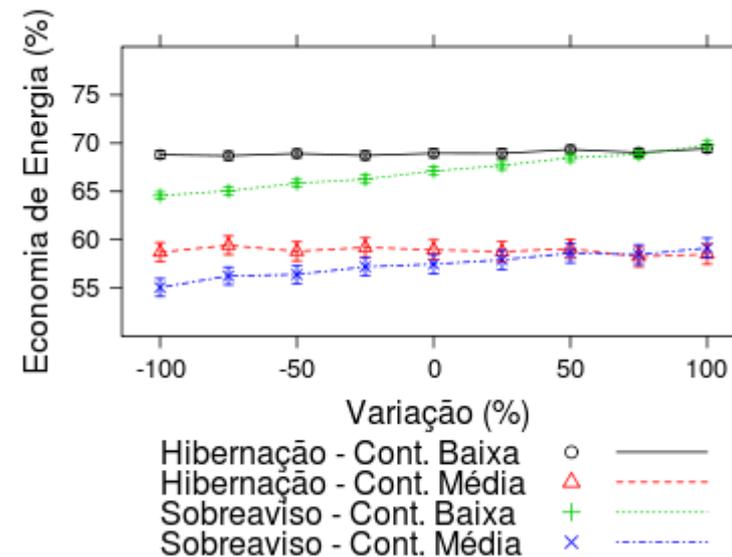
Impactam em até 3% a economia gerada pela estratégia Hibernação e em até 6% a economia gerada pela estratégia Sobreaviso

Impactam apenas a economia gerada pela estratégia Sobreaviso

Latência e potência variam nas mesmas proporções



Latência e potência variam em proporções inversas. Latência segue o eixo das abcissas

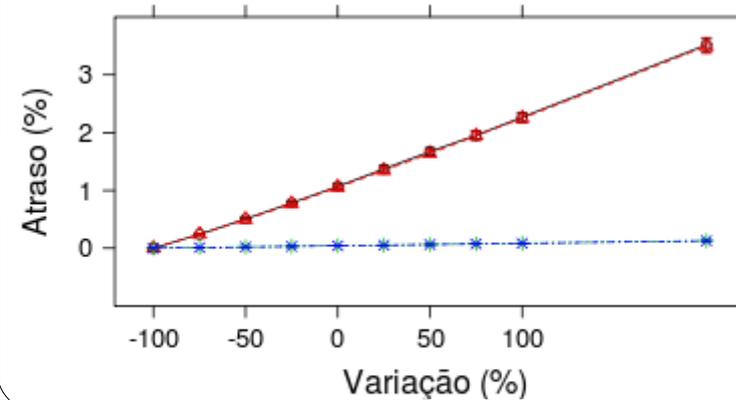


Análise de sensibilidade

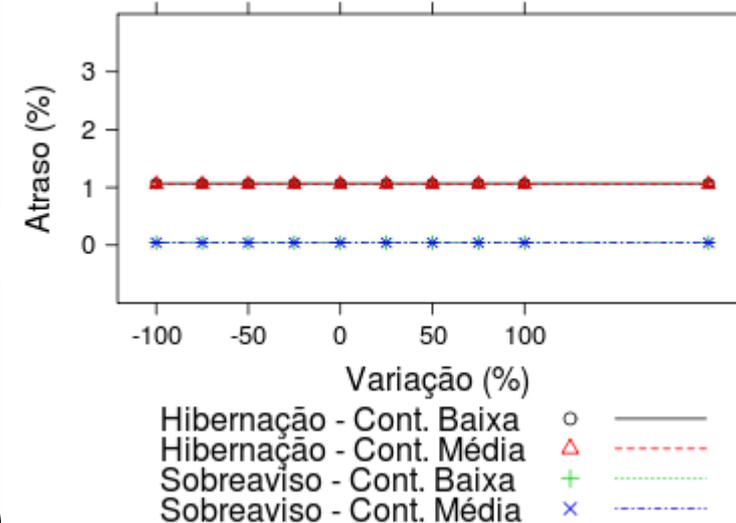
Não impactam o atraso gerado pela estratégia Sobreaviso. Aumentam em até 2% o atraso gerado pela estratégia Hibernação

Não impactam o atraso

Potência fixa e latência varia



Latência fixa e potência varia



Conclusão

- Até 65,53% de economia de energia com o uso de estratégias de dormência
- A estratégia Sobreaviso gera menor atraso no tempo de resposta
- O uso de estratégia de dormência pela grade gera menos transições do que o usuário fazer uso da estratégia localmente

Conclusão

- EA permite aumentar a economia de energia em até 3%
- TI permite reduzir o número de transições e o atraso gerado no tempo de resposta

Obrigado!

lesandrop@lsd.ufcg.edu.br
fubica@dsc.ufcg.edu.br

Referências

Mujtaba Talebi and Thomas Way. Methods, metrics and motivation for a green computer science program. SIGCSE Bull., 41:362–366, March 2000

Energy Star: http://www.energystar.gov/index.cfm?c=power_mgt.pr_pm_step1

Kamal Sharma and Sanjeev Aggarwal. Energy aware scheduling on desktop grid environment with static performance prediction. In Proceedings of the 2009 Spring Simulation Multiconference, SpringSim '09, pages 105:1–105:8, San Diego, CA, USA, 2009. Society for Computer Simulation International.

M. Lammie, P. Brenner, and D. Thain. Scheduling grid workloads on multicore clusters to minimize energy and maximize performance. In 10th IEEE/ACM International Conference on Grid Computing, 2009, pages 145 –152, 2009.

Lammie et al. (2009) Scheduling Grid Workload on Multicore Clusters to Minimize Energy and Miximize Performance

Condor Project. Condor version 7.4.4, 2010. Disponível em: <http://www.cs.wisc.edu/condor/>. Último acesso em dezembro de 2010.

Referências

Alexandru Iosup, Ozan Sonmez, Shanny Anoep, and Dick Epema. The performance of bags-of-tasks in large-scale distributed systems. In Proceedings of the 17th international symposium on High performance distributed computing, HPDC '08, pages 97–108, New York, NY, USA, 2008. ACM.

Joshua Reich, Michel Goraczko, Aman Kansal, and Jitendra Padhye. Sleepless in seattle no longer. In Proceedings of the 2010 USENIX conference on USENIX annual technical conference, USENIXATC'10, pages 17–17, Berkeley, CA, USA, 2010. USENIX Association.

Lesandro Ponciano and Francisco Brasileiro. On the impact of energy-saving strategies in opportunistic grids. In Energy Efficient Grids, Clouds and Clusters Workshop, proceedings of the 11th ACM-IEEE International Conference on Grid Computing (Grid 2010), pages 282 – 289, Bruxelas, Bélgica, 2010. ACM-IEEE.

Lesandro Ponciano, Francisco Brasileiro, Jaíndson Santana, Marcus Carvalho, and Matheus Gaudencio. Usando as estratégias sobreaviso e hibernação para economizar energia em grades computacionais oportunistas. Revista Brasileira de Redes de Computadores e Sistemas Distribuídos, 2011. (Aceito para publicação)

Referências

Lesandro Ponciano, Jaindson Santana, Marcus Carvalho, Matheus Gaudencio, and Francisco Brasileiro. Análise de estratégias de computação verde em grades computacionais oportunistas. In Anais do XXVIII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos, pages 307–320, Porto Alegre, Brasil, may 2010. SBC.

Saurabh Kumar Garg and Rajkumar Buyya. Exploiting heterogeneity in grid computing for energy-efficient resource allocation, 2009.

Canonical Ltd. Power management in ubuntu. Disponível em: <https://wiki.ubuntu.com/power-management-in-Ubuntu>. Último acesso em janeiro de 2011.

Microsoft Corporation. Windows power management. Disponível em: <http://www.microsoft.com/whdc/archive/winpowmgmt.mspx>. Último acesso em janeiro de 2011.

Referências das Figuras

- A figura apresentada na página 12 foi obtida em:
computingaustralia.com.au