

# 33 Agosto 2007  
**LINUX**  
MAGAZINE

# LINUX

A REVISTA DO PROFISSIONAL DE TI

# MAGAZINE

#33 08/07

R\$ 13,90  
€ 7,50



VEJA COMO O LINUX SE BENEFICIA DOS NOVOS  
PROCESSADORES COM MÚLTIPLOS NÚCLEOS

# DUAL CORE

0.33

- » O Apache é mais rápido em um dual core? p.34
- » Programação paralela com OpenMP p.44
- » CPUs multicore disponíveis no mercado p.37
- » Processos paralelos no kernel p.50

## ZFS NO LINUX? p.68

Conheça o super sistema de arquivos  
da Sun, já disponível no pingüim

## SOA p.30

Nosso novo colunista explica por que  
se fala tanto nesse conceito

## VEJA TAMBÉM NESTA EDIÇÃO:

- » Segunda aula da certificação LPIC-2 p.53
- » CACTI, um novo gerenciador de redes p.65
- » iPAQ ressuscitado pelo Linux p.54
- » Caracteres estranhos? Perl resolve! p.73

# FRACTO

Hospedagem de sites

Registros de Domínios

Servidor VPS

Servidor Dedicado

Servidor Virtual

Collocation

IP Dedicado

0800

Planos DDD/DDI

VPN IP

Frame Relay

Ligações Locais

[www.fracto.com.br](http://www.fracto.com.br)

SP (11) 3523-1580

3835-6990

RJ (21) 3723-1580

## SOLUÇÕES

COMPLETAS PARA SUA EMPRESA  
CRESCER A CADA DIA.

**Agente  
Autorizado  
Intelig**  
Telecom



## Expediente editorial

### Diretor Geral

Rafael Peregrino da Silva  
peregrino@linuxmagazine.com.br

### Editor-chefe

Tadeu Carmona  
tcarmona@linuxmagazine.com.br

### Editor

Pablo Hess  
phess@linuxmagazine.com.br

### Revisão

Arali Lobo Gomes  
agomes@linuxmagazine.com.br

### Editor de Arte

Renan Herrera  
rherrera@linuxmagazine.com.br

### Assistente de Arte

Igor Daurício  
isilva@linuxmagazine.com.br

### Centros de Competência

*Centro de Competência em Software:*

Oliver Frommel: ofrommel@linuxnewmedia.de  
Kristian Kiffling: kkiffling@linuxnewmedia.de  
Peter Kreussel: pkreussel@linuxnewmedia.de  
Marcel Hilzinger: hilzinger@linuxnewmedia.de  
Andrea Müller: amueller@linuxnewmedia.de

*Centro de Competência em Redes e Segurança:*

Achim Leitner: aleitner@linuxnewmedia.de  
Jens-Christoph B. Jbrendel@linuxnewmedia.de  
Hans-Georg Eber: hgesser@linuxnewmedia.de  
Thomas L. Leichtenstern@linuxnewmedia.de  
Max Werner: mwerner@linuxnewmedia.de

### Anúncios:

Rafael Peregrino da Silva (Brasil)  
anuncios@linuxmagazine.com.br  
Tel.: +55 (0)11 4082 1300  
Fax: +55 (0)11 4082 1302

Osmund Schmidt (Alemanha, Áustria e Suíça)  
anzeigen@linux-magazine.com

Brian Osborn (Outros países)  
ads@linux-magazine.com

### Assinaturas:

www.linuxnewmedia.com.br  
assinaturas@linuxmagazine.com.br

### Na Internet:

www.linuxmagazine.com.br – Brasil  
www.linux-magazin.de – Alemanha  
www.linux-magazine.com – Portal Mundial  
www.linuxmagazine.com.au – Austrália  
www.linux-magazine.ca – Canadá  
www.linux-magazine.es – Espanha  
www.linux-magazine.pl – Polônia  
www.linux-magazine.co.uk – Reino Unido  
www.linux-magazin.ro – Romênia

### Gerente de Circulação

Cláudio Guilherme dos Santos  
csantos@linuxmagazine.com.br

Apesar de todos os cuidados possíveis terem sido tomados durante a produção desta revista, a editora não é responsável por eventuais imprecisões nela contidas ou por consequências que advêm de seu uso. A utilização de qualquer material da revista ocorre por conta e risco do leitor.

Nenhum material pode ser reproduzido em qualquer meio, em parte ou no todo, sem permissão expressa da editora. Assume-se que qualquer correspondência recebida, tal como cartas, emails, faxes, fotografias, artigos e desenhos, são fornecidos para publicação ou licenciamento a terceiros de forma mundial não exclusiva pela Linux New Media do Brasil, a menos que explicitamente indicado.

Linux é uma marca registrada de Linus Torvalds.

Linux Magazine é publicada mensalmente por:

Linux New Media do Brasil Editora Ltda.

Av. Fagundes Filho, 134  
Conj. 53 – Saúde  
04304-000 – São Paulo – SP – Brasil  
Tel.: +55 (0)11 4082 1300  
Fax: +55 (0)11 4082 1302

Direitos Autorais e Marcas Registradas © 2004 - 2007:  
Linux New Media do Brasil Editora Ltda.

Distribuição: Dismag

Impressão e Acabamento: Parma

ISSN 1806-9428

Impresso no Brasil



# Coerência e coesão

## Prezados leitores da Linux Magazine,

O Software Livre e de Código Aberto (SL/CA) surgiu a partir de uma comunidade de entusiastas, sem mercado próprio, embora muitos já utilizassem SL/CA profissionalmente. É verdade que desde bem cedo já se começou a pensar em formas de ganhar dinheiro com esses softwares. Muitas tentativas foram feitas, obviamente com variados graus de sucesso.

Entretanto, apenas recentemente se consolidou a idéia de que o SL/CA constitui um mercado de fato, pronto para ser explorado por quem deseja obter lucros. Como consequência, muitos integrantes da comunidade de usuários e desenvolvedores permanecem iludidos com a idéia de que o SL/CA não pode ser explorado por empresas privadas para a geração de lucro, chegando até mesmo a crer que a principal licença livre, a GPL, proíbe a comercialização de softwares.

É importante esclarecer para essas pessoas que há diversos modelos de negócios associados ao SL/CA, e que os mesmos têm como objetivo gerar lucro, sem no entanto ferir os princípios do Software Livre – aliás, muito pelo contrário, eles costumam ser importantes formas de fomento ao desenvolvimento desses softwares.

Apenas para exemplificar – os casos são realmente incontáveis –, os membros da comunidade criticam o fato de distribuidores Linux cobrarem dinheiro para fazer acordos com fabricantes de hardware, mas esquecem-se de que é necessário trabalho – e trabalho tem um custo – para se criar uma imagem de sistema operacional adequada ao hardware em questão, e também para mantê-la atualizada e segura.

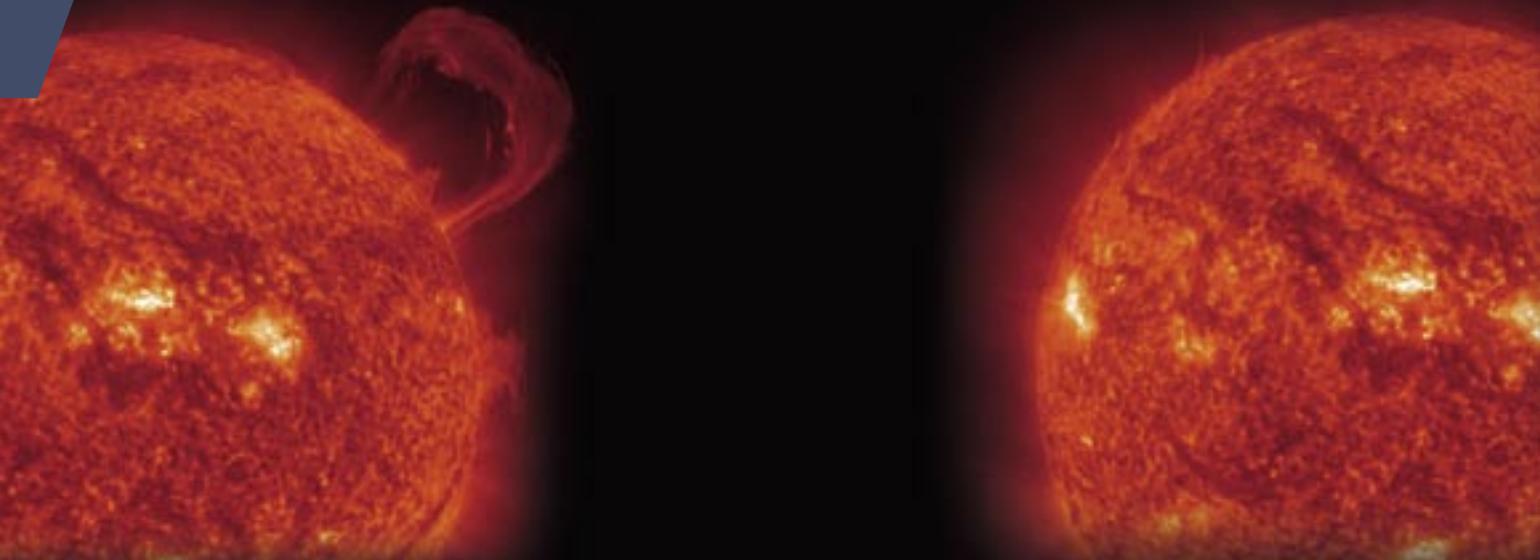
Há que se ter em mente o fato de que ainda que o software seja livre, os serviços providos em torno dele são pagos, pois de fato têm um custo. Os membros da comunidade devem sentir-se livres para, em vez de apenas criticarem aqueles responsáveis por aumentar a percepção do mercado a respeito do SL/CA, montarem eles próprios um negócio que enriqueça esse ecossistema, cobrando, para isso, a quantia que acharem adequada.

Esse comportamento prejudicial da comunidade pode ser comparado a quem vende um pão mas proíbe o comprador de passar manteiga nele; o SL/CA seria o pão, e a manteiga representaria a agregação de valor ao produto – prestação de serviços, principalmente.

A tecnologia do Software Livre e de Código Aberto já se mostrou capaz de produzir excelentes softwares. Para alcançar uma maior difusão e influenciar ainda mais o mercado de TI, ela necessita apenas que seus apoiadores – a comunidade e as empresas já envolvidas – não tenham ideais conflitantes.

Pablo Hess  
Editor





## CAPA

### **Espírito fraterno** 33

Com o fim da corrida dos gigahertz, os fabricantes de processadores agora apostam na força do paralelismo. Mas os programas devem estar preparados para essa mudança.

### **Apache na vitrine** 34

Os benchmarks confirmam: apenas aplicativos com uso intensivo de CPU e bem paralelizados aproveitam múltiplos núcleos.

### **Gêmeos** 37

Os processadores de múltiplos núcleos estão conquistando tanto servidores quanto desktops e até mesmo notebooks. Conheça os termos que definem essa tendência e veja as listagens de processadores multi-núcleo.

### **Programas orquestrados** 44

Com o fim da corrida por frequências mais altas, o mesmo ocorre também com a aceleração dos programas. O atual problema dos desenvolvedores é adequar seus softwares para aproveitarem os benefícios dos processadores multinúcleo.

### **O delegado** 50

Uma das missões do kernel Linux é a distribuição de processos e threads da maneira mais proporcional possível nos núcleos da CPU. Algumas funções da API de escalonamento ajudam o programador de aplicativos a influenciar o kernel da melhor forma.



## COLUNAS

|                       |           |
|-----------------------|-----------|
| <b>Augusto Campos</b> | <b>08</b> |
| <b>Charly Kühnast</b> | <b>10</b> |
| <b>Klaus Knopper</b>  | <b>12</b> |
| <b>Zack Brown</b>     | <b>14</b> |

## NOTÍCIAS

|  |           |
|--|-----------|
| <b>Segurança</b>                             | <b>16</b> |
| ◆ Firefox, SeaMonkey e Thunderbird           |           |
| ◆ VIM  |           |
| ◆ FreeType                                   |           |
| ◆ PHP  |           |
| ◆ Samba                                      |           |
| ◆ SquirrelMail                               |           |
| ◆ file                                       |           |
| ◆ GIMP                                       |           |
| ◆ PPTPD                                      |           |
| ◆ IPsec                                      |           |
| ◆ libpng                                     |           |
| <b>Geral</b>                                 | <b>18</b> |
| ◆ Notebook popular mais fácil                |           |
| ◆ Cookies do Google vão se autodestruir      |           |
| ◆ Apple adquire CUPS                         |           |
| ◆ Firefox alcança 30% dos usuários na Europa |           |
| ◆ Justiça condena crackers no Pará           |           |
| ◆ Linux legaliza uso de máquinas de bingo    |           |
| ◆ Da GPLv2 para GPLv3, a passos curtos       |           |
| ◆ Fim da linha para o PHP 4                  |           |

## CORPORATE

|  |           |
|--|-----------|
| <b>Notícias</b>  | <b>20</b> |
| ◆ Sun amplia oferta de virtualização de armazenamento  |           |
| ◆ Intel também mostra seus números                     |           |
| ◆ Código Aberto cresce nas aplicações críticas         |           |
| ◆ Red Hat encerra trimestre fiscal com receita recorde |           |
| ◆ Bull lidera novo consórcio                           |           |
| ◆ Fujitsu traz novidades em servidores blade           |           |
| <b>Entrevista: Ifractal</b>                            | <b>22</b> |
| <b>Artigo: GPLv3</b>                                   | <b>24</b> |
| <b>Entrevista: Microsoft e Propriedade Intelectual</b> | <b>26</b> |
| <b>Entrevista: PSC da Fujitsu</b>                      | <b>28</b> |
| <b>Coluna: Edgar Silva</b>                             | <b>30</b> |
| <b>Coluna: Cezar Taurion</b>                           | <b>32</b> |



## Tutorial

|  |           |
|--|-----------|
| <b>A libertação dos iPAQ</b>   | <b>54</b> |
| Somente Software Livre permite a retro-computação, ou seja, a reutilização de hardware considerado obsoleto pelo mercado por meio da adição de softwares atuais. |           |



|  |           |
|--|-----------|
| <b>LPI nível 2: Aula 2</b>   | <b>58</b> |
| O primeiro passo da inicialização de um sistema Linux é o carregamento do kernel. Nesta aula vamos conhecer algumas particularidades do kernel e dos sistemas de arquivo do Linux. |           |

## SYSADMIN

|   |           |
|---|-----------|
| <b>Monitorar é preciso</b>  | <b>65</b> |
| Entre as várias atividades desenvolvidas pelos administradores de redes está a de monitorar roteadores, switches, servidores e seus respectivos serviços a fim de garantir taxas aceitáveis de desempenho e disponibilidade. O Cacti é uma boa opção de ferramenta de monitoramento que pode ajudar muito nessa tarefa. |           |



|   |           |
|---|-----------|
| <b>Sem limites</b>  | <b>68</b> |
| Um sistema de arquivos virtualmente impossível de ser esgotado, com múltiplas ferramentas que permitem, inclusive, a montagem automática de RAID's Nível 0, Nível 1 e Z. Não é uma promessa de ficção científica: esse sistema existe e foi inventado pela Sun. |           |

## PROGRAMAÇÃO

|   |           |
|---|-----------|
| <b>Sopa de letrinhas</b>  | <b>73</b> |
| Quando caracteres estranhos aparecem no código ou nos dados de um programa, os programadores Perl precisam de uma solução que evite as complicações do Babel. |           |



## SERVIÇOS

|                              |           |
|------------------------------|-----------|
| <b>Editorial</b>             | <b>03</b> |
| <b>Emails</b>                | <b>06</b> |
| <b>Linux.local</b>           | <b>78</b> |
| <b>Eventos</b>               | <b>80</b> |
| <b>Índice de anunciantes</b> | <b>80</b> |
| <b>Preview</b>               | <b>82</b> |

*Emails para o editor*

# Permissão de Escrita

Se você tem dúvidas sobre o mundo Linux, críticas ou sugestões que possam ajudar a melhorar a nossa revista, escreva para o seguinte endereço: **cartas@linuxmagazine.com.br**. Devido ao volume de correspondência, é impossível responder a todas as dúvidas sobre aplicativos, configurações e problemas de hardware que chegam à Redação, mas garantimos que elas são lidas e analisadas. As mais interessantes são publicadas nesta seção.

## Linux em laptops?

Eu adoraria ver um artigo cobrindo o Linux em laptops. Vocês poderiam falar das principais distribuições e descrever o suporte delas para recursos como:

- ◆ docking stations;
- ◆ suspender para RAM e para o disco;
- ◆ controle de ventoinhas;
- ◆ escalonamento de frequência do processador;
- ◆ leitores de CD/DVD e disquete hotplug;
- ◆ USB e firewire;
- ◆ hotplug de USB e firewire;
- ◆ resolução nativa de LCDs;
- ◆ redes sem fio, incluindo WPA, WPA2, WEP e Bluetooth.

Seria possível cobrir melhor a questão do Linux em laptops?

**Jessé Assumpção**  
Votorantim - SP

*Agradecemos sua sugestão. Concordamos que os problemas relacionados a laptops e notebooks são muito importantes. Em números passados, já cobrimos assuntos como hotplug e hibernação, e estamos planejando análises de computadores*

*portáteis. Uma de nossas dificuldades, no entanto, é que cada laptop é diferente dos outros, pois realmente não há um padrão que os fabricantes sigam. Como consequência, não existe uma solução única para todas as máquinas em conjunto com todas as distribuições. Uma fonte interessantíssima de informações a respeito de portáteis é nossa coluna fixa do Klaus Knopper. Ele frequentemente responde perguntas de nossos leitores internacionais em relação a laptops, e costuma enfatizar um conhecimento mais profundo do funcionamento de sistemas Linux. ■*



# nova regra do mercado

# dividir para multiplicar



DO BANCO DE DADOS À INTERFACE, O SOFTWARE LIVRE DO GOVERNO DO PARANÁ ESTÁ MULTIPLICANDO RESULTADOS

Independência tecnológica. Confiança. Segurança. Evolução permanente. Codificação auditável. Protocolos confiáveis. Sistemas robustos e escaláveis. Ampla rede de suporte técnico. Interoperabilidade. Personalização. Racionalização de custos.

Quando resolveu apostar no software livre o Governo do Paraná sabia o que estava fazendo. A parceria com a comunidade software livre tem produzido resultados de alto impacto social e econômico.

Do banco de dados à interface gráfica, as ferramentas e aplicações de código aberto utilizadas pela CELEPAR - Informática do Paraná estão multiplicando as soluções de TI no ambiente de governo e na sociedade. Avanços que são divididos com a comunidade através da liberação dos códigos fonte.

Compartilhar conhecimentos e somar esforços. Esta é a regra que o Paraná usa para crescer e criar oportunidades.



QUEM CONHECE, SABE. QUEM NÃO CONHECE, PRECISA VER.

Veja alguns exemplos das soluções desenvolvidas pela Celepar:

**EXPRESSO** sistema integrado de correio eletrônico, agenda, fluxos de trabalhos (workflow) e catálogos. Permite o compartilhamento e a busca de informações corporativas, independente da plataforma ou de limites geográficos, técnicos ou organizacionais.

**PLATAFORMA PARANÁ** (framework) conjunto de ferramentas, métodos e padrões para a produção de sistemas. Maior produtividade com arquitetura tecnológica pré-definida, organização de acervos e reutilização de componentes.

**HABILITAÇÃO E VEÍCULOS** (DETRAN-PR) sistemas de alta complexidade que gerenciam os processos de expedição, renovação, multas, pontos e situação de 3,5 milhões de condutores e igual número de veículos.

**XOOPS** - eXtensible Object Oriented Portal System, ferramenta dinâmica e interativa para a criação de portais, comunidade virtuais, sítios de notícias, intranets e weblogs de pequena e grande escala.

**DIÁ-Á-DIA EDUCAÇÃO** portal colaborativo com recursos didáticos, banco de imagens, dicionários, sons e vídeos, tradutores e links desejáveis. Possui área para publicação de conteúdos e pesquisas nas diversas áreas do conhecimento.



**CELEPAR**  
INFORMÁTICA  
do PARANÁ

Software Livre: o Paraná usa e abusa

\*A Licença Pública Geral para a Administração Pública é um decreto governamental que permite e dá cobertura legal para o uso, distribuição, adaptação, e distribuição de todos os softwares produzidos pelo Governo do Estado do Paraná.

central@celepar.pr.gov.br 55(41) 3350-5000 www.celepar.pr.gov.br  
© Linux New Media do Brasil Editora Ltda.



Projetos em dia

# Augusto Campos

Como levar o gerenciamento de projetos à administração de sistemas.

por Augusto Campos

O gerenciamento de projetos é um conjunto de disciplinas que tem estado em voga nas organizações brasileiras nos anos recentes. Em última análise, ele busca aumentar a taxa de sucesso em projetos, adotando uma série de métodos e melhores práticas cujo foco é garantir que os projetos sejam bem planejados, executados e acompanhados, alcançando o custo, escopo, qualidade e outros atributos conforme planejado, no prazo determinado.

Assim como diversos outros conceitos da Administração, o gerenciamento de projetos pode ser muito melhor empregado se for adotado na organização como um todo. Nas organizações que usam essas

desse tipo de processo – e, certamente, o sucesso nessa implantação irá motivá-lo a estudar sobre o assunto, e pode ser a semente da adoção do gerenciamento de projetos na sua organização.

Uma das ferramentas de Código Aberto mais populares na gestão de projetos é o *dotProject*[1]. O *dotProject* pressupõe algum entendimento das técnicas associadas a projetos, e pode ser muito bem empregado no ambiente da administração de redes e do suporte de informática, em equipes de vários tamanhos (até mesmo quando de uma pessoa só) nas definições e acompanhamento de projetos de implantação de sistemas, de mudanças de procedimentos e muitos outros casos. Pode ainda cobrir aspectos como o conjunto de pessoas envolvidas, as listas de tarefas, os cronogramas (incluindo os famosos gráficos de Gantt), a comunicação e o compartilhamento no âmbito dos projetos.

Caso você prefira dar passos menores, pode considerar a adoção de uma ferramenta como o *ProjectBench*[2], que alia o apoio a atividades de gerência de falhas – as típicas atividades de suporte e *help-desk* – ao enfoque de projetos, facilitando as tarefas de acompanhamento de projetos, especialmente em organizações que dão suporte aos sistemas que elas mesmas desenvolvem.

De uma forma ou de outra, a adoção de técnicas de gestão sempre deve ser adaptada à cultura de sua organização e às suas demandas. Iniciativas bem sucedidas de implantação de novas técnicas sempre acabam sendo notadas, e refletem sobre a qualidade, a produtividade e até a motivação da equipe. Esperamente, você não vai se arrepender! ■

## A adoção de técnicas de gestão sempre deve ser adaptada à cultura de sua organização e às suas demandas.

técnicas, é provável que isso se reflita diretamente na área de suporte e de infra-estrutura (especialmente administração de redes e sistemas) – possivelmente, alguns dos integrantes dessas equipes serão até mesmo gerentes de alguns projetos, ou usualmente receberão ou preencherão documentos formais de projetos.

Mas você não precisa esperar a sua organização adotar formalmente o gerenciamento de projetos. A iniciativa de se informar sobre o tema pode começar com uma visita a uma boa livraria e à profusão de websites que explicam essas técnicas – em especial da forma definida pela *PMI* (*Project Management Institute*), que são inclusive as bases das normas *ISO* e *ABNT* sobre esse assunto.

E, mesmo sem buscar grande formação no tema, você certamente pode dar um razoável primeiro passo para agregar valor aos seus processos internos e serviços se adotar ferramentas de suporte ao gerenciamento de projetos. Isso facilita a criação e o acompanhamento de cronogramas, diagramas de alocação de equipes e recursos, projeção de custos e outros produtos típicos

### Mais Informações

[1] dotProject: <http://www.dotproject.net/>

[2] Project Bench: <http://projectbench.sourceforge.net/>

### O autor

**Augusto César Campos** é administrador de TI e, desde 1996, mantém o site [BR-linux.org](http://BR-linux.org), que cobre a cena do Software Livre no Brasil e no mundo.



# ALÉM DAS SOLUÇÕES E TREINAMENTOS RECONHECIDOS NO MERCADO, O ATENDIMENTO E SUPORTE QUE A SUA EMPRESA PRECISA

“

Nossa equipe de TI sempre elogia o pessoal de suporte da Red Hat Brasil, o que contribui ainda mais para o bom relacionamento das duas empresas. Estamos, inclusive, avaliando a aquisição de novas subscrições para comportar a ampliação de nossos negócios e trazer mais tranquilidade para nossos clientes internos e externos.

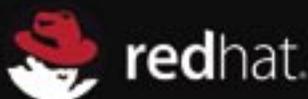
**Reinaldo Gonçalves Melero**

Gerente de Informática  
Copagaz Distribuidora de Gás Ltda

”

**www.br.redhat.com**  
**11 3529.6018**

Solicite uma versão para teste:  
[www.redhat.com/rhel/details/eval](http://www.redhat.com/rhel/details/eval)



P3Scan

# Charly Kühnast

Buscar vírus em emails normalmente é função do servidor SMTP ou de alguma máquina ligada a ele. Que tal mover essa proteção para o outro lado – isso é, para as conexões clientes com os servidores SMTP e POP?

por Charly Kühnast

O P3Scan[1] é um proxy de email que se posta diante na frente dos *daemons* SMTP e POP3, aceitando conexões de clientes que desejem baixar ou se livrar de emails. Ele repassa os comandos dos clientes e verifica os emails atrás de conteúdo maléfico antes de encaminhar os emails propriamente ditos. O P3Scan evita dependências exóticas, e baseia-se apenas na biblioteca *pcre-devel*, presente na maioria das distribuições.

Obviamente o P3Scan pode utilizar softwares antivírus. Nesse caso, escolho o *ClamAV*, mas também é possível utilizar o *F-Prot*, *F-Secure*, *Kaspersky* e provavelmente outros produtos, contanto que possuam um cliente em linha de comando. O P3Scan também pode integrar o *SpamAssassin* e o *DSPAM*, oferecendo a capacidade de eliminar propagandas indesejadas dos emails.

O *Iptables* oferece a possibilidade de utilizar o P3Scan como um proxy de emails transparente. Os usuários não tomam qualquer conhecimento da existência do programa, ao menos enquanto os emails que chegarem permanecerem livres de vírus. Pode-se facilmente instalar e configurar o P3Scan em um roteador Linux. Eu uso o P3Scan numa porta alta – o padrão é 8110 – utilizando o *Iptables* para enviar todas as conexões POP3 para a porta exigida:

```
iptables -t nat -I PREROUTING ! -i eth0 -p tcp -s
➤ 192.168.1.0/24 --dport 110 -j REDIRECT --to-
➤ ports 8110
```

Depois disso, o P3Scan aceitará conexões POP3, passando todos os comandos POP – tais como *RETR*, *DELE* etc. – para o servidor de destino e buscando mensagens legítimas no servidor. Ele também passará as mensagens pelo antivírus e, caso necessário, pelo filtro de spam.

## Configuração

O P3Scan é controlado pelo arquivo `/etc/p3scan/p3scan.conf`, e as seguintes entradas são críticas:

**targetip:** Se o P3Scan for usado como proxy transparente, esse valor deve ser `0.0.0.0`. Caso contrário, digite o IP do servidor “real” para o qual o P3Scan repassará as conexões clientes;

**bytesfree = <n>:** São necessários ao menos *n* bytes livres em disco; caso contrário, o P3Scan fechará. Note que o programa gera diversos processos-filhos (dez, por padrão); no pior caso, todos podem precisar lidar com grandes anexos ao mesmo tempo;

**scanner = <comando>:** Entre aqui o comando que inicia o antivírus. Se você usar o *ClamAV*, essa linha fica:

```
scanner = /usr/bin/clamscan --no-summary
```

**viruscode =:** O P3Scan avalia o código de retorno do antivírus para determinar se o email está infectado ou limpo. Normalmente, o antivírus retorna o (zero) se a mensagem estiver limpa, e 1 (um) se contiver um vírus. Alguns antivírus utilizam outros códigos de retorno. Para fazer o P3Scan avaliar esses códigos, é necessário acrescentar uma linha ao final do arquivo. Se seu antivírus retorna um valor de 1, 5 ou 13 para “Contaminado!”, a linha deve ser `viruscode 1, 5, 13`. O mesmo princípio se aplica aos códigos de retorno diferentes dos que indiquem “Limpo!”, mas a linha, nesse caso, começa com `goodcode =`.

**overwrite = /usr/bin/p3pmail:** Essa linha elimina códigos HTML dos emails. Ela impede que os clientes carreguem imagens automaticamente ao ler seus emails, o que seria perigoso e permitiria que o spammer soubesse que a mensagem foi recebida e aberta. ■

## Mais Informações

[1] P3Scan: <http://p3scan.sourceforge.net>

### O autor

**Charly Kühnast** é administrador de sistemas Unix no datacenter Moers, perto do famoso rio Reno, na Alemanha. Lá ele cuida, principalmente, dos firewalls.





# Itautec

## SEJA LIBRIX NA RUA, SEJA LIBRIX EM CASA, SEJA LIBRIX NO TRABALHO.

Agora, além do Librix (Linux da Itautec), a sua empresa pode contar com o melhor e mais estável pacote de hardware e software do mercado, testado e homologado pela Itautec.

Toda a **liberdade** que você precisa para trabalhar com mais mobilidade, usando a internet sem fio, e ainda operar com software livre.

É mais **segurança**, porque a Itautec oferece suporte técnico especializado via internet ou pelo telefone, serviços de tuning e configuração e ainda atendimento nacional on site.

Tem alta **tecnologia** para os aplicativos como editor de textos, planilha eletrônica, editor de imagens e apresentações. É mais facilidade e maior flexibilidade no seu dia-a-dia. Na hora de trabalhar, não se sinta preso. Seja Librix.



Foto ilustrativa.

### Itautec InfoWay Minitorre

IDEAL PARA ENTRETENIMENTO  
E ESTAÇÃO DE TRABALHO.

- Código da oferta: IN-598
- Processador Intel® Celeron® D 331 (256 KB L2 cache, 2.66 GHz, 533 MHz)
  - Librix - Distribuição Linux Itautec
  - 512 MB de memória
  - HD de 80 GB
  - Floppy
  - Teclado e mouse
  - Caixas acústicas e microfone
  - **MONITOR NÃO INCLUSO**

**36x R\$ 26,40**  
pelo cartão BNDES\*  
ou R\$ 799,00 à vista



DIPZ

Foto ilustrativa.

### Servidor Itautec LR100

PERFORMANCE E SEGURANÇA  
PARA A SUA EMPRESA.

- Código da oferta: IN-599
- Processador AMD Opteron 1210
  - Sem sistema operacional
  - 1 GB de memória
  - HD de 160 GB
  - DVD-RW (leitor e gravador de CD e DVD)
  - Floppy
  - Auto Manager

**36x R\$ 51,19**  
pelo cartão BNDES\*  
ou R\$ 1.549,00 à vista



PRESENTE EM MAIS  
DE 2.700 CIDADES.

[www.itautecshop.com.br](http://www.itautecshop.com.br)

COMPRA DIRETAMENTE DO FABRICANTE

**0800 121 444**

De 2ª a 6ª, das 8h às 20h. Sábado, das 9h às 18h.

Ofertas válidas até 8/8/2007 ou enquanto durarem os estoques. Celeron, Celeron Inside, Centrino, o logotipo Centrino, Core Inside, Intel, o logotipo Intel, Intel Core, Intel Inside, o logotipo Intel Inside, Intel Viiv, Intel vPro, Itanium, Itanium Inside, Pentium, Pentium Inside, Xeon e Xeon Inside são marcas comerciais ou marcas registradas da Intel Corporation ou de suas subsidiárias nos Estados Unidos ou em outros países. \*Financiamento para pessoa jurídica através do cartão BNDES, com taxa de 1,03% a.m. Necessário possuir o cartão de crédito citado, sujeito à confirmação da disponibilidade da linha de crédito para as localidades e limite para operação. Consulte nossa Central de Atendimento para informações sobre outras condições de financiamento para pessoa física ou jurídica pelo telefone 0800-121-444. A velocidade de comunicação de 56 Kbps depende e pode variar de acordo com o tipo e a qualidade da linha telefônica utilizada. Para possibilitar o acesso à internet são necessários uma linha telefônica ou banda larga e um provedor à sua escolha. Preços com impostos incluídos para São Paulo. Frete não incluído. Demais características técnicas e de comercialização estão disponíveis em nosso site e no Televendas. Fica ressalvada eventual retificação das ofertas aqui veiculadas. Quantidade: 10 unidades de cada. Empresa/produto beneficiado pela Lei de Informática. Fotos meramente ilustrativas.

Pergunte ao Klaus!

# Klaus Knopper

O criador do Knoppix responde as mais diversas dúvidas dos leitores.  
por Klaus Knopper

## Monitoramento

Receio que meu sistema Linux às vezes esteja enviando ou recebendo mais dados do que eu preciso. Tenho medo de que ele possa estar recebendo código invasor ou enviando spam.

É possível monitorar o que sai e entra durante um período específico de tempo, jogar esses resultados num arquivo e analisá-lo depois? Uso o Firefox e o KMail no Suse 10.2 com ADSL via Ethernet.

### Resposta

Se seu sistema estiver seriamente comprometido, o invasor será capaz de colocar nele versões dos utilitários administrativos contaminadas com cavalos-de-troia destinadas a ocultar acessos não autorizados. Uma solução completa para o problema de prevenção de invasões requer planejamento cuidadoso e adiantado; entretanto, se seu kit de ferramentas de administração não tiver sido substituído por um *rootkit*,

diversos utilitários estão disponíveis para auxiliar a descoberta de atividades suspeitas.

Pode-se usar o *iptraf* para obter uma lista dos bytes transmitidos por número da porta (figura 1) e um *netstat -tup* (figura 2) para identificar suas conexões clientes de seu computador. Ambos os comandos requerem permissões de root para mostrarem todas as informações relevantes.

Eu faço isso frequentemente para otimizar o firewall e a modelagem do tráfego. Além disso, se você suspeita que algo esteja errado com seu computador, como um servidor de emails indesejado agindo como proxy para spam, use o comando *netstat -tulp* (novamente como root) para identificar quais servidores em seu computador estão escutando em quais portas (figura 3). Esse comando é capaz de detectar servidores secretos de email rodando no sistema; ainda assim, é possível que um servidor de email comum rodando na porta 25 (SMTP) esteja sendo usado como *relay* para spam, caso não esteja configurado corretamente.

Se você está desconfiado, é melhor usar uma ferramenta de detecção de intrusões, como o *Snort*, *Tiger* ou *TCT*, localmente, e usar um *live CD* para eliminar a influência de um possível kernel comprometido.

Ocasionalmente, o tráfego extra ou conexões adicionais à Web no *netstat* podem ser causados pelo “pré-carregamento inteligente” de páginas com links no navegador, dependendo das configurações e dos plugins ativados. Além disso, algumas páginas web podem eventualmente baixar novos *banners* de anúncio.

Um efeito colateral sem relação técnica: se você verificar atividade de disco estranha no início da manhã, isso pode ser resultado dos scripts normais que atualizam a documentação e alimentam a base de dados de softwares como o *locate* ou o *beagle*. Use o *top* ou o *htop* para conferir a lista dos processos mais ativos. ■

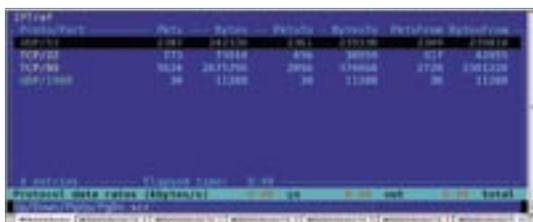


Figura 1 O *iptraf* lista os bytes transmitidos pelo número da porta.



Figura 2 O comando *netstat -tup* identifica suas conexões clientes.

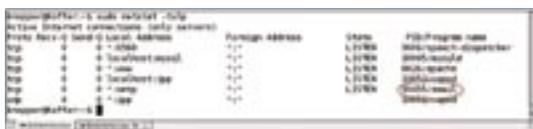


Figura 3 O comando *netstat -tulp* identifica quais servidores estão escutando em quais portas.

## O autor

**Klaus Knopper** é o criador do Knoppix e co-fundador do evento LinuxTag. Atualmente ele trabalha como professor, programador e consultor.



# Suas conexões simultâneas de streaming já chegaram no limite?

Streaming Plug In, conexão simultânea ilimitada.



## CONEXÕES SIMULTÂNEAS ILIMITADAS DE STREAMING

|                 | Windows Media I | Windows Media II | FLASHCOMM I | FLASHCOMM II | FLASHCOMM III |
|-----------------|-----------------|------------------|-------------|--------------|---------------|
|                 | Ilimitada       | Ilimitada        | 10 conexões | 50 conexões  | Ilimitada     |
| Espaço em Disco | 300MB           | 700MB            | 200MB       | 400MB        | 600MB         |
| Taxa de Transf. | 100GB           | 300GB            | 50GB        | 150GB        | 300GB         |
| Preço           | R\$90,00        | R\$190,00        | R\$70,00    | R\$180,00    | R\$290,00     |

### Você tem problemas em mostrar seu áudio e vídeo a todos os seus clientes?

O Streaming da Plug In é a melhor forma de mostrar o áudio ou vídeo para todos os seus clientes, porque ele tem uma diferença essencial: pacotes de Streaming, com conexões simultâneas ilimitadas, que garantem dessa maneira sua tranquilidade, permitindo entre outras coisas, transmissão ao vivo de Rádio e TV através da Internet!

**Streaming Plug In. Para nós, limite não existe.**

- **Revenda** - Linux, Windows e Mista
- **Streaming** - Conexões Ilimitadas
- **E-mail Marketing** - Ações Segmentadas

Ligue e contrate:  
**4003-1001**

Contrate online:  
**www.plugin.com.br**

**PlugIn**

As novidades do kernel 2.6.22

# Pablo Hess

Novas pilhas WLAN e FireWire, além das já tradicionais dezenas de novos drivers. Confira o que traz a versão mais recente do Linux.

por Pablo Hess

Com mais de meio milhão de linhas de código alteradas, acrescentadas ou apagadas, a versão 2.6.22 do kernel Linux estabelece um novo recorde na série 2.6. Exatos 74 dias após a liberação da última versão estável (2.6.21), o kernel teve retrabalhadas suas pilhas WLAN (que cuida dos aspectos de redes sem fio) e FireWire. Além disso, foram incluídos novos drivers para as placas de rede sem fio do projeto de laptops educacionais OLPC, foram incorporados drivers Ivtv para placas de captura de vídeo, e também diversos drivers para placas PVR WinTV, da Hauppauge. Não menos impressionante é a inclusão de suporte a mais uma arquitetura de processadores, Blackfin.

Para as configurações, foi desenvolvida a interface Cfg80211, capaz de interagir perfeitamente com a nova plataforma, e que provavelmente substituirá as ultrapassadas Wireless Extensions.

## Outros destaques

Também estreando no kernel 2.6.22 há uma nova pilha FireWire, que cobre praticamente todas as funções da versão anterior, deixando de fora (e em maus lençóis) apenas algumas funções mais exóticas.

O novo alocador de memória do kernel, com o curioso nome de SLUB, visa a oferecer menor overhead e maior desempenho que o atual SLAB, e deve substituí-lo gradativamente.

Grande parte das alterações da nova versão do Linux lidam com o suporte à arquitetura de processadores embarcados Blackfin, da Analog Devices. Além disso, outro processador, o novo modelo Micrel/Kendin KS8695 da família ARM de CPUs, também passa a ter suporte no Linux.

O suporte ao Playstation 3 continua melhorando, porém ainda faltam itens importantes para o uso em produção.

Por incrível que pareça, até o Windows® Vista afetou o novo kernel: o driver LDM para discos dinâmicos do Windows foi alterado para oferecer compatibilidade com a versão dessa ferramenta contida no Vista.

Há centenas de outras alterações relevantes, que infelizmente não cabem nesta coluna. Elas envolvem novos drivers e revisões de drivers pré-existentes nas áreas mais diversas, como redes, armazenamento, escalonamento do processador e do disco, monitoramento de hardware e dispositivos de vídeo.

Convido os leitores a testar a nova versão e avaliar se as melhorias estão prontas para o uso em produção. ■

## O autor

**Pablo Nehab Hess** é editor da Linux Magazine e tem mestrado em genética. Após conhecer o Linux através da computação científica, interessou-se imediatamente pela administração de sistemas e redes, o que praticou em alguns laboratórios na universidade.



Por incrível que pareça, até o Windows® Vista afetou o novo kernel

## Nova pilha WLAN

O principal objetivo por trás da nova pilha WLAN, segundo os desenvolvedores, é facilitar o suporte por parte dos fabricantes de hardware.

Embora boa parte da culpa pelo mau suporte a dispositivos WLAN no Linux seja dos fabricantes, o kernel Linux não está isento de responsabilidade. Até o advento da nova pilha Mac80211 não havia uma plataforma de código WLAN básica com a qual todos os desenvolvedores concordassem, e por isso muitos drivers ficavam de fora do kernel, pois era necessário reimplementar tais funcionalidades em cada driver, individualmente.

Todavia, a turma do kernel ainda não integrou nenhum driver à nova plataforma, mas algumas equipes de desenvolvimento de distribuições já o fizeram com sucesso. O driver iwlmwifi, ainda em desenvolvimento, utilizará a Mac80211 para oferecer suporte – sem necessidade de firmwares proprietários – aos chips Intel PRO/Wireless 3945ABG e 4965AGN.

# IBM FORUM Brasil

## Development Conference

O maior fórum de negócios, visão de futuro e tecnologia da IBM Brasil

A Linux New Media apóia o *IBM FORUM Brasil - Developer Conference*, o maior encontro de TI da IBM!

Venha conhecer como funcionam os softwares IBM na prática e participar de cursos e laboratórios com os melhores profissionais do mercado. Participando do evento, você pode ainda obter até **3 certificações profissionais IBM!**

### Passes Simples

**R\$ 500,00\*\***

03 certificações\*  
Cursos  
Laboratórios  
Material de apoio  
Tradução simultânea  
Almoço no local

Acesse o site e  
conheça outros  
pacotes especiais!

### Cursos

- Software Configuration Management
- Gerência de Portfólio de Projetos
- Desenvolvimento de SOA: Passo a passo
- Rational para Mainframe

*Entre no site e confira a agenda completa*

### Laboratórios

- Acelerando a entrega do Software com Gerência de Build & Release
- Testando Web services e aplicações SOA com Rational
- Aprimorando a qualidade de software com gerência de testes e automação dos testes

*E muito mais!*

### Certificações



Acesse [www.ibm.com/br/navcode](http://www.ibm.com/br/navcode)  
e digite o código **IBMFORUMDEV**

# 13/09

WTC Hotel  
São Paulo

Apoio



**LINUX NEW MEDIA**  
The Pulse of Open Source



Patrocínio

IBM developerWorks **Live!**

# Firefox, SeaMonkey e Thunderbird

Foi descoberta uma vulnerabilidade na qual o protocolo APOP permite que agressores remotos adivinhem os três primeiros caracteres de uma senha através de ataques *man-in-the-middle* que usam colisões MD5 e IDs de mensagens especialmente criados.

Esse problema de escala de projeto afeta todos os produtos que utilizam APOP, incluindo *Thunderbird*, *Evolution*, *Mutt*, *fetchmail* e *SeaMonkey*.

Múltiplas vulnerabilidades dos mecanismos de layout do Mozilla Firefox, Thunderbird e SeaMonkey permitem que agressores remotos causem uma negação

de serviço através de vetores relacionados a ponteiros inválidos, corrupção de fila, questões de sinalização de valores e outros problemas. Múltiplas vulnerabilidades no mecanismo *JavaScript* permitem que agressores remotos causem uma negação de serviços e possivelmente executem código arbitrário via vetores que causam corrupção de memória. (CVE-2007-1558, CVE-2007-2867, CVE-2007-2868) ■

Referência no Mandriva: MDKSA-2007:105

Referência no Red Hat: RHSA-2007:0353

Referência no Ubuntu: USN-469-1

## VIM

A *sandbox* do *vim* permite acesso a funções perigosas, como *writefile*, *feedkeys* e *system*, que podem permitir que agressores executem comandos de shell auxiliados por usuários, através de *modelines*. (CVE-2007-2438) ■

Referência no Mandriva: MDKSA-2007:101

Referência no Red Hat: RHSA-2007:0346-3

Referência no Suse: SUSE-SR:2007:012

Referência no Ubuntu: USN-463-1

## FreeType

A *FreeType* não verifica corretamente o número de pontos das fontes *TrueType*. Se um usuário for convencido a usar uma fonte especialmente criada, o agressor conseguiria executar código arbitrário com os privilégios do usuário. (CVE-2007-2754) ■

Referência no Gentoo: GLSA-200705-22

Referência no Ubuntu: USN-466-1

## PHP

Foi descoberta uma falha no manipulador de comandos *ftp* no *PHP*. Os comandos não eram corretamente filtrados quanto a caracteres de controle, e um agressor conseguiria enviar

comandos *ftp* arbitrários através de argumentos especialmente criados. Um estouro de *buffer* no manipulador de requisições SOAP no PHP poderia permitir que agressores remotos enviassem uma requisição SOAP especialmente criada e executassem código arbitrário com os privilégios do servidor web. Um estouro de *buffer* no *filter factory* de usuários no PHP poderia permitir que um agressor local criasse um script especial e executasse código arbitrário com os privilégios do servidor web. O *PEAR installer* não validava os caminhos de diretório, então, se um usuário instalasse um pacote *PEAR* malicioso, um agressor poderia sobrescrever arquivos arbitrários. (CVE-2007-2509, CVE-2007-2510, CVE-2007-2511, CVE-2007-2519) ■

Referência no Debian: DSA-1295, DSA-1296

Referência no Mandriva: MDKSA-2007:102, MDKSA-2007:103

Referência no Red Hat: RHSA-2007:0348, RHSA-2007:0348, RHSA-2007:0355

Referência no Ubuntu: USN-462-1

## Samba

O *Samba* não se desfazia totalmente dos privilégios de root ao traduzir *SIDs*. Um usuário remoto autenticado seria

capaz de realizar operações *SMB* durante uma pequena janela de oportunidade e obter privilégios de root.

O *Samba* contém diversas falhas ao processar parâmetros RPC codificados em *NDR*. Ele não consegue sanar corretamente a entrada de procedimentos remotos fornecida através de *RPCs* Microsoft. (CVE-2007-2444, CVE-2007-2446, CVE-2007-2447) ■

Referência no Debian: DSA-1291

Referência no Gentoo: GLSA-200705-15

Referência no Mandriva: MDKSA-2007:104

Referência no Slackware: SSA:2007-134-01

Referência no Suse: SUSE-SA:2007:031

Referência no Ubuntu: USN-460-1

## SquirrelMail

Múltiplas vulnerabilidades do tipo *cross-site scripting* (*XSS*) no filtro *HTML* do *SquirrelMail* 1.4.0 até 1.4.9a permitem que agressores remotos injetem scripts web ou *HTML* arbitrários através de *data: URI* num anexo *HTML* de email, ou vários conjuntos de caracteres não-ASCII que não são corretamente filtrados quando visualizados com o Microsoft Internet Explorer. (CVE-2007-1262) ■

Referência no Debian: DSA-1290-1  
 Referência no Mandriva: MDKSA-2007:106  
 Referência no Red Hat: RHSA-2007:0358

## file

Um estouro de inteiros no programa *file* 4.20, ao ser executado em sistemas de 32 bits, poderia permitir que agressores auxiliados por usuários executassem código arbitrário através de um arquivo grande que dispare um estouro que ignora a declaração `assert()`. ■

Referência no Gentoo: GLSA-200705-25  
 Referência no Mandriva: MDKSA-2007:114  
 Referência no Red Hat: RHSA-2007:0391-3

## GIMP

Um estouro de buffer baseado em pilha na função `set_color_table` de `sunras.c` no plugin *SUNRAS* do *GIMP* 2.2.14 permite que agressores remotos auxiliados por usuários executem código arbitrário através de um arquivo *RAS* especialmente criado. (CVE-2007-2356) ■

Referência no Gentoo: GLSA-200705-08  
 Referência no Mandriva: MDKSA-2007:108  
 Referência no Red Hat: RHSA-2007:0343  
 Referência no Suse: SUSE-SR:2007:011  
 Referência no Ubuntu: USN-467-1

## PPTPD

Foi descoberta uma vulnerabilidade no *PPTPD*, um *daemon* para o protocolo PPTP para Linux. A falha poderia permitir que agressores remotos enviassem um pacote especialmente criado e interrompessem túneis PPTP estabelecidos, levando a uma negação de serviço. (CVE-2007-0244) ■

Referência no Debian: DSA-1288-1 pp-tpd  
 Referência no Gentoo: GLSA-200705-18  
 Referência no Suse: SUSE-SR:2007:010

## IPSec

Uma falha foi descoberta no servidor de troca de senhas IPSec *racoon*. Agressores remotos poderiam enviar um pacote especialmente

criado e interromper túneis IPSec já estabelecidos, levando a uma negação de serviço. (CVE-2007-1841) ■

Referência no Gentoo: GLSA-200705-09  
 Referência no Mandriva: MDKSA-2007:084  
 Referência no Red Hat: RSA-2007:0342  
 Referência no Suse: SUSE-SR:2007:008  
 Referência no Ubuntu: USN-450-1

## libpng

Foi descoberta uma vulnerabilidade na *libpng*. A função `png_handle_tRNS` em `pngutil.c`, na *libpng* permite que agressores remotos causem uma negação de serviço através de uma imagem PNG em preto-e-branco com um valor incorreto de *CRC tRNS*. (CVE-2007-2445) ■

Referência no Gentoo: GLSA-200705-24  
 Referência no Mandriva: MDKSA-2007:116  
 Referência no Red Hat: RHSA-2007:0356  
 Referência no Slackware: SSA:2007-136-01

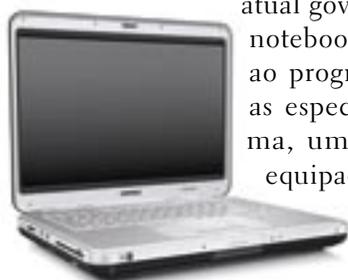
## Postura das principais distribuições Linux quanto à segurança

| Distribuição | Referência de Segurança   | Comentários  |
|--------------|---|--|
| Debian       | Info: <a href="http://www.debian.org/security">www.debian.org/security</a><br>Lista: <a href="http://lists.debian.org/debian-security-announce">lists.debian.org/debian-security-announce</a><br>Referência: DSA-... 1  | Alertas de segurança recentes são colocados na homepage e distribuídos como arquivos <i>HTML</i> com links para os patches. O anúncio também contém uma referência à lista de discussão.   |
| Gentoo       | Info: <a href="http://www.gentoo.org/security/en/glsa">www.gentoo.org/security/en/glsa</a><br>Fórum: <a href="http://forums.gentoo.org">forums.gentoo.org</a><br>Lista: <a href="http://www.gentoo.org/main/en/lists.xml">www.gentoo.org/main/en/lists.xml</a><br>Referência: GLSA: ... 1 | Os alertas de segurança são listados no site de segurança da distribuição, com link na homepage. São distribuídos como páginas <i>HTML</i> e mostram os comandos necessários para baixar versões corrigidas dos softwares afetados.  |
| Mandriva     | Info: <a href="http://www.mandriva.com/security">www.mandriva.com/security</a><br>Lista: <a href="http://www1.mandrivalinux.com/en/flists.php3#2security">www1.mandrivalinux.com/en/flists.php3#2security</a><br>Referência: MDKSA-... 1  | A Mandriva tem seu próprio site sobre segurança. Entre outras coisas, inclui alertas e referência a listas de discussão. Os alertas são arquivos <i>HTML</i> , mas não há links para os patches.   |
| Red Hat      | Info: <a href="http://www.redhat.com/errata">www.redhat.com/errata</a><br>Lista: <a href="http://www.redhat.com/mailling-lists">www.redhat.com/mailling-lists</a><br>Referência: RHSA-... 1   | A Red Hat classifica os alertas de segurança como "Erratas". Problemas com cada versão do Red Hat Linux são agrupados. Os alertas são distribuídos na forma de páginas <i>HTML</i> com links para os patches.  |
| Slackware    | Info: <a href="http://www.slackware.com/security">www.slackware.com/security</a><br>Lista: <a href="http://www.slackware.com/lists">www.slackware.com/lists</a> (slackware-security)<br>Referência: [slackware-security] ... 1  | A página principal contém links para os arquivos da lista de discussão sobre segurança. Nenhuma informação adicional sobre segurança no <i>Slackware</i> está disponível.  |
| Suse         | Info: <a href="http://www.novell.com/linux/security">www.novell.com/linux/security</a><br>Lista: <a href="http://www.novell.com/linux/download/updates">www.novell.com/linux/download/updates</a><br>Referência: suse-security-announce Referência: SUSE-SA ... 1                         | Após mudanças no site, não há mais um link para a página sobre segurança, contendo informações sobre a lista de discussão e os alertas. Patches de segurança para cada versão do <i>Suse</i> são mostrados em vermelho na página de atualizações. Uma curta descrição da vulnerabilidade corrigida pelo patch é fornecida. |

1 Todas as distribuições indicam, no assunto da mensagem, que o tema é segurança.

# ▶ Notebook popular mais fácil

O Governo brasileiro oficializou a utilização do FAT (Fundo de Amparo ao Trabalhador) para aquisição de equipamentos de Informática cobertos pelas especificações do Programa Computador para Todos. A maior novidade da mais recente etapa desse que é, sem dúvida, um dos programas de maior êxito do atual governo, é a inclusão de notebooks de até R\$ 1,8 mil ao programa. Para obedecer as especificações do programa, um notebook deve ser equipado com sistema operacional e aplicativos de Código Aberto,



frequência mínima de processamento de 1,4 GHz, memória principal mínima de 512 MB, unidade de disco rígido interna com capacidade mínima de 40 GB e ao menos um gravador de CD e leitor de DVD (*combo*). A tela deve ter tamanho mínimo de 14 polegadas TFT colorido (matriz ativa).

Outras especificações: a interface de som deve ser compatível com Sound Blaster PCI 16 Bits, *Plug & Play*. O sistema multimídia do notebook deve possuir dois alto-falantes estéreo embutidos ao equipamento, assim como três interfaces de comunicação USB 2.0, interface de comunicação sem fio (rede Wi-fi) e adaptador 110/220V. De acordo com as normas, o notebook não deve pesar, também, mais de três quilos. ■

## ▶ Cookies do Google vão se autodestruir

Em um lance digno de filmes do estilo “Missão Impossível”, o Google anunciou que, a partir de agora, os *cookies* ligados a seus sites irão se autodestruir, apagando-se do disco rígido, em dois anos. O Google tomou essa decisão após ouvir especialistas das áreas de segurança e privacidade digital, que sugeriram a substituição da atual política de permanência de cookies da empresa, que hoje faz com que um arquivo desse tipo seja programado para permanecer em disco rígido até 2038.

“Concluimos que seria bom para a privacidade a redução significativa do tempo de vida dos nossos cookies” declarou Peter Fleischer, do conselho de privacidade do Google, ao divulgar a recente medida da empresa. Cookies são arquivos de texto que são armazenados no disco rígido de computadores que acessam uma determinada página – serviços do Google, no caso – e que guardam informações do usuário acerca de suas configurações personalizadas para uma determinada página, preferências, buscas mais recentes, dentre outros elementos. ■



## ▶ Firefox alcança 30% dos usuários na Europa

O navegador Firefox vem conquistando audiência de modo crescente na Europa, segundo a consultoria XiTiMonitor. Ela afirma que o navegador responde por 27,8% do mercado para exibição e utilização de conteúdo Web na Europa, atingindo participação de quase 50% em países como Eslovênia e Finlândia. Esses índices de presença agressiva do Firefox, na Europa, contrastam com sua situação nos Estados Unidos, onde sua presença é até 50% menor.

De acordo com dados referentes a julho, o Firefox atingiu 27,8% dos navegadores usados na Europa, aumento de 3,6 pontos percentuais em comparação com o mês de março. A pesquisa também afirmou que o Internet Explorer, da Microsoft, acumulou 66,5% de participação, com Opera, Safari e Netscape completando os cinco primeiros com 3,5%, 1,7% e 0,3%, respectivamente.

A pesquisa recolheu dados e traçou suas linhas de informação com a consulta de 96 mil sites, localizados em 32 países europeus – incluindo Irlanda, Ucrânia, Finlândia e Grécia. Eslovênia, Finlândia e Eslováquia tiveram a maior penetração do Firefox, com o navegador de código aberto atingindo 47,9%, 45,4% e 40,4%, na ordem, durante a semana entre 2 e 8 de julho. Outras fortes participações do Firefox estão na Alemanha (38%), Polônia (39,6%) e Irlanda (38,6%). ■

## ▶ Apple adquire CUPS

O projeto CUPS anunciou, por meio de seu site ([www.cups.org](http://www.cups.org)), que a Apple, conhecida empresa e representativo rótulo da indústria de tecnologia, contratou Michael Sweet, principal desenvolvedor do CUPS, além de adquirir o código-fonte do sistema. Segundo Sweet, “o CUPS ainda será um Software Livre, governado pela licença GPL2, e eu continuarei a desenvolvê-lo dentro da Apple”. Consta, no entanto, segundo fontes da mídia especializada, que a Apple não comprou apenas o código-fonte do CUPS, mas também registrou as marcas Common UNIX Printing System, CUPS e o logotipo como sendo de sua propriedade.

A Apple já havia, em 2002, licenciado o CUPS para ser o sub-sistema de impressão do Mac OS X. ■

## ▶ Justiça condena crackers no Pará

A Justiça Federal do Pará condenou um total de 65 pessoas por crimes de lavagem de dinheiro pela Internet e estelionato, além de fraude ideológica e documental. Os acusados, agora formalmente condenados, acessavam de forma fraudulenta contas correntes e de poupança de clientes dos bancos Caixa Econômica Federal (CEF), Banco do Brasil, Banco Itaú e Banco do Estado de Goiás. ■

## ▶ Linux legaliza uso de máquinas de bingo

As máquinas de bingo e video-pocker apreendidas pela Receita Federal no Estado do Rio de Janeiro estão sendo transformadas em computadores de baixo custo para utilização em projetos sociais. Reaproveitando o hardware original das máquinas – placa-mãe, monitor e fonte – e utilizando Linux como sistema operacional na reinstalação das máquinas, além de adaptar mouse e teclado ao gabinete, o projeto, coordenado pelo Professor Marcos Paulo Monteiro, consegue gastar apenas R\$ 30,00 por máquina, já pronta para acesso à Internet. ■

## ▶ Da GPLv2 para GPLv3, a passos curtos

Segundo informações da Palamida ([www.palamida.com](http://www.palamida.com)), empresa que vem acompanhando a migração de projetos de Software Livre da versão 2 da licença GPL (GPLv2) para a sua versão 3 (GPLv3), até o momento apenas 194 projetos migraram sua versão de licença. Esse número representa pouco mais de 1% dos softwares que fazem uso do formato General Public Licence (GPL).

A lista de conversões para a GPLv3 inclui poucos projetos de grande relevância. O mais famoso é o Projeto Samba, servidor de arquivos que compatibiliza redes Linux com servidores de arquivos que utilizam SMB (*Server Message Block*), o protocolo de compartilhamento de diretórios de propriedade da Microsoft. Nomes como o compilador GCC, a interface gráfica KDE e o próprio kernel Linux ainda preferem permanecer sob a versão 2 da licença GPL. ■

## ▶ Fim da linha para o PHP 4

Após exatos três anos de desenvolvimento, a Fundação PHP ([www.php.net](http://www.php.net)) finalmente lançou, oficialmente, a versão 5 de sua linguagem. Apesar dos contínuos acréscimos e melhorias do PHP 4 durante o processo de desenvolvimento da versão 5, executada nesse mesmo triênio, a fundação resolveu descontinuar a versão 4, sob a alegação de que “o PHP 5 é mais rápido, estável e alcança maior produtividade em relação ao seu antecessor”. Além disso, o PHP 6 já está a caminho. O suporte ao PHP 4.4 será dado pela equipe do projeto até o dia 8 de agosto de 2008. ■

# LINUX IMPACTA

O Maior e Melhor Centro de Treinamento de Linux do Brasil

"O mercado de Linux movimentará US\$35 bilhões"\*

### Preparatório para a Certificação LPI

Linux LPI 101 - Fundamentos | Linux LPI 101- Implementação e Adm.  
Linux LPI 102 - Gerenciamento e Manutenção | Linux LPI 102 - Impl.  
de Infra-estrutura de Redes

**Avançado:** Linux Shell Script | LDAP | Apache | Samba | Firewall

Camiseta + 26% de desconto nas Certificações Linux

Faça Provas LPI no Centro de Exames  
Prometric-Impacta. Saiba o Resultado na Hora!

Participe das Palestras Gratuitas. Inscreva-se no site: [www.impacta.com.br](http://www.impacta.com.br)



**IMPACTA**  
TECNOLOGIA  
Divisão Networking



Av. Paulista, 1009 - 9º andar  
Tel. (11) 3254-2200 | [www.impacta.com.br](http://www.impacta.com.br)

\*Fonte: IDC. Número previsto para o ano de 2008

# ► Sun amplia oferta de virtualização de armazenamento

A Sun Microsystems, um dos grandes players do mercado de virtualização de recursos, adicionou ao seu portfólio de soluções de infra-estrutura de armazenamento de informações o sistema de discos Sun StorageTek 9990V. O anúncio consolida a parceria de sucesso com a Hitachi Data Systems.

O novo dispositivo amplia as opções de escalabilidade da família de produtos Sun StorageTek 9900, que já conta com os modelos 9985 e 9990. Além de ganho de desempenho, o modelo incorpora o software de Provisionamento Dinâmico, que melhora a utilização e permite adicionar capacidade de armazenamento sob demanda ao hardware central, de maneira simples e transparente.

A Sun e a Hitachi Data Systems possuem um bom histórico de consolidação de parceria tecnológica, sempre voltado à resolução de problemas envolvendo o armazenamento de informações e a redução do Custo Total de Propriedade (TCO) nos ambientes de grande porte de seus clientes.

“Auxiliamos milhares de companhias dos mais diversos segmentos, em todo o mundo, endereçando diferentes desafios em gerenciamento de dados. Agora estamos satisfeitos por trazer ao mercado novas possibilidades com a funcionalidade de provisionamento dinâmico”, afirmou Nigel Dessau, vice-presidente de Marketing e Operações de Storage da Sun Microsystems.

Ao mesmo tempo em que anuncia um incremento na sua linha de produtos de storage, a Sun também demonstra apoio mais ativo ao projeto OpenSolaris e ao seu desdobramento nas aplicações de storage da empresa. O suporte ao sistema de arquivos ZFS – sistema de 128 bits que suporta até 18 bilhões de bilhões ( $18.4 \times 10^{18}$ ) de vezes mais dados do que os atuais formatos de 64 bits – já foi corretamente portado para o OpenSolaris, bem como o de uma grande variedade de outros formatos de armazenamento, como o iSCSI e o SAM-FS (*Storage Archive Manager*). ■

## ► Intel também mostra seus números

Em 18 de julho último a Intel Corporation anunciou – seguindo o seu calendário fiscal – a receita do segundo trimestre deste ano, que chegou ao valor de US\$ 8,7 bilhões. A renda operacional da empresa chegou a US\$ 1,35 bilhões; a renda líquida alcançada foi de USD 1,3 bilhões e o lucro por ação (EPS – sigla em inglês) chegou a US\$ 0,22.

“A execução operacional da Intel continuou a fortalecer-se resultando em destacadas linhas de produtos e um crescimento sólido, ano após ano, da receita”, declarou o presidente e CEO da Intel, Paul Otellini. “Estamos felizes de que nossos esforços para aumentar a eficiência da companhia estejam resultando em um crescimento do lucro, além do crescimento da receita”. ■

## ► Código Aberto cresce nas aplicações críticas

Segundo recente estudo internacional encomendado pela Unisys Corporation e conduzido pela Forrester Consulting, líder em pesquisas de tecnologia e mercado, cresceu entre as empresas européias e norte-americanas que avaliaram ou estão usando Código Aberto, a percepção dos benefícios desse tipo de software. A melhoria na percepção geral do mercado corporativo em relação aos sistemas e aplicações de Código Aberto, se deu, sobretudo, em áreas consideradas de grande importância, como geração de receitas e melhoria do atendimento ao cliente.

Boa parte dos quase 500 entrevistados – tomadores de decisões de TI, entre os quais altos executivos – também utiliza o Código Aberto para aplicações de missão crítica, embora demonstre preocupação com a disponibilidade de serviços para aproveitar plenamente o valor das soluções desse tipo de software.

Os motivos mencionados com mais frequência para a adoção de tecnologias de Código Aberto foram a capacidade de utilizar software sem pagar por licenças (62% dos entrevistados) e a flexibilidade dos padrões abertos (quase 80% dos pesquisados), com um maior suporte aos padrões abertos por parte dos desenvolvedores, o uso do código sem restrições e a fuga da dependência de um único fornecedor.

Mais da metade dos entrevistados (58% na América do Norte e 51% no Reino Unido e na Europa Continental) já fazem uso de software de Código Aberto em aplicações de missão crítica; mais de 79% o empregam na infra-estrutura de aplicativos, como bancos de dados, servidores Web e de aplicação.

A versão integral do relatório pode ser obtida em [http://www.unisys.com/eprise/main/admin/corporate/doc/Forrester\\_research-open\\_source\\_buying\\_behaviors.pdf](http://www.unisys.com/eprise/main/admin/corporate/doc/Forrester_research-open_source_buying_behaviors.pdf). ■



# SERVIDOR VIRTUAL

## PELO MENOR PREÇO DO MERCADO

- **TRANSFERÊNCIA ILIMITADA SEM REDUÇÃO DE VELOCIDADE**
- **SEM TAXA DE INSTALAÇÃO**
- **20GB DE ESPAÇO EM DISCO**
- **256 MB DE RAM**
- **LINUX OU WINDOWS**
- **DISTRIBUIÇÕES: CENTOS, UBUNTU, DEBIAN E SLACKWARE.**  
CONSULTE SOBRE OUTRAS.
- **ACESSO DE ROOT**
- **R\$ 199,00/MÊS**

SOLICITE AGORA MESMO

[WWW.XPGPRO.COM.BR/LM](http://WWW.XPGPRO.COM.BR/LM)



## Red Hat encerra trimestre fiscal com receita recorde

A Red Hat, conhecida empresa que detém os direitos da distribuição Linux do mesmo nome e de outras tecnologias de Código Aberto, e que possui parceiros de peso como IBM e Intel, encerrou o primeiro trimestre fiscal do ano de 2008 com uma receita total de US\$118,9 milhões – montante 42% maior em comparação ao mesmo período do ano passado e 7% maior em relação ao trimestre anterior. A receita com subscrição de software foi de US\$103,0 milhões, 44% superior ao trimestre equivalente do ano anterior e 7% maior frente ao quarto trimestre fiscal de 2007. O atual ano fiscal da empresa terminará em 29 de fevereiro de 2008.

A receita líquida desse trimestre da Red Hat foi de US\$16,2 milhões, ou o equivalente a US\$0,08 por ação da empresa, um ganho em relação aos US\$0,07 por ação – totalizando US\$13,8 milhões – negociados no mesmo trimestre do ano passado. O rendimento líquido ajustado não-GAAP do período foi de US\$33,7 milhões, ou US\$0,16 por ação, após ajustes de compensação de ações e despesas com impostos. ■



**redhat**

## Bull lidera novo consórcio

O Grupo Bull, um dos maiores nomes em Tecnologia da Informação na Europa, está juntando forças com vários dentre os maiores nomes em HPC da França, Alemanha, Espanha e Reino Unido para lançar o consórcio ParMA (Programação Paralela para Arquiteturas com Vários Núcleos). O principal objetivo do novo consórcio é ajudar a comunidade HPC (Computação de Alto Desempenho) a usufruir dos benefícios dos processadores mais rápidos e desenvolver tecnologias inovadoras flexíveis e abertas que aproveitem todas as vantagens das arquiteturas de vários núcleos.

Atualmente, os fabricantes de processadores oferecem arquiteturas com vários núcleos, enquanto a maioria das aplicações, grande parte delas seqüenciais, só desfrutam da potência de um único núcleo. O projeto ParMA, a partir dessa constatação, refinará e definirá, em um primeiro momento, a prioridade dos requisitos reais do mercado. Uma primeira versão da tecnologia ParMA será apresentada no primeiro ano do projeto, visando experimentos e conformação de mercado com uma ampla variedade de aplicações. Uma segunda versão será liberada no segundo ano do projeto para mais testes. ■

## Fujitsu traz novidades em servidores blade

A Fujitsu anuncia performance aprimorada e consumo de energia reduzido para seu servidor blade Primergy® BX600 S3. O ecossistema da solução é composto pelo chassi blade Primergy BX600 S3 e pelo servidor blade quadcore (ou seja, com quatro processadores executando funções de processamento) Primergy BX620 S4, que recentemente estabeleceu um recorde mundial em *benchmark* para um servidor com duplo soquete.

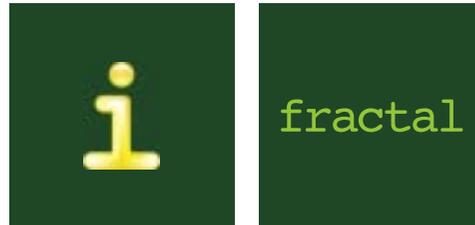
O chassi blade Primergy BX600 S3 atua como a interface entre os servidores blade no chassi e todos os demais componentes do sistema, tomando-se a plataforma ideal para consolidação e virtualização. Seu novo *midplane* triplica a velocidade de *throughput* de I/O, o que é essencial para infra-estruturas virtualizadas de alto desempenho, além de oferecer interfaces para 60 cabos de 1 Gigabit e conexões *Fiber Channel* de ponta a ponta, com velocidades de até 4 Gigabits por segundo. ■

Entrevista com Felipe Waltrick, Diretor de Tecnologia da iFractal

# Linux no ponto

A iFractal desenvolveu um inovador sistema de controle de ponto e acesso no SENAC Amazonas, e explica as vantagens do Código Aberto para essa implementação.

por Pablo Hess



**Linux Magazine»** Como foi o envolvimento da iFractal com o SENAC do Amazonas?

**Felipe Waltrick»** O SENAC precisava de um novo sistema de controle de ponto para seus funcionários, e também de controle de acesso para seus alunos. A licitação lançada por eles exigia que a solução fosse baseada em Software Livre, e por isso nós nos diferenciamos dos outros concorrentes potenciais, que não integravam soluções em Software Livre.

**LM»** Como é a solução desenvolvida?

**FW»** É um equipamento de ponto e controlador de acesso, porém com uma diferença. Em vez de cada aluno ou funcionário possuir um cartão, que pode ser perdido e

utilizado por terceiros até mesmo para fins ilegais, o leitor de ponto que integramos efetua uma leitura biométrica de impressões digitais.

Há diversas formas de implementar isso. Nossa primeira opção envolvia o armazenamento das digitais em um único servidor, que seria consultado pelas catracas, através da rede, a cada acesso. Essa abordagem logo foi excluída, pois a dependência de um servidor ou da rede era indese-

equipamentos de ponto são autônomos, não dependendo da rede. Em caso de queda de energia, eles continuam funcionando normalmente, pois têm no-breaks. Se a rede sair do ar, a catraca continua verificando e registrando os acessos, e somente quando for restaurada a atividade normal os dados serão repassados ao servidor.

Como há diversas catracas (sete, para ser mais preciso), é necessária uma sincronização das bases de dados



**Foi de fundamental importância termos acesso ao protocolo de comunicação dos equipamentos.**



**Figura 1** Felipe Waltrick, Diretor de Tecnologia da iFractal

jável; seria muito desagradável, em caso de falta de energia ou queda da rede, ter que esperar a restauração dos serviços para retomar os acessos normais ao edifício.

Utilizamos então uma outra técnica. As digitais são cadastradas e armazenadas localmente nas catracas, que comportam, cada uma, alguns milhares de cadastros. Dessa forma, os

de digitais entre os equipamentos, o que é feito através de funções embutidas nos mesmos, que exportam e exportam as bases de dados entre si. Assim, além da distribuição das bases de dados, é feito um backup da base consolidada.

Com essa infraestrutura funcionando corretamente e de forma estável, implementamos sobre ela nosso software.

Nessa etapa, foi de fundamental importância termos acesso ao protocolo de comunicação dos equipamentos, pois o acesso aos mesmos é feito através da rede, por TCP/IP.

O equipamento, no mesmo momento, consulta a tabela de restrições para verificar se a pessoa em questão tem permissão para acessar o edifício naquele dia e horário, e então permite ou impede o acesso da mesma.

Assim que é feito o acesso, os dados sobre o evento (ID da pessoa, nome, hora e quaisquer outros dados interessantes) são enviados para o servidor em que funciona nossa aplicação web. Nesse mesmo momento, essas informações já podem ser visualizadas e controladas on-line através de consultas ao servidor web.

Em nossa solução, empregamos o FreeBSD como sistema operacional, PostgreSQL como banco de dados, Ruby para a comunicação de alto nível, Java nas funções de controle de tempo, PHP como linguagem web e C para interagir com o hardware do equipamento.

**LM» Houve alguma dificuldade decorrente do uso de Software Livre?**

**FW»** Sim, mas conseguimos transpô-la com certa facilidade. Graças à exigência do SENAC em relação ao uso de Software Livre, o fornecedor dos equipamentos de ponto eletrônico precisava que a integração de seus produtos fosse realizada em Código Aberto. Caso contrário, o fornecedor perderia a venda.

Em virtude disso, eles facilitaram bastante nosso trabalho, enviando-nos amostras dos equipamentos e também o código-fonte e especificações do protocolo usado. Para nós, isso faz total sentido, já que eles (os fornecedores) são os únicos a fabricar esses produtos, e portanto devem incentivar sua adoção em qualquer ambiente, sem temer a abertura de suas especificações.

Segundo o fornecedor, eles já haviam procurado outros integradores para realizar essa tarefa, mas nenhum se comprometera a usar Software Livre.

**LM» Quais foram os benefícios do uso de Código Aberto nessa implementação?**

**FW»** O fornecedor do hardware distribuiu uma API que já trata as funções de baixo nível, porém ela é feita para sistemas Windows®. Como recebemos do fornecedor o próprio protocolo e as especificações do equipamento, conseguimos uma pequena inovação, que na realidade deixou o responsável no SENAC bastante impressionado: quando era feito o acesso, a pequena tela do equipamento costumava mos-

instantaneamente, caso seja possível. Se não é adotado o modo off-line, isso é, aquele em que as consultas são feitas apenas localmente, até que seja reestabelecido o estado funcional da rede ou da energia.

Somente com nossa implementação o gerente ou diretor pode acompanhar a frequência dos alunos e funcionários, pelo servidor web, em tempo real, e isso é altamente desejável.

**LM» Qual foi a importância, do ponto de vista do cliente, do uso de Código Aberto no projeto?**

**FW»** Boa pergunta, pois, do ponto de vista de quem desenvolve, as licenças necessárias não têm custo extra, já que tudo será repassado ao cliente através do preço da solução.



**O cliente tem posse do código-fonte que comprou, o que significa que não será mais dependente do fornecedor.**

trar ao usuário algum número críptico, e implementamos a funcionalidade de exibir o nome da pessoa cadastrada para aquela digital. Ou seja, quando o aluno passa pela catraca, ele não vê seu número de matrícula, e sim seu nome, na telinha.

Além disso, de posse do protocolo, conseguimos também acelerar todo o processo de validação do acesso e comunicação com o servidor.

Modéstia à parte, nossa implementação, ao final, foi bastante superior às que simplesmente utilizaram a API que o fornecedor entrega. O modo on-line implementado pela API, por exemplo, não é de fato on-line, pois ele utiliza um *buffer* na máquina local e realiza a comunicação com o servidor de forma assíncrona, portanto potencialmente mais lenta. Enquanto isso, da forma como fizemos, a comunicação ocorre

Portanto, o único afetado é realmente o cliente. Usando Software Livre, ele não é obrigado a pagar qualquer licença de sistema operacional, banco de dados ou servidor web. Além disso, como o sistema desenvolvido também tem seu código aberto, o cliente tem posse do código-fonte que comprou, o que significa que não será mais dependente do fornecedor – no caso, de nós.

**LM» Mas essa postura não parece ir contra o seu próprio trabalho?**

**FW»** De forma alguma. Costumamos dizer que temos clientes e não reféns. Não queremos que ele seja nosso cliente apenas por falta de opção ou à força. Ou seja, fornecemos um produto e um serviço a nosso cliente, e ele deve ter a liberdade de escolher se nossa qualidade é boa ou não para suas necessidades e expectativas. ■

Mudando para manter seu significado

# Novidades na GPLv3

A Licença Pública Geral GNU, GPL, atingiu sua terceira versão, com o objetivo de esclarecer pontos que permaneciam obscuros para alguns.

por Alexandre Oliva



A GNU GPL é a licença de Software Livre preferida entre os desenvolvedores. Mais de dois terços de todos os softwares livres a utilizam. Exemplos proeminentes são o GCC, o GNU Emacs e o kernel Linux (não confundir com sistema operacional GNU/Linux, com programas sob várias licenças). A grande maioria desses softwares permite relicenciamento sob a GPLv3, lançada em 29 de junho último, daí a importância de saber o que ela traz de novo.

## A idéia da GPL

Software é livre quando o usuário tem quatro liberdades fundamentais:

- ◆ Executar o programa para qualquer propósito;
- ◆ Estudá-lo e adaptá-lo às suas necessidades;
- ◆ Redistribuí-lo, da forma como foi recebido;
- ◆ E modificá-lo e distribuir as modificações.

A GPL foi a primeira licença *copyleft* a usar direito autoral não

só para conceder permissões, mas também para garantir que o programa permaneça livre, como explica seu preâmbulo. Já na GPLv1, a primeira versão da licença, explicitava-se o respeito do autor às liberdades do usuário, porém condicionando o gozo de algumas das liberdades ao respeito às liberdades dos demais usuários do software. Viola a licença quem distribui o programa impondo restrições adicionais a quem o receba, direta ou indiretamente.

Potenciais restrições, há muitas. Negar acesso ao código-fonte (a forma mais adequada para modificar o software) ou usar termos de licenciamento potencialmente mais restritivos são duas já explicitamente proibidas na GPLv1. Na GPLv2, foi adicionada a cláusula “liberdade ou morte”, explicando que qualquer obrigação de restringir a liberdade de outros implica não poder distribuir o programa. Isso coibia a distribuição sob licenças restritivas de patentes, ou sob quaisquer outras condições que

impedissem a distribuição “sem restrições adicionais”, proibição implícita na GPLv1, cláusula 6.

## “Restrições adicionais” na GPLv3

A GPLv3 agrega vários exemplos de restrições adicionais que ferem o princípio do preâmbulo das GPLs de que os direitos concedidos na GPL acompanham a obrigação de não negá-los a outros usuários do programa.

Distribuir software e depois limitar seus usos através de patentes nele implementadas é uma restrição adicional que, pela lei norte-americana, a GPLv2 já proibia de forma implícita. Na GPLv3, a licença de patente passou a ser explícita. O uso de patentes para impor restrições adicionais às liberdades de qualquer outro usuário com relação ao programa, modificado ou não, viola a licença.

Leis draconianas que criminalizam a desativação de DRM (*Digital*

*Restrictions Management*, ou Gestão Digital de Restrições—são mediadas técnicas de controle de acesso à informação em formato digital), mesmo para usos permitidos por lei, também não podem ser usadas para impedir usuários de modificar o software GPL.

O uso de assinaturas digitais para impedir a execução de versões modificadas do software GPL que acompanha dispositivos como o vídeo-cassete digital TiVo também é restrição adicional. Para quem acredita que os termos “sem restrições adicionais” da GPLv2 não proibiam a prática da *Tivoização*, a GPLv3 esclarece que o usuário deve receber, juntamente com o código-fonte, informação de instalação necessária para o pleno gozo das liberdades. Vale ressaltar que software em ROM em princípio não caracteriza restrição adicional, e a GPLv3 permite explicitamente essa prática.

Terceirização não exige as partes do cumprimento das obrigações da GPL. Novell e Microsoft fecharam um acordo em que uma distribui o software e a outra impõe restrições, através do licenciamento restritivo de supostas patentes. Nessas condições, Novell seria impedida de distribuir o software, mas a GPLv3 abre uma exceção para que ela possa distribuir software GPLv3 em nome da Microsoft, assim estendendo a licença de patentes para todos. A exceção não vale para acordos posteriores.

Ou seja, a GPLv3 não impõe de fato novas restrições. Ao contrário, deixa claro que novas formas de restrição não fogem do escopo de “sem restrições adicionais”.

## Mais vantagens

Explicitando todos esses casos de restrições adicionais não permitidas, a GPLv3 traz muito mais seguran-

ça para o usuário e mais facilidade para o titular do programa coibir tais infrações da licença. Mas há várias outras novidades da GPLv3 para trazer tranqüilidade ao usuário e ao desenvolvedor.

Quem recebe Software Livre em protocolos P2P, como *bittorrent*, pode estar violando a GPLv2, pois está também distribuindo o programa recebido, provavelmente desacompanhado de seus fontes. A GPLv3 permite explicitamente essa forma de distribuição, sem exigir aceitação dos termos da licença, desde que a informação sobre como obter os fontes correspondentes seja disponibilizada junto à oferta do binário. A GPLv3 também flexibiliza a exigência de oferta de fontes em mídia tangível, permitindo a oferta dos fontes para download, inclusive em servidores mantidos por terceiros. Só não exige o distribuidor dos binários da responsabilidade de providenciar os fontes para o usuário, caso o servidor sugerido deixe de oferecê-los.

A GPLv3 introduz compatibilidade com a conhecida licença *Apache 2.0*, assim como com a menos conhecida *GNU Affero GPL*, uma licença que exige a oferta dos fontes mesmo para usuários remotos do software, como por exemplo software que roda num servidor na Internet.

Advogados vão apreciar o ar de “legalês” da GPLv3, a redução de ambigüidades, a internacionalização alcançada, evitando o uso de termos específicos da lei norte-americana de direito autoral, a permissão para adicionar termos de ausência de garantia e os dois mecanismos de perdão de infração acidental: cumprimento da licença em até 30 dias após a primeira notificação de infração, ou cumprimento por 60 dias consecutivos após a última infração, se não houver notificação antes do fim do prazo.

## Balanço positivo

Ainda que certa cautela seja justificável até que o entendimento da nova licença se consolide, temer a GPLv3 só faz sentido para quem pretendia modificar ou distribuir softwares sob a licença GPL e impor restrições adicionais aos demais usuários, esperando, com isso, que prevalecessem interpretações da GPLv2 favoráveis a seus propósitos.

O espírito da GPL sempre foi garantir que o software permaneça livre para todos os seus usuários, inclusive para exploração comercial. No entanto, é importante ressaltar que aquelas pessoas que não compartilham esse espírito de liberdade são tão livres para se abster de modificar ou distribuir software sob a GPL quanto sempre foram. Ou seja, pode-se simplesmente usar os binários de qualquer programa sob essa licença, sem alterar seu código-fonte.

Com a GPLv3, o usuário, programador ou empresário tem ainda mais garantias de que nenhum dos desenvolvedores ou distribuidores do software vá impedi-lo de executar, modificar ou distribuir o software, ao menos enquanto os primeiros estiverem dispostos a também respeitar as liberdades dos demais usuários do software. Se suas atividades dependem do uso desse software, essa segurança adicional há de ser muito bem-vinda. ■

## Mais informações

[1] URL oficial deste documento:  
<http://www.fsfla.org/svnwiki/blogs/1xo/pub/gplv3-novidades.pt.html>

## O autor

**Alexandre Oliva** é engenheiro, mestre em Ciências da Computação e funcionário da Red Hat, além de co-fundador e conselheiro da Free Software Foundation Latin America.

Entrevista com Roberto Prado, Gerente de Estratégias da Microsoft

# Propriedade intelectual e Linux

A Microsoft anunciou que o Linux viola diversas de suas patentes. Conversamos com Roberto Prado, Gerente de Estratégias da empresa, sobre o significado disso no Brasil.  
por Pablo Hess

**Microsoft®**

**Linux Magazine»** No Brasil, o software está sujeito unicamente à legislação de direitos autorais. Como a Microsoft pretende agir no país, frente ao Linux, em relação a esse tema?

**Roberto Prado»** Nós investimos em pesquisa e desenvolvimento, então não podemos negar que acreditamos em propriedade intelectual. Porém, reco-

de uso de nossos produtos. Como todos podem ver, essa tem sido a estratégia da Microsoft, com diversos acordos tendo sido firmados nos últimos meses.

A terceira opção é o confronto legal, no qual não há acordo e ambos os envolvidos se julgam de posse de evidências suficientes a seu favor.

**LM»** Os acordos que tratam da interoperabilidade passam também pela propriedade intelectual. Porém, no Brasil isso não seria necessário. Por que, então, aplicar esses acordos no Brasil?

**RP»** Temos que nos focar em algumas distribuições (aquelas com que temos acordos); não podemos escrever um guia genérico de interoperabilidade, pois ele não funcionaria.

Diversos softwares, como o Samba, por exemplo, têm particularidades em cada distribuição, e por isso é importante buscarmos a interoperabilidade de forma específica para elas.



*Todos sabemos que não é interessante um sistema que só conversa consigo mesmo.*

nhecemos que há cenários que teremos que reinventar, e por esse motivo criamos a iniciativa *Shared Source*.

Temos debatido isso na Microsoft Brasil, e chegamos à conclusão de que há três posturas possíveis a adotar, comuns a qualquer indústria.

A primeira abordagem é não fazer nada quanto a violações de nossa propriedade intelectual. Continuaríamos investindo em pesquisa e desenvolvimento, e daí se originariam boas e más idéias. Alguns desses frutos talvez viessem a ser imitados ou clonados, e não tomaríamos qualquer atitude para impedir isso.

A segunda postura é estabelecer acordos para a concessão do direito

**LM»** Qual é o objetivo em comum dos acordos firmados pela Microsoft até o momento?

**RP»** Os acordos oferecem oportunidades positivas para os dois lados envolvidos. Nosso concorrente comum é o *mainframe*, pois a arquitetura x86 está ganhando mercado dos antigos RISC – também por conta de tecnologias como a virtualização.

Até o momento nenhum cliente meu recomendou que eu não fizesse acordos de interoperabilidade, nem tive retornos negativos dos acordos já feitos. Todos sabemos que não é interessante um sistema que só conversa consigo mesmo.

**LM»** A alegação das 235 patentes da Microsoft supostamente violadas pelo Linux, então, teve como objetivo principal estimular acordos?

**RP»** Achei essa matéria um pouco infeliz, mas infelizmente não posso falar da declaração, pois não a vi. O intervalo entre o que saiu na mídia e a realização dos acordos foi muito curto. Além disso, o impacto de declarações como essa é sempre negativo, então não creio que esse tenha sido, de fato, o objetivo.

De qualquer forma, esse assunto não é dominado por muitas pessoas da área de tecnologia, tanto que é tratado apenas por advogados. ■

# Guia de TI

Soluções em Tecnologias Abertas

LINUX NEW MEDIA  
The Pulse of Open Source



## Garanta já sua vaga para o Guia de TI 2008!

Cadastre-se agora e apareça  
gratuitamente na maior  
e mais completa lista  
de empresas que oferecem  
soluções de TI baseadas  
em tecnologias abertas.

Cadastre a sua solução gratuitamente!  
[www.guiadeti.com.br](http://www.guiadeti.com.br)

**Cadastre-se:**  
11 4082-1300

**Publicidade:**  
11 4082-1300

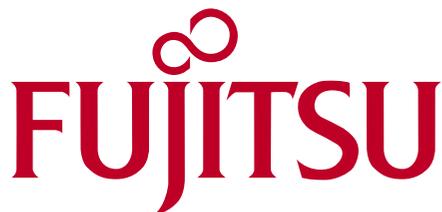
**Assinatura:**  
11 4082-1300

Platform Solutions Center da Fujitsu

# Testes abertos

A Fujitsu cria um laboratório e oferece aos clientes e ISVs todo um ambiente para o teste de plataformas e soluções.

por Pablo Hess



Janet Goulden - [www.sxc.hu](http://www.sxc.hu)

Funcionários responsáveis pela homologação de softwares já têm um trabalho mais fácil. Inaugurado em maio, o *Platform Solutions Center* da Fujitsu funciona no andar térreo do prédio da empresa, no centro da movimentada cidade de São Paulo. Edson Siqueira, diretor de vendas da companhia e nosso entrevistado na edição 30 da **Linux Magazine**, apresentou o conceito envolvido na criação desse verdadeiro laboratório de testes feito para o cliente.

“Tradicionalmente, o cliente solicita a nós, fornecedores, o hardware, enquanto adquire de outro fornecedor o sistema operacional, suas aplicações etc.”, explicou Edson. A Fujitsu, no entanto, já estava um pouco à frente nesse quesito, fornecendo, além do hardware, os equipamentos de *storage*, parte do *middleware* – apoiada nos principais *players* do mercado de bancos de dados, como Oracle e BEA, por exemplo – e, por último, até mesmo o sistema operacional – Red Hat e Suse – caso o cliente desejasse.

No entanto, a redução dos pontos de contato do cliente com os

inúmeros fornecedores não era suficiente para livrá-lo completamente de complicações ao formular e executar um projeto envolvendo essas partes.

Um ISV (*Independent Software Vendor*), ao ser consultado a respeito do que é necessário para que o projeto do cliente possa ser executado, geralmente não possui métricas muito definidas. Como resultado, o ISV pode apenas oferecer termos de comparação com projetos semelhantes executados por outros clientes: “No cliente X, um projeto semelhante usou o sistema operacional Y, um hardware com tais especificações e o banco de dados Z”. Segundo Edson, “é claro que, por questões de risco, o ISV é obrigado a superdimensionar os requisitos de infra-estrutura para sua solução. O cliente, diante disso, opta por economizar em outras partes do projeto, o que se traduz em aumento do risco para ele, seja por uma estrutura superdimensionada ou mesmo subdimensionada.”

Nas palavras do executivo, “o PSC atua justamente nesse pon-

to, com o objetivo de responder todas essas questões, alinhadas aos projetos individuais, para arquiteturas abertas. Os ISVs geralmente dispõem de ambientes de pequeno porte para testes, pois tais ambientes têm custo bastante alto. A idéia é que os ISVs tragam para o ambiente da Fujitsu, com máquinas de médio e grande porte, como as que serão encontradas no cliente, suas soluções em desenvolvimento. Com isso, podem realizar *benchmarks* e, principalmente, criar métricas, como o tempo para a entrega do sistema ou o volume de transações por minuto. Em suma, o objetivo do PSC é minimizar o risco para o cliente, ao mesmo tempo garantindo a ele uma implementação facilitada, e possibilitando ao ISV a formulação de métricas para suas soluções”.

A postura da Fujitsu favorece o sucesso de ISVs pequenos e iniciantes tanto quanto o daqueles já estabelecidos no mercado, contribuindo assim para um mercado mais dinâmico, calcado, sempre, sobre padrões abertos. ■



Você está  
**assustado**  
 com os altos custos  
 de armazenamento  
 usando tecnologia  
 Fibre Channel  
 e iSCSI?

# A resposta é AoE!

ATA-over-Ethernet = armazenamento **Rápido, Confiável e Simples.**

[www.coraid.com](http://www.coraid.com)



## EtherDrive® SRxxxx

- Appliances RAID rápidas e flexíveis com slots para discos SATA hot swap
- Conheça em nossa linha completa de Appliances EtherDrive® Storage e VirtualStorage, além de nossos Gateways



1. Storage Ethernet de 10 Gigabits sem o overhead do TCP/IP!
2. Expansível sem limites, com o menor preço possível!!
3. Se você quiser mais espaço ... basta comprar mais discos – simples assim!!!

Visite-nos em [www.coraid.com](http://www.coraid.com)



SOA: só tecnologia resolve?

# Edgar Silva

Você sabe o que é necessário para que a Arquitetura Orientada a Serviços realmente tenha efeitos positivos?  
por Edgar Silva

A Arquitetura Orientada a Serviços – SOA – é a grande *buzzword* do momento no Brasil. Onde quer que se vá, setor privado ou governo, deparamos-nos com essa sigla, que pode representar o tudo ou o quase nada, dependendo da sua percepção. Este artigo pretende posicioná-lo em relação a alguns aspectos muitas vezes negligenciados devido à preocupação com os nós da gravata.

Você já assistiu a trinta palestras e leu cem revistas falando de SOA, mas eu pergunto: você já tem opi-

executivos é direcionado ao *core business* da empresa; mas como fica o controle sobre o conhecimento? Como saber se a empresa não está desenvolvendo agora algo que já foi pago para ser desenvolvido, ou como garantir a integridade se mudarmos algo sem saber as dependências e a rastreabilidade dos componentes envolvidos na solução? Para resolver questões como essas, a indústria criou o RAS – *Reusable Asset Specification*, padrão para controle de qualidade de software e rastreamento de interligações e dependências entre os componentes de uma solução.

O RAS visa a controlar tudo que uma corporação desenvolve, de um simples diagrama UML a um script *shell*. Tudo é um ativo, e não apenas códigos fora do controle da empresa. O RAS permite administrar o inventário de software e visualizar claramente o que a empresa possui, de onde parte e onde quer chegar.

Baseado nos princípios do RAS, é possível aplicar métricas de impacto de mudanças, ou seja, ao alterar regras de negócios de um componente, quais serão os pontos afetados? Você já imaginou o quanto é importante saber de antemão que a tabela X exige mudanças no requisito Y? Isso não é mais sonho, já temos toda uma especificação definida pela OMG – *Object Management Group* (mesmo órgão que define o *Corba*) –, na qual podemos nos basear para criar nossas próprias ferramentas ou adquirir alguma no mercado.

Várias siglas e conceitos novos pairam no ar ao se adotar SOA, mas dificilmente observo a adoção da SOA em estágios passo-a-passo; as empresas querem “orquestrar” serviços, sem saber exatamente quais componentes possuem e como eles se relacionam. RAS é um ótimo caminho para trazer uma visão mais ampla para essas empresas. Ao pensar em SOA, devemos evitar olhar apenas a tecnologia, atentando-nos para o controle, integração, qualidade e governança de software. ■

## Como saber se a empresa não está desenvolvendo agora algo que já foi pago para ser desenvolvido?

não formada e conseguiria conceituar esse termo? Antes de responder a você mesmo ou ler a centésima primeira revista, vamos, primeiro, verificar algumas motivações econômicas para SOA:

- ▶ Empresas adquirem outras empresas;
- ▶ Escritórios ao redor do mundo colaboram entre si;
- ▶ Companhias reduzem custos por meio da terceirização.

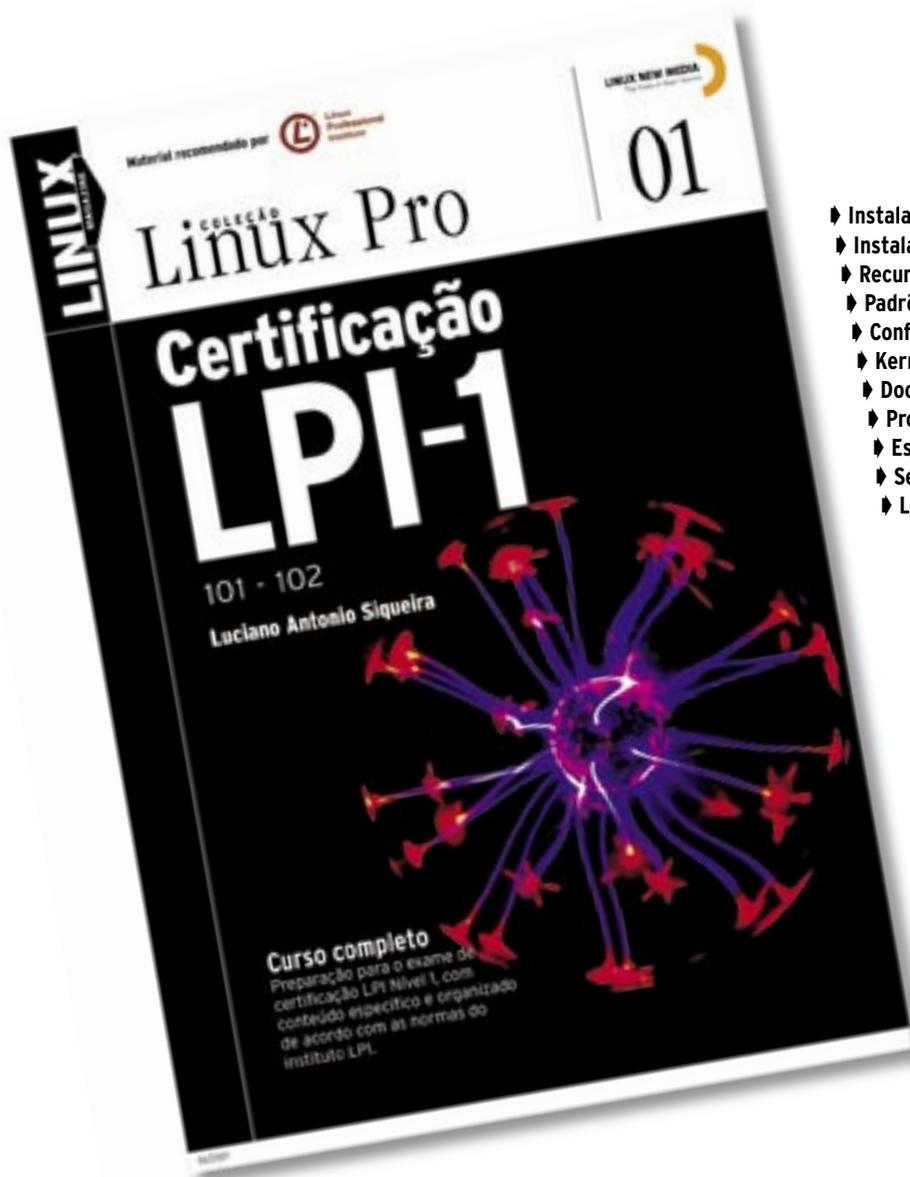
Quando tudo era centralizado, as coisas já não eram simples e, com os fatores expostos acima, se torna cada vez mais natural criar canais comuns, mapear todos os processos, executar, editar e auditar regras de negócios, que precisam de toda a resiliência para estabelecer um diferencial competitivo de mercado, fazendo-se necessário a adoção da SOA pelas empresas. Mas só tecnologia resolve? Vejamos alguns princípios básicos dentro das corporações que, às vezes, são esquecidos, como a governança de software.

Se uma empresa é focada no mercado financeiro, a diretoria opta por investir em capacitação para essas áreas de negócio e terceiriza todo o desenvolvimento de soluções. É errado? Não, pois o ponto de vista dos

### O autor

Edgar Silva é *Solutions Architect* e *JBoss Sales Engineer* da Red Hat Brasil, além de também ministrar palestras no Brasil e no exterior sobre *Java*.

## Prepare-se para a principal certificação profissional do mercado Linux!



- ▶ Instalação e configuração de hardware
- ▶ Instalação de programas (pacotes Debian e RPM)
- ▶ Recursos avançados do Shell e comandos GNU/Linux
- ▶ Padrões Linux e manutenção do sistema
- ▶ Configurações do servidor X
- ▶ Kernel e carregamento do sistema
- ▶ Documentação oficial e sistemas de impressão
- ▶ Programação shell script
- ▶ Estruturas de redes e servidores (Apache, ssh, [x]inetd, etc)
- ▶ Segurança: restrições de acesso a recursos
- ▶ Livro recomendado pelo LPI



Estude para a prova de acordo com o conteúdo programático estabelecido pelo LPI.

Em breve nas livrarias ou no site [www.linuxmagazine.com.br](http://www.linuxmagazine.com.br)

Código Aberto e a Cauda Longa

# Cezar Taurion

O Código Aberto e o conceito Cauda Longa revelam similaridades.

por Cezar Taurion

Estava relendo alguns pedaços do livro “A Cauda Longa” (*The Long Tail*), de Chris Anderson, e comecei a pensar na relação deste conceito com a indústria de software. A Cauda Longa está impulsionando grandes transformações em vários mercados, como o de mídia e o fonográfico. Juntando esse conceito ao de Código Aberto, veremos grandes transformações também na indústria de software. Gostaria de debater com vocês algumas destas idéias.

O conceito de Cauda Longa propõe que determinados negócios podem obter uma parcela significativa de sua receita pela venda cumulativa de grande

ro suficiente) podem, agora, ser desenvolvidos em Código Aberto e entrar no mercado. Os custos de comercialização desses softwares também tende a ser zero, pois não são necessárias hordas de vendedores, mas simples downloads e marketing viral (blogs e outros meios de disseminação de informação).

Temos, então, campo para explorar o mercado da Cauda Longa no software. E já temos exemplos bem sucedidos disso: os plug-ins do Eclipse podem atender a um mercado bem menor, que não seria alcançado caso fosse necessário um caro e sofisticado software monolítico, com 30 a 40 funções de plug-ins aglutinadas. A página de plug-ins do Eclipse certamente possui algum plug-in que atenda às suas necessidades, mesmo que elas sejam muito específicas. Uma consulta que fiz no início de julho me trouxe 1.487 plug-ins disponíveis.

A modularização do Linux também permite criar variantes para processadores de pouca disseminação no mercado. Existe algum outro sistema operacional que rode em tantos processadores como o Linux? Claramente o Linux também atende o mercado da Cauda Longa.

O modelo comercial tradicional prende a indústria no mercado de grandes volumes, que por sua vez precisa de base significativa de clientes para se sustentar.

O mercado de software tradicional vai morrer? Na minha opinião, um enfático não! Acredito que iremos conviver em um contexto onde os modelos de massa e de Cauda Longa vão compartilhar os palcos. Continuaremos a ter produtos de software bem abrangentes em funcionalidades e que venderão suficientemente bem para manter um mercado próprio. Mas também teremos muitos outros produtos baseados em Código Aberto que se encaixarão muito bem no contexto da Cauda Longa. ■

## Existe algum outro sistema operacional que rode em tantos processadores como o Linux?

número de itens, vendidos em pequenas quantidades. Isso é possível porque a Internet abre oportunidades de acesso que antes não existiam. É um modelo diferente do mercado de massa, onde poucos artigos são vendidos em quantidades muito grandes. Na indústria de livros, música e mídia faz todo o sentido. Por exemplo, a Amazon relata que 57% de sua receita vem de produtos da Cauda Longa, que não estão disponíveis (e jamais estariam) nas livrarias tradicionais, limitadas pelo caro espaço físico das lojas.

E como o Código Aberto e o conceito de Cauda Longa afetam a indústria de software? No Código Aberto, o custo de desenvolvimento de um software é rateado por uma comunidade de desenvolvedores. Não existe um centro de custo único, portanto, o custo individual é muito pequeno.

Assim, softwares que tinham seu projeto cerceado pelo pequeno tamanho do seu mercado potencial (seu custo de produção não gerava retorno finance-

### O autor

**Cezar Taurion** é gerente de novas tecnologias aplicadas da IBM Brasil. Seu blog está disponível em <http://www-03.ibm.com/developerworks/blogs/page/ctaurion>.



O Linux aproveita a tendência dos múltiplos núcleos

# Espírito fraterno

Com o fim da corrida dos gigahertz, os fabricantes de processadores agora apostam na força do paralelismo. Mas os programas devem estar preparados para essa mudança.

por Jan Kleinert

Com bicicletas é semelhante: dobrar os participantes não é apenas melhor, mas também mais veloz. Quando os dois ciclistas não possuem o mesmo nível de treinamento, é necessário administrar seu desempenho; ou um dos dois trabalha mais que o outro, pedalando praticamente sozinho, ou os dois se alternam como principal fonte de esforço. Esse cenário é verdadeiro para a AMD e a Intel, pois ambas já incluem em seus processadores mais de um núcleo, transformando-os efetivamente em CPUs multiprocessadas.

Entretanto, em relação aos sistemas multiprocessados tradicionais, essa abordagem de múltiplos núcleos possui três importantes vantagens: não são necessárias placas-mãe especiais, o processador tem menor custo – pois é apenas um, e não dois ou mais – e o desempenho é melhor. Os núcleos ficam não apenas fisicamente mais próximos, como também eletricamente, permitindo uma intercomunicação mais veloz. O trabalho lógico também é menor, pois o *cache* pode ser compartilhado dinamicamente entre os núcleos, o que traz ainda mais vantagens.

A preferência pela multiplicação dos núcleos tem relação com a Física: os últimos processadores da Intel, por exemplo, adotavam a arquitetura *Netburst*, que não foi capaz de atingir o antigo objetivo de 4,0 GHz. Quanto mais milhões de transistores são inseridos num processador, maior é o caminho entre os mesmos, aumentando em consequência os efeitos indutivos e capacitivos.

## Watts

A única forma de atingir frequências mais altas em espaços maiores seria aumentando a voltagem – porém, a dissipação de energia cresce com o quadrado dessa distância. Por isso os antigos modelos de núcleo único atingiam os 100 Watts de consumo. O calor devia fluir do processador para o dissipador

de calor, e então para o ambiente – uma transmissão problemática.

Nos servidores *blade*, na computação de alto desempenho e em qualquer situação que exija a aglomeração de processadores essas questões adquirem importância ainda maior.

As CPUs de dois ou até quatro núcleos são capazes de ir contra esse aumento de produção de calor, pois dois núcleos de 2,5 GHz tendem a gerar menos calor que um único de 3,5 GHz.

Todavia, para que o esforço valha a pena, o software deve ser capaz de usá-lo. O Linux possui suporte a *SMP* (Multiprocessamento Simultâneo, na sigla em inglês), e conhece as *Kernel-Threads*, o que na realidade já basta para a maioria dos usos de múltiplos núcleos. Ao executar mais de uma tarefa na máquina, o kernel consegue decidir entre a utilização de dois ou apenas um núcleo. Ele próprio lida com a distribuição de tarefas.

## Multithreading

A forma mais fácil de tirar proveito de múltiplos núcleos é tratada nos artigos sobre o *OpenMP* e sobre *threading*. O artigo a respeito dos *benchmarks* mostra as medições obtidas. E para começar, apresentamos os principais termos que povoam o cenário atual do processamento *multi-core*, como *Hyperthreading*, *Vanderpool*, *Cool'n'Quiet*, *XD* e *NX*, assim como a guerra entre os processadores AMD e Intel. ■

## Índice das matérias de capa

Apache na vitrine pág 34.

Gêmeos pág 37.

Programas orquestrados pág 44.

O delegado pág 50.



O que as CPUs multicore realmente oferecem

# Apache na vitrine

Os benchmarks confirmam: apenas aplicativos com uso intensivo de CPU e bem paralelizados aproveitam múltiplos núcleos.

por Jens-Christoph Brendel

Na tentativa de descobrir a capacidade de desempenho de alguns computadores atuais com CPUs de núcleo duplo, primeiro o *benchmark* HTTP *Siege* recebeu uma incumbência. O programa de testes possui um modo que simula situações reais de carga, introduzindo pausas bem curtas entre o fornecimento de páginas web de um número configurável de usuários virtuais.

## Sem stress

Ao mesmo tempo, a instalação diminui a concorrência dos processos do *Apache* pelos recursos do sistema, especificamente da CPU. O resultado é que o número de núcleos é indiferente para o servidor web. Todo o processamento que ele realiza pode ser feito em apenas um deles. Núcleos adicionais não atrapalham, mas também não ajudam. Dessa forma, com

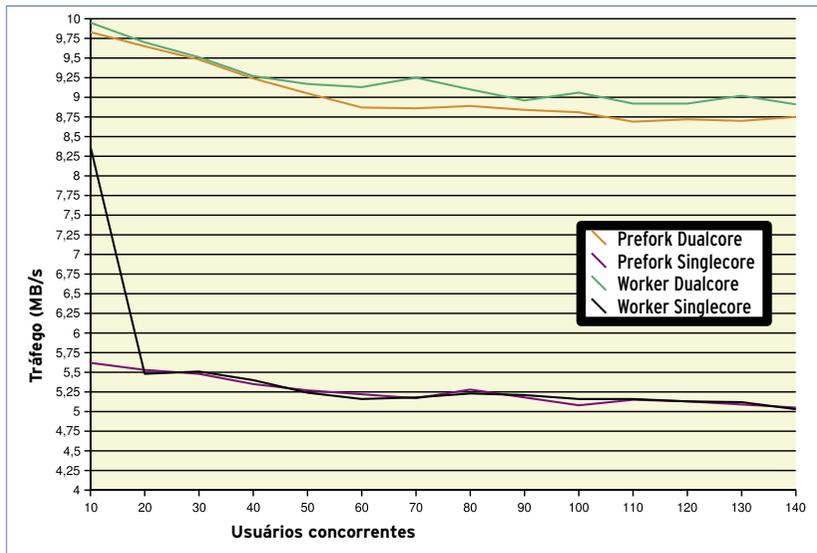
páginas pequenas (2 KB), um número maior de usuários aumenta proporcionalmente a capacidade do servidor. A elevação não foi limitada por nenhum gargalo nas estruturas testadas. A diferença entre um e dois núcleos se perde na turbulência das estatísticas.

Mesmo entre os diferentes módulos de multiprocessamento do Apache

(MPM) não se pode distinguir qualquer diferença: o modo *Prefork*, que reproduz o comportamento do antigo Apache 1.3, e inicia *a priori* um número pré-definido de processos-filhos para execução das requisições, tem escalabilidade idêntica à do modo *Worker*, que ajusta o número de processos de forma variável para a carga.



**Figura 1** A Linux Magazine executou todos os testes deste artigo em um Opteron Dual Core 2215 de 2,61 GHz com 1 MB de cache. O servidor 2-HE é equipado com 2GB de RAM DDR2, disco rígido SATA e *Debian*.



**Figura 2** No modo de *benchmark* do *Siege*, o servidor atinge seu limite. A concorrência dos processos aumenta fortemente, a CPU é exigida e o joio se separa do trigo: um núcleo fornece apenas metade de seu desempenho.

Aumentando para 200 KB ou até mesmo 2 MB as páginas de teste a serem requisitadas, a princípio nada muda. Naturalmente o desempenho do servidor testado aumenta, mas até aí processadores de

núcleo único e duplo permanecem tão próximos — até mesmo acima de 100 usuários — que a diferença não é significativa.

Disso é possível tirar um primeiro ensinamento: a capacidade de cálculo

da CPU é apenas um dos inúmeros recursos do sistema, e nem sempre é o gargalo do desempenho de todo o sistema. Além disso, em qualquer forma de trabalho de escritório, seja e-mail, navegação ou edição de texto, vale o mesmo, já que a pressão só aumenta a lentidão e o computador não fica mais rápido com núcleos múltiplos.

## No limite

Nesse servidor web de exemplo, as taxas se alteram quando são retiradas as pequenas pausas entre as requisições, e o modo de benchmark do *Siege* gera o fluxo máximo de requisições possível. Já são dez usuários solicitando páginas do servidor nesse modo de operação, entre 2.000 e 100.000 vezes por minuto, de acordo com o tamanho da página.

Com isso, o servidor já não está mais frio. Agora o desempenho não sobe mais com o aumento do núme-

# pyconbrasil [3]

encontro brasileiro da comunidade python

30, 31 de agosto e 1º de setembro de 2007  
em Joinville - SC, na SOCIESC.

Temas desta edição:

- Plone / Zope
- Zope3 / Grok
- Django / TurboGears
- Maemo
- Sugar (OLPC/XO)
- Eagle / GTK / QT
- Ensino de Python
- ... e muito mais

Chamada para trabalhos, inscrições e patrocínio, acesse:

<http://pyconbrasil.com.br/>

organização:



apoio:

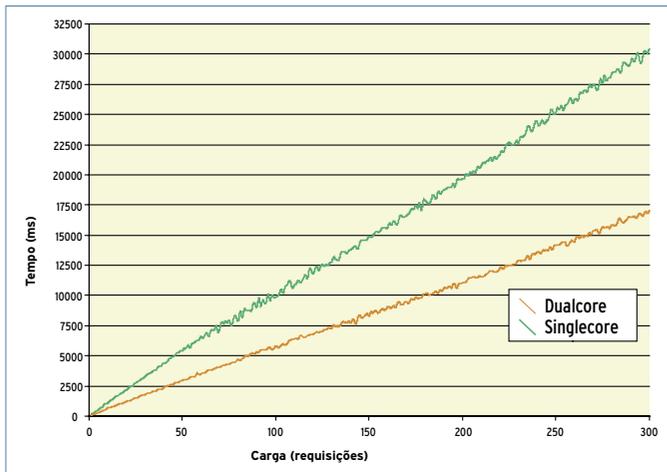


comercialização:



patrocínio:





**Figura 3** Resultado do *Hackbench*, medido a partir do sistema *dualcore* visto na figura 1. O teste de carga da CPU mostra cerca do dobro do desempenho com dois núcleos (curva laranja).

ro de usuários, mas diminui devido à sobrecarga. O grande número de processos compartilhados solicita a CPU de forma muito mais intensa. Como resultado, a diferença no número de núcleos que juntos fazem os cálculos agora é perceptível: apenas um deles trabalha um pouco acima da metade de sua potência (**figura 2**). O módulo *Worker* se sai um pouco melhor com esse regime do que o antigo módulo *Prefork*.

Esses resultados confirmam, como se esperava, de forma particularmente impressionante, que cada benchmark avalia o trabalho da CPU de maneira bem específica. O *Hackbench* [2], por exemplo, é uma ferramenta de hackers do kernel, com a qual é possível medir e otimizar as capacidades do escalonador de processos no kernel. O programa gera um cliente e um servidor, que conversam um com o outro através de um software gerenciador, e, para cada conexão,

por sua fome de poder de processamento. Quem quiser tirar a prova medindo os tempos de renderização com o conhecido renderizador de Código Aberto *Povray* verá, entretanto, que não há qualquer diferença entre um ou mais núcleos.

Aqui, na verdade, roda apenas um processador, pois o sistema operacional também assimila apenas uma CPU. Com isso, o resultado do *mpstat* (do pacote *sysstat*), executado durante a renderização, pode facilmente convencer quem o observar. Fica claro que uma CPU permanece 100% ocupada, enquanto a outra não faz absolutamente nada (**figura 4**). Esse é, portanto, o segundo ensinamento: a carga deve ser dividida, senão os núcleos adicionais não são utilizados.

Na melhor das hipóteses, o aplicativo já está dividido e otimizado para o multiprocessamento. Se não for o caso, sempre podemos deixar que o sistema operacional

| Time     | CPU   | User   | nice | sys  | idlet | irq  | soft | total | idle   |
|----------|-------|--------|------|------|-------|------|------|-------|--------|
| 15:20:44 | cpu/% |        |      |      |       |      |      |       |        |
| 15:20:45 | all   | 49.50  | 0.00 | 0.50 | 0.00  | 0.00 | 0.00 | 0.00  | 50.00  |
| 15:20:45 | 0     | 0.00   | 0.00 | 0.00 | 0.00  | 0.00 | 0.00 | 0.00  | 100.00 |
| 15:20:45 | 1     | 100.00 | 0.00 | 0.00 | 0.00  | 0.00 | 0.00 | 0.00  | 0.00   |

**Figura 4** O *Povray* inicia apenas um processo, e por isso pode trabalhar com apenas uma CPU — uma segunda não lhe serve para nada. Já o renderizador *Blender*, também livre, consegue utilizar uma máquina com múltiplos núcleos.

é iniciado um novo processo. Esse teste mostra, do começo ao fim, o dobro do desempenho ao se usar dois núcleos (**figura 3**).

## Distribuição

Diante dessas observações, sugere-se a hipótese de que núcleos adicionais, em todos os casos, acelerariam sensivelmente aqueles programas famosos

distribua os diversos processos por uma série de CPUs. Isso inclui naturalmente processos externos de aplicativos que rodam em paralelo, e também aqueles do próprio sistema operacional. No entanto, os programas que utilizam apenas um único processo também não se beneficiarão dos vários núcleos. Por exemplo, no caso da renderização, vale a pena mudar para a ferramenta livre *Blender*, que tira proveito de sistemas multi-core ou de vários processadores.

## Resultado

Mais processadores nem sempre significam performance superior. Para que um aplicativo rode de fato com maior velocidade em sistemas de múltiplos núcleos, ele deve depender primariamente do processador e ser paralelizável, ou seja, produzir diferentes processos que possam ser executados simultaneamente. Os programas geralmente utilizados para tarefas normais de escritório não costumam preencher esses requisitos, pois não são consumidores vorazes de processamento, sendo limitados, na maioria das vezes, pela velocidade de leitura dos discos e pelo tempo de espera da interação com o usuário.

Até mesmo servidores nem sempre se encaixam nessa situação, como mostra o benchmark do HTTP. No caso deles, um núcleo a mais não traz acréscimos reais de velocidade, em contraste com o que parece acontecer com aplicações bem distribuídas e de cálculo intensivo, que observam avanços perceptíveis. ■

## Mais Informações

[1] Benchmark HTTP Siege: <http://www.joedog.org/JoeDog/Siege>

[2] Hackbench: <http://developer.osdl.org/craiger/hackbench>

# Gêmeos

Os processadores de múltiplos núcleos estão conquistando tanto servidores quanto desktops e até mesmo notebooks. Conheça os termos que definem essa tendência e veja as listagens de processadores multi-núcleo.

por Jan Kleinert

mike gieson - www.sxc.hu



Um processador de núcleo duplo (*dual core*) possui dois processadores em uma única peça, enquanto um *quad-core* engloba quatro núcleos distintos ou dois núcleos duplos. Contudo, a AMD e a Intel não foram as criadoras dessa tecnologia, mas sim as responsáveis por sua popularização. Elas alimentam seu sucesso com uma competição de modelos com números baseados tanto em detalhes técnicos quanto em fatores de mercado. Este artigo busca orientar o leitor em direção ao modelo mais adequado a suas necessidades.

O multi-processamento para os mais pobres é o *Multi-threading* com suporte de hardware. Com esse método, o processador com um único núcleo é capaz de operar

sobre seus registros de forma independente – uma vez por *thread*. Portanto, ele é híbrido: embora se apresente como um processador de núcleo simples e consiga operar apenas uma instrução por vez, as mudanças de contexto ocorrem mais rapidamente.

A implementação dessa técnica pela Intel recebeu o nome de *Hyper-threading* (HT [1]), e é utilizada em processadores das arquiteturas *Netburst* (Pentium 4, Pentium D, Pentium EE e Xeon) e *Core 2*. Os processadores AMD Athlon 64 X2, assim como os Opterons de núcleo duplo, também informam ao sistema operacional que são compatíveis com essa tecnologia, e a IBM também a emprega, sob o nome de *Multithreading* simétrico, em seus processadores Power 5. Por

último, a Sun permite, através dessa mesma tecnologia, que cada processador *UltraSparc T1* execute quatro threads por núcleo.

Em relação ao software, essas CPUs comportam-se como sistemas multiprocessados. Para isso, é importante que o



Figura 1 O Intel Core 2 Duo alcança bons resultados gerando pouco calor.

sistema operacional possua suporte a SMP (multiprocessamento) e a *kernel threads*. A Intel anuncia ganhos de desempenho

da ordem de 10 a 20% provenientes do uso de sua tecnologia multithreading em conjunto com operações multi-ta-

refa, e de até 33% no caso de programas otimizados pelo compilador para tirar proveito da tecnologia.

**Tabela 1: CPUs de núcleo duplo da Intel**

| Modelo                            | Codiname   | Frequência    | Cache L2 | FSB      | HT <sup>1</sup> | EM64T | XD <sup>2</sup> | VT <sup>3</sup> | Processo | TDP   | Soquete   |
|-----------------------------------|------------|---------------|----------|----------|-----------------|-------|-----------------|-----------------|----------|-------|-----------|
| <b>Pentium D</b>                  |            |               |          |          |                 |       |                 |                 |          |       |           |
| 805                               | Smithfield | 2,66 GHz      | 2 MB     | 533 MHz  | -               | X     | X               | -               | 90 nm    | 95 W  | LGA775    |
| 820                               | Smithfield | 2,8 GHz       | 2 MB     | 800 MHz  | -               | X     | X               | -               | 90 nm    | 95 W  | LGA775    |
| 830/840                           | Smithfield | 3/3,2 GHz     | 2 MB     | 800 MHz  | -               | X     | X               | -               | 90 nm    | 130 W | LGA775    |
| 920-960                           | Presler    | 2,8-3,6 GHz   | 4 MB     | 800 MHz  | -               | X     | X               | X               | 65 nm    | 95 W4 | LGA775    |
| 915-945                           | Presler    | 2,8-3,4 GHz   | 4 MB     | 800 MHz  | -               | X     | X               | -               | 65 nm    | 95 W  | LGA775    |
| <b>Pentium Extreme Edition</b>    |            |               |          |          |                 |       |                 |                 |          |       |           |
| 840                               | Smithfield | 3,2 GHz       | 2 MB     | 800 MHz  | X               | X     | X               | -               | 90 nm    | 130 W | LGA775    |
| 955/965                           | Presler    | 3,47/3,73 GHz | 4 MB     | 1066 MHz | X               | X     | X               | X               | 65 nm    | 130 W | LGA775    |
| <b>Core 2 Duo /Core 2 Extreme</b> |            |               |          |          |                 |       |                 |                 |          |       |           |
| E4300                             | Allendale  | 1,8 GHz       | 2 MB     | 800 MHz  | -               | X     | X               | -               | 65 nm    | 65 W  | LGA775    |
| E6300/6400                        | Conroe     | 1,86/2,13 GHz | 2 MB     | 1066 MHz | -               | X     | X               | X               | 65 nm    | 65 W  | LGA775    |
| E6600/6700                        | Conroe     | 2,4/2,66 GHz  | 4 MB     | 1066 MHz | -               | X     | X               | X               | 65 nm    | 65 W  | LGA775    |
| X6800                             | Conroe     | 2,93 GHz      | 4 MB     | 1066 MHz | -               | X     | X               | X               | 65 nm    | 75 W  | LGA775    |
| <b>Xeon</b>                       |            |               |          |          |                 |       |                 |                 |          |       |           |
| 3040/3050                         | Conroe     | 1,86/2,13 GHz | 2 MB     | 1066 MHz | -               | X     | X               | X               | 65 nm    | 65 W  | LGA775    |
| 3060/3070                         | Conroe     | 2,4/2,66 GHz  | 4 MB     | 1066 MHz | -               | X     | X               | X               | 65 nm    | 65 W  | LGA775    |
| 5020-5050                         | Dempsey    | 2,5-3,0 GHz   | 4 MB     | 667 MHz  | -               | X     | X               | X               | 65 nm    | 95 W  | 771 pinos |
| 5060-5080                         | Dempsey    | 3,2-3,73 GHz  | 4 MB     | 1066 MHz | -               | X     | X               | X               | 65 nm    | 130 W | 771 pinos |
| 5063                              | Dempsey    | 3,2 GHz       | 4 MB     | 1066 MHz | -               | X     | X               | X               | 65 nm    | 95 W  | 771 pinos |
| 5110-5130                         | Woodcrest  | 1,6-2 GHz     | 4 MB     | 1066 MHz | -               | X     | X               | X               | 65 nm    | 65 W  | 771 pinos |
| 5140/5150                         | Woodcrest  | 2,33/2,66 GHz | 4 MB     | 1333 MHz | -               | X     | X               | X               | 65 nm    | 65 W  | 771 pinos |
| 5148                              | Woodcrest  | 2,33 GHz      | 4 MB     | 1333 MHz | -               | X     | X               | X               | 65 nm    | 35 W  | 771 pinos |
| 5160                              | Woodcrest  | 3 GHz         | 4 MB     | 1333 MHz | -               | X     | X               | X               | 65 nm    | 80 W  | 771 pinos |
| 7020                              | Paxville   | 2,66 GHz      | 2 MB     | 667 MHz  | X               | X     | X               | X               | 90 nm    | 165 W | 604 pinos |
| 7030                              | Paxville   | 2,80 GHz      | 2 MB     | 800 MHz  | X               | X     | X               | X               | 90 nm    | 165 W | 604 pinos |
| 7040                              | Paxville   | 3 GHz         | 4 MB     | 667 MHz  | X               | X     | X               | X               | 90 nm    | 165 W | 604 pinos |
| 7041                              | Paxville   | 3 GHz         | 4 MB     | 800 MHz  | X               | X     | X               | X               | 90 nm    | 165 W | 604 pinos |
| 7110M/7120M                       | Tulsa      | 2,6/3,0 GHz   | 4 MB     | 800 MHz  | X               | X     | X               | X               | 65 nm    | 95 W  | 604 pinos |
| 7110N/7120N                       | Tulsa      | 2,5/3,0 GHz   | 4 MB     | 667 MHz  | X               | X     | X               | X               | 65 nm    | 95 W  | 604 pinos |
| 7130M                             | Tulsa      | 3,20 GHz      | 8 MB     | 800 MHz  | X               | X     | X               | X               | 65 nm    | 150 W | 604 pinos |
| 7130N                             | Tulsa      | 3,16 GHz      | 8 MB     | 667 MHz  | X               | X     | X               | X               | 65 nm    | 150 W | 604 pinos |
| 7140M                             | Tulsa      | 3,40 GHz      | 16 MB    | 800 MHz  | X               | X     | X               | X               | 65 nm    | 150 W | 604 pinos |
| 7140N                             | Tulsa      | 3,33 GHz      | 16 MB    | 667 MHz  | X               | X     | X               | X               | 65 nm    | 150 W | 604 pinos |
| <b>Itanium 2</b>                  |            |               |          |          |                 |       |                 |                 |          |       |           |
| 9015                              | Montecito  | 1,4 GHz       | 12 MB    | 400 MHz  | X               | X     | -               | X               | 65 nm    | 104 W | FC-PGA4   |
| 9020                              | Montecito  | 1,4 GHz       | 12 MB    | 533 MHz  | X               | X     | -               | X               | 65 nm    | 104 W | FC-PGA4   |
| 9030                              | Montecito  | 1,6 GHz       | 8 MB     | 533 MHz  | -               | X     | -               | X               | 65 nm    | 104 W | FC-PGA4   |
| 9040                              | Montecito  | 1,6 GHz       | 18 MB    | 533 MHz  | X               | X     | -               | X               | 65 nm    | 104 W | FC-PGA4   |
| 9050/9052                         | Montecito  | 1,6 GHz       | 24 MB    | 533 MHz  | X               | X     | -               | X               | 65 nm    | 104 W | FC-PGA4   |

<sup>1</sup>Hyperthreading | <sup>2</sup>Bit no-execute | <sup>3</sup>Virtualização (Vanderpool) | <sup>4</sup>De acordo com a série, também 130W.

## Dual core e Quad core na Intel

A AMD e a Intel não concordam muito quanto a seus esquemas de números de versões dos processadores, como mostram os sites [2] e [3]. As tabelas 1 a 3 tentam sistematizar a comparação e a busca de processadores de núcleos duplos e quádruplos. Para evitar o risco de imprecisões, os modelos para notebooks foram omitidos.

Já no início de 2005 a Intel oferecia o Pentium Extreme Edition e alguns modelos de Pentium D como *Dual-core* (tabela 1). Em outubro chegaram os primeiros Xeon DP, em janeiro de 2006 veio a linha Core Duo, seguido das linhas Core 2 Duo (figura 1) e Core 2 Extreme em julho. A arquitetura dos últimos é muito superior à Netburst: mesmo operando em frequências bastante inferiores e gerando menores quantidades de calor, a nova arquitetura é capaz de superar a anterior em capacidade



**Figura 2** O Athlon 64 X2 é compatível com os soquetes de 939 pinos dos seus irmãos de núcleo simples.

de processamento. Os grandes e caros modelos *EE* lideram seus respectivos benchmarks.

## Dual core na AMD

Nos processadores AMD64 com núcleo duplo, cada núcleo possui seu próprio cache L2 individual, mas

todos compartilham o controlador de memória através da tecnologia *Hypertransport*[4],[5]. Esses processadores utilizam os mesmos soquetes que aqueles de núcleo simples. As placas-mãe necessitam apenas de uma atualização da BIOS.

Desde janeiro de 2006 são fabricados diversos Athlon FX com dois



## O Thin Client de melhor relação custo benefício do mercado e funcionalidades exclusivas



O TC-NET revoluciona o mercado de Thin Clients ao apresentar um sistema operacional com funcionalidades exclusivas, o TC-OS.

Compatível com a maioria dos servidores de terminal existentes, essa solução agrega redução de custos de aquisição e atualização, facilidade de instalação e manutenção, maior estabilidade, confiabilidade e segurança, baixo consumo de energia e tamanho compacto.



## Sistema Operacional TC-OS

- Recursos de administração centralizados;
- Servidor VNC para permitir o controle remoto de cada thin client;
- Disponível com Flash (64mb ou 128mb) ou na versão boot - PXE;
- Clientes para CITRIX-ICA, Microsoft RDP 5.1, XDMCP e VNC e Tarantella;
- Diversos aplicativos para uso local como **Skype**, navegador, Java, protetor de tela, entre outros (disponíveis apenas na versão TC-NET 2).



**ThinNetworks**  
Soluções Thin Client

Fone: (61) 3366-1333  
vendas@thinnet.com.br  
Skype: thinNetworks  
e thinNetworks1  
www.thinnet.com.br

núcleos. E desde fevereiro de 2007 está à venda o mais novo lançamento da AMD, o Athlon 64 X2 6000+, com 3GHz. Em meados de fevereiro de 2007 a AMD divulgou detalhes da arquitetura de seus processadores Opteron Quadcore nativos, codinome Barcelona.

Juntamente com o ganho de desempenho proveniente da utilização de mais núcleos e um cache L3 comum, além de *pipelines* mais longos, o Barcelona trará um novo gerenciamento de temperatura, em função da revisão de sua microarquitetura. As frequências de cada núcleo são ajustadas separadamente, mesmo que o controlador de memória esteja trabalhando a toda velocidade. De forma geral, a CPU será capaz de desligar as funções desnecessárias. Os processadores projetados com quatro núcleos não deverão aquecer mais que os modelos atuais; ou seja 68, 95 ou 120 watts, dependendo do modelo. A AMD pretende lançar a linha Barcelona no terceiro trimestre de 2007.

## Funções especiais

Quase todos os processadores de núcleo duplo possuem comandos especiais, seja para operação com 64 bits (*EM64T*, *AMD64*), sejam funções matemáticas que

## Quadro 1: Virtualização em hardware

Na utilização do *hyperthreading*, uma grande complicação é a implementação das funções de virtualização. A AMD e a Intel incluem tais funções em parte dos seus processadores, pois, diferentemente dos *mainframes*, a arquitetura IA-32 não dispõe de suporte dos sistemas operacionais ao funcionamento paralelo de múltiplos sistemas operacionais: o kernel e os *drivers* dos sistemas hóspedes como o Linux rodam no anel 0 da CPU (sob o modo privilegiado do kernel). O software de controle do sistema virtual transfere o sistema hóspede para um anel mais alto.

Quando o cliente faz uma solicitação que só funciona no anel 0, o processador responde com uma exceção. O software de virtualização capta essa solicitação e emula a solicitação privilegiada – porém, somente em teoria. Na prática, isso nem sempre funciona, tanto na manipulação do SMP quanto nos comandos de 64 bits. Os processadores Intel manipulam particularmente com a tabela de descritores global a tabela central de gerenciamento de memória. O compartilhamento de carga da tabela de descritores global só funciona corretamente no anel 0, e solução uma exceção compatível com a virtualização; entretanto, o compartilhamento de armazenamento não funciona da mesma forma.

Nas CPUs Intel, o termo *VMX (Virtual Machine Extensions)* – ou *VTI*, no Itanium – indica a presença da tecnologia *Vanderpool*, que oferece suporte à virtualização por *hypervisor* com administração dos recursos do processador, em especial, das exceções. Com isso, o sistema operacional hóspede não percebe que não possui a CPU inteira para si, e então o software de virtualização obtém controle total sobre o sistema hóspede, embora o último ainda funcione no anel 0.

### AMD Pacífica

Da mesma forma que a tecnologia *VMX* da Intel, a AMD inclui em seus processadores o termo *SVM (Secure Virtual Machine Architecture)*. O software de virtualização também funciona em modo *hypervisor*, e com privilégios ainda maiores que aqueles do sistema hóspede rodando no anel 0. Dessa forma, o hóspede acredita ser o único a usar a CPU.

Diferentemente dos líderes do mercado, os processadores AMD 64 têm o controlador de memória integrado. Por essa razão, a tecnologia Pacífica também virtualiza por hardware o controlador de memória. Cada sistema cliente certamente necessitará de uma região de memória, a qual o *hypervisor* deverá transformar em endereços físicos. Em processadores sem Pacífica, essa tarefa é efetuada por software. A *SVM* da Pacífica implementa o modo aninhado de paginação de memória (*nested paging*) que, através de uma tabela de mesmo nome, disponibiliza para cada sistema hóspede um registro próprio (virtual, obviamente) da memória principal.

auxiliem na computação gráfica ou multimídia (*SSSE3*, *SSE3*). A maioria das CPUs consegue impedir que códigos sejam executados em regiões de memória isoladas. A

técnica denominada *XD (Execute Disable)* ou *NX (No Execute)* deverá, de certa forma, proteger os programas de danos durante sua execução.

Tabela 2: CPUs Quadcore da Intel

| Modelo                              | Codinome   | Frequência   | Cache L2 | FSB      | HT <sup>1</sup> | EM64T | XD <sup>2</sup> | VT <sup>3</sup> | Processo | TDP   | Soquete   |
|-------------------------------------|------------|--------------|----------|----------|-----------------|-------|-----------------|-----------------|----------|-------|-----------|
| <b>Core 2 Quad / Core 2 Extreme</b> |            |              |          |          |                 |       |                 |                 |          |       |           |
| Q6600                               | Kentsfield | 2,4 GHz      | 8 MB     | 1066 MHz | -               | X     | X               | X               | 65 nm    | 105 W | LGA775    |
| QX6700                              | Kentsfield | 2,66 GHz     | 8 MB     | 1066 MHz | -               | X     | X               | X               | 65 nm    | 130 W | LGA775    |
| <b>Xeon</b>                         |            |              |          |          |                 |       |                 |                 |          |       |           |
| X3210/3220                          | Kentsfield | 2,13/2,4 GHz | 8 MB     | 1066 MHz | -               | X     | X               | X               | 65 nm    | 105 W | LGA775    |
| E5310/5320                          | Clovertown | 1,6/1,86 GHz | 8 MB     | 1066 MHz | -               | X     | X               | X               | 65 nm    | 80 W  | 771 pinos |
| E5345                               | Clovertown | 2,33 GHz     | 8 MB     | 1333 MHz | -               | X     | X               | X               | 65 nm    | 80 W  | 771 pinos |
| X5355                               | Clovertown | 2,66 GHz     | 8 MB     | 1333 MHz | -               | X     | X               | X               | 65 nm    | 120 W | 771 pinos |

<sup>1</sup> Hyperthreading | <sup>2</sup> Bit no-execute | <sup>3</sup> Virtualização (Vanderpool) | <sup>4</sup> De acordo com a série, também 130W.

Tabela 3: CPUs dual core da AMD

| Modelo                      | Revisão         | Codiname   | Freq           | Cache L2 | Hypertr. <sup>1</sup> | RAM  | AMD64 | NX <sup>2</sup> | VT <sup>3</sup> | TDP   | Largura | Soquete |
|-----------------------------|-----------------|------------|----------------|----------|-----------------------|------|-------|-----------------|-----------------|-------|---------|---------|
| <b>AMD Athlon 64</b>        |                 |            |                |          |                       |      |       |                 |                 |       |         |         |
| EX-60                       | E6              | Toledo     | 2,6 GHz        | 2 MB     | 1 GHz                 | DDR  | X     | X               | -               | 110 W | 90 nm   | 939     |
| EX-62                       | F2              | Windsor    | 2,8 GHz        | 2 MB     | 1 GHz                 | DDR2 | X     | X               | X               | 125 W | 90 nm   | AM2     |
| EX-70-74                    | F3              | Windsor FX | EX 2,6-3,0 GHz | 2 MB     | 1 GHz                 | DDR2 | X     | X               | X               | 125 W | 90 nm   | F       |
| <b>AMD Athlon 64 X2 AM2</b> |                 |            |                |          |                       |      |       |                 |                 |       |         |         |
| 3800+/4200+                 | E4 <sup>4</sup> | Manchester | 2,0/2,2 GHz    | 1 MB     | 1 GHz                 | DDR  | X     | X               | -               | 89 W  | 90 nm   | 939     |
| 4600+                       | E4 <sup>4</sup> | Manchester | 2,4 GHz        | 1 MB     | 1 GHz                 | DDR  | X     | X               | -               | 110 W | 90 nm   | 939     |
| 4400+                       | E6              | Toledo     | 2,2 GHz        | 2 MB     | 1 GHz                 | DDR  | X     | X               | -               | 89 W  | 90 nm   | 939     |
| 4400+/4800+                 | F6              | Toledo     | 2,2/2,4 GHz    | 2 MB     | 1 GHz                 | DDR  | X     | X               | -               | 110 W | 90 nm   | 939     |
| 3800+-5000+                 | F2              | Windsor    | 2,0-2,6 MHz    | 1 MB     | 1 GHz                 | DDR2 | X     | X               | X               | 89 W  | 90 nm   | AM2     |
| 5400+                       | F3              | Windsor    | 2,8 GHz        | 1 MB     | 1 GHz                 | DDR2 | X     | X               | X               | 89 W  | 90 nm   | AM2     |
| EE 3800+-4600+              | F2              | Windsor    | 2,0-2,4 MHz    | 1 MB     | 1 GHz                 | DDR2 | X     | X               | X               | 65 W  | 90 nm   | AM2     |
| EE 3800+-5000+              | E3              | Windsor    | 2,0-2,6 MHz    | 1 MB     | 1 GHz                 | DDR2 | X     | X               | X               | 65 W  | 90 nm   | AM2     |
| EE SFF 3800+                | F2              | Windsor    | 2,0 GHz        | 1 MB     | 1 GHz                 | DDR2 | X     | X               | X               | 35 W  | 90 nm   | AM2     |
| 4000+-5200+                 | F2              | Windsor    | 2,0-2,6 MHz    | 2 MB     | 1 GHz                 | DDR2 | X     | X               | X               | 89 W  | 90 nm   | AM2     |
| 5600+                       | F3              | Windsor    | 2,8 GHz        | 2 MB     | 1 GHz                 | DDR2 | X     | X               | X               | 89 W  | 90 nm   | AM2     |
| 6000+                       | F3              | Windsor    | 3,0 GHz        | 2 MB     | 1 GHz                 | DDR2 | X     | X               | X               | 125 W | 90 nm   | AM2     |
| EE 4000+-4800+              | F2              | Windsor    | 2,0-2,4 MHz    | 2 MB     | 1 GHz                 | DDR2 | X     | X               | X               | 65 W  | 90 nm   | AM2     |
| EE 5200+                    | F3              | Windsor    | 2,6 GHz        | 2 MB     | 1 GHz                 | DDR2 | X     | X               | X               | 65 W  | 90 nm   | AM2     |
| 3600+-5000+                 | G1              | Brisbane   | 1,9-2,6 GHz    | 1 MB     | 1 GHz                 | DDR2 | X     | X               | X               | 65 W  | 65 nm   | AM2     |
| <b>Opteron</b>              |                 |            |                |          |                       |      |       |                 |                 |       |         |         |
| 165-185                     | E6              | Denmark    | 1,8-2,6 GHz    | 2 MB     | 1 GHz                 | DDR  | X     | X               | -               | 110 W | 90 nm   | 939     |
| 865-875                     | E1              | Egypt      | 1,8-2,2 GHz    | 2 MB     | 1 GHz                 | DDR  | X     | X               | -               | 95 W  | 90 nm   | 940     |
| 265-290/865-890             | E6              | Italy      | 1,8-2,8 GHz    | 2 MB     | 1 GHz                 | DDR  | X     | X               | -               | 95 W  | 90 nm   | 940     |
| 260HE-275HE/860HE-875HE     | E6              | Italy      | 1,6-2,2 GHz    | 2 MB     | 1 GHz                 | DDR  | X     | X               | -               | 55 W  | 90 nm   | 940     |
| 1210-1218                   | F2 <sup>5</sup> | Santa Ana  | 1,8-2,6 GHz    | 2 MB     | 1 GHz                 | DDR2 | X     | X               | X               | 103 W | 90 nm   | AM2     |
| 1220SE                      | F2,F3           | Santa Ana  | 2,8 GHz        | 2 MB     | 1 GHz                 | DDR2 | X     | X               | X               | 125 W | 90 nm   | AM2     |
| 1210HE-1218HE               | F3              | Santa Ana  | 1,8-2,6 GHz    | 2 MB     | 1 GHz                 | DDR2 | X     | X               | X               | 65 W  | 90 nm   | AM2     |
| 2210-2218                   | F2 <sup>5</sup> | Santa Rosa | 1,8-2,6 GHz    | 2 MB     | 1 GHz                 | DDR2 | X     | X               | X               | 95 W  | 90 nm   | F       |
| 8212-8218                   | F2 <sup>5</sup> | Santa Rosa | 2,0-2,6 GHz    | 2 MB     | 1 GHz                 | DDR2 | X     | X               | X               | 95 W  | 90 nm   | F       |
| 2212HE-2216HE/8212HE-8216HE | F2 <sup>5</sup> | Santa Rosa | 2,0-2,4 GHz    | 2 MB     | 1 GHz                 | DDR2 | X     | X               | X               | 68 W  | 90 nm   | F       |
| 2220SE/8220SE               | F2,F3           | Santa Rosa | 2,8 GHz        | 2 MB     | 1 GHz                 | DDR2 | X     | X               | X               | 119 W | 90 nm   | F       |

<sup>1</sup> Velocidade do Hypertransport | <sup>2</sup> Bit NX (No execution) | <sup>3</sup> Virtualização (Pacifica) | <sup>4</sup> Parcialmente também como Toledo E6 com metade do cache | <sup>5</sup> 1210-1220, 2210-2220 e 8212-8218/20 também como F3

Nem todas as CPUs suportam virtualização; é na compra de um servidor que a economia pode vir a ser um tiro no pé, no futuro. O **quadro**

**1** explica esse princípio, e as **tabelas 1 a 3** mostram as possibilidades de cada modelo de processador. Energeticamente falando, o processador

não produz nada além de calor. O valor de **TDP** (*Thermal Design Power*) mostra o limite superior de produção de calor. Como todos os avanços na

fabricação dos chips dependem do aumento da frequência, os fabricantes se norteiam pelos truques de economia de energia dos seus processadores projetados para notebooks:

- ▶ Reduzir lentamente a frequência quando o uso do processador é pouco exigido;
- ▶ Lenta redução da tensão;
- ▶ Desligamento da tensão nas unidades sem uso;
- ▶ Colocar em hibernação as CPUs inativas.

Aparentemente, a tecnologia *Cool'n'Quiet* da AMD para processadores de desktop (nos Opterons, *Optimized Power Management*) é mais eficiente que o *EIST/Speedstep* da Intel. No uso de múltiplos núcleos é importante ser capaz de colocar cada núcleo em hibernação independentemente dos outros. É por isso que desde a versão 2.6.18 o kernel espelha os registros `sched_mc_power_savings` e `sched_smt_power_savings` no `sysfs`, em `/sys/devices/system/cpu/cpuX`. Com eles, é possível ativar (1) e desativar (0) a economia de energia individualmente para cada núcleo[6].

O princípio por trás disso é que, estando ativa a economia de energia e havendo uma baixa solicitação do sistema, todas as tarefas são deslocadas para um núcleo apenas, o que permite que os outros núcleos passem para o modo de economia de energia. Em meados de 2006, Suresh Siddha, da Intel, implementou essa função no kernel.

## E o software?

Na prática – e também na teoria, de acordo com a Lei de Amdahl[7] – a eficácia dos *multicore* diminui com o grau de paralelização do software. Além do Linux, os sistemas UNIX, Windows® XP, 2003 e Vista oferecem suporte a esses processadores. Eles dividem os processos pelos núcleos e trabalham paralelamente. O artigo “O delegado”, à [página 50](#) desta edi-

ção da *Linux Magazine*, enfatiza as estratégias de organização do Linux 2.6. Em contrapartida, se há apenas um único processo em andamento, ele deveria ser paralelizado com o uso de threads – com o auxílio, por exemplo, do compartilhamento de memória ou da passagem de mensagens (com *MPI*). Nos sistemas multicore e SMP, a programação ideal para compartilhamento de memória regularia as opções de compilação (no Linux, podem ser usados o *GCC*, o *ICC*, o *PGI*[8] ou outros), diretivas de paralelização e bibliotecas especiais de paralelismo, como a *OpenMP*.

A biblioteca *Intel Threading Building Blocks*[9] para C++ disponibiliza algoritmos e estruturas para a implementação de threads. O *Thread Checker* 3.0[10] para Linux e a análise de performance do *Vtune*[11] para Linux se dedicam a depurar e localizar *dead locks*. Na segunda forma de paralelização – a programação por passagem de mensagens – a comunicação entre processadores não é tão eficiente quanto o compartilhamento de memória. A comunicação livre entre CPUs ocorre, na maioria das vezes, por *IP* (*Ethernet*, *Myrinet*, *Infiniband*). Em contrapartida, esse modelo tem melhor escalabilidade.

## Conclusões e perspectivas

Atualmente, a Intel, com sua tecnologia *Core 2*, supera sua eterna concorrente AMD nos benchmarks em desktops. A *Core 2* demonstra que o melhor é a união de quatro ou mais núcleos em um chip, principalmente quando, em breve, a espessura de seus processos diminuir para 45 nanômetros. Todavia, o aumento de frequência para essas CPUs ainda é um sonho.

Para os desenvolvedores, a Intel tem em suas prateleiras contribuições

para o desenvolvimento do kernel e diversos softwares SMP. Para isso, a fabricante tem em seus porões o cadáver de seu caro Itanium, contra o qual a AMD introduziu com sucesso nos servidores pesados o poderoso Opteron. A AMD também é claramente melhor no quesito de economia de energia. Quem se alegra e lucra com o rápido desenvolvimento e a queda de preços, naturalmente, é o consumidor. O mais novo exemplo são as CPUs multinúcleo, que merecem ser estimuladas até o último bit. ■

## Mais Informações

- [1] Hyperthreading: <http://www.intel.com/technology/hyperthread>
- [2] Buscador de processadores da Intel: <http://processorfinder.intel.com>
- [3] Especificações de comparação de processadores da Intel: <http://www.amdcompare.com>
- [4] AMD Multicore site: <http://multicore.amd.com/br-pt/AMD-Multi-Core.aspx>
- [5] Barramento Hypertransport: <http://www.hypertransport.org>
- [6] Descrição do patch do kernel de Suresh Siddha: <http://tinyurl.com/26wjzd>
- [7] Lei de Amdahl: [http://pt.wikipedia.org/wiki/Lei\\_de\\_Amdahl](http://pt.wikipedia.org/wiki/Lei_de_Amdahl)
- [8] Compilador PGI: <http://www.pgroup.com>
- [9] Intel Threading Building Blocks 1.0 para Windows, Linux e Mac OS: <http://www.intel.com/cd/software/products/asm-na/eng/294797.htm>
- [10] Intel Thread Checker 3.0 para Linux: <http://www.intel.com/cd/software/products/asm-na/eng/291669.htm>
- [11] Intel Vtune Performance 8.0 Analyzer para Linux: <http://www.intel.com/cd/software/products/asm-na/eng/vtune/239145.htm>

# SAIBA POR QUE OS CLIENTES ESCOLHEM E APROVAM A RED HAT

“

Hoje uma nova gama de produtos e serviços da Red Hat está sendo utilizada ou avaliada pela JT, que trabalha com praticamente todos os principais fornecedores de tecnologia de TI no país.

“Cabe ressaltar que este projeto de modernização, em curso desde 2002 e que conta com a participação ativa da Red Hat Brasil, já é considerado um sucesso absoluto e tem se expandido para outros órgãos do sistema judiciário brasileiro.

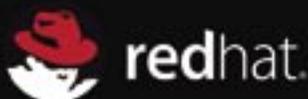
**Alberto Daniel Muller**

Diretor de Serviço de Infra-Estrutura e  
Planejamento do Tribunal Regional do Trabalho

”

**www.br.redhat.com**  
**11 3529.6018**

Solicite uma versão para teste:  
[www.redhat.com/rhel/details/eval](http://www.redhat.com/rhel/details/eval)



# Programas orquestrados

Com o fim da corrida por frequências mais altas, o mesmo ocorre também com a aceleração dos programas. O atual problema dos desenvolvedores é adequar seus softwares para aproveitarem os benefícios dos processadores multinúcleo.

por Michael Hebenstreit

Serge - www.sxc.hu



Com a introdução das chamadas CPUs *multicore*, surge repentinamente para os programadores de desktop um tema interessante, que anteriormente era uma aplicação quase exclusiva de clusters: o desenvolvimento de programas paralelos. A simples pergunta de como um programa utilizará diversas CPUs simultaneamente abre uma verdadeira caixa de pandora, com novos e inesperados problemas: programas seriais corretos rodando em hardware paralelo freqüentemente se comportam de forma inesperada.

Embora seja relativamente fácil rodar diversos programas em um só servidor ocupando as CPUs existentes, para um único programa isso é mais complicado. Por enquanto, não devemos nos preocupar se e como é

possível dividir o trabalho. Por analogia, podemos considerar a construção de uma casa: vários pedreiros trabalham paralelamente para levantar as paredes da casa. Apenas quando elas estiverem prontas outros trabalhadores — bombeiros, eletricitistas etc — poderão fazer sua parte, paralelamente. Por último, vários pintores poderão realizar a pintura da casa ao mesmo tempo.

Na primeira e terceira fases, (construção das paredes e pintura, respectivamente) fala-se em distribuir os dados (paralelismo de dados); a segunda fase é típica da divisão de tarefas (paralelismo de tarefas). A divisão dos dados corresponde, na maioria das vezes, à divisão de um grande *loop*; enquanto funções simples, executadas paralelamente, assumem os diversos trabalhos. Em ambos os casos,

o desenvolvedor deverá assegurar que o programa continue gerando os resultados corretos, que as CPUs sejam irmanamente carregadas e que, na medida do possível, todas as etapas do programa estejam paralelizadas.

## Lei de Amdahl

Suponhamos que um programa com tempo de execução de dez minutos tenha 80% de seu algoritmo paralelizado e rode em um número progressivo de CPUs. Em quatro processadores, a parte paralelizada desse programa ocuparia dois minutos (8/4) além dos dois minutos da parte que roda em série. Portanto, o tempo total de execução do programa seria de quatro minutos, o que corresponde a um fator de aceleração de  $10 \text{ min} / 4 \text{ min} = 2,5$ , apesar de diversos usuários de quatro CPUs esperarem um fator de aceleração de 4,0.

Essa correlação é conhecida como Lei de Amdahl. A eficiência (fator de aceleração/número de CPUs) importa no último caso em somente 62,5% ( $2,5 / 4 \times 100$ ). A relação fica mais desfavorável à medida que aumentamos o número de CPUs. O tempo de execução do programa será sempre afetado pelos dois minutos de execução serial; o maior fator de aceleração alcançável, portanto, é  $10 \text{ min} / 2 \text{ min} = 5$ . No passado, foi dada mais atenção à prova dos princípios para desenvolvimento de programas paralelos do que a sua implementação efetiva.

### Exemplo 1: Um exemplo simples

```
01 #include <stdio.h>
02
03 main()
04 {
05     #pragma omp parallel
06     {
07         printf("Hello World\n");
08     }
09 }
```

A partir dos sistemas Unix com diversas CPUs, iniciou-se o paralelismo baseado em tarefas, com as *Posix threads*. O mesmo conceito também funciona no Windows®, apesar deste usar uma API modificada. Embora ambos sejam relativamente semelhantes, as diferenças são tão grandes que um programa em C precisa ser desenvolvido duas vezes para que se adapte às duas plataformas. Linguagens de script, como Python, escondem do usuário essas diferenças, fazendo com que os aplicativos baseados em threads sejam mais facilmente programáveis.

## Nos clusters, MPI

Um princípio inteiramente diferente de distribuição de dados em um cluster levou ao desenvolvimento do MPI, a *Message Passing Interface*, ou Interface de Passagem de Mensagens. Nessa regra, o mesmo binário é iniciado em diversos computadores simultaneamente. Baseado em uma ID fornecida pela MPI, cada programa decide qual trabalho deverá assumir. A comunicação entre os processos individuais é feita através da troca de informações. Porém, um programa serial existente precisa ser completamente reestruturado para utilizar várias CPUs.

Por volta de 1990, os sistemas paralelos começaram a se expandir comercialmente como sistemas eficazes, principalmente para a área da simulação numérica. Por essa razão, diversos fabricantes de hardware se uniram em um consórcio e lançaram a pedra fundamental do *OpenMP*. As condições iniciais eram as seguintes:

- ▶ Possuir uma interface simples;
- ▶ Permitir paralelização gradual durante o desenvolvimento;
- ▶ Oferecer eficiência;
- ▶ Não depender de um fabricante específico;
- ▶ Grupo-alvo: desenvolvedores de simulação numérica;
- ▶ Códigos seriais não precisam ser alterados.

Daí as extensões do OpenMP seriam efetuadas como pragmas de compilador. Esse procedimento faz com que compiladores que não entendem OpenMP consigam compreender as instruções adicionais como comentários, permitindo assim que o código seja compilado como puramente serial. O principal campo de aplicação da simulação numérica levou a especificação a se concentrar principalmente na paralelização.

O modelo completo cabe em cerca de cem páginas, e pode ser encontrado em [1]. A especificação abrange atualmente C, C++ e Fortran. A implementação fica a cargo de cada fabricante, porém todos costumam basear-se nas regras já existentes para programação com threads.

## Opções de compilação

O exemplo 1 mostra a utilização do OpenMP. A instrução `#pragma omp parallel` é suficiente para paralelizar o programa serial. Dependendo do compilador, pode ser necessário indicar interruptores adicionais ao compilar. O compilador da Intel, por exemplo, requer a opção `-openmp:icc -openmp -o hello_omp hello_omp.c`. Caso contrário, o ICC considera a instrução `pragma`

### Exemplo 2: ID e valor de uma thread

```
01 #include <omp.h>
02 main()
03 {
04     #pragma omp parallel
05     {
06         printf("Hello World do processo %d de %d\n",
07             omp_get_thread_num(), omp_get_num_threads());
08     }
09 }
```

como desconhecida e a ignora. No compilador GNU[2] a opção interruptor se denomina `-fopenmp`:

```
gcc -fopenmp -o hello_omp hello_omp.c
```

O GCC suporta o OpenMP a partir da versão 4.2, presente na versão atual do Fedora.

Na execução, um programa OpenMP verifica o número de threads paralelas suportadas pelo hardware, e com isso calcula quantas threads de software devem ser iniciadas. Um sistema com um chip quadcore corresponde a um sistema com quatro CPUs. Com um desses, o usuário pode simular o comportamento de uma máquina mono-processada através do uso do parâmetro `OMP_NUM_THREADS`. O contrário também é possível: num sistema com apenas uma CPU, esse mesmo parâmetro permite o desenvolvimento de programas OpenMP com mais threads, permitindo que se teste o comportamento paralelo. É claro que, nesse caso, muitos possíveis erros de tempo de execução em sistemas paralelos não são apresentados.

|         | Comandos da thread 1     | Registrador CPU 1 | Memória X | Register CPU 2 | Comandos da thread 2     |
|---------|--------------------------|-------------------|-----------|----------------|--------------------------|
| Tempo ↓ | copia X -> Registrador   | 0                 | 10        | 0              | ...                      |
|         | Soma 1 ao Registrador    | 10                | 10        | 0              | copia X -> Registrador   |
|         | copia o Registrador -> X | 11                | 10        | 10             | Soma 1 ao Registrador    |
|         | ...                      | 11                | 11        | 11             | copia o Registrador -> X |
|         | ...                      | 11                | 11        | 11             | ...                      |

Figura 1 Em caso de operações em uma ordem diferente da planejada, podem ocorrer condições de corrida. Na operação em ordem correta, o programa do exemplo 4 deve gerar como saída o valor 12.

## Exemplo 3: Simulação numérica

```

01 #define TAMANHO 30000
02
03 void Calcula()
04 {
05     int progressoGeral=0,i,j,porcentFeito,achado=0,procid;
06     static double X[TAMANHO];
07
08     for( i = 0; i <= TAMANHO; i ++ )
09     {
10         X[i] = sqrt(exp(cos(i))*exp(sin(i)));
11         for(j=1;j<i;j++)
12             X[i] *= sqrt(exp(cos(j))*exp(sin(i)));
13
14         if( X[i] > 0.5)
15             achado ++;
16         progressoGeral++;
17         porcentFeito = (int)((float)progressoGeral / (float)TAMANHO *100.0 + 0.5);
18
19         if( porcentFeito % 10 == 0 )
20             printf("\n\b\b\b\b\b%3d%",porcentFeito);
21     }
22     printf("\nachado %d\n",achado);
23 }

```

## Dependências

Para executar o exemplo simples `hello_omp`, cada thread gera as mesmas instruções e joga sua saída em um único `Hello World`. Isso é mostrado mais profundamente no [exemplo 2](#). Nele há duas funções de bibliotecas de tempo de execução do OpenMP, além das threads iniciadas e das IDs de thread.

A API do OpenMP oferece também possibilidades de se controlar o número de threads com o uso de travas (*locks*) explícitas. No entanto, o emprego dessas bibliotecas impede a compilação do programa com compiladores sem suporte ao OpenMP, o que não ocorre com os pragmas. Para dividir o trabalho em diversas threads, o OpenMP solicita as estruturas `for` e `sections`.

Como mencionado, a razão para o desenvolvimento do OpenMP é a simulação numérica. Junto a isso, corre constantemente uma rotina dentro de um loop, na maioria das vezes para avaliar valores dentro de um determinado vetor. O simples [exemplo 3](#) demonstra esse princípio, com um código serial que avalia valores de uma seqüência e ao mesmo

tempo verifica quantos desses valores são superiores a 0,5. No loop mais externo ([linha 8](#)) estão os cálculos individuais independentes uns dos outros — podemos realizar duas iterações sem que o resultado de uma influencie o resultado da outra.

## Exemplo 4: Paralelização defeituosa

```

01 void Calcula()
02 {
03     int progressoGeral=0,i,j,porcentFeito,achado=0,procid=0;
04     static double X[TAMANHO];
05
06     #pragma omp parallel for private(i,j,porcentFeito) shared(progre
07     ↳ssogeral,X,achado)
08     for( i = 0; i <= TAMANHO; i ++ )
09     {
10         X[i] = sqrt(exp(cos(i))*exp(sin(i)));
11         for(j=1;j<i;j++)
12             X[i] *= sqrt(exp(cos(j))*exp(sin(i)));
13
14         if( X[i] > 0.5)
15         {
16             achado ++;
17         }
18
19         progressoGeral++;
20
21         porcentFeito = (int)((float)progressoGeral/(float)TAMANHO
22         ↳*100.0 + 0.5);
23
24         if( porcentFeito % 10 == 0 )
25             printf("\n\b\b\b\b\b%3d%",porcentFeito);
26     }
27     printf("\nachado %d\n",achado);
28 }

```

## Variáveis dependentes

Esse não é o caso no loop `j` ([linha 11](#)): para calcular a iteração `j + 1`, é necessário sempre ter rodado `j`. Dependências como essa (*Forward Data Dependency* ou recorrência) impedem que seja possível a paralelização de um loop sem uma reorganização completa. O programador deveria verificar se há uma dependência interna ou externa. Nesse caso, o primeiro loop do OpenMP será assim:

```

#pragma omp parallel
{
    #pragma omp for
    for ( i = 0; i <=TAMANHO; i ++ )
}

```

Ou, em formato reduzido:

```

#pragma omp parallel for
for(i = 0; i <= TAMANHO; i ++ )

```

Isso é suficiente para permitir o cálculo do loop em diversas threads. A região de iteração será simplesmente dividida entre as threads. Nesse caso, em uma



## Exemplo 7: Parcialmente paralelizado

```

01 #pragma omp parallel sections
02 {
03     #pragma omp section
04     A= calcula_a(x1,y1,z1);
05     #pragma omp section
06     B= calcula_b(x2,y2,z2);
07     #pragma omp section
08     C= calcula_c(x3,y3,z3);
09     #pragma omp barrier
10 }
11 D= calcula_d(A,B,C);
12 #pragma omp parallel sections
13 {
14     #pragma omp section
15     E= calcula_e(x4,y4,d);
16     #pragma omp section
17     F= calcula_f(x5,y5,d);
18     #pragma omp section
19     G= calcula_g(x6,y6,d);
20 }

```

quando somente uma thread executa modificações sem sincronizar o valor da variável, de forma que as threads que estão fazendo a leitura obtêm valores ultrapassados. A atividade “thread 1 lê, thread 1 escreve, thread 2 lê” pode gerar um resultado diferente de “thread 1 lê, thread 2 lê, thread 1 escreve”.

## Ferramentas auxiliares

Normalmente, todas as variáveis acessadas por múltiplas threads são compartilhadas, de forma a evitar falhas como essa. Em um programa pequeno esse problema é resolvido

manualmente, mas projetos maiores requerem a adoção de ferramentas de software adequadas, como por exemplo o *Intel Thread Checker*.

Essa ferramenta encontra uma grande gama de conflitos de acesso que surgem em virtude da execução do programa (figura 2). Porém, o programador deve conhecer o caminho de execução dos códigos. A ferramenta

é incapaz de testar segmentos do código que não sejam executados na rodada de testes.

No exemplo, a ferramenta reconhece como problemáticas as linhas que incrementam as variáveis `achado` e `generalProgress`. Porém, ambas as variáveis são necessárias para transportar as informações sobre progresso e locais entre as threads individuais. Portanto, é seguro afirmar que, em qualquer momento, somente uma thread executa modificações nas variáveis. No entanto, o custo da trava necessária para isso é um tempo adicional.

## Trava

O OpenMP reconhece a diretiva `#pragma omp critical` como sinalizador para um segmento crítico. Isso dá a certeza de que somente uma thread executará o código no bloco seguinte. Uma variante está disponível para operações simples, como `++` ou `*=`. Como essas estruturas ocorrem com muita frequência, foi desenvolvido o `pragmas`:

```
#pragma omp atomic
generalProgress++;
```

Diante disso, a vantagem consiste no ínfimo trabalho necessário às operações de travamento que ocorrem automaticamente, imperceptíveis ao usuário.

Para a variável `achado` é oferecida uma forma alternativa. Isso é praticamente insignificante para o tempo de execução dos valores corretos das posições já citadas — é suficiente que esse valor esteja disponível no final do loop. Portanto, poderíamos dividir `achado` em um vetor e alocar para cada thread uma posição do vetor como variável privada. No final do loop, esse vetor deveria ser somado, com a finalidade de manter todas as posições encontradas. Essa ação é chamada de redução, e o OpenMP disponibiliza elementos de linguagem específicos para isso. É suficiente ter a variável `reduction(Operator:Variable)` como sinalizador:

```
#pragma omp parallel for private(i
➤ ,j,percentFeito) shared(progress
➤ oGera1,Xreduction(+:achado)
```

No início da operação, cada thread recebe uma variável privada `achado`, cujo valor é automaticamente 0 (um valor inicial; para a operação `*`, por exemplo, o valor de partida é 1). As threads podem alterar suas cópias privadas sem travas. No final do loop, o *OpenMP-Runtime* totaliza todos os valores e guarda-os em `achado`.

Com essas duas modificações, o programa roda corretamente (exemplo 5), porém não muito rápido. O *Intel thread profiler* é capaz de auxiliar nesse ponto. Atualmente, a análise dos dados de tempo gerados no Linux através de interface gráfica lamentavelmente só é possível no Windows.

O Profiler demonstra, pelos dados de duração levantados, uma perfeita representação individual das threads — principalmente nos casos em que todas as threads exportam códigos úteis

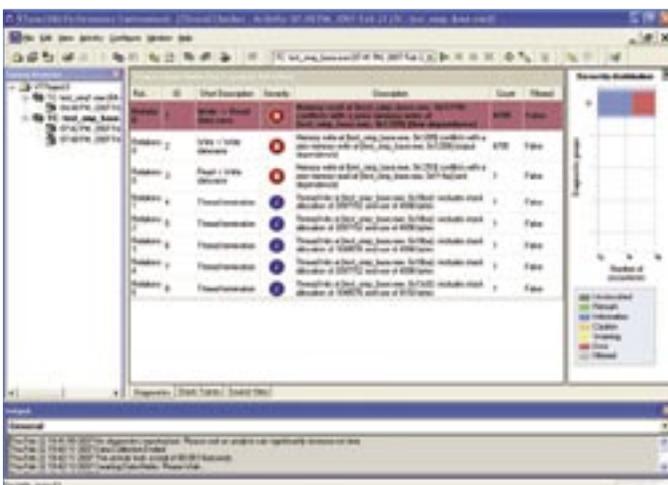


Figura 2 O *Intel Thread Checker* encontra diversos problemas potenciais em sua análise. Infelizmente, sua interface gráfica só está disponível para Windows.

para o cálculo de tempo total e não desperdiçam grande parte de suas vidas com ciclos de espera. Analisando o exemplo, demonstra-se que individualmente as threads executam trabalhos inteiramente diferentes.

## Injusto

Se analisarmos o loop mais detalhadamente, rapidamente encontraremos a razão disso. Quanto maior for o índice *i* no loop externo, mais vezes o loop interno *j* será executado. O OpenMP divide o loop *i* em partes iguais, como mencionado. Já com duas threads, resulta em desequilíbrio.

Para contornar esse tipo de problema, pode-se usar a opção `schedule(dynamic,CHUNK)`, que realiza uma divisão dinâmica de carga em um loop em tempo de execução. O loop é dividido em pedaços do tamanho do `CHUNK`, e cada thread recebe um segmento para trabalhar. O valor deveria ser menor que o tamanho do loop dividido pelo número de threads. Se uma thread terminar a parte que lhe foi destinada, ela recebe um `CHUNK` para trabalhar.

## Travas ocultas

O exemplo acima contém outro problema que, sem as ferramentas adequadas, ficará quase imperceptível. A saída de `printf` contém uma trava oculta (condicional de sistema). Se duas threads escreverem simultaneamente na console (ou seja, `stdout`), então o mais lento dos dois precisará aguardar até que o mais rápido termine sua tarefa. Na realidade, isso causa uma espera desnecessária e torna o programa mais lento. Todavia, não é necessário que o trabalho seja realizado por mais de uma thread.

Para essa finalidade, são oferecidos dois pragmas: `#pragma omp single` cuida para que o bloco rode em apenas uma thread; `#pragma omp master` faz com que apenas a thread o veja essa parte do código. Entretanto, não é permitida a existência dos dois pragmas em um loop

paralelizado. Em função disso, o programador é convencido a optar pela função da API `omp_get_thread_num()`. Com alguma habilidade e o emprego de `#ifdef_OPENMP`, o programa passa a ser compilável de forma serial.

Apesar de `pragma omp for` ser baseado na paralelização do OpenMP, `section` também pode ser usado para efetuar o paralelismo na escala de tarefas. Exemplos típicos são estruturas como a do exemplo 6. Os cálculos de *A*, *B* e *C* podem ocorrer independentemente. Isso também vale para *E*, *F* e *G* — somente é necessário *D* já ter sido calculado. Aqui atua outra vez a lei de Amdahl: ou `calcula_d()` roda de forma serial e restringe a escalabilidade, ou a própria função deverá ser paralelizada. Para uma escalabilidade limitada, são suficientes poucas instruções OMP (exemplo 7).

Na primeira parte do exemplo 7, as diretivas de seção cuidam para que os cálculos das variáveis *A*, *B* e *C* ocorram em paralelo. No final do bloco `section`, é implicitamente colocada uma barreira: antes que o programa continue seu trabalho, todas as threads deverão ter alcançado esse ponto. Uma barreira desse tipo também pode ser feita pela diretiva `#pragma omp barrier` (linha 9). Na maioria das vezes é recomendável a utilização explícita de `barrier` ao invés da conhecida estrutura dos pontos de sincronização implícitos, pois a primeira é menos sujeita a erros e, de uma forma geral, melhora a legibilidade do código.

Outra possibilidade é a utilização explícita da biblioteca OpenMP. O exemplo 8 mostra o quanto é simples para o programador dividir o loop entre as threads existentes. A vantagem em relação à programação de threads explícita está no fato de que o programador praticamente não precisa se preocupar com a produção da thread

## Exemplo 8: Valor de uma thread

```
01 nt=omp_get_num_threads()
02 #pragma omp parallel private(i,id) reduction(+:sum)
   ↳shared(nt)
03 {
04   id = omp_get_thread_num();
05   for(i=id;i<TAMANHO;i+=nt)
06   {
07     sum += calcula(i);
08   }
09 }
```

e outras particularidades específicas do sistema. Com isso, o programa fica mais facilmente portátil — embora perca um pouco do potencial para atingir um desempenho otimizado.

## Muito mais simples

O OpenMP atingiu seu objetivo de portar códigos seriais para máquinas SMP com pouco trabalho. O OpenMP, diferentemente da maioria das bibliotecas de threads, também é portátil. Sua utilização é fácil de aprender, e é igualmente simples cometer erros de gravação. Em função disso, recomenda-se testar repetitivamente cada programa de paralelização em busca de erros que não ocorram em códigos seriais.

O OpenMP continua sendo constantemente desenvolvido. No compilador da Intel, estão disponíveis diversas propostas de melhoria que trazem otimizações, por exemplo, nos âmbitos do escalonamento e da geração de threads. A simplicidade do OpenMP também é motivo para diversas limitações claras em comparação com threads. Dessa forma, cada desenvolvedor deverá decidir se o OpenMP é adequado a seu projeto. Porém, levando em consideração que os sistemas multicore são regra, e não exceção, entre os consumidores, o OpenMP é uma boa opção na maioria casos. ■

## Mais Informações

[1] OpenMP: <http://www.openmp.org>

[2] GCC-Gomp: <http://gcc.gnu.org/projects/gomp>

Quem delega processos no Linux?

# O delegado

Uma das missões do kernel Linux é a distribuição de processos e threads da maneira mais proporcional possível nos núcleos da CPU. Algumas funções da API de escalonamento ajudam o programador de aplicativos a influenciar o kernel da melhor forma.  
por Eva-Katharina Kunst e Jürgen Quade



Para a boa escalabilidade de um sistema multinúcleo, a divisão das tarefas pendentes nos processadores disponíveis é primordial — o chamado balanceamento de carga. A grande trava do kernel (*Big Kernel Lock*) foi introduzida por Linus Torvalds para essa finalidade. Ela cobre a fila de execução, e alcançou um modelo de multi-escalonamento de CPU maduro e sofisticado no kernel 2.6. Nesse modelo, cada CPU administra sua própria fila de execução através do escalonador  $O(1)$  (veja o **quadro 1**).

Caso uma CPU estabeleça uma divisão de carga significativamente desigual entre si mesma e uma outra CPU, ela retira processos da fila de execução de alguma CPU sobrecarregada e os joga em sua própria fila (veja a **figura 1**). Nessa situação, seria ingênuo prestar atenção apenas em uma divisão igualitária dos processos disponíveis. O escalonador precisa fazer muito mais,

incluindo lidar com o trabalho adicional ocasionado pela transferência de uma thread de um processador para outro — a migração de processo — através do inescapável curso do *cache* do processador. Essa perda de desempenho depende fortemente da arquitetura de multiprocessadores utilizada.

## Zoológico de processadores

A menor complexidade é encontrada em arquiteturas *Hyperthreading* (*Symmetric Multithreading*, *SMT*). Nelas, estão disponíveis para o software dois núcleos de processador independentes — ou seja, dois processadores lógicos. Como ambos os núcleos dividem a memória principal e o cachê do processador, existem dependências significativas entre os processadores lógicos. A migração de uma thread de um núcleo lógico para a parte oposta correspondente, portanto, exige um custo. Além disso, com essa arquitetura, o ganho de desempenho também é comparativamente pequeno.

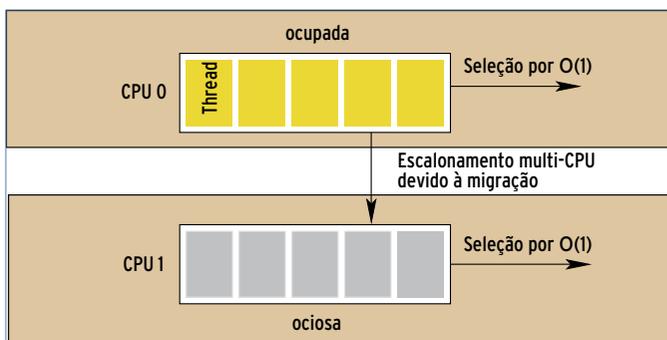
O multiprocessamento simétrico clássico (*SMP*), construído, por exemplo, com base nos processadores multinúcleo atuais, como Athlon X2 e Core 2 Duo, é diferente. A

cada mudança de processo, as informações de processos que porventura ainda se encontrem na CPU tornam-se obsoletas. Os sistemas *NUMA* são ainda maiores, com processadores que conseguem acessar mais rapidamente o espaço de memória destinado a eles do que a memória administrada por outras CPUs. Na realidade, é utilizada frequentemente uma combinação das diferentes arquiteturas básicas. Um sistema *NUMA*, por exemplo, consiste em diversos nós, cada um dos quais constituído de vários processadores (*SMP*). Até mesmo as CPUs *Hyperthreading* podem ser utilizadas.

## Domínios de escalonamento

Para lidar com as diferentes arquiteturas de multiprocessadores, o Linux definiu os conceitos de *Domínios de Escalonamento* e *Grupos de Escalonamento*. Um domínio de escalonamento consiste em vários grupos de escalonamento. Um grupo representa, por sua vez, uma CPU ou um domínio de escalonamento secundário. Um sistema de dois processadores, por exemplo, é modelado pelo Linux em um domínio de escalonamento com dois grupos, um para cada CPU. Enquanto isso, um computador com processador *Hyperthreading* é dividido em um domínio e dois grupos.

Um sistema heterogêneo, ao contrário, consiste em vários domínios, onde os grupos de escalonamento



**Figura 1** Cada CPU mantém uma lista própria dos processos nas filas de execução. Sob uma divisão desigual de carga significativa, o escalonador transfere processos entre as filas de execução.

**Tabela 1: Flags de domínio do Escalonador**

| Símbolo                       | Significado   |
|-------------------------------|---|
| <b>SD_WAKE_IDLE</b>           | Esse sinalizador é relevante quando o kernel desperta um processo de cálculo adormecido. Se o sinalizador for ativado para o respectivo domínio de escalonamento, o escalonador transfere o processo ativado para uma outra CPU dentro do domínio, caso ela esteja ociosa.  |
| <b>SD_WAKE_AFFINE</b>         | Tem efeito quando um processo de cálculo adormecido é ativado. Se o processo ativado não estiver mais no <i>cache</i> , o kernel o transfere para a CPU que executa o código do escalonador, ou então a <i>thread</i> permanece no processador destinado a ela.   |
| <b>SD_WAKE_BALANCE</b>        | Esse sinalizador é relevante quando um processo de cálculo adormecido é ativado. O processo de cálculo ativado é então transferido para a CPU atual, ou seja, aquela que executa o código do escalonador, caso ela esteja ociosa.   |
| <b>SD_BALANCE_FORK</b>        | O sinalizador tem significância quando uma tarefa chama a chamada de sistema <i>fork</i> . Nesse caso, o escalonador distribui equilibradamente os processos de cálculo entre os domínios seguintes, os quais ativaram o sinalizador.   |
| <b>SD_BALANCE_EXEC</b>        | Influencia o escalonador, quando uma tarefa chama a chamada de sistema <i>exec</i> . O escalonador procura a camada mais alta na árvore do domínio que ativou esse sinalizador. Depois, ele procura no interior desse domínio a CPU com menor carga e envia o processo para ela.  |
| <b>SD_BALANCE_NEWIDLE</b>     | O sinalizador é relevante quando uma fila de execução está vazia (o processador está ocioso). Se uma fila ficar vazia, o escalonador procura o domínio mais alto que estabeleceu esse sinalizador. No interior do domínio, é buscada a CPU mais carregada. O escalonador transfere um dos processos de cálculo dessa CPU para a CPU ociosa (transferência de processos de cálculo).   |
| <b>SD_SHARE_CPUPOWER</b>      | Esse sinalizador é tipicamente estabelecido no nível <i>SMT</i> . Se a CPU ativa estiver ociosa, ocorre uma migração de processo. Uma particularidade: caso o outro núcleo do processador <i>Hyperthreading</i> esteja executando um processo de cálculo com maior prioridade, o escalonador se ocupa da execução dos processos de cálculo com menor prioridade. Com isso, toda a capacidade de cálculo fica disponível ilimitadamente para o processo de tempo real. |
| <b>SD_SHARE_PKG_RESOURCES</b> | É estabelecido tipicamente no nível <i>SMP</i> . Serve unicamente para inicializar a variável <i>cpu_power</i>  |
| <b>SD_SERIALIZE</b>           | É normalmente utilizado no nível <i>//NUMA//</i> . Se for ativado, distribui irremediavelmente o kernel e realiza um <i>Soft-IRQ</i> .  |

do contêiner superior (domínios) não representam processadores, mas outros domínios. Um exemplo de topologia construída dessa forma é mostrado na [figura 3](#).

Dentro do domínio de escalonamento, o kernel cuida para que haja uma divisão de carga balanceada entre os grupos, procedimento este que pode ser configurado. Assim, o usuário pode determinar na inicialização se, no momento da chamada do sistema, *exec* ou *fork* devem ser balanceados, se isso deve acontecer na chamada a uma thread que está dormindo ou somente quando a CPU estiver rodando vazia (veja a [tabela 1](#)).

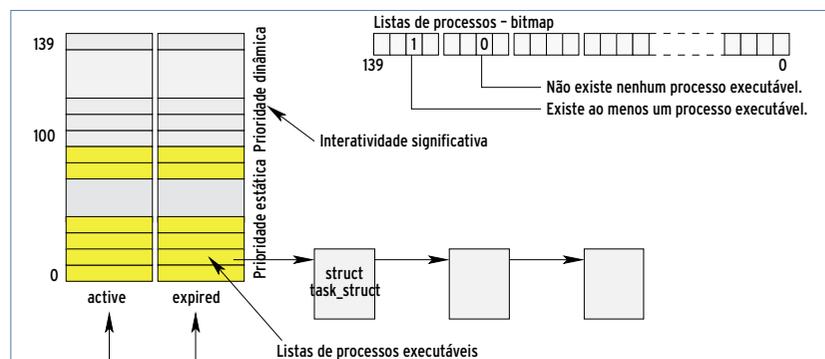
Para também distribuir a carga em um processo de cálculo intensivo, que não está adormecido e não chama *exec* ou *fork*, deve ser especificado para cada domínio de escalonamento um intervalo individual, dentro do qual o escalonador possa realizar um

novo balanceamento em cada caso. Isso acontece através do novo software *IRQ SCHED\_SOFTIRQ* e da função *run\_rebalance\_domain()*.

Em princípio, o Linux reconhece três tipos de domínio: SMT, SMP e NUMA. Logo na inicialização, o sistema se estabelece sobre uma topologia de domínio. Assim que o kernel identifica uma CPU *Hyperthreading*, ele chama para a inicialização do domínio correspondente

a macro *SD\_SIBLING\_INIT()*. Um nó NUMA inicializa *SD\_NODE\_INIT()*. As macros são específicas para cada arquitetura, e se encontram, por exemplo, em *arch/i386/kernel/topology.h*. Os sinalizadores de balanceamento de carga padrão utilizados para os sistemas *x86* estão listadas na [tabela 2](#).

Em princípio, é positivo que domínios SMT sejam balanceados frequentemente pelo kernel, devido às poucas



**Figura 2** Para o escalonamento, o kernel opera duas listas com processos, pertencentes às listas de prioridade normal ou de tempo real.

perdas na migração de processos, enquanto domínios NUMA, ao contrário, são mais raros. Para o balanceamento, naturalmente é preciso definir a carga de cada grupo de escalonamento. Para isso, o kernel escalona a carga de uma CPU com ajuda da variável `cpu_power`, disponível para cada grupo de escalonamento. Enquanto um grupo de escalonamento SMP de dois processadores possui uma `cpu_power` de 2, um processador Hyperthreading pode responder apenas com 1.

Caso o escalonador chegue à conclusão de que é necessária uma migração de processos, ele ativa a thread do kernel de prioridade máxima `migration/X`, disponível no sistema para cada CPU, e conduz a migração propriamente dita. No entanto, são migrados apenas os processos de cálculo que não estão ativos no momento e que não se encontram no cache, dos quais o escalonador então presume que não haja mais nenhuma informação útil para a continuidade do processo no cache da CPU. Depois que essas condições foram cumpridas, em alguns casos a afinidade da CPU evita ainda um deslocamento de processo. Cada processo de cálculo pode propriamente determinar em qual processador deve rodar, e com isso, também, em qual propriamente não deve.

As chamadas de sistema `sched_set_affinity()` e `sched_get_affinity()` alteram a afinidade da CPU. Com isso, o *CPU Blocker* de uma CPU pode

## Tabela 2: Padrões para arquiteturas de multiprocessadores

| SMT                | SMP                    | NUMA            |
|--------------------|------------------------|-----------------|
| SD_SIBLING_INIT()  | SD_CPU_INIT()          | SD_NODE_INIT()  |
| SD_BALANCE_NEWIDLE | SD_BALANCE_NEWIDLE     | SD_BALANCE_FORK |
| SD_BALANCE_EXEC    | SD_BALANCE_EXEC        | SD_BALANCE_EXEC |
| SD_WAKE_IDLE       | -                      | SD_WAKE_BALANCE |
| SD_WAKE_AFFINE     | SD_WAKE_AFFINE         | -               |
| SD_SHARE_CPU_POWER | SD_SHARE_PKG_RESOURCES | SD_SERIALIZE    |

## Quadro 1: Escalonamento de 1 Processador

A seleção do processo a ser recebido pela CPU (local) na seqüência é estabelecida pelo Linux por controle de prioridades. As prioridades disponíveis, que somam no total 140, são divididas em dois domínios: o das prioridades estáticas e o das prioridades dinâmicas. A diferença é que o escalonador pode alterar a prioridade no domínio dinâmico – de 100 até 139. A prioridade de um processo de cálculo no domínio estático – de 0 a 99 – não é modificada pelo escalonador. Essas prioridades são previstas para tarefas com demanda de tempo real, enquanto processos de cálculo normais recebem tipicamente uma prioridade do domínio dinâmico (figura 2).

Para cada prioridade há duas listas, nas quais estão inseridas as tarefas executáveis e threads – as listas *active* e *expired*. O escalonador escolhe o primeiro processo com a prioridade mais alta da lista *active*. Caso o processo, após o decorrer de um intervalo (ou fatia, no linguajar dos desenvolvedores) de tempo, não esteja finalizado ou tenha adormecido, ele o pendura na lista *expired*, da qual o escalonador multi CPU retira, por afinidade, o processo de cálculo a ser migrado.

Junto com o escalonamento *round-robin* descrito, o Linux emprega também, no domínio das prioridades estáticas (prioridades de tempo real), a seleção do próximo processo de cálculo segundo o princípio “Quem chega primeiro é o primeiro a se servir” (First Come First Serve).

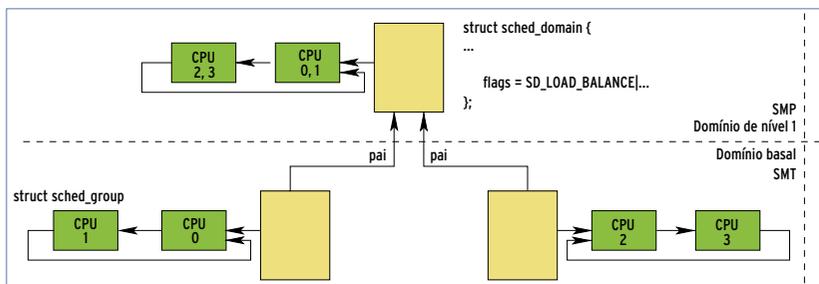
ser determinado, e o desempenho do cache otimizado. Além disso, as threads pertencentes a um grupo podem ser organizadas. Isso também leva as threads a dividirem seus dados em segmentos, e dessa forma alguns deles podem ser carregados no cache da CPU, em alguns casos, para aumento do desempenho.

## Isolamento para tempo real

A configuração de afinidades de CPU em ambiente de aplicações de tempo real é especialmente proveitosa. Na verdade, o kernel oferece a possibilidade de excluir um ou mais processadores do balanceamento através da opção de inicialização:

```
isolcpus=<CPU-Nummer>,...,<CPU-Nummer>
```

A única forma de se transferir um processo de cálculo para uma CPU isolada e ao mesmo tempo excluir dela um processo é a chamada de sistema `sched_set_affinity()`. Também se pode reservar uma CPU isolada para aplicações de tempo real, e dessa forma pode-se contar com



**Figura 3** O Linux constrói uma arquitetura com dois processadores *Hyperthreading* com três domínios de escalonamento: dois domínios basais e um domínio separado de nível 1, que contém todas as quatro CPUs. Cada domínio basal contém um processador com dois núcleos lógicos.

## Exemplo 1: getcpu.c

```

01 #include <stdio.h>
02 #include <stdlib.h>
03 #include <unistd.h>
04 #define __USE_GNU
05 #include <sched.h>
06 #include <errno.h>
07
08 static unsigned int getcpu()
09 {
10     unsigned int i, cpuid;
11     FILE *fp;
12
13     fp=fopen( "/proc/self/stat", "r" );
14     if( fp==NULL ) {
15         perror( "/proc/self/stat" );
16         exit( -1 );
17     }
18     fscanf( fp, "%*d %*s %*s" );
19     for( i=3; i<39; i++ ) // 0 39o. campomostra a ID da CPU
20         fscanf( fp, "%d", &cpuid );
21     return cpuid;
22 }
23
24 static unsigned int getaffinity( cpu_set_t *mask )
25 {
26     CPU_ZERO( mask );
27     if( sched_getaffinity( getpid(), sizeof(*mask), mask ) ) {
28         perror("sched_setaffinity");
29         exit( -2 );
30     }
31     return mask->__bits[0]; // XXX - soh um hack
32 }
33
34 int main( int argc, char **argv )
35 {
36     cpu_set_t mask;
37     int cpuid;
38
39     cpuid = getcpu();
40     printf("CPU ativa: %d\n", cpuid );
41     printf("Mascara de afinidade: 0x%x\n", getaffinity( &mask ) );
42     CPU_CLR( cpuid, &mask ); // Liberar a CPU ativa.
43     printf("Reiniciar a afinidade: 0x%x\n", mask.__bits[0] );
44     if( sched_setaffinity( getpid(), sizeof(mask), &mask ) ) {
45         perror("sched_setaffinity"); // Eh um sistema UP???
46     }
47     printf("Mascara de afinidade: 0x%x\n", getaffinity( &mask ) );
48     printf("CPU ativa %d\n", getcpu() );
49
50     return 0;
51 }

```

comportamentos de tempo essencialmente determinísticos.

O código de aplicativo no **exemplo 1** mostra como a afinidade deve ser lida, configurada, e também como fazer para forçar um processo de migração. Uma descrição das macros e funções utilizadas é fornecida em [2]. Esse código, entretanto, não deve fazer parte de um sistema produtivo.

Para simplificar, ele assume alguns compromissos: ele é adaptado para uma máquina de dois processadores. Além disso, ele não utiliza a chamada de sistema `get_cpu()`, introduzida com o kernel 2.6.19. Essa chamada controla a CPU sobre a qual rodam os processos atuais. Como a chamada é nova, ela ainda não está presente na biblioteca C padrão.

A chamada de função universal `syscall()` pode ajudar apenas em arquiteturas i386. Em uma arquitetura x86-64, `get_cpu()` é implementado como a chamada `vsyscall`, uma chamada que roda sozinha no espaço do usuário – sem transferência para o espaço do kernel. Entretanto, `vsyscall` requer outros mecanismos de chamada. O código de exemplo extrai a informação sobre qual CPU abriga o processo ativo do arquivo `/proc/self/stat`. A informação está no 39º campo (**exemplo 1, linha 19**).

O **exemplo 1** informa a CPU em que ele é executado. Depois, é forçada uma migração de processo, na qual ele apaga a CPU atualmente utilizada do campo de bit dos processadores em afinidade. A nova saída da CPU confirmará a transferência do processo.

## Trabalho manual

Resta finalmente destacar que diversos métodos – aperfeiçoados a cada versão do kernel – se encarregam de uma melhor distribuição de carga entre os processadores. Especialmente quem coloca altos requisitos para o comportamento de tempo dos processos pode colocar o escalonador no caminho certo com ajuda da afinidade da CPU. ■

## Mais informações

[1] Domínios de escalonamento e balanceamento de carga: [http://140.114.71.71/kerneltracing/slides/scheduling\\_domain.ppt](http://140.114.71.71/kerneltracing/slides/scheduling_domain.ppt)

[2] Documentação da biblioteca Glibc: <http://www.rdrs.net/document/c/gnuclibrary/html/CPU-Affinity.html>

## Os autores

**Eva-Katharina Kunst**, jornalista, e **Jürgen Quade**, professor da Faculdade de Niederrhein, são fãs do código aberto desde os primórdios do Linux. Juntos eles publicaram um livro com o título "Linux Treiber entwickeln" (Desenvolvimento do driver Linux), sobre o kernel 2.6.

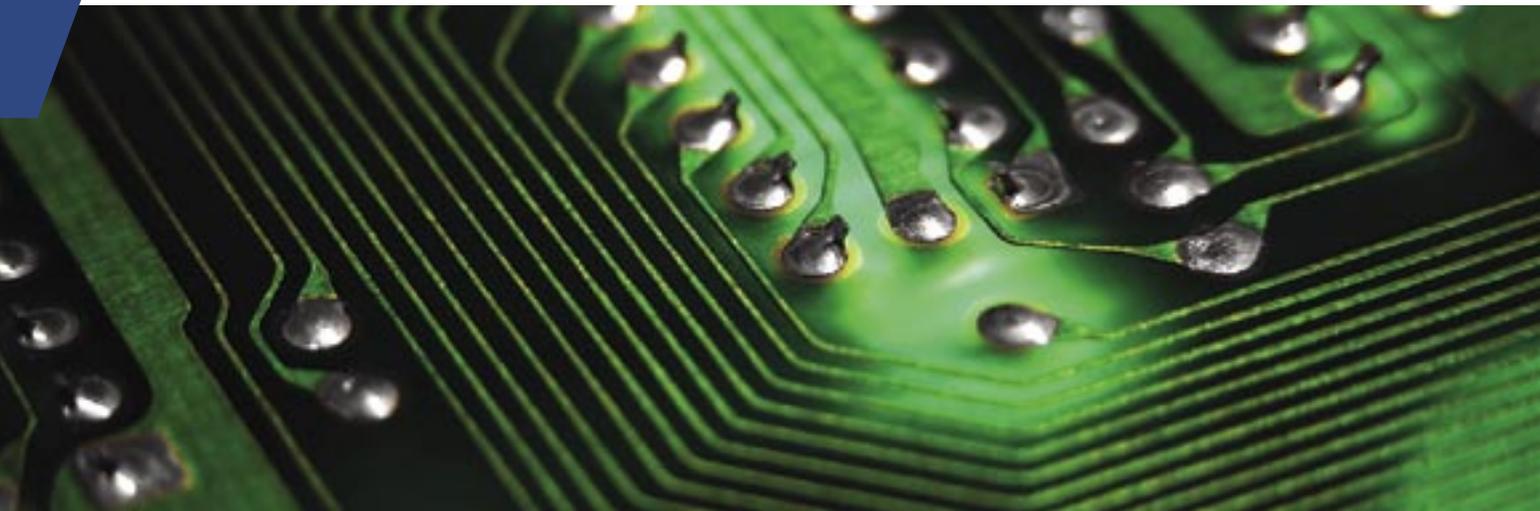
Transforme seu velho iPAQ em um novo dispositivo

# A libertação dos iPAQ

Somente software livre permite a retro-computação, ou seja, a reutilização de hardware considerado obsoleto pelo mercado por meio da adição de softwares atuais.

por **Alessandro de Oliveira Faria (Cabelo)**

Christopher Potter – sxc.hu



Neste artigo mostramos como proporcionar aos iPAQ Pocket PC, usando Linux e Software Livre, recursos disponíveis nos modernos desktops convencionais. Ao instalar Linux em seu iPAQ, o equipamento passa a contar com diversos softwares livres disponíveis no universo Open Source como, por exemplo, editores de textos, planilhas de cálculo, players de áudio e vídeo, dentre outros utilitários. Isso é possível graças a uma característica ímpar do Software Livre: a liberdade para modificar e recompilar software para outras plataformas.

Para a elaboração desta matéria, utilizamos um iPAQ h395, no qual o sistema operacional nativo é o Windows CE. O processo de libertação do iPAQ é composto de 4 fases:

- ◆ Backup do WinCE;
- ◆ Gravação do bootloader;
- ◆ Transferência do sistema Linux;

◆ Sobreposição do sistema operacional.

Vantagens de utilizar Linux no seu iPAQ:

- ◆ Portabilidade: Diversos modelos de hardware utilizando as mesmas aplicações atuais;
- ◆ Custo: Isenção de royalties ou licenças de uso;
- ◆ Flexibilidade: Ferramentas e utilitários usados em PCs de mesa (desktops) funcionam em modestos palmtops.

A distribuição utilizada foi o *Familiar Linux* [1] 0.8.4., baseada no Debian e compatível com diversos modelos iPAQ e Jornada. O kernel utilizado é o 2.4.19-rmk6-pxa1-hh41; a interface padrão é a *Opie* 1.2.1 ou, se preferir, *GPE*. O Familiar Linux tem suporte a portas seriais, USB, cartões PCMCIA, Cartões SD, rede Wi-fi, PPP sobre USB (módulo do kernel *usbnet*) infravermelho e bluetooth. Estão disponíveis também o suporte

a IPv6, Iptables, samba e ssh, entre outros serviços.

A interface *Opie* é baseada no *QT Embedded* sobre o kernel *FrameBuffer* – as aplicações são derivadas do KDE, como *Konqueror* e *Konsole*. Já a Interface *GPE* é baseada no GTK, e utiliza o servidor X reduzido, o *Kdrive*.

Antes de iniciar a instalação, vale a pena mencionar que a alteração do sistema operacional implica riscos. Sendo assim, não nos responsabilizamos por danos causados ao equipamento – ao executar os procedimentos deste documento esteja ciente de que erros podem acontecer.

## Instalação

Para a execução do processo de instalação do Familiar Linux precisamos, além de um iPAQ em condições de funcionamento, dos seguintes equipamentos:

- ▶ Base de sincronismo do PocketPC: acompanhada do respectivo cabo serial e USB;
- ▶ Computador: com Linux e os pacotes *minicom*, *SynCE* e *SynCE-KDE-Pack* instalados.

O iPAQ utilizado foi um exemplar do modelo h3950 com 64MBytes de RAM, 32 MBytes de ROM, um slot para cartão Mini SD e um processador PXA250 400 MHz. A base de sincronismo foi utilizada para efetuar a comunicação entre o iPAQ e o PC via comunicação serial, ou usando PPPoE sobre USB. Minicom é o software utilizado para transferir a imagem para o iPAQ, utilizando o protocolo *Ymodem*. Por fim, os pacotes *SynCE* e *SynCE-KDE-Pack* são utilizados para transferir o programa *BootBlaster* (em breve saberemos para o que ele serve) para o iPAQ.

Em primeiro lugar, baixaremos a imagem correspondente ao modelo do seu iPAQ em [2]. Nessa etapa, devemos escolher a versão do Familiar Linux que será utilizada – no meu caso, selecionei a versão estável 0.8.4.

Logo após, selecione o modelo do seu iPAQ. Caso esteja confuso, basta obter mais informações sobre o seu equipamento em [3]. Por último, selecione a interface gráfica OPIE ou GPE. Como sou “fã de carteirinha” do KDE, selecionei o OPIE, construído com *Qtopia*, que deriva do *QT*, que é a base do KDE.

Depois de toda configuração escolhida, clique em download. Vale a pena mencionar que esta imagem pode ser obtida diretamente em [4]. Basta escolher nesse link a pasta correspondente ao seu iPAQ e baixar a imagem conforme a sua preferência (OPIE ou GPE).

## O BootBlaster

*BootBlaster* é um programa com várias funções: salvar a imagem do WinCE (improvável você pre-

cisar dela novamente!), gravar o novo bootloader (imagem de boot para o kernel) na memória flash e verificar o novo bootloader. Para baixar este programa, entre no link do modelo de seu palmtop em [5] e selecione o arquivo *BootBlasterXXXX-2.6.exe* (onde XXXX é o modelo de seu iPAQ).

Devemos também baixar a imagem do bootloader que instalaremos no iPAQ. Baixe o arquivo com o nome *bootldr-xxxx.bin*.

## Pacotes suplementares

Os pacotes *synce*, *synce-kde-pack* e *minicom* da distribuição utilizada por você devem ser baixados e instalados para dar continuidade aos procedimentos deste documento. No nosso caso, utilizei os seguintes pacotes na minha distribuição OpenSuse 10.2:

- ▶ `synce-0.9.0-1.i386.rpm`
- ▶ `synce-kde-pack-0.9.1-6.i586.rpm`
- ▶ `rzs-0.12.20-871.i586.rpm`
- ▶ `minicom-2.2-12.i586.rpm`

O pacote *rzs* é necessário para o *minicom* trabalhar com o protocolo *YMODEM*, que utilizaremos na posterior transferência da imagem.

## Salvando a imagem do Windows CE

Claro está que, depois de instalado o Linux em seu iPAQ, você **nunca mais** voltará a utilizar o Windows CE. Mas, em todo caso, vamos fazer um backup da imagem.

Se você utiliza software proprietário, utilize o programa *ActiveSync* (figura 1) para enviar os arquivos *bootblaster* e *bootldr* para a pasta *My Documents* do iPAQ. Outra face da moeda: se você deseja utilizar apenas software livre para efetuar a transferência, efetue os procedimentos listados a seguir.



Figura 1 O caminho mais fácil – e proprietário – é utilizar o próprio *ActiveSync* do Windows para transferir arquivos.

Em primeiro lugar, conecte o cabo USB e verifique se tudo esta funcionando corretamente, utilizando o comando `dmesg`, como no exemplo abaixo:

```
cabelo@lapcognitec01:~> dmesg
drivers/usb/serial/usb-serial.c:
↳USB Serial support registered for
↳generic
usbcore: registered new driver
↳usbserial_generic
drivers/usb/serial/usb-serial.c:
↳USB Serial Driver core
drivers/usb/serial/usb-serial.c:
↳USB Serial support registered for
↳PocketPC PDA
drivers/usb/serial/ipaq.c: USB
↳PocketPC PDA driver v0.5
ipaq 1-1.2:1.0: PocketPC PDA
↳converter detected
usb 1-1.2: PocketPC PDA converter
↳now attached to ttyUSB0 <== Veja
usbcore: registered new driver ipaq
```

Se tudo estiver em perfeito funcionamento, veremos que o iPAQ se comunica pela porta `/dev/ttyUSB0`. Executaremos agora, como superusuário, o gerenciador de conexões para PocketPCs (*dccm*).

```
lapcognitec01:/home/cabelo # dccm
```

Caso seu iPAQ esteja protegido por senha, utilize o comando `dccm -p password`, como mostrado no exemplo abaixo:

```
lapcognitec01:/home/cabelo # dccm
->-p gisele
```

Defina a porta de comunicação com o comando `synce-serial-config`:

```
lapcognitec01:/home/cabelo #
->synce-serial-config ttyUSB0
You can now run synce-serial-
->start to start a serial
->connection.
```

Agora inicialize a comunicação PPP sobre USB com o comando `synce-seria-start`. Se tudo funcionar corretamente, você obterá um resultado similar ao exemplo abaixo:

```
lapcognitec01:/home/cabelo #
->synce-serial-start
Serial connection established.
Using interface ppp0
Connect: ppp0 /dev/ttyUSB0
local IP address 192.168.131.102
```

```
remote IP address 192.168.131.201
Script /etc/ppp/ip-up finished
->(pid 6072), status = 0x0
```

Para trabalhar com o iPAQ, no modo console, utilize os comandos seguintes:

- ▶ `prun`: Executa um programa (no iPAQ).;
- ▶ `pmkdir`: Cria uma pasta no iPAQ;
- ▶ `pmv`: Move ou renomeia arquivos;
- ▶ `pcp`: Copia arquivos do iPAQ;
- ▶ `prmdir`: Exclui um pasta;
- ▶ `prm`: Apaga arquivos;
- ▶ `pls`: Exibe o conteúdo de uma pasta.;

No exemplo abaixo, veja como copiar os arquivos `bootldr-pxa-xxx` e `BootBlaster3900-xxx` no modo console:

```
lapcognitec01:/home/cabelo # pcp
->bootldr-pxa-2.21.12.bin [:/My
->Documents/bootldr.bin"
lapcognitec01:/home/cabelo # pcp
->BootBlaster3900-2.6.exe]
->Documents/Bootblaster.exe"
```

A seguir, iremos executar o Bootblaster utilizando a linha de comando, em um terminal:

```
lapcognitec01:/home/cabelo # prun
->"/My Documents/Bootblaster.exe"
```

Se o pacote `synce-kde-pack` foi instalado, você pode utilizar o recurso `rapip` para, nada mais, nada menos, que integrar os comandos `synce` com o Konqueror, fazendo com que o navegador enxergue seu iPAQ como uma pasta do sistema. Com a habilitação deste recurso, basta arrastar os arquivos para o iPAQ (**figura 2**).

Agora, copiaremos a imagem do Windows CE, utilizando o programa `BootBlaster.exe` para essa tarefa.

Para efetuar o backup da imagem e do boot, clique no menu `FLASH`, e logo após em `Save Bootldr.gz Format` e `Save Wince.gz Format`, conforme mostrado na **figura 3**.

Este procedimento é demorado, durando por volta de seis minutos. Ao seu término serão criados, na pasta `My Documents` do iPAQ, os arquivos `asset_image.gz`, `saved_bootldr.gz` e `wince_image.gz`.

Copie esses arquivos para uma pasta em seu computador para um dia lembrar como o seu iPAQ era limitado... Após a conclusão do backup, não esqueça de apagar estes arquivos do iPAQ para liberar memória. Agora estamos prontos para instalar o `Bootldr`.

## Instalando o Bootldr

Consideramos esta operação a mais arriscada de todos os procedimentos deste artigo pois, se ocorrer algum erro nesta etapa (como, por exemplo, acabar a bateria), teremos uma probabilidade muito grande de destruir o bootloader. Portanto, preste bastante atenção.

Execute novamente o `BootBlaster.exe` e escolha a opção `FLASH` do menu principal, seguida do item `PROGRAM`. Uma mensagem de advertência será exibida informando para não resetar ou desligar o equipamento durante a operação.

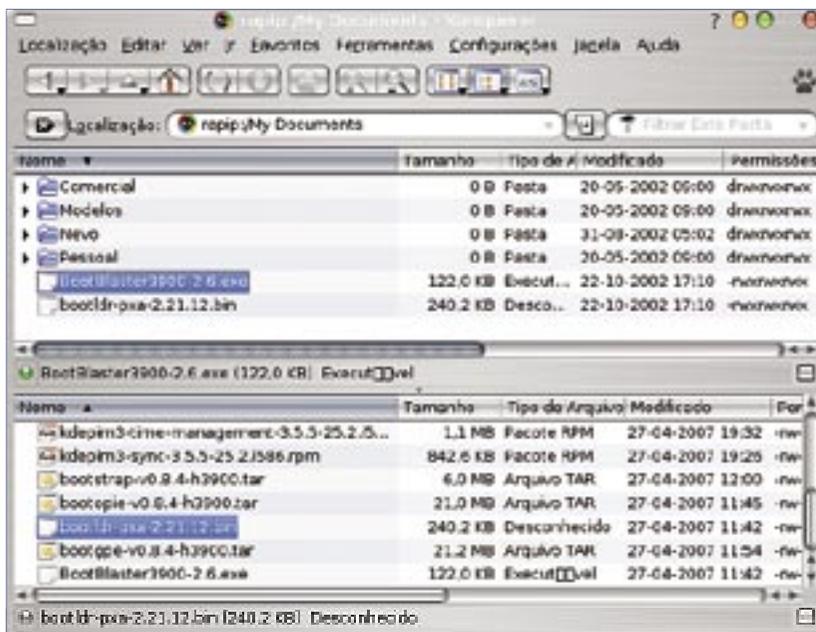


Figura 2 Utilize o `rapip` para integrar comandos do `SynCE` ao `Konqueror`.



**Figura 3** Execute o *BootBlaster.exe* dentro de seu iPAQ para executar o backup da imagem e o boot do aparelho.

Confirme o recebimento da advertência e selecione o arquivo a ser gravado na bootloader (bootldr-xxx.bin). Após a seleção, o programa descompacta o arquivo (se necessário) e o grava na memória flash.

Para garantir toda a operação, entre novamente no item *flash* do menu principal e selecione a opção *verify*. Este processo simplesmente verifica a integridade dos dados gravados no bootloader.

Ao concluir a instalação do bootloader, conferimos se o novo bootloader está em perfeito funcionamento. Basta pressionar o botão central de navegação junto com o botão reset (pequeno orifício) localizado na parte inferior do iPAQ. Se a instalação foi completada com êxito, o nosso pingüim Tux aparecerá na tela.

## Gravando a imagem no seu iPAQ

Chegou o momento da substituição do sistema operacional Windows CE pelo Familiar Linux. Conecte o cabo serial ao seu desktop e, no bootloader, escolha a opção *Serial Bootloader*.

*Console*. Aconselho a utilizar a conexão via porta serial e a base de sincronização para enviar a imagem Linux ao iPAQ.

Execute o comando `minicom -s` para configurar os parâmetros de comunicação. No menu principal [*configuration*], selecione a opção *Serial port setup*.

Quando utilizamos a conexão USB, o cabo de conexão USB utiliza a porta `/dev/ttyUSB0` em *Serial Device*. Agora se estiver utilizando o cabo serial, basta utilizar `/dev/ttyS0`. Em *Bps/Par/Bits*, especifique o valor `115200 8N1`. Desabilite a opção *Hardware e Software Flow Control*. Ao salvar as alterações, entre na opção *Exit* para iniciar a comunicação.

Se tudo estiver funcionando corretamente, obteremos um console onde devemos digitar o comando `load root`. Ao executar o comando, o iPAQ entrará em modo de recebimento via YMODEM. Veja no exemplo abaixo:

```
boot> load root
partition root is a jffs2
->partition:
expecting .jffs2 or wince_image.gz.
After receiving file, will
->automatically uncompress .gz images
loading flash region root
```



**Figura 4** iPAQ com Linux: 100% funcional e com cara de KDE

using ymodem  
ready for YMODEM transfer...

Bom, agora iremos iniciar o transplante... Pressione `CRTL+A+Z` e selecione a opção *SEND FILE* pressionando a *tecla S*. Selecione o protocolo *YMODEM* e escolha a imagem, navegando pelo sistema de arquivos – a imagem selecionada será transferida para o iPAQ. Após a transferência, o conteúdo da memória Flash será removido e a nova imagem automaticamente será gravada e verificada.

Após o término do processo, basta pressionar o botão *Boot FlashROM* no iPAQ e aproveitar seu PocketPC, que acaba de ser ressuscitado com a liberdade e robustez do Software Livre. Vale a pena conferir o vídeo do iPAQ funcionando com o Familiar Linux em [6]. ■

### Mais Informações

- [1] Site oficial do Familiar Linux: <http://familiar.handhelds.org>
- [2] <http://familiar.handhelds.org/releases/v0.8.4/install/download.html>
- [3] <http://en.wikipedia.org/wiki/IPAQ>.
- [4] <http://familiar.handhelds.org/releases/v0.8.4/install/files/>
- [5] <http://familiar.handhelds.org/releases/v0.8.4/install/files/>
- [6] <http://www.youtube.com/watch?v=izmztt2bhs8>

### Mais Informações

**Alessandro de Oliveira Faria (Cabelo)** é analista de Negócios da NETi Tecnologia (<http://www.netitec.com.br/alessandro>) e colaborador da comunidade Viva o Linux (<http://www.vivaolinux.com.br>).

# LPI nível 2: Aula 2

O primeiro passo da inicialização de um sistema Linux é o carregamento do kernel. Nesta aula vamos conhecer algumas particularidades do kernel e dos sistemas de arquivo do Linux.  
por Luciano Siqueira



**Linux  
Professional  
Institute**

## Tópico 202: Inicialização do sistema

### 2.202.1 Inicialização do sistema e processo de boot

O primeiro passo da inicialização de um sistema Linux é o carregamento do kernel.

Se a partição raiz for montada com sucesso, o primeiro processo disparado é o *init*, e justamente por isso o *init* possui número de processo – *pid* – igual a 1. Apesar de não ser utilizado por todas as distribuições, o *init* é o processo inicial padrão exigido na certificação LPI, e por isso nos ateremos a ele.

O *init* é responsável por invocar todos os serviços do sistema, através dos *scripts de inicialização*.

A localização dos *scripts de inicialização* varia de acordo com o padrão utilizado pela distribuição. No Slackware, que utiliza o padrão BSD, os *scripts* encontram-se em */etc/rc.d/*. Em sistemas como o Debian, que utilizam o padrão System V, os *scripts* encontram-se em */etc/init.d/*.

Quais *scripts* serão disparados dependerá do runlevel determinado no arquivo */etc/inittab*, pela linha:

```
id:4:initdefault:
```

Essa linha indica que o runlevel padrão é o 4. Num sistema Debian, serão disparados os *scripts* localizados no diretório */etc/rc4.d/*. Neste padrão

de inicialização existe um diretório que corresponde a cada nível execução.

Dentro de cada um dos diretórios existem links simbólicos para os *scripts* em */etc/init.d*. Constam apenas links para *scripts* que devem ser disparados no runlevel correspondente.

Num sistema que obedece ao padrão BSD, como o Slackware, os *scripts* de inicialização são disparados por *scripts* mestres disparados pelo *init*. Por exemplo, no runlevel 4, além dos *scripts* padrão (*/etc/rc.d/rc.S* e */etc/rc.d/rc.M*, que invocam *scripts* adicionais) será executado o */etc/rc.d/rc.4*, que se encarregará de disparar os *scripts* responsáveis por abrir o gerenciador de login gráfico.

### Administração de *scripts*

Apesar de poderem ser manipulados diretamente, muitas distribuições trazem ferramentas para ativar/desativar *scripts* na inicialização e ativar/desativar sua execução em tempo real. Alguns dos comandos não estão disponíveis em todas as distribuições.

### *insserv*

O comando *insserv* é utilizado para ativar ou desativar serviços durante a inicialização. Para que um serviço seja iniciado durante o boot, é usado o comando *insserv* da seguinte forma:

```
insserv named
```

Isso fará com que o daemon *named* seja iniciado em todo boot. Para

deixar de iniciar o serviço em todo boot, o próprio *insserv* é usado da seguinte forma:

```
insserv -r named
```

O arquivo de configuração do *insserv* encontra-se em */etc/insserv.conf*. Neste arquivo constam as dependências necessárias para cada serviço em particular.

### *chkconfig*

O *chkconfig* é uma espécie de interface para o *insserv*. Também é utilizado para ativar e desativar serviços, porém oferece alguns recursos adicionais. Utilizando *chkconfig --list serviço*, se for um serviço controlado pelo *xinetd*, será mostrado se o serviço está liberado ou não (*finger on* ou *finger off*). Se for um serviço de boot, será mostrado em quais runlevels é ativado: *anacron 0:off 1:off 2:on 3:on 4:on 5:on 6:off*.

Para ativar ou desativar um serviço, é utilizado o argumento *level*, da seguinte forma:

```
chkconfig --level 345 ntpd off
```

Esse comando fará com que o serviço *ntpd* seja desativado nos runlevels 3, 4 e 5.

Serviços controlados pelo *xinetd* são imediatamente ativados ou desativados pelo *chkconfig*, sem necessidade de reiniciar o *xinetd*. Outros serviços não são afetados imediatamente, mas somente no próximo boot. A inicialização ou a

interrupção imediata destes deve ser feita manualmente.

### *startproc*

O comando `startproc` checa por todos os processos do programa especificado como argumento e o inicia somente se nenhum processo for encontrado, para evitar redundância. O programa fornecido deve conter o caminho completo para o mesmo. A opção `-p caminho` pode ser utilizada para especificar o caminho para o arquivo que contém o PID do programa especificado.

### *checkproc*

O `checkproc` é utilizado para mostrar o(s) número(s) PID para o programa fornecido como argumento. O programa deve ser fornecido com seu caminho completo. Se também for fornecido o caminho do arquivo que contém o PID do programa, o `checkproc` mostra a linha de comando ou o nome base do programa.

### *killproc*

Semelhante ao programa `kill`, o `killproc` envia sinais para processos. Porém, o `killproc` toma como argumento principal o caminho completo para um programa. Todos os processos que correspondem ao programa indicado receberão o sinal indicado.

### *startpar*

O comando `startpar` é utilizado para iniciar scripts de inicialização em paralelo, no lugar de iniciá-los em série.

## 2.202.2 Recuperação do sistema

As duas possíveis causas para uma falha na inicialização do sistema é um serviço mal configurado ou uma falha no sistema de arquivos. Para ambos existem métodos eficientes de recuperação.

Tanto o Lilo quanto o Grub oferecem um prompt para editar e adicionar argumentos ao kernel que será carregado. Esses argumentos podem ser passados tanto nos arquivos de

configuração dos carregadores de boot quanto diretamente no momento em que o kernel será carregado, logo depois das mensagens do BIOS:

```
linux-2.6.18 S
```

Essa linha especifica o kernel chamado `linux-2.6.18`, entrando no runlevel *Single mode*. Para que nenhum script de inicialização seja disparado durante a inicialização, é utilizada a opção `init` no prompt do Lilo ou do Grub, antes do kernel ser carregado:

```
linux-2.6.18 init=/bin/bash
```

Dessa forma, no lugar de invocar o programa `init`, o único programa invocado logo após o kernel ser carregado será o interpretador de comandos `bash`. Antes que seja possível fazer reparos diretamente no arquivo `/etc/inittab` ou no(s) script(s) problemáticos, é necessário remontar o dispositivo raiz com a opção `rw` (leitura e escrita), pois o kernel, no primeiro momento, monta a partição raiz apenas para leitura:

```
mount -o remount,rw /
```

Pode ser utilizado um editor de texto padrão, como o `vi`, para efetuar as alterações em arquivos de configuração. Depois de realizadas as alterações nos arquivos, como alterar o runlevel padrão no `/etc/inittab` para que o próximo boot encaminhe o sistema para *single mode* (de forma que seja possível prosseguir a manutenção) ou evitar que o modo gráfico inicie, podemos remontar a partição raiz de volta para somente leitura:

```
mount -o remount,ro /
```

### *Falha no sistema de arquivos*

Falhas de leitura ou escrita podem impedir que o sistema inicie por completo. Uma mensagem como essa pode aparecer durante o processo de inicialização:

```
*** An error occurred during the
↳file system check."
*** Dropping you to a shell; the
↳system will reboot"
*** when you leave the shell."
Give root password for
↳maintenance(or type Control-D for
↳normal startup):
```

Após fornecer a senha de root, será possível utilizar ferramentas como o `fsck` para corrigir os erros no sistema de arquivos:

```
fsck /dev/sda1
```

### *Recuperação do setor de boot*

No caso do kernel do sistema se tornar inacessível, será necessário utilizar uma mídia de boot alternativa, como um live CD, para conseguir um sistema funcional.

Carregado o sistema alternativo, é necessário montar a partição raiz do sistema com problemas, para reinstalar o carregador de boot:

```
mount /dev/hda1 /mnt
chroot /mnt /bin/bash
grub-install /dev/hda
ou
lilo -v
```

Este exemplo assume que os arquivos do kernel (imagem do kernel e `initrd`) estejam na mesma partição problemática. Caso não estejam, será necessário montar a partição onde se encontram esses arquivos antes de executar o comando `grub-install` ou `lilo`.

## Tópico 203: Sistema de arquivos

### 2.203.1 Trabalhando com o sistema de arquivos Linux

No Linux, todo tipo de mídia ou partição de disco rígido é acessada através do processo de montagem.▶

Este procedimento pode ser realizado automaticamente ou manualmente, através do comando `mount`.

### *mount*

O comando `mount` é encarregado de anexar uma partição contendo um sistema de arquivos a um diretório na árvore, de forma que seu conteúdo torne-se acessível.

Forma mais comum de utilização do comando `mount`:

```
mount -t tipo dispositivo
↳ diretório
```

A opção `-t` determina qual é o sistema de arquivos utilizado na partição ou dispositivo a ser montado, especificado pelo termo `dispositivo`. O último argumento, `diretório`, indica qual diretório será o ponto de montagem da partição ou dispositivo.

O Linux é compatível com uma grande diversidade de sistemas de arquivo. O `mount` é capaz de descobrir a maioria dos sistemas de arquivos, sem necessidade de fornecer a opção `-t`. Porém, alguns tipos de sistemas de arquivos, como sistemas de arquivo remotos, precisam ser especificados. Quando usado sem qualquer argumento, o `mount` mostra os dispositivos montados.

Essas informações também podem ser consultadas em `/etc/mtab` e em `/proc/mounts`. O arquivo `mtab` já não é mais o padrão em sistemas Linux, mas ainda é mantido por questões de compatibilidade.

Conteúdo do arquivo `/etc/mtab`:

```
/dev/sda1 / xfs rw 0 0
proc /proc proc rw 0 0
sysfs /sys sysfs rw 0 0
usbfs /proc/bus/usb usbfs rw,devg
↳ d=10,devmode=0666 0 0
/dev/sda5 /home xfs rw 0 0
/dev/hda5 /mnt/hda5 xfs rw 0 0
none /sys/fs/fuse/connections
↳ fusectl rw 0 0
```

Algumas distribuições apenas mantém um link simbólico em `/etc/mtab` para o arquivo `/proc/mounts`. O conteúdo de `/proc/mounts` é bastante semelhante aos demais.

A única diferença entre o `/etc/mtab` e o `/proc/mounts` demonstrados é a diretiva `rootfs` em `/proc/mounts`. O `rootfs` refere-se a partição temporária criada para o inird, durante o boot.

Se o dispositivo ou partição está previamente configurado no arquivo `/etc/fstab`, será necessário fornecer ao `mount` apenas o caminho completo para a partição ou para o ponto de montagem. Para desmontar o dispositivo, utiliza-se o comando `umount`. Para o `umount`, é necessário fornecer somente o caminho completo para o dispositivo ou ponto de montagem, mesmo que a partição montada não conste em `/etc/fstab`.

### *fstab*

Durante os procedimentos de carregamento do sistema, é o arquivo `/etc/fstab` que determina os pontos de montagem dos dispositivos. Cada linha corresponde a um dispositivo, contendo os seguintes termos separados por tabulações ou espaços:

1. Dispositivo
2. Ponto de montagem (swap se for uma partição de swap)
3. Tipo de sistema de arquivos
4. Opções
5. Dump (0 ou 1). Determina se o dispositivo deverá ser considerado pelo comando `dump`. Se ausente, 0 é considerado.
6. `fsck` (1 ou 2). Determina a ordem da checagem feita pelo `fsck` durante o boot. Para a partição raiz (`/`), deve ser 1. Se ausente, 0 é presumido e a checagem não será feita no boot.

### *Swap*

Por padrão, o espaço de swap é criado durante a instalação do sistema e ativado a cada boot. Uma entrada de partição swap no `fstab` corresponde a seguinte linha:

```
/dev/sda2 swap swap defaults 0 0
```

Note que não há ponto de montagem para partições swap, porque não se trata de um sistema de arquivos do ponto de vista tradicional. Onde deveria estar indicado o ponto de montagem, consta apenas o termo `swap`.

Dependendo das necessidades do sistema, pode ser necessário criar partições swap adicionais. O sistema pode tornar-se lento caso alguma aplicação exija mais memória do que o computador dispõe, seja ela física ou swap. Neste cenário, o espaço de swap deve ser aumentado ou deve ser criado uma ou mais partições swap adicionais. Até o kernel 2.4.10, o número máximo de áreas de swap era 8. Após essa versão, o número máximo aumentou para 32.

O comando utilizado para configurar uma partição como swap é o `mkswap`. Sua sintaxe é simples, sendo necessário fornecer como argumento apenas o caminho completo para a partição a ser utilizada.

O `mkswap` apenas configura a partição para ser utilizada como swap, mas não a ativa. Para ativar uma partição, usa-se o comando `swapon`. Como ocorre no o `mkswap`, é necessário apenas fornecer o caminho completo para a partição, mas é possível fornecer algumas opções:

- ◆ `-a`: Ativa todas as partições marcadas como swap em `/etc/fstab`, exceto aquelas que possuem a opção `noauto`. As partições swap contidas em `/etc/fstab` que já estiverem ativas serão ignoradas;
- ◆ `-e`: Usada com `-a` faz com que partições não encontradas sejam ignoradas sem aviso;
- ◆ `-L label`: Consulta o arquivo `/proc/partitions` e usa a partição correspondente a `label`;
- ◆ `-p 0-32767`: Define a prioridade da área de swap. Quanto maior o número, maior a prioridade. No arquivo `/etc/fstab` é indicado com a opção `pri=0-32767`. Áreas de maior prioridade são utilizadas por

completo antes que áreas de menor prioridade sejam utilizadas. Áreas de mesma prioridade são utilizadas em paralelo. Podem ser utilizados valores negativos para estabelecer prioridades ainda menores;

- ▶ -s: Mostra o uso de cada espaço de swap, exatamente igual ao conteúdo do arquivo `/proc/swaps`;
- ▶ -U `uuid`: Consulta o arquivo `/etc/partitions` e utiliza a partição de `uuid` correspondente.

Para desativar um espaço de swap, utiliza-se o comando `swapoff`. Pode ser usado fornecendo como argumento o caminho completo para a partição ou com a opção `-a`. No último caso, todos os espaços de swap encontrados em `/etc/fstab` e `/proc/swaps` serão desativados.

Arquivos comuns também podem ser utilizados como áreas de swap, exatamente como se fossem partições. Para criar um arquivo com propósito de swap, pode ser utilizado o comando `dd`:

```
# dd if=/dev/zero of=swapfile
↳bs=1024k count=256
```

Os procedimentos seguintes são os mesmos, apenas trocando o caminho completo da partição pelo arquivo recém criado.

### Sync

Quando trabalhando com montagem e desmontagem de sistemas de arquivos, é importante ter em mente que alguns dados correm risco de serem perdidos. Para obter maior velocidade, os dados são mantidos num cache antes que sejam realmente gravados no dispositivo. Se um dispositivo for desmontado antes que esses dados em cache sejam gravados, os mesmo serão perdidos. Apesar de ser muito difícil de ocorrer, uma vez que o kernel faz a sincronização de dados em disco automaticamente, é recomendável executar o comando `sync` toda vez que uma desmontagem for realizada ou um dispositivo móvel – como um pendrive – for removi-

do. Dependendo de quantos dados estiverem armazenados no cache, o `sync` pode demorar até aproximadamente 30 segundos para finalizar a gravação.

## 2.203.2 Manutenção de um sistema de arquivo Linux

Sistemas Linux são capazes de trabalhar com praticamente todos os tipos de sistemas de arquivos. Os mais utilizados, no entanto, são o `ext3` – sucessor do `ext2` – e `reiserfs`. Outros sistemas de arquivos, como o `XFS`, também são bastante utilizados, mas nos ateremos à criação e manutenção do `ext2`, `ext3` e `reiserfs`.

### Criando um sistema de arquivos

Após criar uma partição através de um particionador, como o `fdisk` ou o `cfdisk`, é necessário criar um sistema de arquivos na mesma para que possamos gravar e ler arquivos. O comando padrão para essa tarefa é o `mkfs`, que invoca comandos como `mkfs.ext2`, `mkfs.ext3` ou `mkfs.reiserfs`, dependendo do sistema de arquivos indicado pela opção `-t`. Para o `ext2` ou `ext3`, também é possível usar o comando `mke2fs`.

Um sistema de arquivos pode ser criado apenas fornecendo como argumento o caminho completo para a partição:

```
# mke2fs /dev/hda6
mke2fs 1.38 (30-Jun-2005)
Filesystem label=
OS type: Linux
Block size=4096 (log=2)
Fragment size=4096 (log=2)
856480 inodes, 1712922 blocks
85646 blocks (5.00%) reserved for
↳the super user
First data block=0
53 block groups
32768 blocks per group, 32768
↳fragments per group
16160 inodes per group
Superblock backups stored on blocks:
    32768, 98304, 163840, 229376,
↳294912, 819200, 884736, 1605632
```

```
Writing inode tables: done
Writing superblocks and filesystem
↳accounting information: done
```

```
This filesystem will be
↳automatically checked every 21
↳mounts or
180 days, whichever comes first.
↳Use tune2fs -c or -i to override.
```

No exemplo foi criada uma partição `ext3`, que não tem suporte a `journaling`. Para criar uma partição `ext3`, que utiliza `journaling`, poderiam ser usados três comandos diferentes: `mkfs -t ext3 /dev/hda6`, `mkfs.ext3 /dev/hda6` ou `mke2fs -j /dev/hda6`.

### Verificando e corrigindo

O comando `dumpe2fs` é utilizado para mostrar diversas informações de baixo nível sobre um sistema de arquivos `ext2` ou `ext3`. No entanto, sua utilização sem argumentos mostra muito mais informações que um administrador comumente precisaria ver. A opção `-h` oferece um visão geral sobre o sistema de arquivos:

```
Filesystem volume name: <none>
Last mounted on: <not available>
Filesystem UUID:
↳0b7d5787-34c1-43d4-bed3-
↳8995fa0158cd
Filesystem magic number: 0xEF53
Filesystem revision #: 1
↳(dynamic)
Filesystem features: filetype
↳sparse_super
Default mount options: (none)
Filesystem state: clean
Errors behavior: Continue
Filesystem OS type: Linux
(...)
```

Essas informações também podem ser obtidas usando `tune2fs -l /dev/hda6`. Se o sistema de arquivos utilizar `journaling`, aparecerá o termo “`has_journal`” na linha `Filesystem features`. A principal função do comando `tune2fs`, no entanto, é de ajustar algumas configurações de sistemas de arquivo `ext2` e `ext3`, por exemplo: ▶

- ▶ **-c max-montagens:** Determina o número máximo de montagens até que a partição seja automaticamente verificada pelo `fsck`. Se for `0` ou `-1`, o número de vezes que a partição foi montada será ignorado;
- ▶ **-i intervalo[d|m|w]:** Determina o período máximo de tempo até que a partição seja verificada automaticamente. A letra `d` especifica dias, `m` meses e `w` semanas. O valor zero determina que a verificação por tempo deve ser ignorada;
- ▶ **-j:** Converte uma partição `ext2` para `ext3`.

É altamente recomendável que a verificação por tempo ou número de montagens seja utilizada, para evitar possíveis perdas de dados.

O comando correlato ao `tune2fs` para sistemas de arquivos `reiserfs` é o `reiserfstune`. O `reiserfstune` é utilizado principalmente para alterar propriedades de `journaling`, mas também pode alterar opções como o `uuid` do sistema de arquivos (`-u UUID`) ou o `label` do sistema de arquivos (`-l LABEL`). Note que a opção `-l` tem diferentes efeitos para o `tune2fs` e para o `reiserfstune`.

Apesar de não serem freqüentes, sistemas de arquivos Linux podem apresentar erros de médio a longo prazo, especialmente em servidores muito exigidos. Muitos desses erros podem ser contornados com o uso adequado de programas específicos. Mesmo erros físicos no dispositivo costumam ter remédio, até que sejam substituídos por um novo.

A ferramenta padrão de reparo em sistemas de arquivos Linux é o `fsck`, que é apenas um atalho para comandos específicos para cada sistema de arquivos, como `fsck.ext2` e `fsck.reiserfs`. Para `ext2` e `ext3`, existe ainda a variante `e2fsck`.

A maneira correta de verificar um sistema de arquivos é fazê-lo com a partição desmontada. Se a partição a ser verificada for a raiz (`/`), um `live CD` pode ser utilizado.

A sintaxe básica do `fsck` é:

```
fsck /dev/hda1
```

No exemplo, será verificada a primeira partição do disco conectado ao canal IDE primário.

O `fsck` irá consultar o arquivo `/etc/fstab` para descobrir o tipo de sistema de arquivos da partição indicada para verificação via linha de comando. Caso a verificação esteja feita via `live CD` ou a partição a ser verificada não conste no `/etc/fstab`, o sistema de arquivos deve ser especificado à opção `-t`:

```
fsck -t reiserfs /dev/hda1
```

ou:

```
fsck.reiserfs /dev/hda1
```

Opções comuns do comando `fsck`:

- ▶ **-A:** Consulta o arquivo `/etc/fstab` e tenta checar todos os sistemas de arquivos encontrados. Se também for especificado um ou mais sistemas de arquivos (separados por vírgula) com a opção `-t`, apenas os sistemas de arquivos do tipo especificado serão verificados. Se o tipo de sistema de arquivos for precedido por `!`, partições com este tipo de sistema de arquivos serão ignoradas;
- ▶ **-s:** Faz a checagem das partições em série. Por padrão, o `fsck` roda em modo interativo. A opção `-s` evita que haja confusão entre mensagens e perguntas de partições;
- ▶ **-N:** Não faz nenhuma modificação, apenas mostra o que seria feito;
- ▶ **-P:** Checa a partição raiz em paralelo com as demais;
- ▶ **-R:** Usada a opção `-A`, indica que a partição raiz deve ser ignorada (caso esteja montada com opção de escrita).

As opções a seguir podem variar de acordo com o sistema de arquivos a ser analisado/corrigido. Consulte a página manual do `fsck` para cada

sistema de arquivos para assegurar-se das opções:

- ▶ **-a:** Corrigir erros automaticamente, sem fazer nenhuma pergunta;
- ▶ **-n:** Não faz nenhuma correção, apenas informa o erro na saída padrão;
- ▶ **-r:** Realiza os reparos em modo interativo. Pode causar confusão se várias checagens estão acontecendo em paralelo;
- ▶ **-y:** Essa opção determina que o `fsck` adote a alternativa padrão para qualquer problema que encontrar.

### debugfs

Para fazer alterações de baixo nível no sistema de arquivos `ext2` e `ext3`, existe o comando `debugfs`. O comportamento padrão do `debugfs` é abrir um `prompt` para a partição solicitada, onde podem ser utilizados comandos específicos do `debugfs`. Uma utilização bastante comum desse comando é recuperar arquivos apagados em partições `ext2`:

```
# debugfs -w /dev/hda6
debugfs: lsdel
Inode Owner Mode Size
└─Blocks Time deleted
  12    0 100644 32 1/
└─1 Thu May 10 17:42:13 2007
1 deleted inodes found.
debugfs: undelete<12>relatorio.txt
```

No exemplo acima, o `debugfs` abriu a partição `/dev/hda6` com permissão de escrita (`-w`). No `prompt` do `debugfs`, o comando `lsdel` mostrou que o arquivo de `inode 12` foi apagado. O comando `undelete` foi utilizado para anexar o `inode 12` ao arquivo de nome `relatorio.txt`.

Semelhante ao `debugfs`, partições `reiserfs` desfrutam da ferramenta `debugreiserfs`. Sem argumentos, o `debugreiserfs` tem o mesmo comportamento que o `dumpe2fs -h` ou o `tune2fs -l` para partições `ext2/ext3`. O `debugreiserfs` não é capaz de recuperar arquivos apagados.

### 2.203.3 Criando e configurando opções para sistemas de arquivos

Além de criar, configurar e fazer a manutenção de sistemas de arquivos em partições de disco, este tópico contempla a montagem automática de dispositivos e a criação e gravação de imagens para mídias de CD-ROM.

#### Montagem automática

É comum que computadores em rede compartilhem espaço em disco num servidor e a maneira mais usual de fazê-lo é montar o compartilhamento remoto durante o boot das estações. No entanto, se por algum motivo a conexão entre as máquinas cair, o compartilhamento não estará mais acessível.

Uma solução para este problema é utilizar um sistema de montagem sob demanda. Dessa forma, toda vez que o usuário acessar um diretório

pré-determinado, o compartilhamento remoto será montado, caso já não esteja. Essa solução também pode ser adotada para mídias removíveis, evitando a utilização manual do comando `mount`.

#### autofs

Existem diversos tipos de ferramentas de montagem automática para Linux. Cada uma delas tem suas vantagens e desvantagens, ficando reservado ao gosto pessoal do administrador a escolha de uma em detrimento da outra, uma vez que ao usuário final as diferenças são praticamente imperceptíveis.

Dentre os vários sistemas de montagem automática que podem ser utilizados, veremos como utilizar o `autofs`.

O primeiro arquivo de configuração é o `/etc/auto.master`. Nele são especificados quais diretórios serão monitorados.

Conteúdo de um arquivo `/etc/auto.master` simples:

```
/misc /etc/auto.misc --
timeout=10
```

Este arquivo é chamado de mapa mestre (*master map*). Pode haver mais de uma linha no mapa mestre, cada uma especificando um diretório a ser monitorado. O mapa mestre é único e é consultado durante a inicialização do sistema. O script de inicialização (geralmente `/etc/init.d/autofs` ou `/etc/rc.d/autofs`) checa o arquivo `/etc/auto.master` e passa cada linha encontrada como argumento para o comando `automount`, que permanecerá ativo e monitorando os diretórios especificados no mapa mestre. Podem existir várias instâncias do comando `automount`, uma para cada diretório monitorado.

Para cada um desses diretórios monitorados, deve haver um arquivo chamado arquivo de mapa, especificado no segundo campo do mapa



# Certificação Linux Número 1 no Mundo



**LPIC-1:** reconhecida no mundo todo como a certificação inicial para profissionais de Linux

**LPIC-2:** uma certificação avançada em Linux, largamente reconhecida como uma "HOT CERT" do mercado, que proporciona os mais altos salários entre os profissionais de Linux

**LPIC-3:** a primeira certificação profissional enterprise-level em Linux, disponível a partir de janeiro de 2007

**OSPRES:** um programa único de progresso na carreira para TODOS os profissionais de Open Source



Saiba mais,  
faça-nos uma visita  
[www.lpi.org/americalatina](http://www.lpi.org/americalatina)

mestre (no exemplo, `/etc/auto.misc`). O terceiro campo é opcional e pode conter opções que são passadas para o `automount`. No caso do exemplo, `--timeout=10` determina que após uma inatividade de 10 segundos no diretório monitorado, os subdiretórios que porventura estiverem montados serão desmontados.

No arquivo de mapa é que são especificados os diretórios que serão montados automaticamente. A sintaxe de cada linha no arquivo de mapa é:

```
key [-options] location
```

Ainda seguindo o mapa mestre mostrado anteriormente, quando algum evento no diretório `/misc` for detectado, o arquivo `/etc/auto.misc` será consultado. Exemplo de `/etc/auto.misc`:

```
cdrom -fstype=iso9660,ro :/dev/hdc
ernesto -fstype=nfs
↳192.168.11.1:/home/ernesto
```

Neste caso, quando o usuário tentar entrar nos diretórios `/misc/cdrom`, `/misc/ernesto` ou em qualquer subdiretório dentro deles, os diretórios serão montados automaticamente. O primeiro campo de cada linha (*key*) determina o nome do ponto de montagem dentro do diretório monitorado. O segundo campo (*-options*), começado com um hífen, determina as opções de montagem e tipo de sistema de arquivo. O terceiro campo (*location*) especifica o dispositivo ou local de rede a ser montado. No exemplo, a primeira linha especifica o dispositivo `/dev/hdc` na máquina local a ser montado automaticamente em `/misc/cdrom` e a segunda linha especifica o compartilhamento NFS `/home/ernesto` na máquina `192.168.11.1`, a ser montado em `/misc/ernesto` na máquina local.

É possível, ainda, especificar uma fonte remota no lugar de um

arquivo de mapa local, como uma base de dados NIS ou um diretório LDAP. Essas alternativas estão disponíveis invocando o comando `automount` diretamente.

Outro recurso importante nos arquivos de mapa são os caracteres curinga. O caracter `*` no campo *key* corresponde a qualquer ocorrência no diretório monitorado. O caracter `&` no campo *location* será substituído pelo campo *key* da mesma linha.

Utilização de caracteres curinga no arquivo `/etc/auto.misc`:

```
* -fstype=nfs 192.168.11.1:/home/&
```

Quando um usuário ou programa tentar acessar, `/misc/ademar` ou `/misc/osires`, as seguintes entradas serão invocadas, respectivamente:

```
ademar -fstype=nfs
↳192.168.11.1:/home/ademar
osires -fstype=nfs
↳192.168.11.1:/home/osires
```

### Criação de imagens de CD

Uma imagem de CD é um arquivo contendo dados prontos para serem passados diretamente para a mídia.

A criação de uma imagem a partir de uma mídia que já contenha dados é simples. Uma das maneiras de executar essa tarefa é utilizar o comando `dd`:

```
dd if=/dev/cdrom of=imagem.iso
```

ou:

```
dd < /dev/cdrom > imagem.iso
```

A imagem gerada poderá ser copiada para um CD através do `cdrecord`:

```
cdrecord dev=ATA:0,0,0 imagem.iso
```

O comando mostrado assume que o gravador simula um dispositivo SCSI, de LUN `o,o,o`. Em

kernels mais recentes (versão 2.6 em diante), é possível simplesmente especificar o caminho do dispositivo:

```
cdrecord dev=/dev/cdrecord imagem.
↳iso
```

Para criar imagens de CD a partir de arquivos, é utilizado o comando `mkisofs`. A maneira mais simples de utilizar o `mkisofs` é simplesmente indicar o arquivo imagem a ser criado e o(s) diretório(s) de origem dos arquivos:

```
mkisofs -o imagem.iso diretório/um
↳diretório/dois ...
```

Todos os arquivos nos diretórios especificados serão criados na raiz da imagem. Para especificar locais diferentes dentro da imagem, utiliza-se a opção `-graft-points`:

```
mkisofs -o imagem.iso -graft-
↳points parte_um=dir_um parte_
↳dois=dir_dois
```

Outras opções importantes do comando `mkisofs`:

- ▶ `-J`: Gera registros *joliet* além dos nomes de arquivos regulares do ISO9660. Importante se o CD será lido em máquinas com Windows NT ou Windows 95;
- ▶ `-R`: Gera imagens do tipo *Rock Ridge*, que preservam atributos especiais de arquivos Unix (links, permissões, etc);
- ▶ `-udf`: Inclui compatibilidade UDF (*Universal Disk Format*) à imagem gerada.

### Considerações Finais

Lembrem-se dos procedimentos de manutenção, principalmente de como passar opções ao Kernel. Os comandos de criação e manutenção de sistemas de arquivos serão abordados em detalhes. ■

Cacti a alternativa ao Nagios e ao Zabbix

# Monitorar é preciso

Entre as várias atividades desenvolvidas pelos administradores de redes está a de monitorar roteadores, switches, servidores e seus respectivos serviços afim de garantir taxas aceitáveis de desempenho e disponibilidade. O Cacti é uma boa opção de ferramenta de monitoramento que pode ajudar muito nesta tarefa.

por Adriano Matos Meier



SYSADMIN

O número e o tamanho das redes de computadores crescem à medida que a quantidade de informação e de usuários que precisam dessas informações aumentam. Entretanto, esse crescimento e as diferentes tecnologias utilizadas acabam tornando a mais comum rede de computadores difícil de ser monitorada.

Em função disso, diversas ferramentas são desenvolvidas com a idéia de facilitar o trabalho dos gerentes de TI e dos administradores de redes. Entre os critérios para escolha dessas ferramentas podemos citar o custo, a plataforma utilizada, a forma de distribuição da licença, abrangência do monitoramento, facilidade de instalação, configuração e utilização, confiabilidade e documentação, dentre outros.

Muitos são os motivos que levam os administradores a se preocupar com essa gerência – entre eles, conhecer melhor a própria rede, garantindo uma melhor segurança, desempenho e disponibilidade, registro e armazenamento de ocorrência de eventos, além, é claro, de controle de recursos.

## Características da ferramenta

O Cacti [1] é uma ferramenta de monitoramento de rede completa distribuído sob a licença GPL. Surgiu como uma opção de frontend para o RRDTool (criado por Tobias Oetiker, mesmo criador do MRTG), que é o responsável por armazenar os dados coletados através do SNMP ou de scripts e posteriormente gerar os gráficos. Possui suporte ao SNMPv3 (SNMP versão 3), o que o torna mais seguro, pois as consultas são realizadas somente ao informar um nome de usuário e senha válidos.

O Cacti disponibiliza um ambiente de operação agradável e acessível via interface web. Informações como a quantidade de tráfego nos enlaces, uso da CPU, memória e discos rígidos, número de usuários ativos e processos em

execução, além de várias outras, são exibidas sob a forma de gráficos estatísticos (figura 1). Possui controle de acesso por nível de usuário e todas as informações são armazenadas em um banco de dados MySQL.

Sua arquitetura prevê a possibilidade de expansão através de plugins, disponíveis no fórum oficial [2] ou

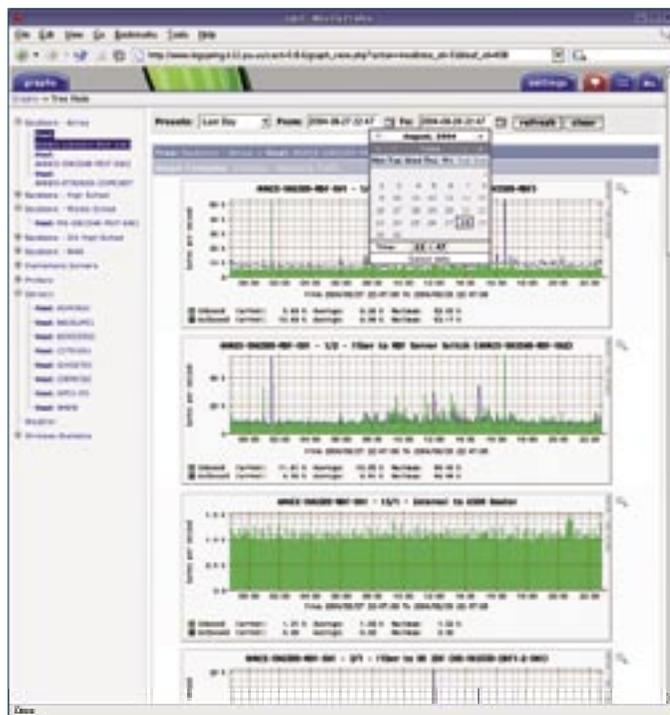


Figura 1 Relatórios de tráfego gerados a partir do Cacti.



Figura 2 Monitoramento de recursos do servidor com o Cacti.

em [3]. Os plugins adicionam novas funcionalidades ao programa, tornando-o ainda mais completo. Também é possível desenvolver seus próprios *plugins* e métodos de monitoramento, se necessário.

Por possuir diversos *templates* (modelos) para tipos de *hosts* e informações a serem monitoradas, o processo de cadastramento dos dispositivos torna-se bastante simples. Ao contrário do Nagios, onde basicamente tudo é feito via edição dos arquivos de configuração, no Cacti todo o processo de cadastro de dispositivos, criação de gráficos e demais itens é feito através da interface web, não sendo necessário um *restart* para que as novas configurações entrem em vigor.

## Instalando e configurando

A versão mais recente do Cacti (0.8.6j) pode ser obtida no seu site oficial [1]. Se você utiliza Debian é possível instalá-lo através dos repositórios APT – porém, antes de instalar os seus pacotes, é interessante que já tenham sido instalados o Apache, o PHP (e o seu módulo para o Apache), o banco de dados MySQL, o Net-SNMP e o RRDTool.

Uma vez instalados e funcionais todos esses elementos, vamos criar a base de dados do Cacti:

```
# mysqladmin --user=root create
➔ cacti
```

Posteriormente vamos informar as permissões de acesso para o banco de dados criado:

```
# mysql --user=root
mysql> grant all on cacti.* to
➔ cacti@localhost identified by
➔ 'senha_do_user_cacti'; flush
➔ privileges;
```

Supondo que todos os elementos estejam instalados, realizamos a implementação do Cacti:

```
# apt-get install cacti
```

A instalação dos pacotes será iniciada imediatamente após seu download. Durante essa etapa serão feitas algumas perguntas referentes a pré-configuração de alguns softwares, dentre elas o hostname do servidor MySQL, o nome da base de dados do Cacti, o nome de usuário e a senha de acesso a base e o tipo de servidor web utilizado.

Terminado o processo de instalação é necessário criar a estrutura do banco de dados através do comando:

```
# zcat /usr/share/doc/cacti/
➔ cacti.sql.gz | mysql -u cacti --
➔ password=senha_do_user_cacti cacti
```

Em seguida, devemos dar permissão de acesso recursivamente aos diretórios *rra* e *log* e ao arquivo *poller.php* para o usuário do Apache :

```
# chown -R www-data:www-data /
➔ usr/share/cacti/site/rra
# chown -R www-data:www-data /
➔ usr/share/cacti/site/log
# chown www-data:www-data /usr/
➔ share/cacti/site/poller.php
```

Por fim, vamos informar os parâmetros de acesso ao arquivo de configuração *config.php*.

```
$database_default = "cacti";
$database_hostname = "localhost";
$database_username = "cacti";
$database_password = "cacti";
```

Ao fim dessa etapa acesse-o através da URL. No primeiro acesso deve-se selecionar a opção *New Install*. Nessa etapa o sistema verifica se todos os pré-requisitos ao funcionamento foram atendidos. Ao clicar em *Finish* deve-se informar o usuário e senha *admin*, informando, posteriormente, uma nova senha.

## Plugin Architecture

Para utilizar de plugins do Cacti primeiramente deve ser instalado o *Plugin Architecture* disponível em [3]. Atualmente ele é compatível com as versões 0.8.6i e 0.8.6j do Cacti. Após baixar o arquivo, deve-se descompactá-lo através do comando:

```
# tar -xvzf cacti-plugin-arch.tar.gz
```

No diretório gerado, irão constar os subdiretórios *files-0.8.6i* e *files-0.8.6j*. Para instalar o *Plugin Architecture* basta copiar todos os arquivos e diretórios, presentes no subdiretório correspondente a versão do Cacti, para a pasta onde o mesmo está instalado, sobrescrevendo os arquivos existentes. Feito isso, é necessário configurar novamente os parâmetros de acesso ao banco de dados no arquivo *config.php*, como explicado anteriormente.

## Adicionando plugins

Um plugin bastante útil é o *Monitor*, onde é possível visualizar o *status* atual de todos os equipamentos monitorados. Outro é o *Ntop* – com o qual é possível abrir o *Ntop* [4] (ferramenta de monitoramento da rede local) no *frame* principal do Cacti (se o mesmo estiver instalado, não necessariamente no mesmo servidor).

A instalação desses e outros plugins é bastante simples. Basta copiá-los

de [3], descompactá-los em um subdiretório no diretório `plugins` (por exemplo, `plugins/monitor`) do Cacti e adicioná-los ao arquivo `config.php`, ficando da seguinte forma:

```
$plugins = array();
$plugins[] = 'monitor';
$plugins[] = 'ntop';
```

Em *Settings/Misc* deve ser informada a URL do Ntop. Também deve ser liberado o acesso para cada usuário em *User Management*, marcando o item *View Ntop*. Agora, ao acessar o Cacti, surgirão dois botões (*Monitor* e *Ntop*) na barra superior da ferramenta.

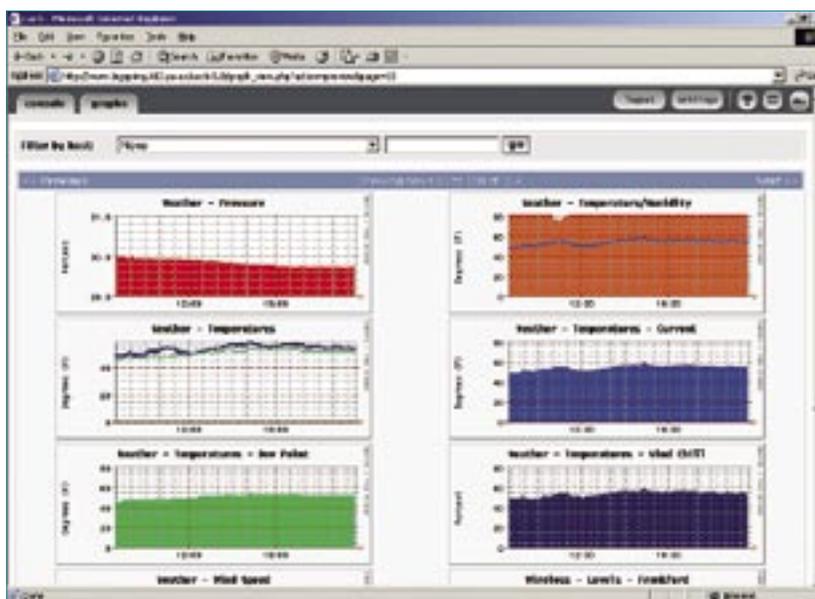
## Cadastrando dispositivos

Em um exemplo simples, vamos adicionar um roteador aos dispositivos. O cadastro é feito no item *Management/Devices*. Clicando no botão *Add* abre-se a tela de cadastro. Os dados informados foram:

```
Description - Roteador
Hostname - 192.168.1.1
Host Template - Generic SNMP-
enable Host
SNMP Community - public (padrão)
SNMP Version - Version 2
SNMP Port - 161 (padrão)
```

Ao clicar no botão *Create* o Cacti automaticamente já coleta algumas informações, via SNMP. Uma vez cadastrado o dispositivo, é necessário informar como será a coleta de dados e, posteriormente, criar os gráficos. É nesse momento que a utilização de templates facilita o trabalho.

Em nosso caso, como selecionamos *Generic SNMP-enable Host*, todas as interfaces foram detectadas automaticamente, pois nele já está informado o método da coleta dos dados e o tipo de gráfico a ser gerado. Para criar os gráficos basta clicar no link *Create Graphs*



**Figura 3** Monitoramento de temperatura e umidade em um servidor equipado com Cacti. Essa funcionalidade requer, obviamente, hardware preparado para a captação desse tipo de informação, sobretudo em se tratando de placas-mãe.

for this Host, selecionar quais interfaces temos interesse em monitorar e clicar novamente em *Create*. As atualizações dos gráficos podem ser verificadas clicando no botão “Graphs”, onde é possível escolher diferentes períodos de tempo e modos de visualização.

Além deste, outros templates valiosos são *Cisco Router*, *Local Linux Machine* e *Windows 2000/XP Host*, cada qual com seu método de coleta de dados e tipos de gráficos.

Existem diversos outros plugins interessantes, como o *Discovery*, que detecta automaticamente dispositivos não monitorados em uma sub-rede, mas que estão com o SNMP habilitado.

## Considerações finais

Um dos pontos fracos do Cacti é o fato de ele não possuir um mecanismo de envio de alertas (e-mail, por exemplo) caso certos eventos aconteçam – fato que pode ser resolvido pela inclusão desses recursos no código-fonte do programa (grandíssima vantagem do Código Aberto!) por qualquer programador competente.

O agendamento para a coleta das informações é feito via *crontab*.

Informações sobre como criar os dispositivos e gráficos podem ser obtidas na documentação da ferramenta, disponível no site oficial [1]. No site encontra-se também uma versão para a plataforma Microsoft Windows®.

Por fim, é importante mencionar que não é intenção do artigo julgar o Cacti melhor ou mais importante que as demais ferramentas citadas, mas apenas apresentar uma solução alternativa. ■

## Mais informações

- [1] Página do Cacti: <http://cacti.net>
- [2] Fórum Oficial: <http://forums.cacti.net>
- [3] Cacti Users: <http://cactiusers.org>
- [4] Página do Ntop: <http://ntop.org>

## O autor

Adriano Matos Meier ([adriano@senai-sc.ind.br](mailto:adriano@senai-sc.ind.br)) é Técnico em Redes de Computadores. Atualmente é analista de suporte no SENAI de Santa Catarina, onde é responsável pelos gateways corporativos.

ZFS – O novo filesystem da Sun

# Sem limites

Um sistema de arquivos virtualmente impossível de ser esgotado, com múltiplas ferramentas que permitem, inclusive, a montagem automática de RAID níveis 0, 1 e Z. Não é uma promessa de ficção científica: esse sistema existe e foi criado pela Sun.

por **Alexandre Borges**



Em 2004, a Sun abriu o código do Solaris para que a comunidade de desenvolvedores pudesse contribuir com o seu desenvolvimento e também pudesse trocar experiências de maneira mais ampla. Essa abertura tentava repetir o movimento extremamente bem sucedido que o Linux iniciou no mundo de TI e que mudou, de maneira profunda, o modo como se encara – sobretudo no mundo corporativo – o progresso e uso de softwares nos dias de hoje.

Esta abertura de código e mudança de pensamento dentro da Sun resultou no projeto *OpenSolaris*. A primeira parte do código do Solaris que a empresa disponibilizou foi a revolucionária ferramenta de *troubleshooting* chamada *Dtrace*, que vem ajudando decisivamente os administradores a solucionarem os problemas de performance causados por aplica-

ções rodando em ambiente Solaris. A seguir o próprio código do sistema operacional foi aberto, e a empresa começou a se alinhar com a inevitável tendência iniciada por seguidores do Linux e do código aberto.

## Variedades

Em todos os sistemas Unix e suas variantes, a preocupação com performance e tolerância a falhas sempre está presente. Sem dúvidas, o ponto crítico para o sucesso de um projeto é a escolha de qual tipo de sistema de arquivos será utilizado para a instalação e execução de suas aplicações.

Dependendo do sistema operacional utilizado, não se tem muita opção em relação ao sistema de arquivos a escolher: é isso o que vemos no mundo Windows, com seu FAT32 e as versões v.4 e v.5 do NTFS.

Isso ocorre de maneira nitidamente diferente (e melhor, por quê não?) nos sistemas Linux, com seu amplo suporte para diversos tipos e conceitos de filesystems, como o *ext2*, *ext3*, *ReiserFS* v.3 e v.4, *vfat*, entre muitos outros.

O sistema de arquivo em disco comumente usado pelo Solaris é o *UFS (Unix File System)*, que une performance com segurança na integridade dos dados. A Sun sempre ofereceu o UFS como seu sistema de arquivos para disco de boot, assim como também para discos e *storages* que abrigam aplicações.

Existe um documento muito interessante que o leitor pode consultar, e que compara a performance do UFS, *ext3* e *ReiserFS* v.3 [1].

Seguindo a avalanche de tantas boas novidades do Solaris 10, a Sun decidiu também investir na concepção de um novo file system que pudesse melhorar a performance e

usabilidade para trabalhar com arquivos no Solaris, elevando o mesmo a um novo patamar de performance de I/O para as aplicações críticas e competindo com os grandes filesystems de mercado: o ZFS (Zetabyte File System).

## Surge o ZFS

O ZFS surgiu por volta de 2004, ainda quando o Solaris 10 estava no estágio de versão beta. Nessas versões iniciais, o ZFS era apenas distribuído internamente dentro da Sun, no formato de dois pacotes para instalação à parte (*SUNWzfsr* e *SUNWzfsd*), pois havia alguns problemas mais sérios (que fazem parte de qualquer processo de criação de software). Oficialmente, o ZFS apenas veio a aparecer “bundled” no Solaris 10, release 06/06.

O ZFS tem muitas *features* interessantes, que simplificam muito o dia-a-dia de um administrador:

- ▶ ZFS é um filesystem de 128 bits, podendo assim endereçar e suportar bilhões de terabytes em dados;
- ▶ O gerenciamento de volume, antes feito com ferramentas como o *Solaris Volume Manager* (antigo *Solstice Disk Suite*) e *Veritas Volume Manager* (produto da Veritas – atual Symantec), já está integrado com o ZFS, e pode-se criar sistemas de arquivos do tipo ZFS sobre volumes (chamados de “pools”) em RAID 0, RAID 1 e RAID Z (este último é um tipo de RAID 5 melhorado na operação de escrita) de forma trivial;
- ▶ Suporte para *Sun Cluster 3.2 e Zones* em Solaris 10;
- ▶ Todo filesystem ZFS tem propriedades associadas e em sua maioria elas podem ser alteradas on-line de modo que não é necessário desmontá-lo para que as mudanças tornem-se ativas;
- ▶ Não existe mais a necessidade de alterar-se o arquivo */etc/vfstab*

(o equivalente do */etc/fstab* no Linux) para que sejamos capazes de montar sistemas de arquivos ZFS no momento da inicialização da máquina até porque, ao criar um sistema ZFS, sua montagem é automática;

- ▶ Não há mais a necessidade de trabalhar com *slices* (partições) nos discos usados para ZFS. É possível, se desejado, utilizar um ou diversos discos inteiros para volumes (*pools*) em ZFS. O Solaris usa, para isto, um layout de partições *EFI*. Deve-se ressaltar que ainda é factível usar partições para construção de pools que receberão sistemas de arquivos ZFS, entretanto esta não é a opção mais utilizada em sistemas críticos;
- ▶ Todas as operações de escrita são feitas usando a técnica *copy-on-write*, tornando os dados sempre consistentes e dispensando o uso do comando *fsck* no ZFS;
- ▶ Suporte completo para ACLs (que podem ser visualizadas e modificadas facilmente com os comandos *ls* e *chmod* ao invés de utilizar os tradicionais *getfacl* e *setfacl*);
- ▶ Facilidade para importar e exportar pools entre sistemas Self-healing (autocorreção) para dados em pools no arranjo RAID 1 e RAID Z através do uso de algoritmos de checksum.

## Usando o ZFS

Vamos explicar algo sobre os comandos e conceitos utilizados pelo ZFS, para que o leitor possa fazer sua própria avaliação em relação as qualidades desse sistema.

Para fazer testes com ZFS, o leitor poderá utilizar qualquer um dos seguintes cenários:

- ▶ Instalar o Solaris 10 (no mínimo release 11/06 – questão de estabilidade) em uma máquina com processadores SPARC, utilizan-

do discos internos ou externos para testes com ZFS;

- ▶ Instalar o Solaris 10 (release 11/06) em máquinas com arquitetura x86, também com discos internos ou externos para testes.;
- ▶ Instalar o Solaris 10 (release 11/06) no VMware (que agora tem sua versão gratuita para Linux e Windows) ou no VirtualBox. A vantagem de sistemas virtualizados é que o leitor pode criar muitos discos virtuais e fazer quaisquer testes necessários com o ZFS;
- ▶ Usar o Solaris 10 em qualquer das situações acima, contudo ao invés de usar discos virtuais, usar arquivos em disco. (usando o comando *mkfile*).

Utilizaremos em nossos exemplos o Solaris 11, com a opção de arquivos em disco via *mkfile* em alguns exemplos e discos virtuais (via VMware) em outros exemplos. Todos os exemplos enunciados aqui podem ser reproduzidos, facilmente, em qualquer ambiente equipado com alguns requisitos mínimos.

Em geral, quando pensamos na criação e uso de filesystems, sempre vêm à mente os seguintes passos:

- ▶ particionamento de discos (usando os comandos *format* em Solaris ou *fdisk* em Linux) para dedicar uma partição para cada filesystem;
- ▶ criação do sistema de arquivos em si (comandos *newfs* no Solaris ou *mkfs* no Linux) de modo que seja possível construir uma estrutura de arquivos e diretórios;
- ▶ Tornar disponível o filesystem para uso geral (comando *mount* em ambos os ambientes).

A utilização do ZFS é muitíssimo mais simples: ele não trabalha com o conceito de particionamento usando essas etapas, ou seja, não é necessário criar uma partição à parte para um sistema criado a partir de ZFS e nem, muito menos, é preciso montar (comando *mount*)

explicitamente o filesystem, já que ele é montado automaticamente na criação e no boot da máquina.

## Criando pools e filesystems

Antes de tudo, vamos criar alguns arquivos em disco:

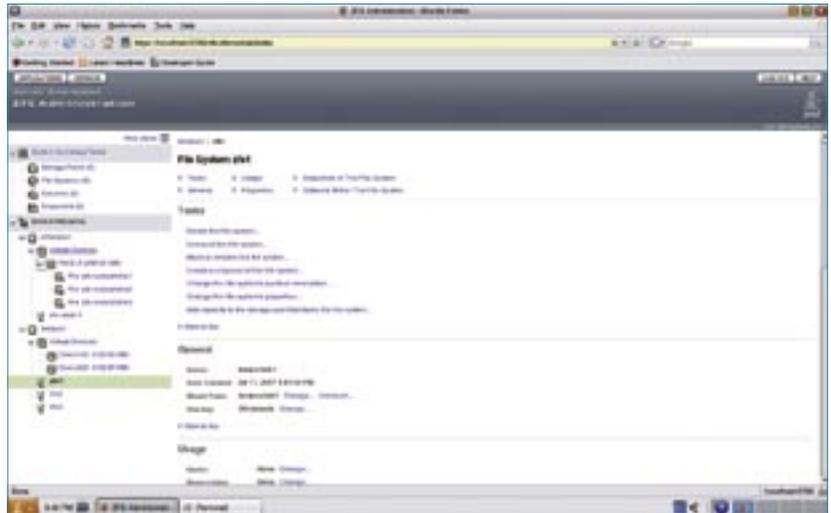
```
bash-3.00# mkdir /zfs-teste
bash-3.00# mkfile 100m /zfs-teste/
↳zfsfile1
bash-3.00# mkfile 100m /zfs-teste/
↳zfsfile2
bash-3.00# mkfile 100m /zfs-teste/
↳zfsfile3
bash-3.00# mkfile 100m /zfs-teste/
↳zfsfile4
bash-3.00# mkfile 100m /zfs-teste/
↳zfsfile5
```

Agora, também para referência, podemos listar quais discos temos em nossa máquina de teste:

```
bash-3.00# format
Searching for disks...done

AVAILABLE DISK SELECTIONS:
  0. c0d0 <DEFAULT cyl 1563 alt
↳2 hd 255 sec 63>
   /pci@0,0/pci-ide@7,1/ide@0/
↳cmdk@0,0
  1. c0d1 <DEFAULT cyl 202 alt
↳2 hd 64 sec 32>
   /pci@0,0/pci-ide@7,1/ide@0/
↳cmdk@1,0
  2. c1d1 <DEFAULT cyl 202 alt
↳2 hd 64 sec 32>
   /pci@0,0/pci-ide@7,1/ide@1/
↳cmdk@1,0
Specify disk (enter its number): ^D
```

Agora podemos criar um pool. Um pool pode ser interpretado como um volume, usando quaisquer tipos de RAID (0, 1 e 5), onde é possível criar quantos filesystems ZFS forem requeridos, sem qualquer tipo de restrição ou particionamento. É como se fosse um espaço livre para uso. Assim:



**Figura 1** Os principais comandos de administração de infraestrutura ZFS estão reunidos em uma única interface web.

```
bash-3.00# zpool create testpool
↳mirror c0d1 c1d1
bash-3.00# zpool list
NAME                SIZE  USED
↳AVAIL  CAP  HEALTH  ALROOT
testpool            192M  112K
↳192M    0%  ONLINE  -
bash-3.00# zpool status
pool: testpool
state: ONLINE
scrub: none requested
config:
NAME  STATE  READ WRITE CKSUM
testpool ONLINE  0    0    0
mirror ONLINE  0    0    0
c0d1  ONLINE  0    0    0
c1d1  ONLINE  0    0    0
errors: No known data errors
```

Os comandos acima possuem alguns pontos interessantes, que merecem ser comentados:

- ▶ o pool pode ter qualquer nome que o usuário deseje lhe dar. Usamos, neste caso, o nome `testpool`;
- ▶ pode-se criar o pool com qualquer configuração de RAID. Para RAID 0, não especifique nada. Para RAID 1 especifique `mirror` e para RAID 5 especifique `raidz`;
- ▶ o overhead do pool é mínimo (112 kb), sobrando muito espaço disponível para uso;
- ▶ podemos verificar a saúde do pool (coluna `health`) utilizando o

comando `zpool list` ou seu respectivo `status` usando o comando `zpool status`, assim como possíveis erros de checksum e estatísticas de leitura e escrita. Os estados possíveis de disco ou pool são `ONLINE`, `DEGRADED`, `FAULTED`, `OFFLINE`, `UNAVAILABLE`.

Vale notar também que, se possuíssemos um sistema de arquivos (ufs, por exemplo) construído nos discos que usamos, deveríamos ter inserido a opção `-f`, de modo a forçar a criação. Observe o comando abaixo:

```
bash-3.00# zpool create testpool
↳-f mirror <disco1> <disco2>
```

O leitor pode também verificar se algum dentre os pools disponíveis apresenta problemas com o comando:

```
bash-3.00# zpool status -x
```

Como já temos nosso pool (`testpool`), podemos criar quantos filesystems forem necessários neste espaço:

```
bash-3.00# zfs create testpool/
↳zfs1
bash-3.00# zfs create testpool/
↳zfs2
bash-3.00# zfs create testpool/
↳zfs3
bash-3.00# zfs list
```

| NAME             | USED | AVAIL | REFER |
|------------------|------|-------|-------|
| ➔ MOUNTPOINT     |      |       |       |
| testpool         | 186K | 160M  |       |
| ➔ 22K /testpool  |      |       |       |
| testpool/zfs1    | 18K  | 160M  | 18K   |
| ➔ /testpool/zfs1 |      |       |       |
| testpool/zfs2    | 18K  | 160M  | 18K   |
| ➔ /testpool/zfs2 |      |       |       |
| testpool/zfs3    | 18K  | 160M  | 18K   |
| ➔ /testpool/zfs3 |      |       |       |

Vale a pena ressaltar alguns aspectos desse último exemplo:

- ▶ todos os sistemas ZFS ocupam o mesmo espaço do pool testpool (160 Mb), tendo um overhead muito pequeno;
- ▶ todos também são montados automaticamente;
- ▶ o processo de criação de filesystems é tão fácil quanto criar diretórios. Aliás, a montagem do sistema de arquivos é feita em uma hierarquia (/testpool/zfs1) que é criada automaticamente. Não é preciso, em nenhum momento, usar o comando `mkdir` para criar tal ponto de montagem.

## Gerenciamento via Web do ZFS

O ZFS apresenta uma boa surpresa para quem é apreciador de interfaces gráficas e deseja administrar os filesystems ZFS (pools, discos, etc.) através de uma interface web: é o seu gerenciador de discos online (figura 1).

Basta digitar a seguinte URL no seu browser:

▶ <https://localhost:6789/zfs>

O login deve ser realizado, em seguida, como usuário *root*. A figura 2 mostra a tela inicial da administração via browser do ambiente ZFS.

Se não for possível acessar esta interface de gerenciamento, deve-se verificar se o servidor da interface está no ar:

```
bash-3.00# /usr/sbin/smcwebserver
➔ status
```

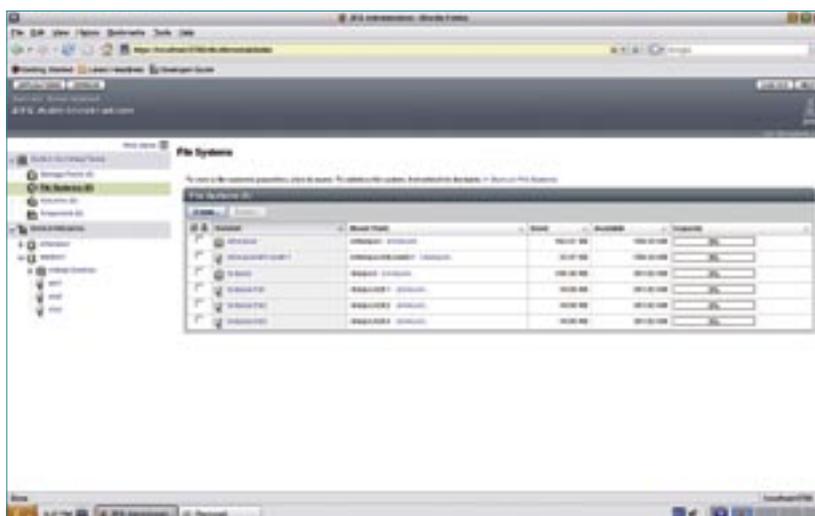


Figura 2 Tela inicial do aplicativo via navegador do sistema de arquivos ZFS.

Se o servidor está fora do ar, é possível inicializá-lo da seguinte forma:

```
bash-3.00# /usr/sbin/smcwebserver
➔ start
```

É conveniente manter o mesmo servidor ativo para os próximos boots:

```
bash-3.00# /usr/sbin/smcwebserver
➔ enable
```

## Renomeando filesystems

Veza ou outra são utilizados nomes não muito apropriados para a criação de um sistema de arquivos, e isso pode prejudicar uma administração mais fácil no ambiente Solaris. Para renomear um sistema de arquivos ZFS, no entanto, basta utilizar o seguinte comando:

```
bash-3.00# zfs rename testpool/
➔ zfs3 testpool/linuxmagazine
```

## Apagando pools e filesystems

Já que aprendemos a criar pools e filesystems e alterar suas propriedades, também devemos aprender a apagá-los. Começando pelo ZFS, podemos

apagar elementos do sistema – ou o próprio sistema – assim:

```
bash-3.00# zfs destroy testpool/
➔ zfs1
bash-3.00# zfs destroy testpool/
➔ zfs2
bash-3.00# zfs destroy testpool/
➔ linuxmagazine
```

Em algumas situações, não será possível destruir o filesystem já que ele pode estar sob o status *busy*. Neste caso, pode ser repetido o mesmo comando com a opção *-f*:

```
bash-3.00# zfs destroy -f
➔ testpool/linuxmagazine
```

## ZFS no Linux

O ZFS já foi portado para o Linux e o dono deste bom trabalho é Ricardo Correa. O projeto usa a tecnologia *FUSE* (Filesystem on Userspace), que foi introduzida no kernel 2.6.14 e que trabalha muito bem com uma implementação do NTFS chamada *ntfs-3g*.

O projeto ZFS/FUSE ainda está no início e tem alguns pequenos problemas: entretanto este é o caminho natural pelo qual todo software passa. O autor não está preocupado ainda com questões de performance, mas está trabalhando para deixar o código suave, com todas

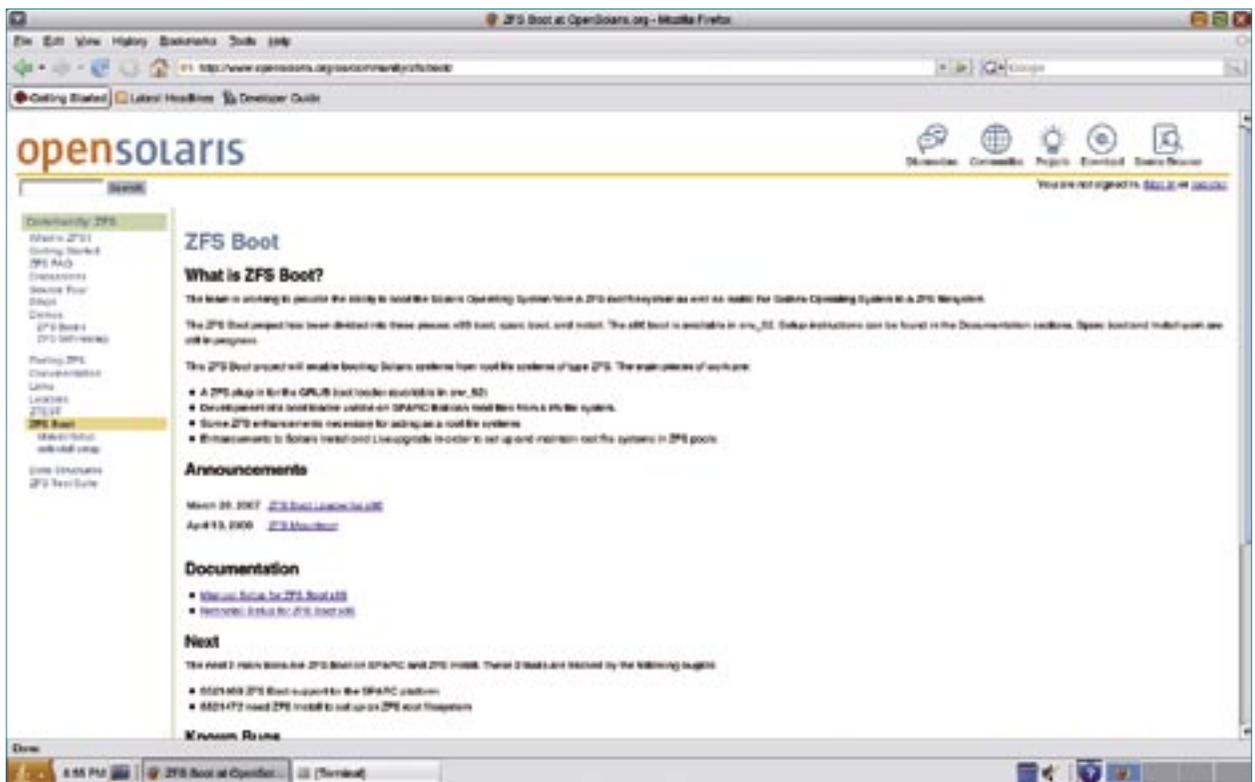


Figura 3 Página oficial do ZFS Boot.

as características do ZFS funcionando bem, para somente depois pensar nos aspectos de velocidade.

Implementar o ZFS no Linux (Fedora 7 é utilizado nos passos seguintes), é muito simples. No site do ZFS on Linux/FUSE, faça o download do código-fonte mais recente, o qual no momento em que este artigo foi escrito é o arquivo `zfs-fuse-0.4.0_beta1.tar.bz2` [2].

Feito isto, siga os seguintes passos para compilar e instalar o ZFS on Linux/FUSE:

```
bash-3.00# yum install fuse-*
bash-3.00# yum install scons
bash-3.00# cd /
bash-3.00# bunzip2 zfs-fuse-0.4.0_
↳beta1.tar.bz2
bash-3.00# tar xvf zfs-fuse-0.4.0_
↳beta1.tar.bz2
bash-3.00# cd /zfs-fuse-0.4.0_
↳beta1/src
bash-3.00# scons
bash-3.00# scons install
```

Execute, a seguir, o daemon responsável pelo serviço de ZFS:

```
bash-3.00# cd /zfs-fuse-0.4.0_
↳beta1/src/zfs-fuse
bash-3.00# ./run.sh
```

## Conclusão

A Sun realmente tem desenvolvido muitos projetos interessantes e, sem dúvidas, o ZFS é um deles. Outros projetos estão em curso, como o próprio OpenSolaris, onde é possível fazer download do próprio código fonte do Solaris e compilar uma versão particular [3]. Outra boa iniciativa é o projeto *OpenSparc*, que traz a versão aberta do processador T1 da Sun, que opera com 8 cores [4].

O ZFS não era suportado no disco de boot do Solaris 10 – ou seja, o filesystem que estava sob os arquivos e diretórios de sistema ainda devia ser UFS. Essa restrição, porém, já foi contornada e o leitor pode verificar o procedimento correto desta implementação em [5] (figura 2). Outra aplicação muito interessante é o *ZFS Test Suite*, que verifica e testa todas as características do ZFS e traz um detalhado relatório [6].

Sem dúvidas, o ZFS é um filesystem que veio para ficar. A tendência é que ele fique cada vez mais maduro, mais sólido e robusto, tanto em sua plataforma original – o Solaris – quanto em suas plataformas “por adoção”, como o Linux. ■

## Mais Informações

- [1] Whitepaper da Sun sobre o ZFS: [http://www.sun.com/software/whitepapers/solaris10/fs\\_performance.pdf](http://www.sun.com/software/whitepapers/solaris10/fs_performance.pdf)
- [2] ZFS on FUSE/Linux: <http://zfs-on-fuse.blogspot.com/>
- [3] Código fonte do Solaris: <http://www.opensolaris.org/os/downloads/on/>
- [4] OpenSparc: <http://www.opensparc.net/>
- [5] Alterações no boot do Solaris 10: <http://www.opensolaris.org/os/community/zfs/boot/>
- [6] ZFS Test Suite: <http://www.opensolaris.org/os/community/zfs/zfstestsuite/>



Cheryl lee - www.sxc.hu

Organização de conjuntos de caracteres com Perl

# Sopa de letrinhas

Quando caracteres estranhos aparecem no código ou nos dados de um programa, os programadores Perl precisam de uma solução que evite as complicações do Babel.

por Michael Schilli

No princípio havia a tabela ASCII – 128 caracteres que permitiam a escrita de textos em língua inglesa, incluindo alguns caracteres presentes em qualquer máquina de escrever, como % e \$, e também alguns caracteres de controle como quebras de linha e página, ou ainda um apito. Era apenas uma questão de tempo até que falantes de outras línguas desejassem escrever caracteres acentuados, cedilhas, tremas etc. A primeira forma de tratar esse problema foi agrupar a todos no próximo conjunto de 128 caracteres. Todos esses 256 caracteres foram numerados de 0 a 255 e codificados em computadores com 8 bits (1 byte) de dados.

## Exemplo 1: latin1

```
01 #!/usr/bin/perl -w
02 use strict;
03 my $s = "ü";
04 print "utrema=$s", "\n";
```

Daí surgiu o padrão ISO 8859, também conhecido como *Latin 1*.

## Línguas diferentes

Tudo começou com o ISO-8859-1, mas ao longo do tempo vieram outras variantes, chegando até o valor atual de ISO-8859-15, que ainda inclui o caractere do euro. Entretanto, a maioria dos navegadores web atuais não utiliza o padrão ISO-8859-1 para decodificar os conteúdos que adotam a codificação ISO-8859-1. Em vez disso, eles empregam o padrão Windows-1252, que inclui alguns caracteres adicionais, como o euro. Obviamente o resto do mundo não quis ficar para trás, e a corrida para exibir conjuntos de caracteres cada vez mais complexos começou.

Esquemas de codificação para línguas asiáticas, como *Shift-JIS* e *BIG5*, por exemplo, foram criados. Porém, os desenvolvedores logo perceberam que isso não os estava levando a lugar algum, e assim criaram o *Unicode*, uma enorme tabela que contém todos os caracteres das línguas comuns do mundo.

## UTF 8

O padrão *UTF 8* oferece uma abordagem para a codificação da tabela Unicode em computador. Se tivesse surgido uma súbita ne-

```
mschilli@vrraah:~$ perl -le 'print chr(0xfc)'
ü
mschilli@vrraah:~$ perl -le 'print chr(0xc3), chr(0xbc)'
??
mschilli@vrraah:~$
```

Figura 1 Em um terminal configurado com ISO-8859-1, a saída em Latin1 fica boa, mas a saída em UTF 8 não faz sentido.

```

utf-8 $ perl -le 'print chr(0xfc)'
utf-8 $ perl -le 'print chr(0xc3), chr(0xbc)'
utf-8 $

```

**Figura 2** Em um terminal configurado com UTF 8, a saída em UTF 8 fica boa, mas os caracteres codificados em ISO-8859-1 não são mostrados.

cessidade de se codificar caracteres ASCII em dois ou quatro bytes, os requisitos de memória teriam aumentado vertiginosamente. Para manter a possibilidade de exibir a antiga tabela ASCII com um único byte, a tabela UTF 8 foi projetada para que os primeiros 128 caracteres fossem os mesmos presentes na tabela ASCII.

Todavia, o próximo grupo de 128 caracteres é composto por códigos especiais de mascaramento que indicam que um número específico de códigos adicionais se segue, de forma a identificar inequivocamente qual caractere da tabela Unicode deve ser exibido. Por exemplo, o ç é armazenado sob o código 231 (0xE7) na tabela ISO-8859-1. Se você possuir um texto ISO-8859-1 que contenha um byte com o valor de 0xE7, o caractere obviamente é um ç.

## Um A com til

Na codificação UTF 8, o ü é representado por dois bytes – 195 e 198 (0xC3 e 0xBC). Se você possui um texto UTF 8 e o computador primeiro vê um byte com o valor de 0xC3 seguido por um de 0xBC, ambos serão igualmente mostrados como a letra ü.

Por outro lado, se a codificação não estiver clara e o computador vir um byte 0xC3, a questão será se isso é o primeiro byte de um u com trema em formato UTF 8 ou um byte ISO-8859-1 que representa um caractere completo. Se o byte inicial 0xC3 de uma sequência UTF 8 for interpretado incorre-

tamente como ISO-8859-15, será exibido o caractere que representa o terror dos programadores que brigam com as diferentes codificações – o Å.

Embora esse caractere faça parte de nosso alfabeto, ele não fará sentido se for exibido incorretamente no lugar de um ü, ç ou qualquer outro. Quando ele é visto nessas condições, isso significa que provavelmente o texto está codificado em UTF 8, mas está sendo interpretado como ISO-8859-1.

## Terminal

Se o objetivo for mostrar no terminal a saída de texto de um programa, então o terminal precisa saber como interpretar a cadeia de bytes gerada pelo programa, de forma a localizar os caracteres corretos para exibir.

Para iniciar um terminal X com suporte a UTF 8, pode-se rodar `xterm -u8 +lc`. A primeira opção ativa o suporte a UTF 8, e a segunda desabilita a interpretação de variáveis de ambiente, como LANG, para evitar que elas interfiram nesse processo.

Para abrir um terminal em modo ISO-8859-1, execute o comando `xterm` com a variável LANG definida, como por exemplo:

```
LANG=en_US.ISO-8859-1 xterm
```

As figuras 1 e 2 mostram a saída de dois scripts Perl em um terminal ISO e outro UTF 8, respectivamente. Um único byte com o valor de 0xE7 é corretamente interpretado como ç pelo terminal ISO vermelho. Entretanto, a sequência 0xC3BC é exibida como um Å e a representação ISO de 0XBC, o caractere para 1/4.

Já o terminal UTF 98 verde não exibe byte algum com o valor de 0xFC, mas a sequência 0xC3BC exibe um ü como esperado.

## Latim como padrão

A menos que o Perl receba ordens para se comportar de outra forma, ele interpretará o código-fonte de um script, incluindo todos os seus caracteres, expressões regulares, variáveis e nomes de funções com o código ISO-8859-1.

Se for usado um editor configurado para essa codificação, o ü no texto do programa será exibido com o código 0xFC, como mostra o utilitário `hexdump` (figura 3).

Porém, o exemplo 2 foi criado no editor vim com o comando `set encoding=utf-8`. As marcações vermelhas na saída do hexdump na figura 3 mostram que o trema no texto do código-fonte do programa foi codificado em dois bytes: C3 e BC.

A linha 3 do exemplo 2 é responsável por fazer o Perl interpretar o código-fonte como UTF 8 (`use utf8`). Isso garante que o texto ü, representado por 0xC3BC no código, exiba um único caractere – o equivalente Unicode a ü. Como resultado, `length($s)` não retornará um valor de 2, mas apenas 1.

A seguir, a linha 6 especifica a forma de saída padrão para UTF 8 (`binmode(STDOUT, ":utf8")`). Isso assegura que o Perl mostrará saídas Unicode no formato UTF 8

### Exemplo 2: utf8

```

01 #!/usr/bin/perl -w
02 use strict;
03 use utf8;
04 my $s = "ü";
05
06 binmode STDOUT, ":utf8";
07 print "utrema=$s", "\n";

```

### Exemplo 3: espiar

```
01 #!/usr/bin/perl -w
02 use strict;
03 use utf8;
04 use Data::Hexdumper;
05 use Encode qw(_utf8_off is_utf8);
06
07 my $s = "ü";
08
09 if( is_utf8($s) ) {
10     print "sinal UTF-8 estah 'ligado'.\n";
11 }
12
13 print "Comprimento: ", length($s), "\n";
14 _utf8_off($s);
15 print "Comprimento: ", length($s), "\n";
16
17 print hexdump(data => $s), "\n";
```

### Exemplo 4: isotest.cgi

```
01 #!/usr/bin/perl -w
02 use strict;
03 use CGI qw(:all);
04
05 print header(
06     -type => 'text/html',
07     -charset => 'iso-8859-1');
08
09 print "O simbolo do Euro eh ",
10     chr(0x80), "\n";
```

nado antes, o UTF 8 é apenas uma forma mais eficaz de se codificar a tabela Unicode.

## Entrada e saída

Quando um programa em Perl lê ou escreve dados, o programador precisa especificar o formato de entrada ou saída para os dados. Para ler as linhas de um arquivo codificado em UTF 8, pode-se usar o truque do `binmode` do exemplo 2 ou especificar a disciplina de linha usando um comando `open` com três parâmetros:

```
open ARQUIVO, "<:utf8", "arquivo.
➔ txt"
```

Se o programa depois ler uma linha do arquivo com `<ARQUIVO>` e atribuir o resultado a um escalar, pode-se ter certeza de que a cadeia de caracteres é Unicode, e o Perl tomará nota desse fato internamente. ➤

e, portanto, que o terminal receberá os dados UTF 8 que espera, exibindo-os corretamente. Sem a chamada de `binmode`, o Perl tentaria converter o texto de saída para Latin 1, o que faria sentido em terminais ISO-8859-1, mas não em um UTF 8.

E tudo começa a piorar se o caractere Unicode não puder ser convertido para Latin 1, como o *katakana* japonês “me” que possui um número Unicode de 30E1. Nesse caso, pode-se ver um aviso *Wide character in print*. O uso de `binmode(STDOUT)` para especificar a disciplina de linha para o arquivo de saída impede que o Perl tente converter a saída, fazendo com que ele mostre um texto UTF 8 cru. Em terminais UTF 8, essa é exatamente a estratégia a ser usada.

## Referência

A página de manual `man iso-8859-1` detalha a codificação Latin 1. Referindo-se ao número octal 347, por exemplo, ou ao valor hexadecimal E7, pode-se ver uma entrada para *LATIN SMALL LETTER C WITH CEDILLA*, ou um ç.

Se for necessário acessar a tabela Unicode, o arquivo `unicore/UnicodeData.txt`, normalmente em

`/usr/lib/perl5/5.8.x/`, é o local certo para procurar. Novamente, há uma entrada para o número 00E7: *LATIN SMALL LETTER C WITH CEDILLA*.

A figura 4 mostra que, logo abaixo do número para o ü minúsculo na tabela Unicode está o número para o mesmo caractere na tabela ISO-8859-1, que é o mesmo nas duas tabelas, pois os criadores do Unicode modelaram os primeiros 256 bytes de acordo com o padrão ISO-8859-1.

Note que o número Unicode não representa a codificação Unicode do caractere. Por exemplo, ü, o caractere Unicode com o número 00FC, é representado como C3BC em UTF 8. Como mencio-

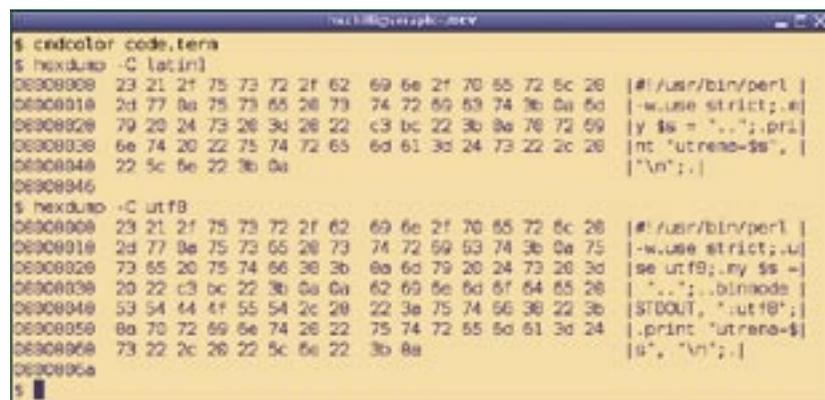


Figura 3 Um script Perl escrito em ISO-8859-1 usará o código 0xFC para representar o ü do código-fonte. Se o código-fonte for escrito em UTF 8, o ü é representado por uma seqüência dos bytes 0xC3BC.

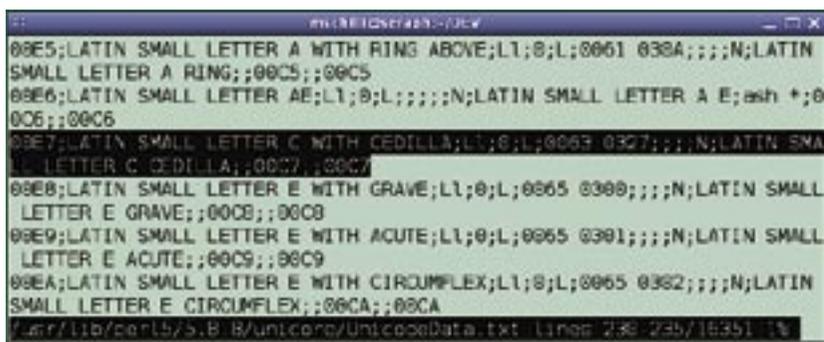


Figura 4 A entrada do caractere ü na tabela Unicode do Perl.

Sem a disciplina de linha, a entrada seria interpretada como ISO-8859-1, e o Perl colocaria os bytes crus em um escalar de cadeia de caracteres sem marcá-lo como UTF 8. O mesmo princípio se aplica à saída. Um `>:utf8` ou `> >:utf8` no segundo parâmetro do `open` especifica a disciplina de linha da saída para o modo UTF 8, e um `print FILE $string` exibirá como saída a cadeia de caracteres formatada em UTF 8 sem alterações. Como alternativa, pode-se usar `binmode` para modificar o manipulador do arquivo.

## O último véu

O exemplo 3 mostra como o Perl lida internamente com cadeias de caracteres Unicode. Devido ao pragma `use utf8` especificado antes, o caractere ü (digitado em UTF 8 no `vim`) é identificado e gerenciado como uma cadeia de caracteres Unicode.

Para permitir que isso ocorresse, o Perl atribuiu um sinal interno que

### Exemplo 5: isotest2.cgi

```
01 #!/usr/bin/perl -w
02 use strict;
03 use CGI qw(:all);
04
05 print header(
06   -type => 'text/html',
07   -charset => 'utf-8');
08
09 binmode STDOUT, ":utf8";
10 print "O simbolo do Euro eh ",
11   "\x{20AC}.\n";
```

pode ser consultado (`is_utf8()`) e manipulado (`_utf8_off()`) com o módulo `Encode`.

A saída de `espilar` na figura 5 mostra que a cadeia UTF 8 realmente possui comprimento de 1. Se o sinal for apagado com `_utf8_off()`, o comprimento da cadeia de repente cresce para dois caracteres.

A saída do módulo do CPAN `Data::Hexdumper` informa que a cadeia agora está armazenada internamente como `0xC3BC` – e que realmente é UTF 8.

O exemplo 4 mostra como um script CGI promete entregar ao navegador texto codificado em ISO-8859-1, mas depois envia um caractere de euro com o código de `0x80`, o que está de acordo com o padrão Windows-1252[1]. Como se pode ver na figura 6, o navegador gene-

rosamente concorda em mostrar o caractere do euro.

Se o script do servidor especificasse ISO-8859-1 em seu cabeçalho, veríamos um ponto de interrogação preto no lugar do símbolo do euro. Esse símbolo possui o código `0xA4` na tabela ISO-8859-1. Se o código for modificado para refletir isso, o navegador novamente mostrará o símbolo do euro corretamente.

## Não tão generoso

A biblioteca cliente web de Perl, `LWP`, não é tão generosa. O exemplo 5 mostra a saída como UTF 8 e até especifica o cabeçalho corretamente.

O caractere de euro na cadeia é representado por seu número de série Unicode `\x{20AC}`. Entretanto, há diversos aspectos a observar no cliente da aplicação web, se o texto da página web for codificado como

### Exemplo 6: clienteweb

```
01 #!/usr/bin/perl -w
02 use strict;
03 use LWP::UserAgent;
04
05 my $ua = LWP::UserAgent->new(
06   parse_head => 0
07 );
08
09 my $resp = $ua->get(
10   "http://perlmeister.com/cgi/isotest.cgi");
11
12 if($resp->is_success()) {
13   my $text = $resp->decoded_content();
14   binmode STDOUT, ":utf8";
15   print "$text\n";
16 }
```

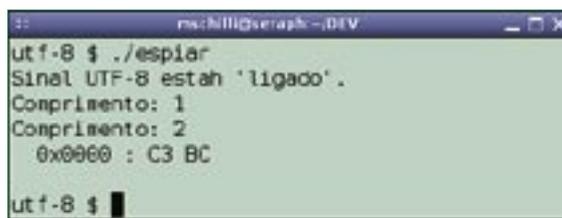


Figura 5 Em uma cadeia Unicode, um caractere multi-byte realmente possui comprimento 1. Se removermos a propriedade Unicode da cadeia, o Perl interpretará os bytes individualmente.



Figura 6 O navegador mostra o símbolo do euro.

UTF 8 no lado do servidor. A idéia é usar a biblioteca LWP para acessar a página a partir do servidor web e, se tudo funcionar direito, para armazenar uma cadeia Unicode em Perl. O exemplo 6 demonstra a técnica.

Devido a uma falha conhecida na biblioteca LWP (ou em `HTML::HeadParser`, para ser mais preciso), o Perl exibe um horrendo aviso “*Parsing of undecoded UTF-8 will give garbage when decoding entities*” ao retornar UTF 8. Isso pode ser evitado especificando-se a opção `parse_head => 0` na chamada do construtor do `UserAgent`[2].

Para permitir que o Perl guarde o texto UTF 8 retornado em uma cadeia Unicode, é necessário evitar usar o típico método `content()` para extrair o texto da página a partir do objeto `HTTP::Response`. O melhor é utilizar `decoded_content()` em seu lugar.

Esse método usa o campo `charset` da resposta do servidor web para adivinhar como decodificar seu conteúdo. Enquanto o cliente continuar honrando a disciplina de linha para `STDOUT`, não há nada para impedir a exibição correta em um terminal configurado para o modo UTF 8.

## Conclusões

Caminhar pelos mundos da codificação sempre foi um problema. Mas se você preferir não restringir a disponibilidade de seu software para uma fração do mercado, é bom adotar seriamente uma estratégia de internacionalização. ■

## Mais Informações

[1] Windows-1252:  
<http://en.wikipedia.org/wiki/Windows-1252>

[2] Falha conhecida na LWP:  
<http://www.mail-archive.com/libwww@perl.org/msg06330.html>

## O autor

**Michael Schilli** trabalha como desenvolvedor de software para o Yahoo!, em Sunnyvale, Estados Unidos. É autor de “Perl Power”, publicado pela editora Addison-Wesley. Sua homepage é <http://perlmeister.com>.

# V Encontro Linuxchix Brasil

Participe do Evento !!

Visite o site [www.linuxchix.org.br](http://www.linuxchix.org.br) e confira as palestras e mini-cursos deste ano!

Data: 07 e 08 de setembro.

Cidade: Brasília/DF

Local: Faculdade IESB - Asa Sul

Apoio e Patrocínio:



Ministério do Planejamento  
Orçamento e Gestão



# Linux.local

*O maior diretório de empresas que oferecem produtos, soluções e serviços em Linux e Software Livre, organizado por Estado. Sentiu falta do nome de sua empresa aqui? Entre em contato com a gente:*

**11 4082-1300** ou [anuncios@linuxmagazine.com.br](mailto:anuncios@linuxmagazine.com.br)

**Fornecedor de Hardware = 1**  
**Redes e Telefonia / PBX = 2**  
**Integrador de Soluções = 3**  
**Literatura / Editora = 4**  
**Fornecedor de Software = 5**  
**Consultoria / Treinamento = 6**

SERVIÇOS

| Empresa                           | Cidade                | Endereço  | Telefone      | Web                         | 1 | 2 | 3 | 4 | 5 | 6 |
|-----------------------------------|-----------------------|---|---------------|-----------------------------|---|---|---|---|---|---|
| <b>Ceará</b>                      |                       |   |               |                             |   |   |   |   |   |   |
| F13 Tecnologia                    | Fortaleza             | Rua Coronel Solon, 480 – Bairro de Fátima Fortaleza - CE - CEP 60040-270                      | 85 3252-3836  | www.f13.com.br              |   | ✓ | ✓ |   | ✓ | ✓ |
| <b>Espírito Santo</b>             |                       |   |               |                             |   |   |   |   |   |   |
| Linux Shopp                       | Vila Velha            | Rua São Simão (Correspondência), 18 – CEP: 29113-120  | 27 3082-0932  | www.linuxshopp.com.br       |   | ✓ | ✓ |   | ✓ | ✓ |
| Megawork Consultoria e Sistemas   | Vitória               | Rua Chapot Presvot, 389 – Praia do Cantão – CEP: 29055-410 sl 201, 202                        | 27 3315-2370  | www.megawork.com.br         |   |   |   | ✓ |   | ✓ |
| Spirit Linux                      | Vitória               | Rua Marins Alvarino, 150 – CEP: 29047-660   | 27 3227-5543  | www.spiritlinux.com.br      |   |   |   | ✓ |   | ✓ |
| <b>Minas Gerais</b>               |                       |   |               |                             |   |   |   |   |   |   |
| Instituto Online                  | Belo Horizonte        | Av. Bias Fortes, 932, Sala 204 – CEP: 30170-011   | 31 3224-7920  | www.institutoonline.com.br  |   |   |   | ✓ |   | ✓ |
| Linux Place                       | Belo Horizonte        | Rua do Ouro, 136, Sala 301 – Serra – CEP: 30220-000   | 31 3284-0575  | corporate.linuxplace.com.br |   |   | ✓ | ✓ |   | ✓ |
| Microhard                         | Belo Horizonte        | Rua República da Argentina, 520 – Sion – CEP: 30315-490                                       | 31 3281-5522  | www.microhard.com.br        |   | ✓ | ✓ | ✓ |   | ✓ |
| TurboSite                         | Belo Horizonte        | Rua Paraíba, 966, Sala 303 – Savassi – CEP: 30130-141   | 0800 702-9004 | www.turbosite.com.br        |   | ✓ |   |   |   | ✓ |
| <b>Paraná</b>                     |                       |   |               |                             |   |   |   |   |   |   |
| iSolve                            | Curitiba              | Av. Cândido de Abreu, 526, Cj. 1206B – CEP: 80530-000   | 41 252-2977   | www.isolve.com.br           |   | ✓ | ✓ |   |   | ✓ |
| Mandriva Conectiva                | Curitiba              | Rua Tocantins, 89 – Cristo Rei – CEP: 80050-430   | 41 3360-2600  | www.mandriva.com.br         |   |   |   | ✓ | ✓ | ✓ |
| <b>Rio de Janeiro</b>             |                       |   |               |                             |   |   |   |   |   |   |
| NSI Training                      | Rio de Janeiro        | Rua Araújo Porto Alegre, 71, 4º andar Centro – CEP: 20030-012                                 | 21 2220-7055  | www.nsi.com.br              |   |   |   | ✓ |   | ✓ |
| Open IT                           | Rio de Janeiro        | Rua do Mercado, 34, Sl, 402 – Centro – CEP: 20010-120   | 21 2508-9103  | www.openit.com.br           |   |   |   | ✓ |   | ✓ |
| Unipi Tecnologias                 | Campos dos Goytacazes | Av. Alberto Torres, 303, 1º andar - Centro – CEP 28035-581                                    | 22 2725-1041  | www.unipi.com.br            |   |   |   | ✓ | ✓ | ✓ |
| <b>Rio Grande do Sul</b>          |                       |   |               |                             |   |   |   |   |   |   |
| Solis                             | Lajeado               | Rua Comandante Wagner, 12 – São Cristóvão – CEP: 95900-000                                    | 51 3714-6653  | www.solis.coop.br           |   | ✓ | ✓ | ✓ | ✓ | ✓ |
| DualCon                           | Novo Hamburgo         | Rua Joaquim Pedro Soares, 1099, Sl. 305 – Centro  | 51 3593-5437  | www.dualcon.com.br          |   | ✓ |   | ✓ |   | ✓ |
| Datarecover                       | Porto Alegre          | Av. Carlos Gomes, 403, Sala 908, Centro Comercial Atrium Center – Bela Vista – CEP: 90480-003 | 51 3018-1200  | www.datarecover.com.br      |   | ✓ |   | ✓ |   |   |
| LM2 Consulting                    | Porto Alegre          | Rua Germano Petersen Junior, 101-Sl 202 – Higienópolis – CEP: 90540-140                       | 51 3018-1007  | www.lm2.com.br              |   |   |   | ✓ |   | ✓ |
| LnX-IT Informação e Tecnologia    | Porto Alegre          | Av. Venâncio Aires, 1137 – Rio Branco – CEP: 90.040.193                                       | 51 3331-1446  | www.lnx-it.inf.br           |   | ✓ |   | ✓ |   | ✓ |
| Plugin                            | Porto Alegre          | Av. Júlio de Castilhos, 132, 11º andar Centro – CEP: 90030-130                                | 51 4003-1001  | www.plugin.com.br           |   | ✓ |   | ✓ |   | ✓ |
| TeHospedo                         | Porto Alegre          | Rua dos Andradas, 1234/610 – Centro – CEP: 90020-008  | 51 3286-3799  | www.tehospedo.com.br        |   | ✓ | ✓ |   |   |   |
| <b>São Paulo</b>                  |                       |   |               |                             |   |   |   |   |   |   |
| Ws Host                           | Arthur Nogueira       | Rua Jerere, 36 – Vista Alegre – CEP: 13280-000  | 19 3846-1137  | www.wshost.com.br           |   | ✓ |   | ✓ |   | ✓ |
| DigiVoice                         | Barueri               | Al. Juruá, 159, Térreo – Alphaville – CEP: 06455-010  | 11 4195-2557  | www.digivoice.com.br        |   | ✓ | ✓ | ✓ |   | ✓ |
| Dextra Sistemas                   | Campinas              | Rua Antônio Paioli, 320 – Pq. das Universidades – CEP: 13086-045                              | 19 3256-6722  | www.dextra.com.br           |   |   |   | ✓ |   | ✓ |
| Insigne Free Software do Brasil   | Campinas              | Av. Andrades Neves, 1579 – Castelo – CEP: 13070-001   | 19 3213-2100  | www.insignesoftware.com     |   |   |   | ✓ |   | ✓ |
| Microcamp                         | Campinas              | Av. Thomaz Alves, 20 – Centro – CEP: 13010-160  | 19 3236-1915  | www.microcamp.com.br        |   |   |   | ✓ |   | ✓ |
| PC2 Consultoria em Software Livre | Carapicuíba           | Rua Edeia, 500 - 06350-080  | 11 3213-6388  | www.pc2consultoria.com      |   | ✓ |   |   |   | ✓ |
| Savant Tecnologia                 | Diadema               | Av. Senador Vitorino Freire, 465 – CEP: 09910-550   | 11 5034-4199  | www.savant.com.br           |   | ✓ | ✓ | ✓ |   | ✓ |
| Epopeia Informática               | Marília               | Rua Goiás, 392 – Bairro Cascata – CEP 17509-140   | 14 3413-1137  | www.epopeia.com.br          |   |   |   |   |   | ✓ |
| Redentor                          | Osasco                | Rua Costante Piovan, 150 – Jd. Três Montanhas – CEP: 06263-270                                | 11 2106-9392  | www.redentor.ind.br         |   | ✓ |   |   |   |   |
| Go-Global                         | Santana de Parnaíba   | Av. Yojiro Takaoca, 4384, Ed. Shopping Service, Cj. 1013 – CEP: 06541-038                     | 11 2173-4211  | www.go-global.com.br        |   |   |   | ✓ |   | ✓ |
| AW2NET                            | Santo André           | Rua Edson Soares, 59 – CEP: 09760-350   | 11 4990-0065  | www.aw2net.com.br           |   |   |   | ✓ |   | ✓ |
| Async Open Source                 | São Carlos            | Rua Orlando Damiano, 2212 – CEP 13560-450   | 16 3376-0125  | www.async.com.br            |   | ✓ |   |   |   | ✓ |
| Delix Internet                    | São José do Rio Preto | Rua Voluntário de São Paulo, 3066 9º – Centro – CEP: 15015-909                                | 11 4062-9889  | www.delixhosting.com.br     |   | ✓ |   | ✓ |   | ✓ |

| Empresa  | Cidade       | Endereço  | Telefone      | Web                           | 1 | 2 | 3 | 4 | 5 | 6 |
|--|--------------|---|---------------|-------------------------------|---|---|---|---|---|---|
| <b>São Paulo (continuação)</b>                       |              |   |               |                               |   |   |   |   |   |   |
| 4Linux   | São Paulo    | Rua Teixeira da Silva, 660, 6º andar – CEP: 04002-031                                       | 11 2125-4747  | www.4linux.com.br             |   |   |   |   | ✓ | ✓ |
| A Casa do Linux                                      | São Paulo    | Al. Jaú, 490 – Jd. Paulista – CEP 01420-000   | 11 3549-5151  | www.acasadolinux.com.br       |   |   | ✓ | ✓ | ✓ | ✓ |
| Accenture do Brasil Ltda.                            | São Paulo    | Rua Alexandre Dumas, 2051 – Chácara Santo Antônio – CEP: 04717-004                          | 11 5188-3000  | www.accenture.com.br          |   |   | ✓ | ✓ | ✓ | ✓ |
| ACR Informática                                      | São Paulo    | Rua Lincoln de Albuquerque, 65 –Perdizes – CEP: 05004-010                                   | 11 3873-1515  | www.acrinformatica.com.br     | ✓ |   |   |   |   | ✓ |
| Agit Informática                                     | São Paulo    | Rua Major Quedinho, 111, 5º andar, Cj. 508 – Centro – CEP: 01050-030                        | 11 3255-4945  | www.agit.com.br               | ✓ | ✓ |   |   |   | ✓ |
| Altbit - Informática Comércio e Serviços LTDA.       | São Paulo    | Av. Francisco Matarazzo, 229, Cj. 57 – Água Branca – CEP 05001-000                          | 11 3879-9390  | www.altbit.com.br             |   | ✓ | ✓ | ✓ | ✓ | ✓ |
| AS2M -WPC Consultoria                                | São Paulo    | Rua Três Rios, 131, Cj. 61A – Bom Retiro – CEP: 01123-001                                   | 11 3228-3709  | www.wpc.com.br                |   |   | ✓ | ✓ | ✓ | ✓ |
| Big Host   | São Paulo    | Rua Dr. Miguel Couto, 58 – Centro – CEP: 01008-010  | 11 3033-4000  | www.bighost.com.br            | ✓ |   |   |   |   | ✓ |
| Blanes   | São Paulo    | Rua André Ampère, 153 – 9º andar – Conj. 91 CEP: 04562-907 ( próx. Av. L. C. Berrini)       | 11 5506-9677  | www.blanes.com.br             | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Commmlogik do Brasil Ltda.                           | São Paulo    | Av. das Nações Unidas, 13.797, Bloco II, 6º andar – Morumbi – CEP: 04794-000                | 11 5503-1011  | www.commmlogik.com.br         | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Computer Consulting Projeto e Consultoria Ltda.      | São Paulo    | Rua Vergueiro, 6455, Cj. 06 – Alto do Ipiranga – CEP: 04273-100                             | 11 5062-3927  | www.computerconsulting.com.br | ✓ |   | ✓ | ✓ | ✓ | ✓ |
| Consist Consultoria, Sistemas e Representações Ltda. | São Paulo    | Av. das Nações Unidas, 20.727 – CEP: 04795-100  | 11 5693-7210  | www.consist.com.br            |   |   | ✓ | ✓ | ✓ | ✓ |
| Domínio Tecnologia                                   | São Paulo    | Rua das Caruaibeiras, 98 – Metrô Conceição – CEP: 04343-080                                 | 11 5017-0040  | www.dominiotecnologia.com.br  | ✓ |   |   |   |   | ✓ |
| EDS do Brasil  | São Paulo    | Av. Pres. Juscelino Kubitschek, 1830 Torre 4 - 5º andar                                     | 11 3707-4100  | www.eds.com                   |   | ✓ | ✓ |   |   | ✓ |
| Ética Tecnologia                                     | São Paulo    | Rua Nova York, 945 – Brooklin – CEP:04560-002   | 11 5093-3025  | www.etica.net                 | ✓ |   | ✓ | ✓ | ✓ | ✓ |
| Getronics ICT Solutions and Services                 | São Paulo    | Rua Verbo Divino, 1207 – CEP: 04719-002   | 11 5187-2700  | www.getronics.com.br          |   |   | ✓ | ✓ | ✓ | ✓ |
| Hewlett-Packard Brasil Ltda.                         | São Paulo    | Av. das Nações Unidas, 12.901, 25º andar – CEP: 04578-000                                   | 11 5502-5000  | www.hp.com.br                 | ✓ |   | ✓ | ✓ | ✓ | ✓ |
| IBM Brasil Ltda.                                     | São Paulo    | Rua Tutóia, 1157 – CEP: 04007-900   | 0800-7074 837 | www.br.ibm.com                | ✓ |   | ✓ | ✓ | ✓ | ✓ |
| iFractal   | São Paulo    | Rua Fiação da Saúde, 145, Conj. 66 – Saúde – CEP: 04144-020                                 | 11 5078-6618  | www.ifractal.com.br           |   |   | ✓ | ✓ | ✓ | ✓ |
| Integral   | São Paulo    | Rua Dr. Gentil Leite Martins, 295, 2º andar Jd. Prudência – CEP: 04648-001                  | 11 5545-2600  | www.integral.com.br           | ✓ |   |   |   |   | ✓ |
| Itautec S.A.   | São Paulo    | Rua Santa Catarina, 1 – Tatuapé – CEP: 03086-025  | 11 6097-3000  | www.itaute.com.br             | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Linux Komputer Informática                           | São Paulo    | Av. Dr. Lino de Moraes Leme, 185 – CEP: 04360-001   | 11 5034-4191  | www.komputer.com.br           | ✓ |   | ✓ | ✓ | ✓ | ✓ |
| Linux Mall   | São Paulo    | Rua Machado Bittencourt, 190, Cj. 2087 – CEP: 04044-001                                     | 11 5087-9441  | www.linuxmall.com.br          | ✓ |   |   | ✓ | ✓ | ✓ |
| Livraria Tempo Real                                  | São Paulo    | Al. Santos, 1202 – Cerqueira César – CEP: 01418-100   | 11 3266-2988  | www.temporeal.com.br          |   |   |   | ✓ | ✓ | ✓ |
| Locasite Internet Service                            | São Paulo    | Av. Brigadeiro Luiz Antonio, 2482, 3º andar – Centro – CEP: 01402-000                       | 11 2121-4555  | www.locasite.com.br           | ✓ |   |   |   | ✓ | ✓ |
| Microsiga  | São Paulo    | Av. Braz Leme, 1631 – CEP: 02511-000  | 11 3981-7200  | www.microsiga.com.br          |   |   | ✓ | ✓ | ✓ | ✓ |
| Novatec Editora Ltda.                                | São Paulo    | Rua Luis Antonio dos Santos, 110 – Santana – 02460-000                                      | 11 6979-0071  | www.novateceditora.com.br     |   |   |   | ✓ | ✓ | ✓ |
| Novell América Latina                                | São Paulo    | Rua Funchal, 418 – Vila Olímpia   | 11 3345-3900  | www.novell.com/brasil         |   |   | ✓ | ✓ | ✓ | ✓ |
| Oracle do Brasil Sistemas Ltda.                      | São Paulo    | Av. Alfredo Egídio de Souza Aranha, 100 – Bloco B – 5º andar – CEP: 04726-170               | 11 5189-3000  | www.oracle.com.br             |   |   |   |   | ✓ | ✓ |
| Proelbra Tecnologia Eletrônica Ltda.                 | São Paulo    | Av. Rouxinol, 1.041, Cj. 204, 2º andar Moema – CEP: 04516-001                               | 11 5052- 8044 | www.proelbra.com.br           | ✓ |   | ✓ | ✓ | ✓ | ✓ |
| Provider   | São Paulo    | Av. Cardoso de Melo, 1450, 6º andar – Vila Olímpia – CEP: 04548-005                         | 11 2165-6500  | www.e-provider.com.br         |   |   | ✓ | ✓ | ✓ | ✓ |
| Red Hat Brasil                                       | São Paulo    | Av. Angélica, 2503, 8º andar Consolação – CEP: 01227-200                                    | 11 3124-6000  | www.redhat.com.br             |   |   | ✓ | ✓ | ✓ | ✓ |
| Samurai Projetos Especiais                           | São Paulo    | Rua Barão do Triunfo, 550, 6º andar – CEP: 04602-002  | 11 5097-3014  | www.samurai.com.br            |   |   | ✓ | ✓ | ✓ | ✓ |
| SAP Brasil   | São Paulo    | Av. das Nações Unidas, 11.541, 16º andar – CEP: 04578-000                                   | 11 5503-2400  | www.sap.com.br                |   |   | ✓ | ✓ | ✓ | ✓ |
| Simplex Consultoria                                  | São Paulo    | Rua Mourato Coelho, 299, Cj. 02 Pinheiros – CEP: 05417-010                                  | 11 3898-2121  | www.simplexconsultoria.com.br |   |   | ✓ | ✓ | ✓ | ✓ |
| Smart Solutions                                      | São Paulo    | Av. Jabaquara, 2940 cj 56 e 57  | 11 5052-5958  | www.smart-tec.com.br          |   |   | ✓ | ✓ | ✓ | ✓ |
| Snap IT  | São Paulo    | Rua João Gomes Junior, 131 – Jd. Bonfiglioli – CEP: 05299-000                               | 11 3731-8008  | www.snapit.com.br             |   |   | ✓ | ✓ | ✓ | ✓ |
| Stefanini IT Solutions                               | São Paulo    | Av. Brig. Faria Lima, 1355, 19º – Pinheiros – CEP: 01452-919                                | 11 3039-2000  | www.stefanini.com.br          |   |   | ✓ | ✓ | ✓ | ✓ |
| Sun Microsystems                                     | São Paulo    | Rua Alexandre Dumas, 2016 – CEP: 04717-004  | 11 5187-2100  | www.sun.com.br                | ✓ |   | ✓ | ✓ | ✓ | ✓ |
| Sybase Brasil  | São Paulo    | Av. Juscelino Kubitschek, 510, 9º andar Itaim Bibi – CEP: 04543-000                         | 11 3046-7388  | www.sybase.com.br             |   |   |   |   | ✓ | ✓ |
| The Source   | São Paulo    | Rua Marquês de Abrantes, 203 – Chácara Tatuapé – CEP: 03060-020                             | 11 6698-5090  | www.thesource.com.br          |   |   | ✓ | ✓ | ✓ | ✓ |
| Unisys Brasil Ltda.                                  | São Paulo    | R. Alexandre Dumas 1658 – 6º, 7º e 8º andares – Chácara Santo Antônio – CEP: 04717-004      | 11 3305-7000  | www.unisys.com.br             | ✓ |   | ✓ | ✓ | ✓ | ✓ |
| Utah   | São Paulo    | Av. Paulista, 925, 13º andar – Cerqueira César – CEP: 01311-916                             | 11 3145-5888  | www.utah.com.br               |   |   | ✓ | ✓ | ✓ | ✓ |
| Visuelles  | São Paulo    | Rua Eng. Domicio Diele Pacheco e Silva, 585 – Interlagos – CEP 04455-310                    | 11 5614-1010  | www.visuelles.com.br          |   |   | ✓ | ✓ | ✓ | ✓ |
| Webnow   | São Paulo    | Av. Nações Unidas, 12.995, 10º andar, Ed. Plaza Centenário – Chácara Itaim – CEP: 04578-000 | 11 5503-6510  | www.webnow.com.br             | ✓ |   | ✓ | ✓ | ✓ | ✓ |
| WRL Informática Ltda.                                | São Paulo    | Rua Santa Ifigênia, 211/213, Box 02– Centro – CEP: 01207-001                                | 11 3362-1334  | www.wrl.com.br                | ✓ |   | ✓ | ✓ | ✓ | ✓ |
| Systech  | Taquaritinga | Rua São José, 1126 – Centro – Caixa Postal 71 – CEP: 15.900-000                             | 16 3252-7308  | www.systech-ltd.com.br        | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

# Calendário de eventos

| Evento                                | Data                          | Local              | Website              |
|---------------------------------------|-------------------------------|--------------------|----------------------|
| Linux Park                            | 15 de Agosto                  | Belo Horizonte, MG | www.linuxpark.com.br |
| Linux Park                            | 23 de agosto                  | Brasília, DF       | www.linuxpark.com.br |
| PyCon Brasil                          | 30 de agosto a 1º de setembro | Joinville, SC      | pyconbrasil.com.br   |
| 5º Encontro Nacional Linuxchix Brasil | 7 e 8 de setembro             | Brasília, DF       | www.linuxchix.org.br |
| Linux Park                            | 18 de setembro                | Recife, PE         | www.linuxpark.com.br |
| Linux Park                            | 2 de outubro                  | São Paulo, SP      | www.linuxpark.com.br |
| Linux Park                            | 25 de outubro                 | Curitiba, PR       | www.linuxpark.com.br |
| Latinoware 2007                       | 13 e 14 de novembro           | Foz do Iguaçu, PR  | www.latinoware.org   |

## Índice de anunciantes

| Empresa         | Pág.   |
|-----------------|--------|
| Celepar         | 07     |
| Coraid          | 29     |
| Fracto          | 02, 80 |
| Guia de TI      | 27     |
| IBM             | 15     |
| Impacta         | 19     |
| Itautec         | 11     |
| LinuxChix       | 77     |
| Linux New Media | 31     |
| Linux Park      | 83     |
| Linux World     | 81     |
| LPI             | 63     |
| Plugin          | 13     |
| Pycon           | 35     |
| Red Hat         | 09, 43 |
| Senac           | 84     |
| Thin Networks   | 39     |
| XPG             | 21     |

**Hospedagem de sites - VPS  
IP Dedicado - Soluções de Voz  
Servidores Dedicados - Colocation**



**FRACTO** [www.fracto.com.br](http://www.fracto.com.br)  
(11) 3523-1580

Agente Autorizado Intelig

OPEN Source.  
OPEN Solutions.

OPEN. For Business.

## LinuxWorld Conference & Expo – Worldwide Series

|                      |                               |  |
|----------------------|-------------------------------|--|
| <b>Seoul</b>         | June 20 – 23, 2007            | <a href="http://www.linuxworldkorea.com">www.linuxworldkorea.com</a>   |
| <b>San Francisco</b> | August 6 – 9, 2007            | <a href="http://www.linuxworldexpo.com">www.linuxworldexpo.com</a>     |
| <b>Beijing</b>       | September 3 – 7, 2007         | <a href="http://www.linuxworldchina.com">www.linuxworldchina.com</a>   |
| <b>Stockholm</b>     | September 5, 2007             | <a href="http://www.linuxworldsummit.se">www.linuxworldsummit.se</a>   |
| <b>London</b>        | October, 24 – 25, 2007        | <a href="http://www.linuxworldexpo.co.uk">www.linuxworldexpo.co.uk</a> |
| <b>Utrecht</b>       | October 30 – November 1, 2007 | <a href="http://www.linuxworldexpo.nl">www.linuxworldexpo.nl</a>       |



World's leading Trade Event for Linux and Open Source in business. *do Brasil Editora Ltda.*



INTERNATIONAL MEDIA SPONSOR

# Na Linux Magazine #34...

## DESTAQUE

### Loja online

Muitos já sabem que a melhor solução para servidores web é composta por quatro componentes de Código Aberto: Linux, Apache, MySQL e PHP. Com esses softwares, aliados a um hardware confiável, é possível construir um portal de grande tráfego, com bom desempenho e segurança.

No entanto, pouco se divulga sobre soluções livres para o comércio eletrônico. Na Linux Magazine 34, vamos mostrar como construir um site completo, incluindo um carrinho de compras e configurando todos os serviços essenciais de forma a obter o melhor desempenho sem sacrificar a segurança dos dados. Abordaremos a implementação de tráfego criptografado por SSL com certificados de segurança, e também a configuração dos servidores de banco de dados livres MySQL e PostgreSQL. Então, comece logo a pensar no que você vai vender em seu site. ■



### LPIC-2: Aula 3

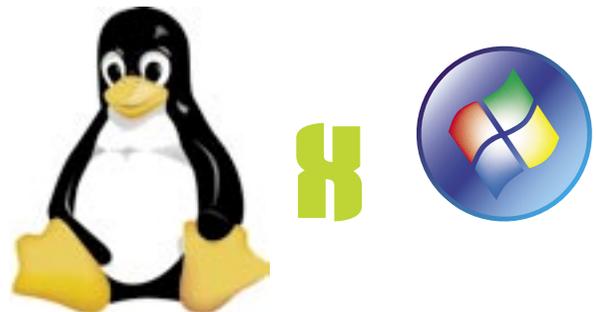
Os dois próximos tópicos abordados serão o 204 e o 207, respectivamente Hardware e Compartilhamento de arquivos e serviços. Temas importantíssimos, como RAID, LVM, Samba e NFS serão discutidos em detalhes. ■

# Na EasyLinux #11...

## DESTAQUE

### Linux versus Windows Vista

O último lançamento da Microsoft alega ser a última palavra em recursos, facilidade e visual. Será mesmo? Veja o que é verdade e o que é mentira nessa comparação detalhada entre o cume da família Windows e a mais amigável das distribuições da família Linux. ■



## FLIPERAMA

### Jogos de Fliperama

Antes mesmo de existirem os computadores domésticos como os conhecemos, os fliperamas já cativavam uma legião de fãs. Na próxima edição, veremos como fazer para rodar os jogos de fliperama e arcade que fizeram sucesso nas décadas de 80 e 90. ■

## FLIPERAMA

### Second Life

O mundo virtual do Second Life chegou aos usuários de Linux. Se você já o conhece, aprenda a viver nele sem precisar de outro sistema operacional. Se você ainda não o conhece, a próxima Easy Linux explicará o que é esse mundo virtual do qual todos estão falando. ■

LINUX NEW MEDIA  
The Pulse of Open Source



SNOW666  
LINUX PARK

Inscreva-se agora pelo site!!  
**Vagas limitadas!**  
[www.linuxpark.com.br](http://www.linuxpark.com.br)



# A importância do Software Livre para o mercado brasileiro de TI

15/08/2007 - Hotel Mercure - Belo Horizonte - MG

23/08/2007 - Hotel Blue Tree Park - Brasília - DF

PATROCÍNIO



IBM developerWorks **Live!**

LINUX  
New Media do Brasil



APOIO



PROMOÇÃO



REALIZAÇÃO



ORGANIZAÇÃO



# Penso Certificação LPI, pensou Senac

A certificação LPI está entre as mais procuradas no mundo por profissionais da área de Tecnologia da Informação, segundo o site Certcities.com, especializado no tema.

As provas do LPI são independentes de distribuição e baseadas no Linux Standard Base, um conjunto de normas que mantém a compatibilidade entre as diferentes versões e distribuições do sistema operacional.

O Senac São Paulo estabeleceu parceria com o LPI, tornando-se um LATP LPI Approved Training Partner, Centro Autorizado de Treinamento LPI para oferecer os cursos preparatórios para a certificação LPI.

Conheça os novos cursos oferecidos pelo Senac e torne-se um profissional certificado!

Certificação LPI nível 1 – LPIC1  
Certificação LPI nível 2 – LPIC2



**viva senac**  
são paulo

Mais informações, acesse  
[www.sp.senac.br/certificacoes](http://www.sp.senac.br/certificacoes)  
0800 883 2000