infra
MAGAZINE
Edição 10 :: Ano I
DevMedia

ESPECIAL:
Gerenciamento de usuários com OpenLDAP
Controle a base de usuários por meio do servidor de diretórios e do Squid

Splunk – Parte 1
Analise os dados do seu ambiente de TI em tempo real

MONITORANDO O AMBIENTE DE TI

Conheça as soluções Cacti, Wireshark e Splunk

Cacti
Avaliando dados de switches, roteadores e servidores

Wireshark
Melhore a gerência da rede com um analisador de pacotes

TENHA ACESSO ILIMITADO A TODO CONTEÚDO DO SITE DEVMEDIA



SEJA UM ASSINANTE MVP!

Tenha acesso garantido a qualquer tema que você possa necessitar!

VIDEOS -AULAS

O assinante MVP acessa todas as vídeo-aulas do site DevMedia. Nenhum outro produto dá essa possibilidade. São mais de **7.000 video-aulas.**

CURSOS

Nosso site possui atualmente **+ de 250 cursos online** com a qualidade que você já conhece e confia. Com o MVP você terá acesso a todos eles.

REVISTAS

Assinante MVP acessa todas as nossas revistas na versão digital, incluindo todas as edições anteriores. Quando pintar aquela dúvida você terá mais de **3.500 artigos para consultar!**

ASSINATURA MVP MOST VALUABLE PROFESSIONAL

A assinatura MVP te dá a oportunidade de aumentar seu conhecimento em diversos assuntos, com qualidade e economia.

Invista na sua carreira se torne um profissional ainda mais completo!

POR APENAS R$59,90 POR MÊS

DevMedia

Acesse: www.devmedia.com.br/mvp

Sumário

Infra

06 – Monitoramento com o Cacti
Adote o Cacti para monitorar switches, roteadores e servidores
[*Thiago Nache*]

Infra, Segurança, Web

13 – Gerenciamento de usuários com OpenLDAP
Controle a base de usuários por meio do servidor de diretórios e do Squid
[*Ivani Nascimento e Pathiene Gerstenberger*]

Infra, Segurança

26 – Análise do tráfego de rede com o Wireshark
Como melhorar a gerência da rede utilizando o famoso analisador de pacotes?
[*Rodrigo Sanches Miani*]

Infra

33 – Splunk: Monitorando o ambiente de TI – Parte 1
Saiba como monitorar todo o seu ambiente de TI
[*Wagner Bianchi*]

[Infra] Artigo que apresenta tecnologias relacionadas a questões de Infraestrutura.

[Segurança] Artigo dentro do contexto de Segurança: redes, servidores, sistemas operacionais, infraestrutura em geral.

[Web] Artigos sobre ou que envolvam tecnologias de infraestrutura para WEB.

FÓRUM
DEVMEDIA

O lugar perfeito para você ficar por dentro de tudo o que acontece nas tecnologias do mercado atual



No fórum da DevMedia você irá encontrar uma equipe disponível e altamente qualificada com consultores e colaboradores prontos para te ajudar a qualquer hora e sobre qualquer assunto. Temos as salas de Java, .NET, Delphi, Banco de Dados, Engenharia de Software, PHP, Java Script, Web Design, Automação comercial, Ruby on Rails e muito mais!

ACESSE AGORA
www.devmedia.com.br/forum

DevMedia

infra MAGAZINE

Ano I • Edição 10 • 2013

Edição

Editor

Eduardo Spínola (*eduspinola@gmail.com*)

Sub Editores

Marco Antônio Pereira Araújo (maraujo@devmedia.com.br)

Rodrigo Oliveira Spínola (rodrigo.devmedia@gmail.com)

Produção

Jornalista Responsável Kaline Dolabella - JP24185

Atendimento ao leitor

A DevMedia possui uma Central de Atendimento on-line, onde você pode tirar suas dúvidas sobre serviços, enviar críticas e sugestões e falar com um de nossos atendentes. Através da nossa central também é possível alterar dados cadastrais, consultar o status de assinaturas e conferir a data de envio de suas revistas. Acesse **www.devmedia.com.br/central**, ou se preferir entre em contato conosco através do telefone 21 3382-5038.

Publicidade

publicidade@devmedia.com.br – 21 3382-5038

Anúncios – Anunciando nas publicações e nos sites do Grupo DevMedia, você divulga sua marca ou produto para mais de 100 mil desenvolvedores de todo o Brasil, em mais de 200 cidades. Solicite nossos Media Kits, com detalhes sobre preços e formatos de anúncios.

Fale com o Editor!

É muito importante para a equipe saber o que você está achando da revista: que tipo de artigo você gostaria de ler, que artigo você mais gostou e qual artigo você menos gostou. Fique a vontade para entrar em contato com os editores e dar a sua sugestão!

Se você estiver interessado em publicar um artigo na revista ou no site Easy Java Magazine, entre em contato com o editor, informando o título e mini-resumo do tema que você gostaria de publicar:

Eduardo Spínola - Editor da Revista
eduspinola@gmail.com

DevMedia

Editorial

Monitorar o ambiente de TI é uma atividade essencial a qualquer administrador/gerente. Devido a sua importância, esta é uma tarefa que ganha cada vez mais destaque e o número de soluções que possibilitam a sua realização só faz aumentar, assim como melhorar em qualidade. Neste cenário, analisaremos nesta edição três ferramentas relacionadas ao monitoramento de TI: Cacti, Wireshark e Splunk.

Monitoramento com o Cacti aborda o monitoramento de ativos de rede utilizando o Cacti. Explica-se o que ele é, como é seu funcionamento, como foi escrito, como adicionar devices e criar gráficos. Além disso, explica-se como configurar o Cacti a partir de uma instalação mínima do CentOS 6, instalar e configurar o servidor web e o banco de dados.

Em **Análise do tráfego de rede com o Wireshark** são apresentadas as principais características da ferramenta de captura de pacotes Wireshark. Serão descritos: o funcionamento básico de um analisador de pacotes, os três módulos principais que compõem a arquitetura do Wireshark e os primeiros passos para o uso da ferramenta. Também será mostrado um breve exemplo da captura de pacotes TCP e HTTP usando filtros. Por fim, serão expostas algumas aplicações práticas em que o Wireshark pode ser utilizado com o intuito de melhorar a gerência e administração das redes de computadores.

Em **Splunk: Monitorando o ambiente de TI – Parte 1** analisaremos o Splunk, uma ferramenta que permite coletar, indexar e analisar a informação gerada pelo negócio, também conhecida pelo nome de dados de máquina. A partir da coleta e da indexação de toda a informação de syslog de uma máquina, por exemplo, é possível determinar vários problemas que a TI não fazia nem ideia de que ocorriam com o ambiente.

Por fim, o artigo **Gerenciamento de usuários com OpenLDAP** demonstra como instalar e configurar o OpenLDAP, protocolo utilizado no gerenciamento de identidades na integração de serviços em redes heterogêneas e como solução para centralizar a autenticação de usuários no proxy Squid.

Eduardo Oliveira Spínola
eduspinola@gmail.com
twitter.com/Java_Magazine



Seu **conhecimento vale muito** na DevMedia

Publique artigos e vídeos gratuitamente no site e participe do programa **PayContent.**

Sua participação pode ser premiada: quanto mais você postar, mais chances de ganhar.



Para saber mais acesse:
www.devmedia.com.br/paycontent

DevMedia

Monitoramento com o Cacti

Adote o Cacti para monitorar switches, roteadores e servidores

O Cacti é uma solução PHP e MySQL que utiliza a engine RRDtool (base de dados desenvolvida para armazenar séries de dados numéricos sobre o estado de redes de computadores) para gerar gráficos. Ela possui uma interface gráfica para gerenciamento bastante amigável onde se podem realizar todas as configurações como, por exemplo: de *devices* monitorados, grupos, *templates*, formas de coleta, formas de armazenamento, formas de exibição, etc.

Neste artigo iremos explicar as principais características da ferramenta, como ela funciona e o que se pode fazer para ajudar o administrador de redes a solucionar problemas que possam ocorrer na infraestrutura.

A instalação padrão já possui diversas configurações prontas. Veremos no estudo de caso, um pouco mais à frente, que para monitorar o uso da banda de um switch ou roteador via SNMP são necessárias pouquíssimas configurações. Se você deseja saber mais sobre como o protocolo SNMP funciona, veja o artigo monitorando ativos de rede referenciado na seção **Links** deste artigo.

Com o Cacti é possível monitorar tanto servidores quanto ativos de rede. Porém, recomenda-se que o administrador utilize o Cacti apenas para os ativos de rede. Caso você tenha um ambiente grande de servidores, utilize o Collectd. Neste artigo abordaremos como monitorar os equipamentos de rede de forma abrangente para que se possa agir antes do problema acontecer.

Resumo DevMan

De que se trata o artigo:

Este artigo aborda o monitoramento de ativos de rede utilizando o Cacti. Explica-se o que ele é, como é seu funcionamento, como foi escrito, como adicionar devices e criar gráficos. Além disso, explica-se como configurar o Cacti a partir de uma instalação mínima do CentOS 6, instalar e configurar o servidor web e o banco de dados.

Em que situação o tema é útil:

Este artigo é útil para todos que tenham interesse no monitoramento de ativos de rede. Em particular, para aqueles interessados na configuração de um ambiente de monitoramento.

Monitoramento com o Cacti:

Este artigo é destinado a todos que ainda não conhecem ou conhecem pouco sobre o Cacti. Aqui você entenderá como ele funciona, quais são as tecnologias envolvidas e como ele pode ser configurado. Descreve-se como instalar e configurar componentes do famoso LAMP (Linux, Apache, MySQL e PHP) e como o Cacti faz uso dessas tecnologias.

Estrutura do Cacti

Para entender como o Cacti funciona, podemos dividi-lo em três partes: coleta de informações, forma de armazenamento dos dados no RRD e como esses dados serão exibidos. Todas as partes precisam estar funcionais e integradas, do contrário você terá problemas.

Caso você tenha algum problema com seu gráfico, antes de tudo veja os logs do sistema. O Cacti possui um modo de DEBUG para seu log bastante eficaz. Consegue-se resolver boa parte dos problemas fazendo análise do mesmo. Caso o administrador não consiga entender o problema pelo log, analise primeiro se os dados estão sendo coletados, se eles estão sendo armazenados no RRD e, por último, se a configuração de exibição está coerente com o que está gravado no RRD.

Coleta

Na instalação padrão do sistema, existem algumas formas padrão de coleta que são suficientes para gerar os gráficos dos roteadores e switches. Porém, caso você tenha uma situação específica e quiser criar um gráfico disso, é possível. Por exemplo, você pode gerar um gráfico de conexões TCP na porta 80 do servidor local.

Apesar do Cacti ter sido escrito na linguagem de programação PHP, os scripts podem ser escritos em qualquer linguagem desde que seja executável e a saída seja definida de acordo com o formato esperado. Os *scripts* são chamados de *Data Input Method* e estão localizados em *LOCAL_INSTALACAO/scripts/*. O Cacti disponibiliza, por padrão, *scripts* para gerar gráficos do servidor local, ou seja, com esses *scripts* você não conseguirá gerar gráficos de computadores remotos.

Outra forma de coleta é por meio de *SNMP DATA QUERY*. Este método consiste em um arquivo XML (*Extensible Markup Language*) que contém as OIDs dos gráficos que serão gerados. As configurações para gerar gráficos de estatísticas de interface de rede e utilização de CPU de alguns equipamentos de rede estão prontas e localizadas em *LOCAL_INSTALACAO/resource/snmp_queries/*.

Armazenamento

A forma de armazenamento das informações no RRD são definidas através de um *Data Template*. Nele, informa-se qual o método de coleta, quais serão os tipos de mídias disponíveis (RRA), o tipo do dado e suas características.

Um RRD é composto por algumas camadas. Observe na **Figura 1** cada uma delas:

- PDP (Primary Data Point): É o ponto inicial de um RRD, ou seja, os valores que estão sendo coletados no momento e que serão utilizados para gerar os valores atuais, sendo eles mínimo, máximo e a média;
- CF (Consolidation Function): Consolida vários PDPs gerando média, mínimo e máximo;
- CDP (Consolidated Data Point): Após terem sido consolidados os últimos valores coletados e gerado a média, valor mínimo e valor máximo, gera-se um item que será armazenado como histórico, o CDP;
- RRA (Round-Robin Archive): Armazena uma quantidade fixa de CDPs para gerar o histórico;
- RRD (Round-Robin-Database): Representa o resultado final, que é composto por vários RRAs.

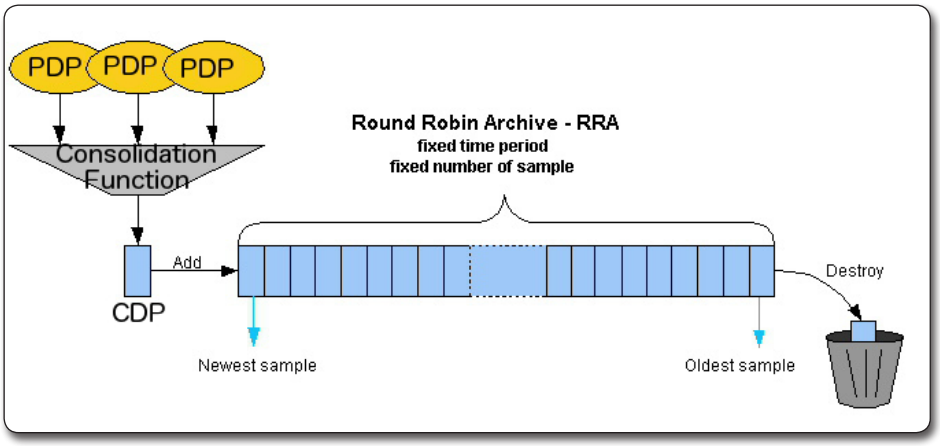


Figura 1. RRD Engine (Adaptada de http://www.loriotpro.com)

Nome	Sistema operacional	Proc	Mem	Endereço IP	Máscara de rede
Cacti1	CentOS 6.3 x86_64	1	1G	10.211.55.11	255.255.255.240

Tabela 1. Informações da máquina virtual

Listagem 1. Instalando o Cacti a partir do código fonte.

```
# yum install –y php php-cli php-snmp php-mysql php-pdo rrdtool net-snmp-utils
# yum install -y wget && wget http://www.cacti.net/downloads/cacti-0.8.8a.tar.gz
# tar xzvf cacti-0.8.8a.tar.gz -C /usr/share/
# ln -s /usr/share/cacti-0.8.8a /usr/share/cacti
```

Exibição

Para exibir um gráfico, o administrador deve associar um *Device* (o nome do servidor) a um *Data Template*. Essa associação gerará um *Data Source* (que é o RRD em si), ou seja, o local onde o Cacti buscará os dados.

Por último, configura-se o Graph Template. Nele informa-se o *Data Source*, quais informações serão exibidas e as características do gráfico.

Graph Templates

Para facilitar a vida do administrador, o Cacti permite que criemos templates. Assim, ao invés de ter que associar vários gráficos a cada novo *device*, criamos um Host Template e depois associamos diversos gráficos a um Host Template e cada novo *device* apenas ao Host Template.

O Cacti já disponibiliza alguns Host Templates de exemplo que podemos utilizar para criação de nossos próprios Host Templates. Durante o estudo de caso faremos uso de um deles.

Hands on

Agora iremos instalar, configurar o Cacti e gerar alguns gráficos de exemplo. O sistema operacional onde a máquina virtual foi criada é o Mac OS X 10.6.8 e teremos o Parallels Desktop como *hypervisor*.

Durante a instalação do CentOS na máquina virtual, escolha o tipo *minimal*. Assim não será personalizado nenhum pacote durante a instalação.

A **Tabela 1** apresenta as informações para instalação e configuração da máquina virtual onde iremos instalar o Cacti.

Instando o Cacti

O Cacti é uma ferramenta simples de instalar e configurar. Basta fazer o download do código fonte, descompactá-lo e instalar o PHP. Para tal, execute a sequência de comandos da **Listagem 1**.

Apache

Como estamos criando gráficos, é natural que precisemos de um servidor web para que o gráfico esteja disponível

remotamente. O httpd já foi instalado como parte dos pacotes anteriores. O processo httpd é um servidor web que será utilizado para possibilitar a visualização dos gráficos. Para configurar este serviço, execute a sequência de comandos da **Listagem 2**. Com estes comandos alteramos o dono dos arquivos do sistema para o usuário que o processo roda, ou seja, apache.

Listagem 2. Configurando o Apache.

```
# chown -R apache: /usr/share/cacti*
# cat > /etc/httpd/conf.d/cacti.conf << EOF
> Alias /cacti   /usr/share/cacti
> <Directory /usr/share/cacti/>
>    Options Indexes FollowSymLinks
>    AllowOverride None
>    Order allow,deny
>    Allow from all
> </Directory>
> EOF
# chkconfig httpd on && service httpd start
```

Feito isso, agora estamos prontos para instalar o banco de dados.

Instalando o MySQL

O Cacti necessita de um banco de dados para armazenar suas configurações e algumas informações. Para isto, devemos instalar o MySQL. Sendo assim, execute a sequência de comandos da **Listagem 3** para instalar e configurar o banco de dados.

Listagem 3. Instalando o MySQL Server.

```
# yum install -y mysql-server
# chkconfig mysqld on && service mysqld start
# /usr/bin/mysqladmin -u root password 'SENHA'
```

Lembre-se de substituir a palavra SENHA pela senha de root que você deseja. É importante que no servidor de produção a senha seja complexa, ou seja, contenha caracteres minúsculos, maiúsculos, especiais e numéricos.

Em seguida, execute o comando *mysql_secure_installation* para configurar o banco. Neste instante serão feitas algumas perguntas no *prompt*, responda conforme a **Listagem 4**.

Listagem 4. Script de configurações de segurança do MySQL.

```
Enter current password for root (enter for none):
Change the root password? [Y/n] n
Remove anonymous users? [Y/n] Y
Disallow root login remotely? [Y/n] Y
Remove test database and access to it? [Y/n] Y
Reload privilege tables now? [Y/n] Y
```

Na primeira resposta digita-se a senha definida anteriormente na **Listagem 3**. Por ser um campo de senha, ele não é exibido. Em seguida, responde-se n para não trocarmos a senha de root, pois ela já foi definida. A terceira resposta, Y, remove o acesso anônimo ao banco de dados. A quarta resposta, também Y, é um requisito de segurança, pois não permite que o root efetue login remotamente no banco de dados. Se este acesso for necessário, utilize o comando GRANT posteriormente. A penúltima resposta é Y para apagar o database test e, por fim, efetua-se o reload dos novos privilégios do banco. Neste caso, a remoção do login remoto do root.

Por fim, criamos o banco de dados do Cacti, liberamos o acesso e criamos as tabelas com as informações iniciais. Para isso, execute a sequência de comandos da **Listagem 5**.

Listagem 5. Criar o banco de dados e importar os dados inicias.

```
# mysqladmin -p -u root create cacti
# echo "GRANT ALL ON cacti.* TO cactiuser@localhost IDENTIFIED BY \"cactiuser\";"
| mysql -p -u root
# echo "FLUSH PRIVILEGES;" | mysql -p -u root
# mysql -pcactiuser -u cactiuser -D cacti < /usr/share/cacti/cacti.sql
```

O comando *mysqladmin* cria o database cacti. Em seguida, utiliza-se o comando *echo* para escrever uma query no MySQL. Por fim, executa-se o script para criar as tabelas do database cacti e inserir os dados iniciais.

SELinux

O SELinux é uma facilidade de segurança disponível nas distribuições baseadas em RPM. Ele é uma camada a mais de segurança para acesso ao sistema operacional, ou seja, ele controla se o soft-ware tem permissão para acessar o arquivo. O RPM é um gerenciador de pacotes, ou seja, ele controla todos os pacotes instalados e suas dependências. Recomenda-se que o SELinux permaneça habilitado, pois isto trará maior segurança ao seu sistema.

Para que o Cacti funcione corretamente com o SELinux, execute a sequência de comandos da **Listagem 6**.

Listagem 6. Configurando o módulo de segurança SELinux.

```
# setsebool httpd_read_user_content=1
# setsebool httpd_tty_comm=1
# setsebool httpd_can_network_connect=1
# setsebool httpd_builtin_scripting=1
# setsebool httpd_enable_homedirs=1
# setsebool httpd_enable_cgi=1
# setsebool httpd_can_network_connect=1
```

Esses comandos são booleanos, ou seja, são flags que se habilitadas permitem que determinados processos possam executar determinada tarefa. Para o sistema operacional Linux, conexões são arquivos, como também os devices, portanto a permissão de acesso também é controlada pelo SELinux.

Caso você tenha problemas para utilizar o SELinux, desabilite-o com o comando *setenforce 0*. Porém, recomenda-se que ele esteja habilitado, pois este é um módulo de segurança para o sistema operacional.

Firewall

Para que você consiga acessar os gráficos, é necessário liberar a porta 80 com o comando *iptables -I INPUT -p tcp --dport 80 -j ACCEPT* e, em seguida, salvar as regras com o comando *service iptables save*. Você pode também desabilitar o firewall com o comando > */etc/syscconfig/iptables* e, em seguida, *service iptables restart*. O firewall estar habilitado ou não é uma questão peculiar a cada ambiente. Se você for desabilitá-lo, certifique-se que existe um firewall, na rede ou sub-rede, filtrando os acessos aos servidores.

Instalação e configuração gráfica do Cacti

Para finalizar a instalação do Cacti, acesse pelo browser o endereço *http://10.211.55.11/cacti/* e clique no botão *Next* na tela inicial da instalação. Feito isso, deverá ser informado se o que estamos fazendo é uma nova instalação ou atualização de versão. Deixe a opção *New Install* selecionada e clique no botão *Next*.

No próximo passo perceberemos que o Cacti tentou identificar se todos os binários necessários para seu funcionamento estão instalados. Observe na **Figura 2** que todos já foram instalados e clique em *Finish*.

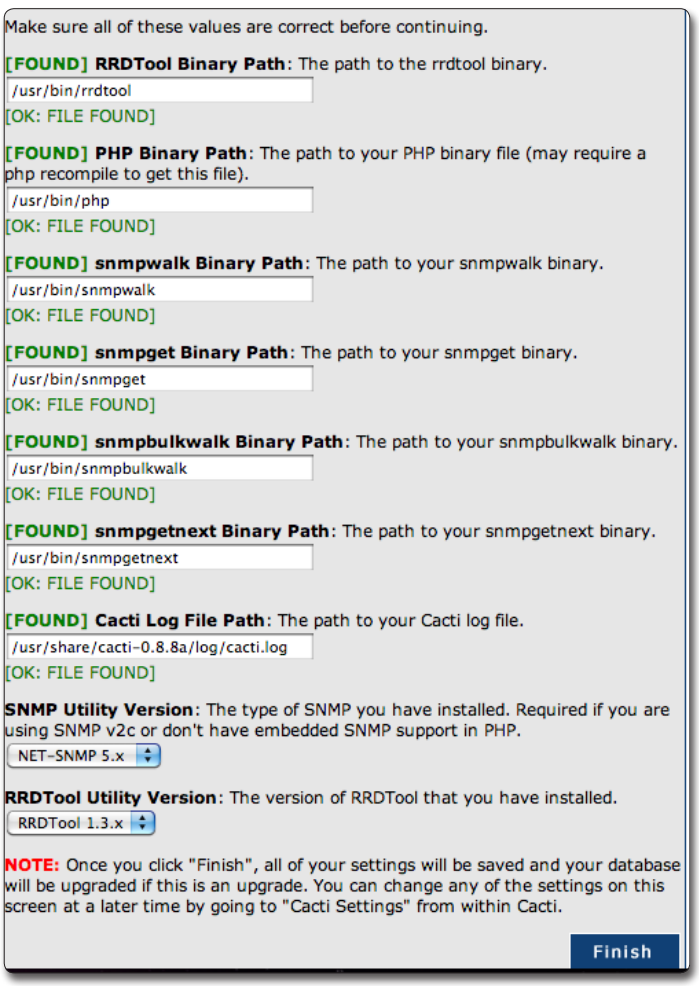


Figura 2. Caminhos para binários do sistema

O arquivo *cacti.sql* inseriu no banco de dados o usuário admin com a senha admin. Será solicitado que você troque a senha nesse primeiro acesso.

Antes de começar a coletar informações e gerar gráficos, devemos configurar o Cacti. Para isso, após acessar o Cacti, no menu esquerdo o administrador encontrará a opção *Settings*. Clique nela e serão exibidas todas as abas de configurações do Cacti.

A primeira aba que será exibida é a aba *General*. Nela encontra-se informação sobre os logs e configurações globais de *SNMP* conforme exibido na **Figura 3**. Não altere nada aqui, mas é importante que você conheça os campos *Poller Logging Level* e o quadro *SNMP Defaults*.

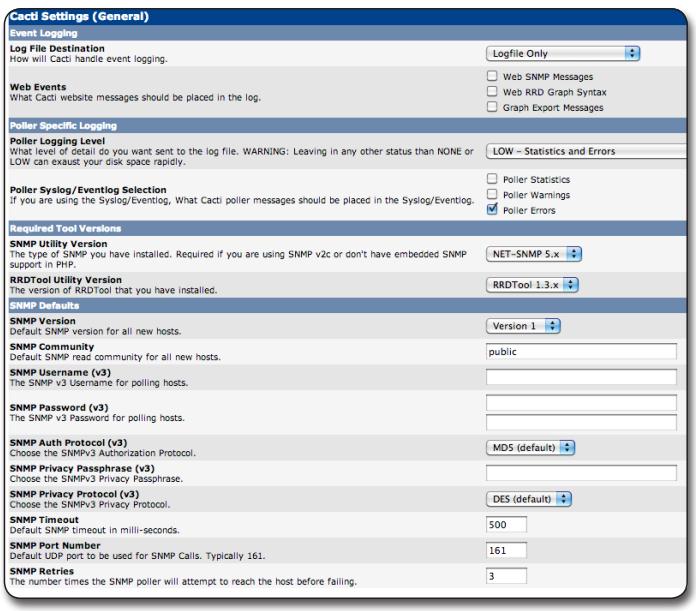


Figura 3. Aba General

No primeiro campo você informa qual o nível de detalhamento do log. Quanto maior o nível, mais informações o Cacti irá gerar. No ambiente de produção que não esteja passando por problemas, recomendamos manter a opção *LOW*. Caso você esteja passando por problemas, altere para *DEBUG* ou *DEVEL*.

Já no quadro *SNMP Defaults*, você define as informações globais de SNMP como, por exemplo: versão e porta. Esta configuração poderá ser sobrescrita quando o administrador adicionar um novo *device*.

Na aba *Paths* não alteramos nada, pois acabamos de sair da instalação e já garantimos que está tudo configurado. Contudo, é importante conhecer esta aba, pois se algum dia precisarmos mudar algum binário do pacote *net-snmp-utils*, o administrador provavelmente a utilizará. Observe na **Figura 4** as configurações atuais.

Agora conheceremos a aba *Poller*. *Poller* é o nome dado ao processo de coleta e armazenamento de dados. Esta é a aba mais importante, pois nela estão as configurações que indicam como o Cacti irá utilizar os recursos do seu servidor. Observe na **Figura 5** as configurações iniciais do Cacti.

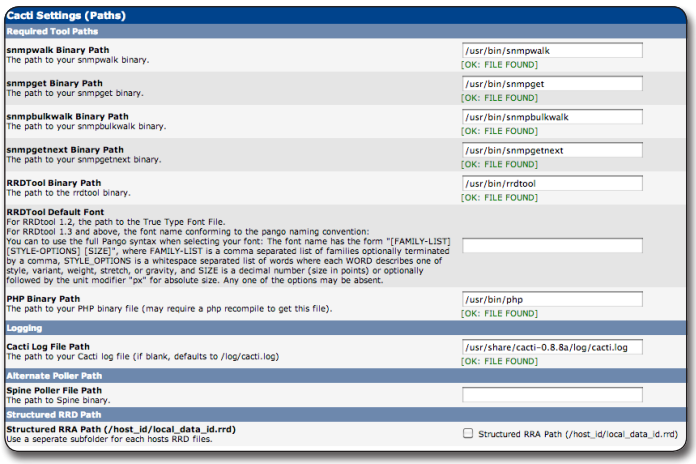


Figura 4. Aba Paths

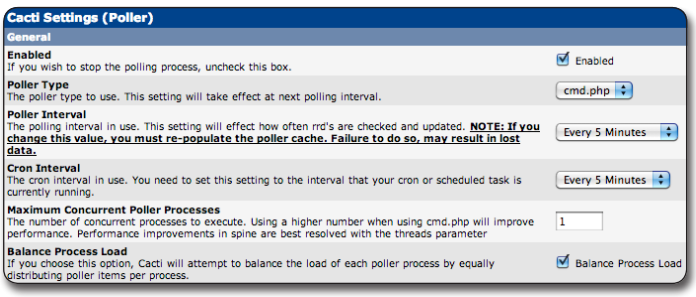


Figura 5. Aba Poller

Para saber o valor de cada um desses campos, verifique as regras a seguir que explicam como configurá-los:

• **Maximum Concurrent Poller Processes:** Quantidade de processos PHP que irão rodar para o Poller. Recomenda-se que o valor seja a quantidade de núcleos de CPU menos um. Isto quer dizer que ficará um núcleo de CPU para o sistema operacional e os outros núcleos serão utilizados pelo Cacti;

• **Poller Interval e Cron Interval:** Informa o tempo de coleta e o intervalo entre elas. Geralmente são iguais e com valor de cinco minutos;

• **Maximum Number of Threads per Process:** Define a quantidade máxima de threads. Este valor vai depender da capacidade do seu processador. Comece com o valor 10, acompanhe o uso da CPU e o tempo total de coleta. Se o tempo de coleta for muito grande e houver recurso de CPU sobrando, aumente a quantidade de Threads.

Aba Visual

Nesta aba estão as configurações de fontes, cores, etc. conforme apresentadas na **Figura 6**. O administrador deve configurar conforme sua preferência. Caso não queira alterar nada, as configurações padrões são suficientes.

Cron

O processo de Poller foi configurado na ferramenta e precisa ser executado no sistema operacional conforme desejado. Para esta tarefa utiliza-se o cron. Este é o agendador de tarefas do Linux, ou seja, é ele quem irá executar o script de Poller do Cacti. Para instalar e configurar o serviço, execute a sequência de comandos da **Listagem 7**.

Nesta listagem, configuramos a coleta (poller) para ser realizada a cada cinco minutos.



Figura 6. Aba Visual

Listagem 7. Instalando e configurando o agendador de tarefas Cron.

```
# yum install -y cronie-noanacron
# echo "*/5 * * * * apache php /usr/share/cacti/poller.php > /dev/null 2>\&1" > /
etc/cron.d/cacti
# service crond start
```

Monitorando o consumo da porta de rede

Para simular um roteador ou switch em nosso servidor, utilizaremos o serviço snmpd do Linux. Nos artigos anteriores, explicamos detalhadamente sobre ele. O pacote **net-snmp-utils** foi instalado anteriormente, no comando *yum install* da **Listagem 1**. Portanto, execute a sequência de comandos da **Listagem 8** para configurar o serviço.

Listagem 8. Configurando serviço SNMPD.

```
# echo 'rocommunity public' > /etc/snmp/snmpd.conf
# chkconfig snmpd on
# service snmpd start
```

Além da interface gráfica, o Cacti também possui uma interface de linha de comando, denominada CLI. Nela é possível realizar diversas ações, sendo inclusive muito útil para atividades em massa.

Para trabalhar com o CLI, execute o comando *cd /usr/share/cacti/cli* e, em seguida, utilize o *ls* para listar todas as ações disponíveis via CLI. Os nomes são bastante intuitivos.

Iremos abordar o script *add_device* para que você entenda como funciona o CLI do Cacti. Existe um *Host Template* padrão que coleta estatísticas das interfaces de rede, então vamos adicionar o novo *device* já neste *template*. Para isso, execute o comando *php add_device.php --list-host-templates*. Este retornará todos os *templates* disponíveis.

Utilizaremos o template "1 Generic SNMP-enabled Host", pois este realiza a coleta via SNMP das estatísticas de rede.

Conhecido o id do template, execute o comando a seguir:

```
# php add_device.php --description=switch1
--ip=localhost.localdomain --avail=pingsnmp --ping_
method=udp --ping_port=53 --template=1
```

Os parâmetros do script são:

- --description – informa o nome dado ao device;
- --ip – informa o endereço do device (em uma situação real, substitui-se o *localhost.localdomain* pelo FQDN ou endereço IP);
- --avail – é a forma de verificar se o servidor está respondendo ou não. Utiliza-se *pingsnmp* para que o Cacti faça um teste de ping e outro teste via SNMP;
- --ping_method – informa o protocolo para o ping;
- --ping_port – informa a porta para o ping;
- --template – informa o ID do Host Template.

O comando executado irá retornar algo como:

```
Success - new device-id: (2)
```

Estude com calma cada um dos scripts. No começo eles podem parecer complicados, mas também podem ser bastante úteis. Realizar tarefas em massa pela interface gráfica é muito mais complicado. Pela interface de linha de comando isto é muito mais simples.

Criando gráficos

Conforme mencionado anteriormente, precisamos associar o *device* criado ao *Data Template*. Deste modo, acesse a interface, no menu *Devices*, clique em *switch1* e depois em *Create Graphs for this host*.

Na próxima tela criamos os gráficos que serão gerados. Para criar o gráfico de tráfego (*IN/OUT*), selecione o combo da interface que você deseja, informe o *graph type In/Out Bits with 95th Percentile* e clique no botão *Create*, conforme **Figura 7**.

Outro gráfico bastante interessante apresentado na **Figura 8** é o *In/Out Bits*. Este modelo de medição de uso de banda é bastante utilizado no mercado para cobrar o consumo de link de internet. Para criá-lo, selecione o combo da interface que você deseja, informe o *graph type In/Out Bits (64-bit Counters)* e clique no botão *Create*.

Primeira coleta

Para a coleta começar a ser realizada, é preciso criar a tabela de Poller. Para isto, execute o comando *php /usr/share/cacti/cli/rebuild_poller_cache.php*.

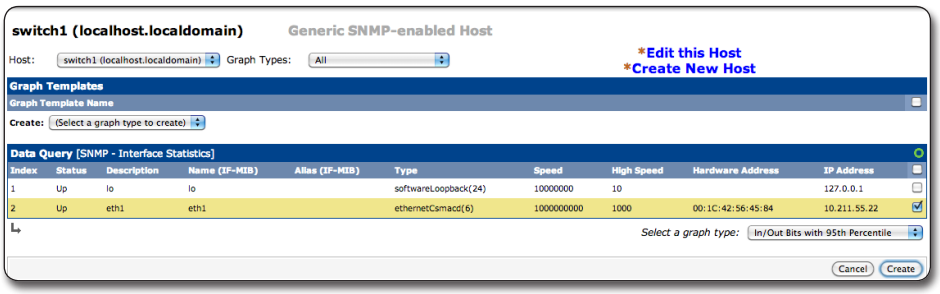


Figura 7. Criar gráfico de consumo de banda

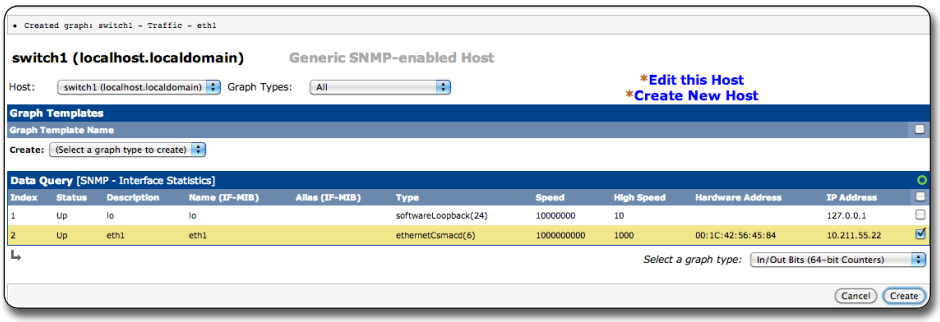


Figura 8. Criar gráfico 95th percentile



Depois de criado o cache do Poller com a execução do script anterior (*rebuild_poller_cache.php*), as coletas serão executadas conforme configurado. Por padrão, os gráficos do servidor do Cacti (chamado localhost) já estão configurados. Para saber se tudo foi feito corretamente, verifique o log do processo de Poller através do comando *tail -f /usr/share/cacti/log/cacti.log*. O resultado deverá ser parecido com:

```
09/02/2012 02:50:02 PM - SYSTEM STATS: Time:1.0234 Method:cmd.php Processes:1
Threads:N/A Hosts:2 HostsPerProcess:2 DataSources:5 RRDsProcessed:5
```

Para que as informações comecem a aparecer no gráfico, devido ao cálculo de média que é feito, o RRD precisa de, no mínimo, três coletas. Portanto, aguarde três coletas e depois acesse o endereço *http://10.211.55.22/cacti/graph_view.php?action=preview* pelo browser para visualizar os gráficos. As **Figuras 9** e **10** ilustram os gráficos gerados para as configurações que fizemos para os itens *Bandwidth In/Out* e *95th percentile*.

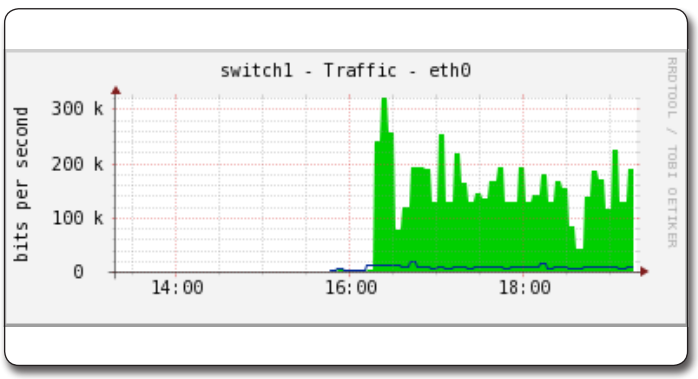


Figura 9. Bandwidth In/Out

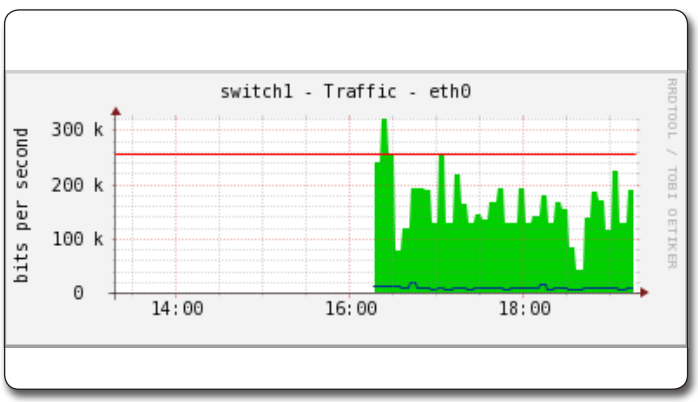


Figura 10. Bandwidth In/Out 95th percentile

Conclusão

Feito todo o trabalho de configuração, o Cacti já estará coletando as informações e armazenando-as. Portanto, aguarde o tempo de três coletas (se você deixou o tempo padrão, aguarde 15 minutos) e acesse pelo navegador o endereço *http://10.211.55.22/cacti/graph_view.php?action=preview*.

O monitoramento de redes é essencial para a infraestrutura de TI. Deste modo, visando segurança e confiabilidade, deve-se aplicar o monitoramento ativo e passivo para garantir o seu correto funcionamento.

Baseado em soluções open source, mostramos neste artigo como é possível monitorar de forma eficiente e eficaz toda a infraestrutura da rede, garantindo assim o bom funcionamento de todos os outros componentes.

Agora seu ambiente já está pronto para gerar gráficos dos seus equipamentos de rede.



Thiago Nache
thiagonbcarvalho@gmail.com

Atua na área de infraestrutura há mais de 8 anos, tem grande experiência em sistema de alta disponiblidade e alta performance. Cursou ensino médio técnico em processamento de dados, graduação em Tecnologia em Redes de Computadores na FIAP e possui certificações Red Hat e ITIL. Tem experiência em empresas de grande porte. Atualmente atua no iG como administrador de sistemas sênior.

Site do Cacti
http://cacti.net

Site do RRDtool
http://oss.oetiker.ch/rrdtool/

Wiki do RRDtool
http://oss.oetiker.ch/rrdtool-trac/

Documentação do RRDtool
http://oss.oetiker.ch/rrdtool/doc/index.en.html

Site do PHP
http://php.net

Artigo sobre o Collectd
http://www.devmedia.com.br/websys.4/webreader.asp?cat=62&revista=infra magazine_5#a-4497

Artigo monitorando ativos de rede
http://www.devmedia.com.br/websys.4/webreader.asp?cat=62&revista=infra magazine_7#a-4893

Gerenciamento de usuários com OpenLDAP

Aprenda a gerir a base de usuários da rede por meio do servidor de diretórios e a aplicar a solução no controle de logins do proxy Squid

Mudanças ocorrem constantemente em um ambiente corporativo, seja por motivo de rotatividade de colaboradores, mobilidade na hierarquia, mudanças de funções e setores, inclusão/retirada de dispositivos, seja por outros motivos. Deste modo, a administração de sistemas torna-se uma tarefa complexa, pois cada um dos fatores mencionados necessita de configuração para que possam operar de acordo com a necessidade do ambiente. Em redes de pequeno porte, muitas vezes é possível utilizar procedimentos manuais para configuração de serviços de rede e administração de usuários. Porém, à medida que a rede cresce, se torna mais complexa e heterogênea, e o gerenciamento de identidades se torna um desafio, trazendo riscos à segurança em função da escolha de senhas frágeis e do vazamento de informações.

Dessa forma, entende-se que todo ambiente de rede necessita armazenar informações de modo a possibilitar o seu gerenciamento. Como exemplo, temos autenticações, grupos de usuários, permissões, cotas de armazenamento, acessos a compartilhamentos, entre outros. Muitas organizações possuem ambientes mistos com várias plataformas (Linux, Windows, Unix) conectadas fisicamente e distribuídas geograficamente, que trocam informações entre si. O problema desse tipo de rede é que para cada plataforma ou rede local existente no ambiente de rede (rede física) será necessário alimentar uma base de dados com as mesmas informações, o que torna a administração complexa, uma vez que será necessário garantir a sincronia e a replicação das informações. Se a solução para a gerência dos dados não for eficaz, torna-se difícil a organização e a distribuição das informações, causando duplicidades, maior custo no suporte e falta de segurança.

Resumo DevMan

De que se trata o artigo:

Neste artigo será demonstrado como instalar e configurar o OpenLDAP, protocolo utilizado no gerenciamento de identidades na integração de serviços em redes heterogêneas e como solução para centralizar a autenticação de usuários no proxy Squid.

Em que situação o tema é útil:

Manter um cadastro centralizado de usuários é um desafio para os administradores de sistema, visto que alguns fatores como rotatividade de colaboradores e mobilidade na hierarquia dificultam a manutenção de informações sobre um determinado recurso de uma organização. No entanto, o OpenLDAP é um protocolo que possibilita a integração de redes heterogêneas por meio de um serviço de diretórios que realiza a centralização da autenticação. Em um projeto de implementação de um serviço de rede como o Squid, que irá atuar em um ambiente com diversos sistemas operacionais, faz-se necessário o controle de acesso dos usuários aos recursos da rede, tarefa na qual o OpenLDAP tem uma importante atuação.

Gerenciamento de usuários com OpenLDAP:

Em um ambiente corporativo, um dos principais requisitos verificados em um sistema ou serviço de rede é a segurança. Tanto o sistema quanto o serviço devem garantir que as informações neles contidas sejam íntegras e que os acessos sejam feitos somente por usuários legítimos, além de fornecer mecanismos para proteção contra acessos não autorizados. Rotatividade de colaboradores, frequente mobilidade de hierarquia e divisão em setores são fatores que dificultam a manutenção de um cadastro de usuários com informações coerentes. Da mesma maneira, a utilização de diversas senhas para o acesso a sistemas ou serviços de rede coloca em risco a segurança da organização. Este artigo aborda a configuração de um servidor OpenLDAP para manter um diretório com as informações dos usuários centralizadas e integrá-lo com um proxy Squid.

Diante disso, o gerenciamento de identidades se torna um fator importante, seja qual for o tamanho da corporação ou da rotatividade de usuários, pois dessa forma há uma grande contribuição na vinculação destes em sistemas de Recursos Humanos (RH) ou cadastros básicos (por exemplo, cadastro de visitantes), permitindo a criação, o bloqueio e a exclusão automática de usuários.

Dentro do contexto de segurança, a prática do gerenciamento de identidades compreende três elementos: diretivas, processos e tecnologias, sendo que as diretivas se referem aos limites e padrões a serem seguidos para cumprir a política da organização e também práticas recomendadas; os processos descrevem as sequências de etapas que levam à conclusão de tarefas; e as tecnologias são as ferramentas automatizadas que auxiliam a organização a atingir os objetivos comerciais de forma eficiente e precisa, respeitando os limites especificados nas diretivas.

Com base nisso, este artigo aborda o aspecto tecnológico do gerenciamento de identidades a partir da instalação e da configuração de um servidor de diretórios utilizando o OpenLDAP integrado com o proxy Squid, de modo a permitir a centralização e o controle dos acessos na rede.

Serviço de Diretórios

Um diretório é um repositório de informações sobre objetos, organizados segundo um critério que facilite a sua consulta. De acordo com o OpenLDAP Software 2.4 Administrator's Guide (ver **Links**), a definição simples para diretório é: "um banco de dados específico, otimizado para busca, leitura e navegação".

Serviços de diretório e bancos de dados compartilham características comuns, tais como buscas rápidas e funções de atualização. Enquanto um serviço de diretório é utilizado mais para leitura do que escrita, em um banco de dados, assume-se que as operações de leitura e de escrita ocorrem mais ou menos com a mesma frequência. Diretórios podem conter informações descritivas baseadas em atributos que podem ser filtrados, ao passo que as transações realizadas devem ser finalizadas totalmente, não podendo ser concluídas de forma parcial, características essas encontradas em bancos de dados que são desenhados para manusear grandes volumes de atualizações e que suportam transações que permitem desistência de operações (*rollback*). Exemplos de diretórios utilizados no cotidiano são as listas telefônicas e os dicionários, pois ambos armazenam informações ordenadas para consulta para facilitar a busca por entradas. No caso da lista telefônica, as entradas são organizadas em ordem alfabética pelo nome da pessoa, no caso do dicionário, por verbete. Na computação, os serviços de diretórios podem buscar informações localmente (por exemplo, ao executar o comando *finger* em um servidor Linux/Unix) ou, ainda, em servidores que mantêm a informação distribuída (por exemplo, uma pesquisa no serviço de resolução de nomes deve retornar o mesmo resultado, independentemente do computador no qual foi realizada a busca).

Dentro do contexto de redes de computadores, seja de pequeno, médio ou grande porte, é justificada a utilização de um serviço de diretório, pois permite a mobilidade dos dados que podem ser agrupados por departamentos ou por unidades, centraliza as informações que ficarão disponíveis a todos os serviços da rede e possibilita delegar a administração de parte do diretório (que pode estar localizado fisicamente em outro servidor) a outros usuários.

LDAP

LDAP, que significa *Lightweight Directory Access Protocol*, tem como função definir o modo como um serviço de diretórios trabalhará, especificando critérios, mecanismos e métodos para armazenar e fornecer informações. Como o nome sugere, é um protocolo leve (*lightweight*) para acessar serviços de diretório, especificamente serviços de diretório baseados em X.500.

O X.500 é um protocolo baseado no conceito OSI, que especifica um modelo para a conexão de serviços de diretórios locais a fim de formar um diretório global distribuído, com suporte a funções de gerenciamento, tais como adição, alteração e remoção de entradas. O protocolo de acesso a diretórios DAP (*Directory Access Protocol*) faz parte das especificações do padrão X.500 e foi desenvolvido para trabalhar junto a todas as camadas do modelo *Open Systems Interconnection* (OSI), com o objetivo de definir o acesso de usuários aos serviços de diretórios que seu padrão provia. Porém, isso tornou sua implementação difícil, gerando aplicações complexas e lentas.

O LDAP foi criado como uma alternativa ao DAP para prover acesso aos serviços de diretórios do X.500, utilizando o modelo da pilha TCP/IP. Por esse motivo, é mais fácil de ser implementado, além de exigir menos recursos da rede e de memória, provendo melhor desempenho.

Posteriormente, foram criados servidores de diretórios para colocar em prática um software provedor de serviços de diretórios específico para o TCP/IP e o LDAP, deixando de lado o X.500. O LDAP foi padronizado pelo *Internet Engineering Task Force* (IETF) em julho de 1993 por meio da *Request For Comments* (RFC) 1487 e atualmente está na versão 3, conforme especificado na RFC 4510.

OpenLDAP

O projeto OpenLDAP é o resultado de um esforço colaborativo para desenvolver um software multiplataforma com ferramentas para implantação de um serviço de diretório LDAP em um ambiente de rede (software servidor, cliente, utilitários). É uma solução utilizada como alternativa às implementações de mercado como Microsoft Active Directory, Novell eDirectory, Sun iPlanet Directory Server e outros.

A configuração do OpenLDAP é feita através do slapd (Stand-Alone LDAP Daemon), que tem como função atender as requisições dos clientes, seja em forma de autenticação ou ainda, resposta a uma consulta no diretório. O slurpd (Stand-Alone LDAP Update Replication Daemon) é utilizado para a replicação do diretório, isto é, propaga as alterações de uma base slapd para outra. Como opções para a segurança, o OpenLDAP fornece suporte à criptografia TLS, SSL e SASL, além de permitir um controle granular por meio de listas de controle de acesso (ACLs).

Dentre as aplicações para uso do OpenLDAP, podemos citar:

- Base de autenticação de usuários da rede para serviços como:
  - Login de rede (utilizado para acessar estações Linux ou estações Windows através do software Samba, que é utilizado para integrar ambientes heterogêneos);
  - E-mail (utilizado para autenticar o usuário no envio e recebimento de mensagens, através dos protocolos POP3, SMTP, IMAP);
  - Acesso à Internet (Proxy);
- Catálogo de e-mails e cadastro de clientes.

Estrutura de serviço de diretório

Antes de realizar a instalação de um serviço de diretório, é importante compreender como as informações são organizadas.

No LDAP, as entradas do diretório são estruturadas de forma hierárquica, como uma árvore, para representar limites geográficos ou organizacionais. Por exemplo, pensando em uma estrutura de árvore, no topo estão as entradas que representam os países, seguidas das entradas de estados e organizações nacionais, e mais abaixo, entradas que representam unidades organizacionais, pessoas, hosts e outras informações que se deseja catalogar.

Uma entrada em um serviço de diretório é uma coletânea de atributos agrupados em um identificador único chamado DN (*Distinguished Name*). Um atributo consiste no par atributo/valor, sendo que um atributo pode ter mais de um valor, como na representação de membros de um grupo; por exemplo, o grupo marketing contém os seguintes membros: Maria, José, João.

A **Tabela 1** apresenta as entradas do modelo da estrutura de diretório.

Entrada	Descrição
C	country (país)
st	stateOrProvinceName (estado)
l	localityName (cidade)
dc	domain component (componente de domínio)
o	organizationName (organização)
ou	organizationalUnit (unidade organizacional)
cn	common name (nome ordinário)
sn	surname (sobrenome)

Tabela 1. Entradas da estrutura de diretórios

A **Figura 1** ilustra uma árvore de diretório considerando uma estrutura geográfica e organizacional. Nesse tipo de estrutura, o topo da árvore contém entradas que representam países e estados. Nesse ponto, a estrutura hierárquica é criada a partir de entradas das organizações nacionais, seguidas das unidades organizacionais que podem representar, por exemplo, unidades de uma empresa, como matriz e filiais, ou ainda divisões internas, como departamentos, grupos, usuários e outros.

Outra maneira para construir um serviço de diretório é utilizar componentes dc (domínio Internet), que refletem a estrutura do DNS. Atualmente, há preferência por esse modelo, pois ele facilita a localização de serviços de diretórios a partir do desmembramento do endereço de Internet, permitindo divisões de hierarquia ou regiões e possibilitando a replicação parcial da base e buscas mais rápidas. A **Figura 2** apresenta um exemplo de diretório baseado na estrutura dc, no qual os primeiros níveis (topo da árvore) representam o nome de domínio da empresa, na mesma ordem de leitura de um domínio DNS.

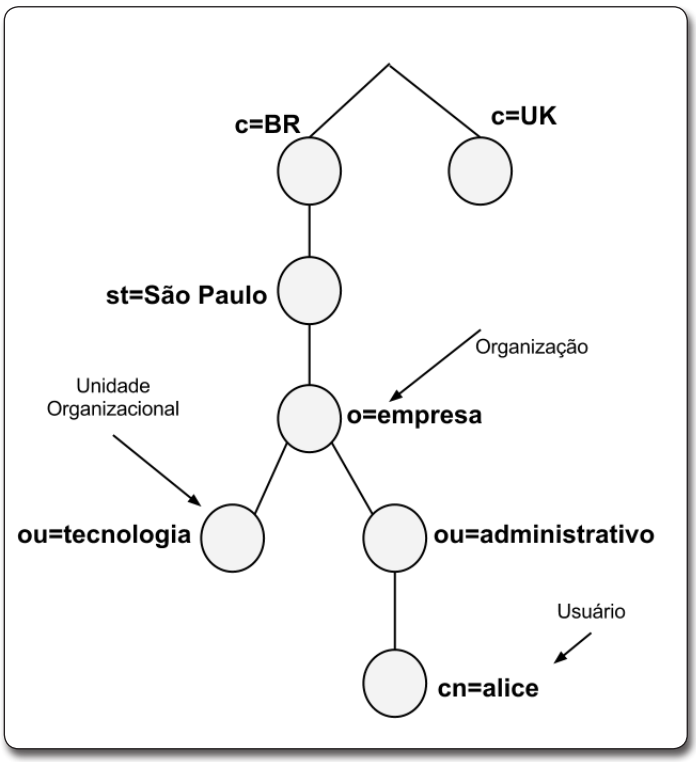


Figura 1. Exemplo de estrutura de diretório

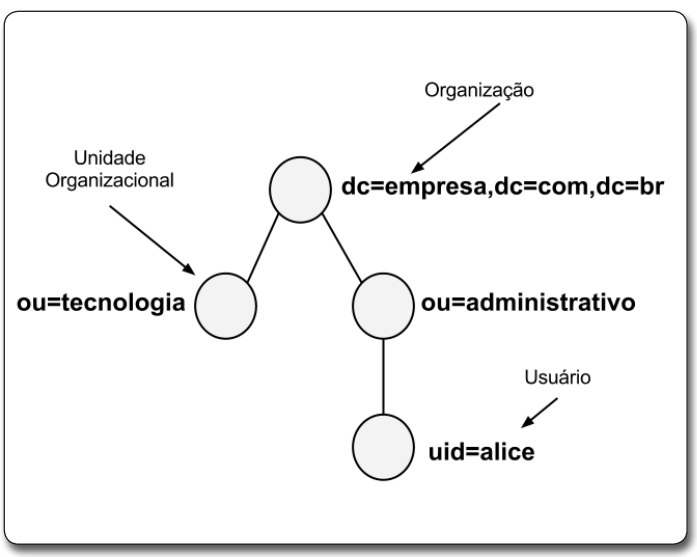


Figura 2. Exemplo de Diretório baseado em estrutura dc

Instalação

Nesta etapa do artigo, será apresentada a instalação dos pacotes utilizados para a configuração do servidor, que será baseado no Sistema Operacional Debian Squeeze, com OpenLDAP versão 2.4.3. Lembrando que, para executar os comandos que serão apresentados, é necessário realizar o login com o usuário root (administrador do sistema).

Para a instalação dos pacotes, será necessário configurar o arquivo */etc/apt/sources.list* com os repositórios externos. A **Listagem 1** apresenta um modelo do arquivo.

Os pacotes que serão instalados são o slapd e o ldap-utils. Os comandos apresentados na **Listagem 2** realizam a pesquisa dos pacotes nos mirrors configurados no */etc/apt/sources.list*, e em seguida é feita a instalação.

Durante a instalação, é solicitada a definição da senha do administrador, conforme demonstrado na **Figura 3**. Na tela seguinte a essa etapa, basta confirmar a senha escolhida.

Listagem 1. Modelo do arquivo /etc/apt/sources.list.

```
deb http://ftp.br.debian.org/debian/ squeeze main
deb-src http://ftp.br.debian.org/debian/ squeeze main
deb http://security.debian.org/ squeeze/updates main
deb-src http://security.debian.org/ squeeze/updates main
# squeeze-updates, previously known as 'volatile'
deb http://ftp.br.debian.org/debian/ squeeze-updates main
deb-src http://ftp.br.debian.org/debian/ squeeze-updates main
deb http://backports.debian.org/debian-backports squeeze-backports main
```

Listagem 2. Pesquisa e instalação dos pacotes.

```
# aptitude search slapd
# aptitude search ldap-utils
# aptitude install slapd ldap-utils
```

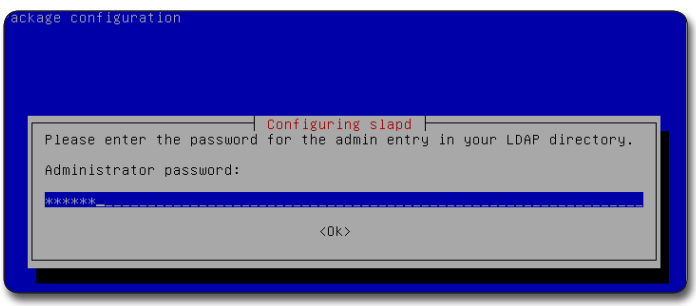


Figura 3. Solicitação da senha do administrador do LDAP

Dadas essas informações, os pacotes serão instalados. É uma boa prática verificar se estes foram instalados corretamente, assim como a versão do protocolo OpenLDAP, se o usuário e grupo utilizados pelo LDAP foram criados e se o serviço foi iniciado. A **Listagem 3** apresenta os comandos utilizados para realizar essa operação.

Configuração

Neste tópico será apresentada a configuração do serviço OpenLDAP, de uma forma tradicional, isto é, por meio do arquivo *slapd.conf*. Em versões mais recentes, o OpenLDAP não traz o arquivo *slapd.conf* dentro do diretório */etc/ldap/*, mas o diretório */etc/ldap/slapd.d/*, que contém uma base chamada *cn=config*. É nesta base onde são armazenadas as informações do serviço de diretório, bem como as informações dos schemas utilizados (identificados pelos arquivos com extensão .schema).

Listagem 3. Verificação dos pacotes, usuário e grupo e processo do serviço LDAP.

```
Pacotes instalados
# dpkg -l | egrep 'slapd|ldap-utils'
ii  ldap-utils                2.4.23-7.2       OpenLDAP utilities
ii  slapd                     2.4.23-7.2       OpenLDAP server (slapd)

Versão do protocolo OpenLDAP
# slapd -V
@(#) $OpenLDAP: slapd 2.4.23 (Jun 16 2011 02:53:39) $
     buildd@murphy:/build/buildd-openldap_2.4.23-7.2-i386-Y1mwvF/openldap-
2.4.23/debian/build/servers/slapd

Usuário e grupo utilizados pelo Ldap
# cat /etc/passwd | grep ldap
openldap:x:104:106:OpenLDAP Server Account,,,:/var/lib/ldap:/bin/false

# cat /etc/group | grep ldap
openldap:x:106:

Verificação do processo do serviço Ldap:
# ps -ef | grep ldap
openldap  1712    1  0 Nov10 ?       00:00:00 /usr/sbin/slapd -h ldap:/// ldapi:/// -g
openldap -u openldap -F /etc/ldap/slapd.d
root     1770  1126  0 00:17 pts/0    00:00:00 grep ldap
```

Ao utilizar o arquivo *cn=config*, a configuração passa a ser online, de modo que, caso seja necessário criar um novo índice ou uma ACL, não será preciso parar o serviço e reiniciá-lo, etapa necessária quando se utiliza o arquivo *slapd.conf*. Entretanto, deve-se observar que a inclusão de uma ACL incorreta, por exemplo, poderá causar indisponibilidade no serviço, pois apesar de ser imediatamente aplicada, a funcionalidade de remoção de entradas ainda não é contemplada na versão 2.4, sendo necessário reiniciar o serviço. Assim, nesse artigo, a configuração do OpenLDAP será realizada por meio do arquivo *slapd.conf*, ficando o *cn=config* como tema para um próximo artigo.

Antes de prosseguir com a configuração, será necessário reconfigurar o pacote slapd utilizando a ferramenta dpkg-reconfigure, de modo a ativar o LDAP com informações, tais como o domínio que será utilizado, tipo de banco que irá armazenar os dados, entre outras.

A **Listagem 4** apresenta os comandos necessários para realizar esse procedimento.

Listagem 4. Reconfigurando o pacote slapd.

```
Parando o serviço:
# /etc/init.d/slapd stop
Removendo os arquivos de base criados durante a instalação:
# rm -f /var/lib/ldap/*

Reconfigurando o pacote slapd:
# dpkg-reconfigure slapd
```

Ao executar o comando *dpkg-reconfigure*, serão solicitadas informações para a nova configuração do slapd. A primeira pergunta é se o administrador deseja omitir a configuração do servidor OpenLDAP. Caso seja escolhido Yes/Sim, nenhuma configuração inicial ou base de dados será criada. Como os arquivos de base foram removidos, nessa tela, a escolha deve ser No/Não, conforme apresentado na **Figura 4**.

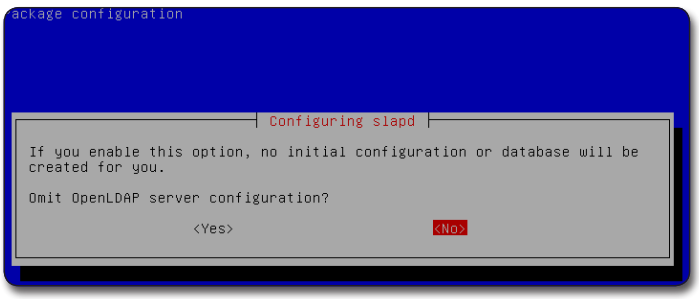


Figura 4. Reconfiguração do pacote OpenLDAP

A próxima pergunta se refere ao domínio que será a base DN (*Distinguished Name*) do diretório do LDAP. Ao preencher os dados dessa tela, é importante observar que deve ser informado o nome de um domínio, conforme ilustrado na **Figura 5**, e não uma entrada do LDAP como "dc=meudominio,dc=com". Ao inserir o domínio, o configurador automaticamente criará a entrada LDAP, desmembrando o que foi digitado em componentes de domínio (dc).

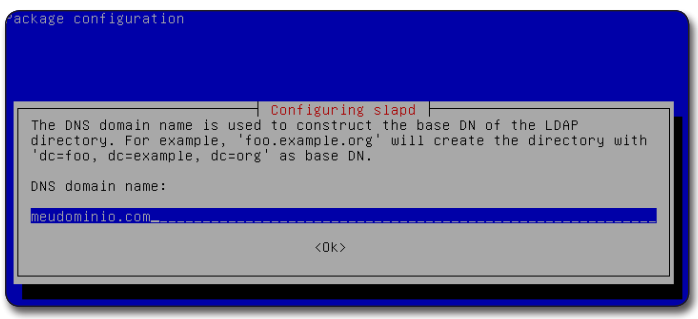


Figura 5. Definição do domínio OpenLDAP

As próximas telas se referem à definição do nome da organização que será utilizada no serviço de diretório, e à senha de root. Para a primeira pergunta, basta informar o nome, como "LDAP Article", por exemplo. Quanto à senha de root, esta se refere ao administrador do serviço de diretório e não do sistema operacional. Assim, é boa prática que a senha seja diferente da senha do SO e que contenha caracteres especiais, letras e números.

Após confirmar a senha escolhida, será solicitada a definição do tipo de banco de dados que será utilizado. Tanto o BDB quanto o HDB são *backends* para armazenamento de dados, com as mesmas opções de configuração. A diferença entre eles está no fato de que o BDB utiliza indexação e cache para acelerar o acesso aos dados enquanto HDB utiliza um banco de dados hierárquico e permite renomear objetos na base de dados. Como o configurador recomenda utilizar o HDB, basta selecionar essa opção conforme ilustra a **Figura 6**.

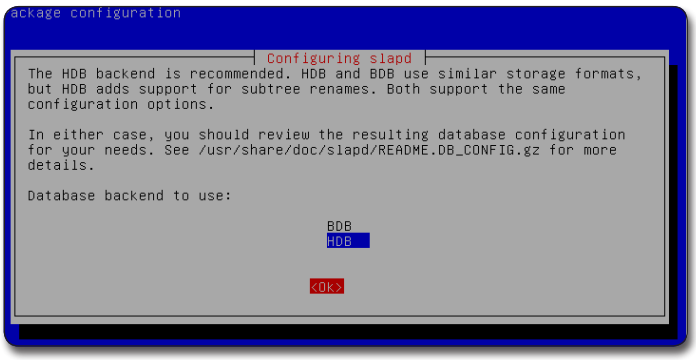


Figura 6. Configuração do banco que será utilizado pelo LDAP

A próxima pergunta está relacionada à desinstalação do pacote slapd. O administrador pode optar por remover completamente o pacote junto com a base de dados, ou remover somente o pacote, mantendo as informações da base. Isso quer dizer que ao utilizar o comando *aptitude purge slapd*, tanto o pacote como todo o conteúdo do banco de dados serão removidos, enquanto utilizar somente *aptitude remove slapd* desinstala somente o pacote, mantendo o conteúdo do banco. Nesse caso, selecione a opção No/Não.

Por fim, será solicitada a confirmação para deixar habilitado o protocolo LDAPv2, conforme apresentado na **Figura 7**. Selecione a opção Yes/Sim, para manter a compatibilidade com clientes que ainda não utilizam o LDAPv3.



Figura 7. Escolha da versão do protocolo utilizado

Ao finalizar a reconfiguração do pacote, a base de dados é recriada e o serviço é iniciado, porém ainda é necessário indicar que o OpenLDAP utilizará as definições do arquivo *slapd.conf* e não do *cn=config*. Esse procedimento é feito a partir da atualização da variável SLAPD_CONF e da cópia do arquivo para o diretório */etc/ldap*. A **Listagem 5** apresenta os comandos para realizar o procedimento.

Diretivas do arquivo de configuração

Conforme dito anteriormente, as configurações para a criação de um serviço de diretórios utilizando o OpenLDAP será feita a

partir do arquivo *sladp.conf*. Assim como a maior parte dos arquivos de configuração no Linux, este arquivo possui comentários que descrevem cada diretiva ou sua sintaxe. Para facilitar o entendimento das diretivas apresentadas, será feito um backup do arquivo original e criado um arquivo limpo, sem comentários. O comando utilizado para esse procedimento é o *egrep*, que removerá linhas que são iniciadas por um comentário (^#), linhas em branco (^$), e em seguida redirecionar a saída para o arquivo que será utilizado para configurar o OpenLDAP, conforme demonstrado na **Listagem 6**.

Listagem 5. Indicando o arquivo de configuração que será utilizado pelo OpenLDAP.

```
Atualização da variável SLAPD_CONF que indica o path do arquivo:
# vim /etc/default/slapd
SLAPD_CONF="/etc/ldap/slapd.conf"

Copiando o arquivo de configuração
# cp /usr/share/sladp/slapd.conf /etc/ldap

Verificando se a cópia foi realizada:
# ls -l /etc/ldap/slapd.conf
-rw-r--r-- 1 root root 4637 Nov 11 05:57 /etc/ldap/slapd.conf
```

Listagem 6. Retirada dos comentários do arquivo de configuração.

```
Cópia do arquivo original
# cp /etc/ldap/slapd.conf /etc/ldap/slapd.conf.original

Remoção dos comentários e linhas em branco
# egrep -v "^#|^$" /etc/ldap/slapd.conf.original > /etc/ldap/slapd.conf
```

Com o arquivo limpo, podemos verificar as diretivas e aplicar as configurações para disponibilizar um serviço de diretórios na rede. O editor de textos utilizado será o vim, porém o usuário poderá editar o arquivo a partir de outros editores de texto. A **Listagem 7** demonstra o arquivo limpo que será utilizado.

A seguir, são apresentadas as principais diretivas configuradas no arquivo *slapd.conf*:

• **allow bind_v2**: Refere-se ao uso da versão 2 do protocolo LDAP e não faz exigência de uso de certificados digitais. É útil manter essa opção para manter compatibilidade com clientes e serviços que utilizam somente essa versão. Caso essa linha não exista no arquivo, basta acrescentá-la;

• **include**: As linhas com essa diretiva estão relacionadas com os arquivos de schema, que são um conjunto de regras que define atributos, classes de objetos e controles, indicando onde cada dado pode ser armazenado. Em uma instalação padrão, os arquivos de schema são armazenados no diretório */etc/ldap/schema* e devem ser definidos linha a linha especificando o seu caminho completo, conforme demonstrado na **Listagem 8**.

Um diretório pode conter mais de um schema e alguns atributos são de preenchimento obrigatório. Por exemplo, ao utilizar o schema **inetOrgPerson**, os atributos **cn** (*common name*) e **sn** (*surname*) obrigatoriamente precisam conter algum valor, enquanto **uid** e **userpassword** (id de usuário e senha, respectivamente) são opcionais, não necessitando ser adicionados. A **Tabela 2** apresenta as especificações fornecidas com os schemas do arquivo *slapd.conf*;

Listagem 7. Conteúdo do arquivo slapd.conf.

```
# vim slapd.conf
allow bind_v2
include         /etc/ldap/schema/core.schema
include         /etc/ldap/schema/cosine.schema
include         /etc/ldap/schema/nis.schema
include         /etc/ldap/schema/inetorgperson.schema
pidfile         /var/run/slapd/slapd.pid
argsfile        /var/run/slapd/slapd.args
loglevel        192
modulepath      /usr/lib/ldap
moduleload      back_hdb
sizelimit       500
tool-threads    1
backend         hdb
access to attrs=userPassword,shadowLastChange
  by dn="cn=admin,dc=meudominio,dc=com" write
  by anonymous auth
  by self write
  by * none
access to dn.base="" by * read
access to *
  by dn="cn=admin,dc=meudominio,dc=com" write
  by * read
database        hdb
suffix      "dc=meudominio,dc=com"
directory   "/var/lib/ldap"
rootdn      "cn=admin,dc=meudominio,dc=com"
rootpw      "senhadificil"
dbconfig    set_cachesize 0 2097152 0
dbconfig    set_lk_max_objects 1500
dbconfig    set_lk_max_locks 1500
dbconfig    set_lk_max_lockers 1500
index       objectClass eq
lastmod         on
checkpoint 512 30
```

Listagem 8. Schemas inclusos no arquivo slapd.conf.

```
include      /etc/ldap/schema/core.schema
include      /etc/ldap/schema/cosine.schema
include      /etc/ldap/schema/nis.schema
include      /etc/ldap/schema/inetorgperson.schema
```

Schema	Descrição
core	Núcleo do OpenLDAP. Obrigatório para que o serviço funcione.
cosine	Contém diversos atributos e classes de objetos que dão suporte a outros schemas, devendo existir para que o schema nis funcione. Exemplos de atributos: autenticação Linux/Windows, dados de e-mail, informações pessoais de usuários tais como nome completo, número do telefone do escritório, ramal, entre outros.
nis	Contém a estrutura de atributos básica para armazenar informações de autenticação de usuários, grupos e hosts Linux/Unix.
inetorg-person	Especifica os atributos utilizados para armazenar informações sobre os usuários.

Tabela 2. Especificações dos schemas utilizados no arquivo slapd.conf

• **pidfile**: No Linux, todos os processos são identificados por um número denominado PID (*Process IDentification*). Essa diretiva indica o arquivo que contém o PID do processo em execução do OpenLDAP;
• **argsfile**: Arquivo que contém os argumentos passados ao servidor no momento de sua execução;
• **loglevel**: Define o nível de verbosidade do registro de atividades (*log*) no OpenLDAP. O nível 0 (*none*) determina que não será gerado log, porém, em caso de problema para iniciar o serviço, é interessante ativar um nível para que sejam gerados registros que podem ser consultados a fim de determinar o problema.

É possível utilizar depuração máxima (-1) ou combinar valores, isto é, somar os valores dos níveis a fim de obter as informações desejadas. Por exemplo, caso deseje habilitar o log para o processamento das configurações (nível 64) e também o processamento de listas ACL (128), basta somar os níveis: 128 + 64 = 192 e alterar a diretiva com esse valor:

```
loglevel 192
```

A **Tabela 3** apresenta alguns níveis de depuração que podem ser utilizados e combinados na diretiva *loglevel*;

Nível	Descrição
-1	Habilita depuração máxima
0	Sem informações de depuração
1	Chamada de funções trace (referentes à API)
2	Depuração e manipulação de pacotes
4	Depuração de funções
8	Gerenciamento de conexões
16	Exibe pacotes de rede recebidos e enviados
32	Processamento de filtros de pesquisa
64	Processamento das configurações
128	Processamento das ACLs
256	Estatísticas de conexões, operações e resultados

Tabela 3. Níveis de depuração da diretiva loglevel

• **modulepath**: Define o diretório dos módulos para os backends (bdb, hdb) utilizados para armazenar os dados do LDAP;
• **moduleload**: Ativa o módulo do backend que será utilizado;
• **sizelimit**: Define a quantidade máxima de entradas retornadas em uma consulta feita na base OpenLDAP. Por padrão, as pesquisas estão limitadas a 500 entradas. Para mais resultados, será necessário aumentar este valor, e ao definir -1, a consulta trará retornos ilimitados;
• **tool-threads**: Define o número de CPUs que serão utilizadas para a indexação da base de dados. Se o servidor tiver mais de um processador, aumente este número;
• **backend**: Permite configurar as opções específicas de cada backend, que nesse artigo é o HDB;
• **access to**: Listas de controle de acesso (ACLs) que definem o que os usuários podem ou não fazer dentro da estrutura do LDAP através de diferentes níveis de acesso. A **Tabela 4** apresenta os privilégios que podem ser concedidos para o usuário.

Nível de acesso	Descrição
None	Sem acesso
Auth	Acesso para autenticação
Compare	Acesso para comparar atributos
Search	Acesso para pesquisas
Read	Acesso para leitura dos resultados das pesquisas
Write	Acesso para modificação de dados dos atributos

Tabela 4. Níveis de acesso para construção de ACLs no LDAP

As ACLs oferecem um controle granular no gerenciamento de identidades no que diz respeito à integração entre os serviços, porém deve-se observar que a configuração deve ser bem calculada para evitar que, ao negar um acesso, esta ação não seja sobrescrita pela próxima regra. A **Listagem 9** contempla as ACLs configuradas no *slapd.conf* e, para efeito de distinção entre uma regra e outra, estão descritas como ACL1, ACL2 e ACL3.

Listagem 9. ACLs configuradas no arquivo slapd.conf.

```
# ACL1
access to attrs=userPassword,shadowLastChange
by dn="cn=admin,dc=meudominio,dc=com" write
by anonymous auth
by self write
by * none

# ACL2
access to dn.base="" by * read

# ACL3
access to *
by dn="cn=admin,dc=meudominio,dc=com" write
by * read
```

A ACL1 se refere às permissões referentes à autenticação (*userPassword*) e alteração de senha (*shadowLastChange*), o que garante os seguintes direitos de acesso a esses atributos:
• Escrita e leitura ao usuário admin:

```
by dn="cn=admin,dc=meudominio,dc=com" write
```

• Autenticação aos usuários anônimos:

```
by anonymous auth
```

• Escrita e leitura aos usuários autenticados (apenas nos dados retornados em suas próprias entradas), permitindo que estes alterem a própria senha:

```
by self write
```

• Para os demais usuários, o acesso é negado:

```
by * none
```

A ACL2 dá direito de leitura à raiz do diretório configurado (dc=meudominio,dc=com):

```
access to dn.base="" by * read
```

No exemplo apresentado, a ACL1 é mais restritiva e impede que a ACL2 dê privilégio de leitura de senhas aos usuários comuns; caso a ordem estivesse invertida, a primeira ACL não teria efeito. E, por fim, a ACL3 garante direito de acesso a todo o diretório para escrita e leitura ao usuário admin e apenas leitura aos demais usuários.

Dando continuidade na descrição das diretivas do *slapd.conf*, ainda temos:

• **database**: Ativa a base de dados escolhida (HDB);

• **suffix**: Campo da estrutura da base LDAP (no exemplo, dc=meudominio,dc=com), que define sob qual hierarquia os dados do serviço de diretório serão armazenados;

• **directory**: Local (*path*) onde os dados do serviço de diretório serão armazenados;

• **rootdn**: Define o usuário que será administrador da base LDAP. Pode ser definido um usuário com o nome Manager ou Admin, uma vez que o usuário root do sistema operacional não é o administrador da base;

• **rootpw**: Define a senha de acesso ao servidor LDAP. Na configuração padrão, a senha é colocada em texto puro no arquivo, porém, por medida de segurança, utiliza-se a ferramenta *slappasswd* (demonstrada mais a frente) para gerar a senha criptografada que será colocada no arquivo;

• **dbconfig**: Especifica parâmetros para configuração do backend utilizado, tais como limite do cache da base em memória, número de objetos que podem ser travados simultaneamente, número de travas (locks) solicitadas e ativadas, número de travamentos (lockers) ativos, etc. Na configuração do limite do armazenamento dos dados em memória (*set_cachesize*), o padrão é 2M, podendo ser alterado caso o servidor tenha memória RAM disponível;

• **index**: Define para quais atributos o slapd deve manter índices de forma a melhorar a busca. A indexação da base deve ser configurada de acordo com o tipo de pesquisa que será feita e as possibilidades de filtros a serem aplicados. A **Tabela 5** apresenta os tipos de indexação ou o tipo de pesquisa a ser realizada;

• **lastmod**: Grava informações de tempo (*timestamp*) em cada modificação e criação de entradas no diretório, permitindo que o cliente verifique quando a informação foi atualizada;

• **checkpoint**: Define quando os buffers da base de dados HDB devem ser gravados no disco, para evitar possíveis perdas. É fornecido um limite em kbytes e outro em minutos, sendo que o primeiro a vencer dispara o checkpoint e grava uma entrada no arquivo de log para registrar o evento.

Indexador	Descrição
eq (equality)	Indexa a informação de acordo com o par exato dos valores dos atributos.
sub (substring)	Indexa a informação para permitir pesquisas por substrings (trechos) dentro dos valores dos atributos.
pres (presence)	Indexa o atributo pela presença de valor, isto é, a busca só terá retorno se o atributo contiver informação.
approx (approximate)	Indexa a informação de acordo com uma combinação aproximada dos valores dos atributos.
none	Não realiza nenhuma indexação.

Tabela 5. Tipos de indexação da base LDAP

Inicializando o servidor de diretório OpenLDAP

Após realizar as configurações, o próximo passo é modificar a permissão do arquivo *slapd.conf*, alterar a propriedade do diretório */etc/ldap* e verificar a sintaxe do arquivo de configuração.

A alteração da permissão do arquivo de configuração visa proteger a senha que foi definida para o administrador da base. Uma vez que durante a instalação os comandos são executados com o usuário root, os arquivos de configuração herdam as permissões desse usuário e se faz necessário utilizar o comando *chown* para alterar para o usuário e grupo **openldap**, utilizados pelo LDAP para iniciar o serviço. A opção *–R* no comando faz com que sejam alterados todos os arquivos e diretórios abaixo do */etc/ldap*, conforme apresentado na **Listagem 10**.

Listagem 10. Alteração do usuário e grupo do diretório /etc/ldap.

```
Alteração da permissão do arquivo slapd.conf
# chmod 600 slapd.conf

Alteração do usuário e grupo para openldap
# chown -R openlap:openldap /etc/ldap

Lista dos arquivos após a modificação:
# ls -l
total 28
-rw-r--r-- 1 openldap openldap  245 Jun 16  2011 ldap.conf
drwxr-xr-x 2 openldap openldap 4096 Jun 16  2011 sasl2
drwxr-xr-x 2 openldap openldap 4096 Nov 11 05:48 schema
-rw------- 1 openldap openldap 1078 Nov 16 05:13 slapd.conf
-rw-r--r—1 openldap openldap 4637 Nov 11 06:37 slapd.conf.original
```

Por fim, é hora de checar a sintaxe e iniciar o serviço. O OpenLDAP disponibiliza diversas ferramentas para a administração do serviço, e para checar a sintaxe do arquivo *slapd.conf*, será utilizado o comando *slaptest*, que verifica possíveis erros dentro do arquivo de configuração. Sua utilização é demonstrada a seguir:

```
# slaptest -f /etc/ldap/slapd.conf -F /etc/ldap
```

As opções do comando *slaptest* são:

• **-f** – utilizado para indicar o arquivo de configuração (*configfile*);

• **-F** – utilizado para indicar o diretório em que essas configurações serão reescritas sempre que forem alteradas (*configdir*).

Caso não haja nenhum erro de sintaxe, o comando exibe a mensagem *config file testing succeeded*, caso contrário, uma mensagem de erro é exibida e deve-se verificar no arquivo a sintaxe ou sequência para a correção. Em caso de sucesso, basta reiniciar o serviço conforme comando a seguir:

```
# /etc/init.d/slapd restart
```

Incluindo registros na base de dados

Diferentemente de um banco de dados, o LDAP não possui um gerenciador que pode ser acionado para ações de manipulação de dados, tais como adição, modificação ou exclusão. Os gerenciadores LDAP existentes não trabalham de forma online, isto é, a modificação de dados é feita por meio de um arquivo texto que contém as informações a serem modificadas. Este arquivo possui um formato específico chamado LDIF (*LDAP Data Interchange Format*), que define os dados necessários para o cadastro de cada tipo de registro.

A seguir, serão apresentados os arquivos LDIF e os comandos necessários para criar um serviço de diretório. Para adição de uma entrada para um domínio, usuário ou grupo, deve-se informar os dados referentes a cada um deles e também quais dados farão parte da entrada, utilizando o **objectClass**, ou classe de objeto, que é um registro que identifica um conjunto de atributos agrupado por tipo, isto é, determina as características opcionais e obrigatórias que um objeto em particular pode ou deve possuir. Por exemplo, durante a criação de um usuário que irá se autenticar no Linux ou Windows, o **objectClass** irá determinar quais itens são obrigatórios ou opcionais para criação da conta.

As definições de cada classe de objeto e cada atributo estão nos arquivos .schema localizados em */etc/ldap/schema* e são responsáveis pela padronização e manutenção da estrutura do Diretório. Por exemplo, para cadastrar a organização, utiliza-se o atributo **organization**; unidade organizacional, utiliza-se o atributo **organizationalUnit**; e para usuários de um modo geral, utiliza-se **inetOrgPerson**.

A **Listagem 11** demonstra a criação do arquivo *meudominio.ldif*, que irá conter os dados básicos, isto é, a estrutura para o topo da árvore do diretório (**organization**) para o domínio *meudominio.com*.

Listagem 11. Criação do arquivo meudominio.ldif.

```
# vim /etc/ldap/meudominio.ldif
dn: dc=meudominio,dc=com
objectClass: dcObject
objectClass: organization
o: LDAP Article
dc: meudominio
```

Após criar o arquivo, é necessário informar ao servidor LDAP onde estão os dados que devem ser inseridos na base. O comando utilizado para essa ação é o *ldapadd*, o qual abre uma conexão com o servidor e autentica o usuário administrador ou outro que tenha permissão de escrita (edição) na base. Na mesma linha, indica-se como argumento o arquivo LDIF criado, que será utilizado para que as entradas sejam adicionadas no diretório, conforme demonstrado na **Listagem 12**.

A **Tabela 6** apresenta os parâmetros utilizados no comando *ldapadd* com uma breve descrição.

Listagem 12. Adicionando o primeiro registro na base.

```
ldapadd -h localhost -W -x -D "cn=admin,dc=meudominio,dc=com" -f /etc/ldap/
meudominio.ldif
Enter LDAP Password:
adding new entry "dc=meudominio,dc=com"
```

Parâmetro	Descrição
-h	Define qual servidor LDAP será utilizado.
-W	Solicita a senha para executar o comando.
-x	Opção utilizada para fazer uma autenticação simples (sem SASL).
-D	Indica o usuário que realizará a operação.
-f	Indica o arquivo fornecido como argumento.

Tabela 6. Parâmetros utilizados no comando ldapadd

Uma vez que a estrutura do topo da árvore do diretório foi criada, é possível inserir dados na base. No entanto, como visto no início deste artigo, um serviço de diretório possibilita organização e hierarquia. Dessa forma, será necessário criar as entradas das unidades organizacionais (*ou*), que irão agrupar os dados cadastrados de forma ordenada e coerente. Para o modelo de Diretório utilizado neste artigo, serão criadas duas *ou*, sendo que a *ou* denominada *Pessoa* irá conter a identificação dos usuários e a *ou Grupos* irá conter grupos. Essa estrutura é representada na **Figura 8**.

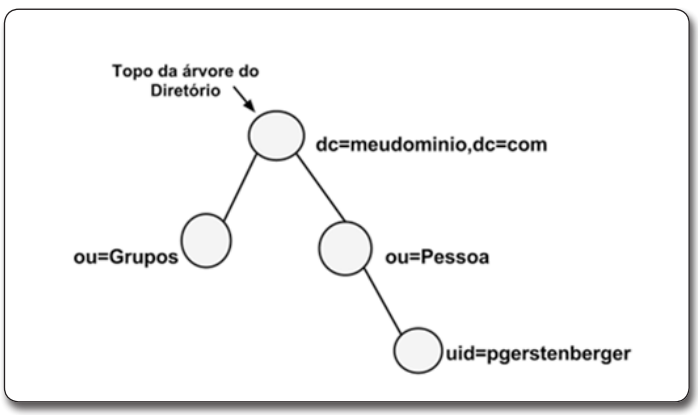


Figura 8. Estrutura do modelo de Diretório a ser implementado

Com o modelo de Diretório que será utilizado, a próxima etapa é criar o arquivo *unidades_ou.ldif* que irá conter os dados das unidades organizacionais Pessoa e Grupos, conforme o conteúdo da **Listagem 13**.

Listagem 13. Criação do arquivo unidades_ou.ldif.

```
dn: ou=Pessoa,dc=meudominio,dc=com
ou: Pessoa
objectClass: top
objectClass: organizationalUnit

dn: ou=Grupos,dc=meudominio,dc=com
ou: Grupos
objectClass: top
objectClass: organizationalUnit
```

Após criar esse arquivo, é necessário incluir as unidades organizacionais no Diretório utilizando o comando *ldapadd*. A **Listagem 14** demonstra a sintaxe para executar essa ação, sendo que o padrão é o mesmo que foi utilizado para criar o topo da árvore, apenas modificando o nome do arquivo que será fornecido como argumento para leitura dos dados que serão cadastrados.

Listagem 14. Adicionando unidades organizacionais no Diretório.

```
# ldapadd -x -h localhost -D "cn=admin,dc=meudominio,dc=com" –W -f /etc/ldap/
unidades_ou.ldif
Enter LDAP Password:
adding new entry "ou=Pessoa,dc=meudominio,dc=com"

adding new entry "ou=Grupos,dc=meudominio,dc=com"
```

Até este ponto foram criadas três entradas no Diretório: uma referente ao topo da árvore, representada pelo domínio **dc=meudominio,dc=com**; a segunda referente à unidade organizacional ou=Pessoa; e a terceira referente à unidade **organizacional ou=Grupos**. Seguindo o modelo da estrutura de Diretório apresentada na **Figura 8**, o próximo passo é criar uma entrada de usuário utilizando uma classe de objeto que contemple os atributos que se deseja disponibilizar.

Para esse exemplo, serão utilizados os atributos da classe de objetos **inetOrgPerson**, localizada no diretório */etc/ldap/schema*. Essa classe é estrutural e não necessita de classes adicionais para que a entrada seja criada, além de definir atributos essenciais para que um registro de usuários associados à organização possa ser criado. A **Listagem 15** apresenta o arquivo da entrada de dados para registro do usuário e o comando para incluir este no Diretório.

Com isso, temos uma estrutura inicial da base LDAP. Agora veremos como realizar as pesquisas para verificar os registros cadastrados.

Consultas

Uma vez que os dados foram cadastrados no Diretório, é necessário realizar a pesquisa para verificar as informações. O comando *ldapsearch* é utilizado para realizar as buscas dentro das bases LDAP e pode ser utilizado de diversas formas, sendo que a mais comum é a pesquisa anônima, isto é, a pesquisa que não exige a identificação do solicitante dos dados. A **Listagem 16** apresenta o resultado de uma busca anônima.

Listagem 15. Arquivo LDIF que contém a entrada de dados do usuário.

```
# vim /etc/ldap/pessoa.ldif
dn: cn=Pathiene Gerstenberger,ou=Pessoa,dc=meudominio,dc=com
objectClass: inetOrgPerson
cn: Pathiene Gerstenberger
givenName: Pathiene
sn: Gerstenberger
uid: pgerstenberger
userPassword: peth1234
mail: pathigerst@gmail.com
description: OpenLDAP

Sintaxe do comando ldapadd para incluir usuário no Diretório
# ldapadd -h localhost -W -x -D "cn=admin,dc=meudominio,dc=com" -f /etc/ldap/
pessoa.ldif
Enter LDAP Password:
adding new entry "cn=Pathiene Gerstenberger,ou=Pessoa,dc=meudominio,dc=com"
```

Listagem 16. Realizando uma busca anônima.

```
# ldapsearch –x
# extended LDIF
#
# LDAPv3
# base <> (default) with scope subtree
# filter: (objectclass=*)
# requesting: ALL
#

# search result
search: 2
result: 32 No such object

# numResponses: 1
```

A **Tabela 7** apresenta os principais parâmetros que o comando *ldapsearch* aceita.

A **Listagem 17** apresenta um exemplo de como buscar os dados do usuário pgerstenberger cadastrado no exemplo de criação de usuário no LDAP.

Autenticação do Squid no OpenLDAP

Na Edição 8 da Infra Magazine, foi apresentado como instalar o servidor Proxy com Squid e retirar relatórios dos acessos à Internet utilizando a ferramenta Sarg. Com o objetivo de aperfeiçoar a instalação do OpenLDAP, este será integrado com o Squid para que os usuários utilizem seus cadastros para fazer login no Squid, permitindo a integração e o controle dos acessos na rede.

Para configurar o Squid com autenticação no LDAP, é necessário indicar no arquivo *squid.conf* que o parâmetro *squid_auth_param* será utilizado para que os usuários sejam autenticados na rede. Porém, antes de antes de iniciar a configuração, deve ser feito um teste em linha de comando para verificar se o Squid já tem conexão com a base LDAP. O comando usado para isso é:

```
# /usr/lib/squid3/squid_ldap_auth -b "dc=meudominio,dc=com" -f "uid=%s" -h 127.0.0.1
```

Opção	Descrição
-x	Autenticação simples, sem utilização de senha.
-h	Host onde o OpenLDAP está rodando. Caso seja omitido, assume o valor que está no arquivo /etc/ldap/ldap.conf.
-b	Base que será utilizada na pesquisa. Caso seja omitida, assume o valor que está no arquivo /etc/ldap/ldap.conf.
-L	Determina o formato de saída (com detalhes ou não) da pesquisa. Se inserido uma vez (-L), resulta no formato LDIFv1; duas vezes (-LL), não exibe comentários; três vezes (-LLL), não exibe também o número da versão. Caso seja omitido, o resultado é exibido em formato estendido.
<filtro>	Condição de pesquisa. Caso seja omitido, assume o filtro padrão 'objectClass=*'.
<atributo>	Atributos que devem ser retornados na pesquisa.

Tabela 7. Parâmetros aceitos pelo ldapsearch

Listagem 17. Exemplo de pesquisa por usuário (uid).

```
# ldapsearch -h localhost -x -b ou=Pessoa,dc=meudominio,dc=com
uid=pgerstenberger
# extended LDIF
#
# LDAPv3
# base <ou=Pessoa,dc=meudominio,dc=com> with scope subtree
# filter: uid=pgerstenberger
# requesting: ALL
#

# Pathiene Gerstenberger, Pessoa, meudominio.com
dn: cn=Pathiene Gerstenberger,ou=Pessoa,dc=meudominio,dc=com
objectClass: inetOrgPerson
cn: Pathiene Gerstenberger
givenName: Pathiene
sn: Gerstenberger
uid: pgerstenberger
mail: pathigerst@gmail.com
description: OpenLDAP

# search result
search: 2
result: 0 Success

# numResponses: 2
# numEntries: 1
```

Onde:

• **-b**: Define o contexto da busca. Em alguns lugares é chamado de "ramo de árvore" porque é onde se define a localização que os usuários serão procurados para se autenticar. No exemplo, foi definido todo o domínio; assim, será feita uma busca completa na árvore LDAP e em toda a sua estrutura;

• **-f**: Define o que será considerado na autenticação. No exemplo, será o uid do usuário;

• **- h**: Indica a localização do servidor OpenLDAP.

Ao executar o comando e teclar <ENTER>, o prompt continuará aberto aguardando a autenticação. Conforme ilustrado na **Figura 9**, deve-se inserir o nome do usuário seguido da senha. Se a conexão for bem sucedida, haverá um retorno OK, caso contrário, haverá o retorno ERR, que pode vir seguido de uma mensagem de erro.

Como o resultado do teste de conexão e autenticação foi bem sucedido, o próximo passo é incluir as linhas apresentadas na **Listagem 18** no arquivo */etc/squid/squid.conf* para indicar que a partir desse momento a autenticação no Squid será realizada pelo LDAP.

As linhas adicionadas no arquivo são:

• **auth_param basic program /usr/lib/squid3/squid_ldap_auth -b "dc=meudominio,dc=com" -f "uid=%s" -h localhost**: Linha que define a conexão do Squid com o LDAP. O parâmetro **auth_param basic program** é utilizado para indicar a lib que será usada para a autenticação dos serviços, que no caso está em */usr/lib/squid3/squid_ldap_auth*, sendo obrigatório adicionar o path completo do arquivo. O parâmetro **uid=%s** foi utilizado no arquivo *squid.conf* para substituir o nome do usuário (pathigerst) que foi inserido no momento da autenticação;

• **auth_param basic children 5**: Especifica a quantidade de processos filhos que serão iniciados durante a autenticação;

• **auth_param basic realm – SQUID AUTENTICANDO NO LDAP:** Determina a mensagem que será apresentada ao usuário quando a autenticação for solicitada;

• **auth_param basic credentialsttl 1 hour:** Informa o tempo de duração que o login será mantido ativo.

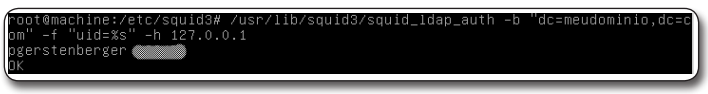


Figura 9. Resultado do teste de conexão do Squid com o servidor LDAP

Listagem 18. Linhas a serem adicionadas no arquivo squid.conf.

```
auth_param basic program /usr/lib/squid3/squid_ldap_auth -b
"dc=meudominio,dc=com" -f "uid=%s" -h localhost
auth_param basic children 5
auth_param basic realm - SQUID AUTENTICANDO NO LDAP
auth_param basic credentialsttl 1 hour
```

Também no arquivo *squid.conf* deve ser configurada a ACL para que as regras, isto é, as linhas que foram adicionadas anteriormente referentes à autenticação, sejam aplicadas, conforme demonstrado na **Listagem 19**.

Listagem 19. ACL para aplicar a autenticação no Squid via LDAP.

```
acl ldap-auth proxy_auth
http_access allow ldap-auth
```

Onde:
• **acl ldap-auth proxy_auth**: Define que a autenticação via proxy por meio do LDAP é obrigatória;
• **http_access allow ldap-auth**: Libera os clientes (browsers) a acessarem o serviço HTTP por meio da autenticação via LDAP.

É importante observar que as linhas mencionadas devem estar antes da linha **http_access deny all**, que bloqueia todas as requisições. Caso a autenticação via LDAP seja inserida após essa linha dentro do *squid.conf*, ela será ignorada.

Com as configurações definidas, é necessário salvar e sair do arquivo de configuração *squid.conf* e carregar as alterações, conforme apresentado na **Listagem 20**.

Listagem 20. Carregando as alterações no squid.conf.

```
# squid -k parse
# squid -k reconfigure
# /etc/init.d/squid restart
```

A primeira linha analisa se as configurações do Squid estão corretas. Em seguida, o comando *squid –k reconfigure* faz com que as alterações no arquivo sejam consideradas. Por fim, o serviço é reiniciado.

Configuração do navegador

Após a configuração do servidor, a próxima etapa é configurar os navegadores das estações clientes para autenticar no LDAP e navegar utilizando o proxy Squid. O exemplo apresentado a seguir é baseado no Firefox e os passos para esse procedimento consistem em acessar o menu *Edit* > *Preferences* e, em seguida, clicar nas opções *Advanced* > *Network* > *Connection* > *Settings*, conforme apresentado na **Figura 10**.



Figura 10. Configurando o Firefox para autenticar no Squid (Etapa 1)

Feito isso, será aberta uma tela para indicar o servidor proxy que será utilizado. Conforme apresentado na **Figura 11**, clique em *Manual proxy configuration*, informe o IP do servidor Squid e a porta de acesso. Em seguida, marque a opção *Use this proxy server for all protocols* e, para finalizar, clique em *OK* em todas as telas.

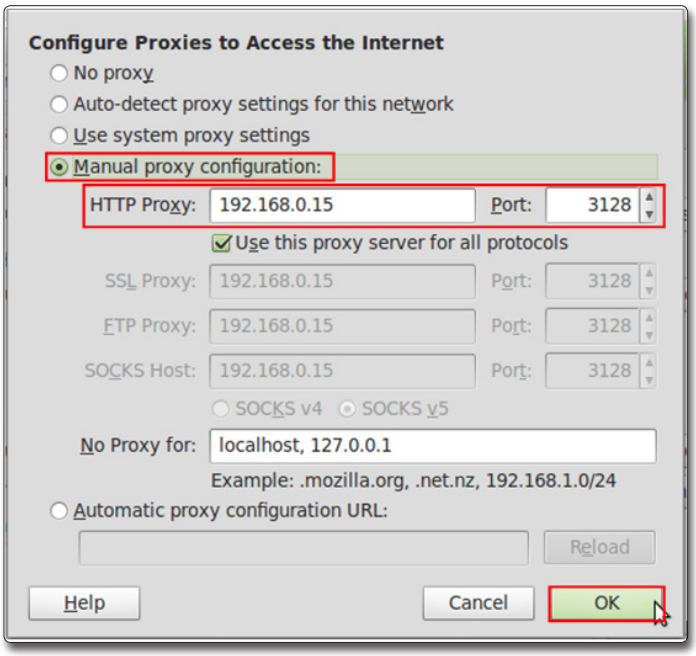


Figura 11. Configurando o Firefox para autenticar no Squid (Etapa 2)

Ao tentar navegar, uma tela de autenticação aparecerá, onde usuário deverá informar seu login e senha, conforme ilustrado na **Figura 12**.

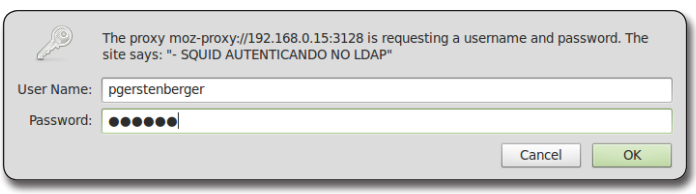


Figura 12. Tela de autenticação para navegação

No servidor onde estão configurados o OpenLDAP e o Squid, é possível ver essa autenticação no log, onde entre as informações de acesso também é mostrado o uid do usuário que está navegando. O administrador pode acompanhar o log de acesso utilizando o comando *tail*. A **Figura 13** ilustra o trecho do log em tempo real.

```
# tail -f /var/log/squid/access.log
```

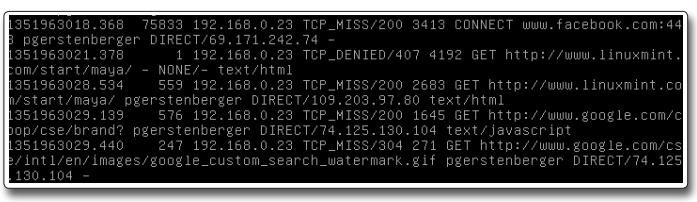


Figura 13. Log de acesso do Squid

Dica de Segurança

Durante a configuração do servidor OpenLDAP, foi definida uma senha simples e deixada em texto claro dentro do arquivo *slapd.conf*, o que pode causar um incidente de segurança, uma

vez que se outro usuário que não o administrador tiver acesso ao arquivo, terá acesso total à base. Pensando nisso, a suíte OpenLDAP oferece o comando *slappasswd*, utilizado para gerar senhas criptografadas que podem ser colocadas em arquivos como *slapd.conf* ou os arquivos com extensão .ldif, impossibilitando a leitura da senha dos usuários.

A **Tabela 8** apresenta os principais parâmetros utilizados para o comando *slappasswd*.

Parâmetro	Descrição
-h (hash)	Define o tipo do algoritmo de criptografia que será utilizado para gerar a senha. As opções aceitas nesse parâmetro são CRYPT, MD5 e SHA, sendo que este último é o padrão.
-v (verbose)	Habilita o modo verbose.

Tabela 8. Parâmetros do comando slappasswd

Conclusão

O protocolo LDAP tem sido cada vez mais utilizado por administradores em razão da facilidade de integração de ambientes heterogêneos, assim como pela sua estrutura leve. A utilização de programas de código aberto (OpenLDAP + Squid) apresenta uma solução de baixo custo com farta documentação encontrada na Internet e em publicações comerciais. Dentro deste contexto, o OpenLDAP pode ser utilizado para gerenciar identidades e integrar informações para que usuários de rede e aplicações tenham acesso aos recursos de maneira padronizada.

No ambiente corporativo, é comum encontrar áreas, departamentos ou até gerências responsáveis por cuidar das informações sobre os funcionários, porém muitas vezes não existe uma integração dessa área com aquela responsável pela administração do controle de acesso aos recursos computacionais da empresa.

Nesse contexto, se faz necessária uma solução que integre as bases de autenticação com os serviços de rede, garantindo que, além da validação do usuário, apenas as aplicações e recursos relacionados com a sua área de atuação estejam disponíveis.

Assim, com o OpenLDAP, é possível configurar um serviço de diretório em uma organização sem custo de licenças, além de ficar a critério do usuário a possibilidade de configurá-lo ou desenvolvê-lo da forma que melhor se adaptar às suas necessidades, seja para autenticar os usuários nos serviços de rede como o proxy demonstrado neste artigo, seja para qualquer outra finalidade na qual seja necessário armazenar e consultar informações de modo centralizado.



Ivani Nascimento
ivani.araujo@gmail.com
Atua no ramo de tecnologia e infraestrutura desde 2006, escreveu artigos para a revista EasyLinux, ministrou mais de 200 horas/aula de formação Linux. Graduada em Tecnologia de Redes de Computadores e Especialista em Segurança da Informação pelo ITA, atua como Analista Linux em uma empresa do mercado financeiro.



Pathiene Gerstenberger
pathigerst@gmail.com
Atua no ramo de tecnologia e infraestrutura desde 2006, escreveu artigos para a EasyLinux e ministrou mais de 200 horas/aula de formação Linux. Bacharel em Ciências da Computação, atualmente cursa MBA em Empreendedorismo e é certificada LPI1-C e NCLA, atua como Analista Linux em uma empresa de compras coletivas.

SUNGALIA, M. Autenticação Centralizada com OpenLDAP – Integrando serviços de forma simples e rápida. Novatec Editora. São Paulo. 2008.

The OpenLDAP Project – Site Oficial.
http://www.openldap.org/

Squid – Site Oficial.
http://www.squid-cache.org/

RFC 1487 – X.500 Lightweight Directory Access Protocol.
http://www.ietf.org/rfc/rfc1487.txt

OpenLDAP Release Roadmap.
http://www.openldap.org/software/roadmap.html

RFC 4510 – Lightweight Directory Access Protocol (LDAP): Technical Specification Road Map.
http://tools.ietf.org/html/rfc4510

OpenLDAP Software 2.4 Administrator's Guide.
http://www.openldap.org/doc/admin24/

Análise do tráfego de rede com o Wireshark

Como melhorar a gerência da rede utilizando o famoso analisador de pacotes?

A rápida difusão do acesso à Internet em diversos dispositivos da rede aumenta a complexidade das atuais redes de comunicação. Consequentemente, a gerência de tais redes também se torna problemática, exigindo dos administradores o conhecimento detalhado de seu ambiente.

A própria experiência adquirida com o trabalho e o uso de ferramentas específicas auxilia a realização dessa complexa tarefa. O uso de ferramentas como Nagios e MRTG, além da possibilidade da coleta de informações utilizando o protocolo SNMP, são alguns exemplos de como a monitoração constante da rede pode ser realizada. Porém, mesmo com a monitoração permanente da rede, determinados problemas somente podem ser solucionados mediante uma investigação detalhada nos pacotes que trafegam na rede. Os analisadores de tráfego são as ferramentas responsáveis por essa investigação.

Wireshark é um software criado com esse propósito, analisar o tráfego de rede. Tais softwares também são conhecidos como analisadores de pacotes ou ainda "sniffers" de rede (ver **Nota do DevMan 1**). Devido a sua eficiência e possibilidade de funcionamento em várias plataformas, o Wireshark se tornou o principal software do gênero.

Os principais objetivos das ferramentas que analisam o tráfego de rede são: monitorar e detectar comportamento anômalo nos diversos equipamentos que compõem uma rede de comunicação. Porém, o uso eficiente de tais ferramentas geralmente requer prática e conhecimento dos protocolos de rede, além da compreensão de como protocolos específicos são usados por diferentes aplicações. Neste artigo, será mostrado um breve histórico sobre o Wireshark, o funcionamento básico e a sua dinâmica de uso. Além disso, serão apresentados alguns exemplos de aplicações práticas dessa ferramenta, com o objetivo de melhorar a gerência e administração de redes.

Resumo DevMan

De que se trata o artigo:

O Wireshark é uma ferramenta de captura de tráfego de rede sob licença GNU GPLv2 que permite visualizar das mais variadas formas os pacotes que trafegam por uma determinada interface de rede. Uma das características da ferramenta é a sua interface gráfica, que permite ao usuário visualizar e selecionar o tipo de tráfego desejado. Devido a sua eficiência e possibilidade de funcionamento em diferentes sistemas operacionais, o Wireshark se tornou o principal software do gênero. Dessa forma, este artigo tem por objetivo descrever as principais características do Wireshark, mostrar o funcionamento básico da ferramenta e fornecer exemplos sobre as diversas aplicações práticas do Wireshark.

Em que situação o tema é útil:

Os pontos destacados neste artigo serão úteis para a compreensão do funcionamento e importância dos analisadores de pacotes, especialmente o Wireshark. Os conceitos e exemplos expostos no texto também permitirão a familiarização teórica e prática do tema pelos profissionais e usuários da área de infraestrutura de tecnologia da informação.

Análise do tráfego de rede com o Wireshark:

Neste artigo serão apresentadas as principais características da ferramenta de captura de pacotes Wireshark. Serão descritos: o funcionamento básico de um analisador de pacotes, os três módulos principais que compõem a arquitetura do Wireshark e os primeiros passos para o uso da ferramenta. Também será mostrado um breve exemplo da captura de pacotes TCP e HTTP usando filtros. Por fim, serão expostas algumas aplicações práticas em que o Wireshark pode ser utilizado com o intuito de melhorar a gerência e administração das redes de computadores.

Breve história sobre o Wireshark

No final de 1997, o estudante de ciência de computação da University of Missouri – Kansas City, Gerald Combs, trabalhava para um pequeno provedor de Internet. Problemas de rede em

tais provedores eram comuns e Gerald buscava ferramentas que pudessem lhe auxiliar na resolução dos mesmos. Porém, naquela época, produtos para análise de protocolos de rede eram caros (custavam em média $1500) e não funcionavam nos sistemas operacionais de sua companhia (Solaris e Linux). Dessa forma, com o intuito de rastrear os problemas ocorridos nos sistemas e aperfeiçoar seus conhecimentos em redes de computadores, Gerald começou a escrever seu próprio analisador de tráfego de rede, batizando o software de Ethereal.



Nota do DevMan 1

Sniffer: A origem do termo "network sniffer" está relacionada ao termo "sniffer", que na língua inglesa, diz respeito aos dispositivos utilizados para detectar substâncias ocultas através do odor, tais como drogas ou explosivos. O termo "network sniffer", em inglês, significa algo como "capturador de rede". Também é comum, por esse motivo, o uso do termo "sniffer de rede" nos países que falam a língua portuguesa.

A primeira versão do Ethereal foi lançada em julho de 1998, após diversas pausas em seu desenvolvimento. O primeiro indicativo de sucesso veio alguns dias após o lançamento da ferramenta, com diversas propostas de correções e melhorias enviadas por diversos colaboradores.

Um colaborador, em especial, foi um dos grandes responsáveis por ajudar a disseminar o Ethereal. No final de 1998, Richard Sharpe, na época lecionando disciplinas de TCP/IP na Australian National University, observou que a ferramenta poderia ser de grande utilidade no ensino de tal disciplina. Assim, ele começou a implementar os protocolos de rede de seu interesse e também buscar outros colaboradores para a melhoria da ferramenta.

Já em 2006, grande parte do código fonte do Ethereal já estava sob a licença GNU/GPL e o software já era bem conhecido nas comunidades Unix e no meio acadêmico, porém alguns problemas com o registro do nome Ethereal forçaram Combs a alterar o nome para Wireshark.

Atualmente, Combs ainda mantém grande parte do código fonte do software, porém, a lista de desenvolvedores que já contribuíram com o código fonte do Wireshark já passa de 600 pessoas.

Funcionamento de um analisador de pacotes

Os analisadores de pacotes funcionam de maneira semelhante. Para melhorar a compreensão do Wireshark, a seguir será mostrada a estrutura básica de uma solução desse tipo.

Um analisador de pacote, como o Wireshark, é definido como um elemento passivo na rede. Ele observa e captura os pacotes de rede enviados e recebidos pelo computador, mas sem interagir com o ambiente, como por exemplo, enviando pacotes. Além disso, também possui a característica de armazenar ou apresentar o conteúdo desses pacotes.

A **Figura 1** mostra a estrutura simplificada de um analisador de pacotes. Ao lado direito da figura estão os protocolos e aplicações comumente executados nos computadores. Já o analisador de pacotes está situado no lado esquerdo e é formado por dois módulos: a biblioteca de captura de pacotes e o analisador de pacotes em si.

Primeiramente o analisador de pacotes deve ser instalado no computador. O próximo passo consiste em monitorar a atividade das camadas de enlace, rede e transporte. Ou seja, algum processo de rede deve ser executado para que haja a troca de pacotes. A biblioteca de captura de pacotes é responsável por receber uma cópia de cada quadro da transmissão e enviar para o analisador de pacotes. O analisador de pacotes, por sua vez, interpreta os quadros e exibe os conteúdos dos pacotes para os usuários.

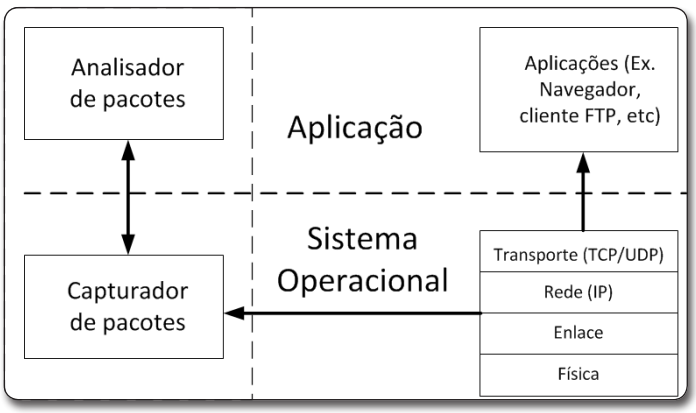


Figura 1. Estrutura básica de um analisador de pacotes.

Por exemplo, vamos supor que estamos interessados em mostrar o conteúdo dos pacotes HTTP em uma determinada troca de mensagens. Após algum acesso ter sido realizado no navegador, o capturador de pacotes irá receber uma cópia dos quadros de transmissão enviados e recebidos. Assumindo que o meio físico é uma rede local, os protocolos das camadas superiores são encapsulados em um quadro Ethernet. Ou seja, até esse momento não existe a diferenciação entre os pacotes (HTTP, TCP, UDP, IP e etc.). A seguir, os quadros são enviados para o analisador de pacotes que irá interpretá-los e diferenciá-los corretamente. Em nosso caso, o analisador irá primeiramente extrair o formato do datagrama IP, depois o segmento TCP e finalmente o HTTP.

Wireshark: Arquitetura

O Wireshark é uma ferramenta de captura de tráfego de rede que permite visualizar das mais variadas formas os pacotes que trafegam por uma determinada interface de rede. Um dos pontos fortes da ferramenta é a sua interface gráfica, que permite ao usuário visualizar e selecionar o tipo de tráfego desejado. Outra característica importante reside no fato da ferramenta ser distribuída sob a licença GNU GPLv2 e poder ser alterada e expandida, por exemplo, através da introdução de novos filtros de pacotes ou do reconhecimento de novos protocolos de rede.

A versão atual estável do Wireshark é a 1.8.4 e está disponível para diversos sistemas operacionais como Microsoft Windows, Mac OS e Linux. A lista de protocolos suportados é extensa e

contém protocolos como: ARP, IP, ICMP, TCP, UDP, DCCP, HTTP e FTP. Assim como protocolos de outras famílias como Novell (IPX, SPX), AppleTalk (LLAP, AARP, DDP, etc.), VoIP (SIP, H323, H245) e P2P (BitTorrent, JXTA) entre outros. A lista completa pode ser encontrada na documentação oficial do Wireshark.

A arquitetura da ferramenta pode ser dividida em três módulos, tal como demonstrado na **Figura 2**. A seguir, cada um desses módulos será apresentado com mais detalhes.

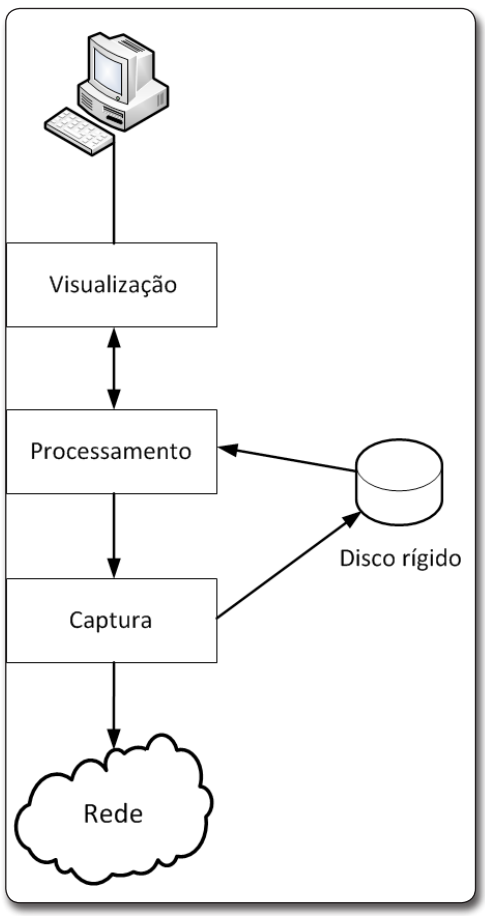


Figura 2. Arquitetura do Wireshark

Visualização

Este módulo é o responsável por toda a interface gráfica disponibilizada ao usuário. O principal objetivo do módulo é permitir o uso intuitivo da ferramenta e que a informação seja entregue ao usuário de forma legível e prática. O usuário pode escolher qual interface de rede será utilizada para capturar o tráfego e também indicar possíveis filtros de captura.

Outra funcionalidade do módulo de visualização do Wireshark é a possibilidade de escolha do modo de apresentação dos resultados obtidos com a captura dos pacotes: tempo real ou final. A principal diferença aqui está no momento em que são processados e apresentados os pacotes. Se for escolhido o modo de tempo real, os pacotes capturados serão visualizados ao mesmo tempo em que a captura está sendo realizada. Ao contrário do que acontece no modo final, onde os pacotes só vão ser processados e visualizados após a captura ser finalizada. A escolha de cada um dos modos afeta a maneira com que o módulo de processamento trabalha.

Com o Wireshark operando no modo em tempo real, conforme o módulo de processamento recebe as informações, ele também disponibiliza visualmente o conteúdo dos pacotes capturados. Os dados mostrados são os seguintes: instante em que o pacote foi capturado desde que a captura foi iniciada, IP de origem, IP de destino, protocolo do pacote e informação complementar. Para facilitar a identificação dos pacotes, dependendo do tipo de protocolo encontrado, uma cor diferente é mostrada na tela. Pacotes inválidos ou não reconhecidos também recebem uma cor que os permite distinguir facilmente dos demais.

O módulo de visualização também pode ser utilizado para a coleta de estatísticas sobre os dados coletados. Existem diversos tipos de estatísticas disponíveis, como por exemplo: o número de pacotes para um determinado destino, o número de pacotes para uma determinada origem, duração dos pacotes e etc. Além disso, os dados capturados podem ser guardados em arquivos (usualmente com a extensão *.pcap) para posteriormente serem carregados e analisados novamente.

Processamento

Este é o módulo central, responsável pelo processamento dos dados obtidos no módulo de captura. O resultado desse processamento é enviado ao módulo de visualização.

O módulo de processamento controla os processos responsáveis pela captura do tráfego de uma determinada interface, a utilização de determinados filtros e a forma como os pacotes serão analisados. Basicamente, é criado um canal de comunicação com o módulo de captura, recebendo os dados da interface alvo e os enviando ao módulo de processamento. Os dados da interface alvo são enviados ao módulo de processamento à medida que são recebidos pelo módulo de captura.

Outra funcionalidade deste módulo é relacionada à biblioteca de dissecadores (“protocol decoders” ou “dissectors”). Nessa biblioteca se encontram os dados relativos à forma com que os pacotes serão analisados. Para cada protocolo existe um dissecador específico, pois cada protocolo organiza o pacote de forma diferente. Existem dois tipos de dissecação: parcial e completa. A parcial é efetuada na primeira análise do pacote e utilizada para obter alguns dados que são enviados rapidamente ao módulo de visualização. Já a completa é realizada quando o módulo de visualização solicita que sejam entregues os dados completos de um determinado pacote.

O módulo de processamento também é responsável por manter um cache dos pacotes capturadores, para que as análises completas e filtragens por pedido de visualização sejam efetuadas de forma mais rápida e eficiente.

Captura

Este módulo é o responsável pela captura dos pacotes que trafegam pela interface de rede. Assim que o processo de captura dos

pacotes é iniciado, ele recebe a informação sobre qual interface será feita a captura e qual o filtro será utilizado, se definido. A captura é realizada com o auxílio de uma biblioteca chamada de PCap (ver **Nota do DevMan 2**), que é instalada juntamente com o Wireshark para a realização das coletas em tempo real. Após o início da captura, o módulo aguarda até receber os pacotes provenientes da biblioteca PCap. Ao recebê-los, o módulo de captura grava os pacotes em um arquivo e envia uma mensagem para o módulo de processamento, indicando que existem novos pacotes para serem processados.



Nota do DevMan 2

PCap: É uma biblioteca de código livre que disponibiliza uma interface de programação, independente do sistema operacional, para realizar a captura de pacotes recebidos ou enviados através das interfaces de rede de um computador.

O intervalo de tempo entre a detecção da existência de novos pacotes e o envio dessa informação ao módulo de processamento deve ser curto o suficiente para tornar a captura dos pacotes aproximadamente em tempo real, mas sem prejudicar o desempenho da ferramenta.

Wireshark: Primeiros passos

Atualmente diversas distribuições Linux já acompanham versões do Wireshark ou possuem pacotes pré-compilados (Ubuntu, por exemplo). Se esse não for o seu caso, então a compilação dos pacotes deve ser realizada. Porém, como o código-fonte do Wireshark utiliza as chamadas Autotools do projeto GNU para gerar os scripts de configuração e os arquivos make, a instalação do Wireshark "manual" pode ser feita utilizando os passos: *configure*, *make* e *make install*.

Para os usuários do Windows, é necessário somente fazer o download do instalador e seguir as instruções. É importante lembrar que a biblioteca PCap também deve ser instalada se o usuário deseja realizar as capturas em tempo real, que é o caso desse artigo.

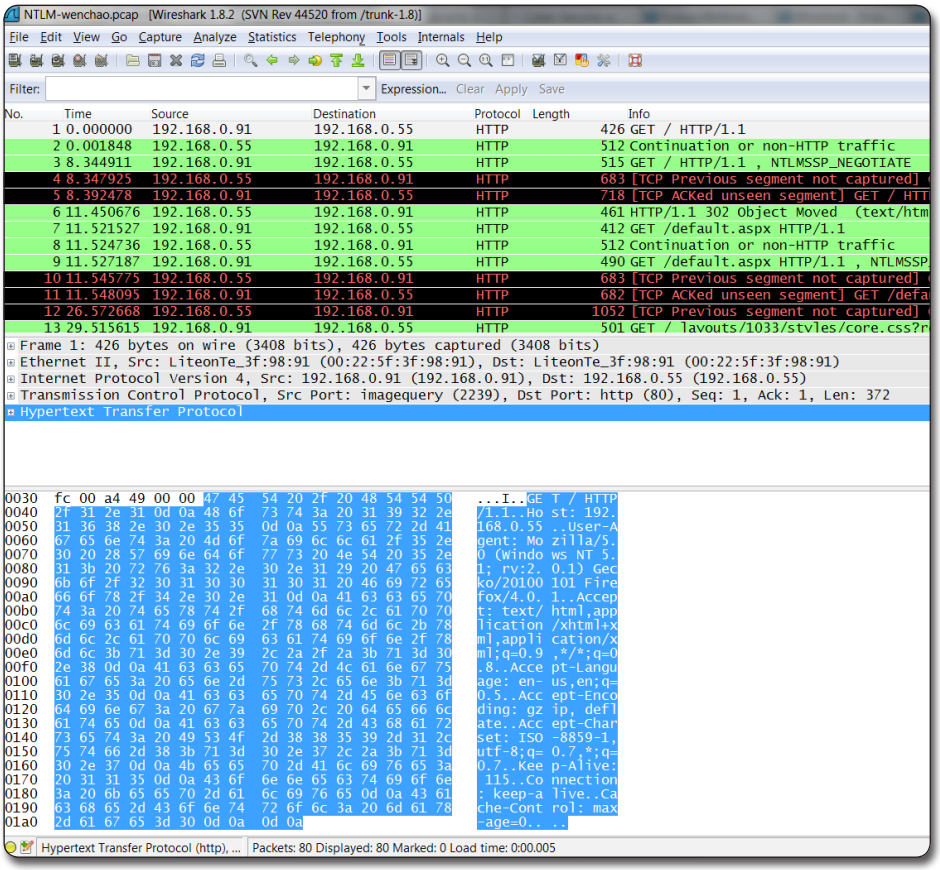


Figura 3. Wireshark – Tela inicial

É importante destacar que qualquer captura do tráfego de rede exige duas importantes configurações: usar a placa de rede no modo promíscuo e executar o Wireshark com privilégios de administrador.

O principal objetivo do modo promíscuo é permitir que a interface de rede receba todos os pacotes que passam por ela, mesmo os que não são destinados a ela. Dessa forma, a captura de pacotes se torna mais eficiente. O próprio Wireshark, quando executado com privilégios de administrador, pode fazer a troca para esse modo (*Capture -> Options -> Capture all in promiscuous mode*).

A tela inicial do Wireshark, mostrada na **Figura 3**, é divida em sete partes:

1. Menu – Utilizado para iniciar as ações. Contém as opções: File, Edit, View, Go, Capture, Analyze, Statistics, Telephony, Tools, Internals e Help;

2. Barra de ferramentas – Fornece acesso rápido aos itens do menu usados com frequência;

3. Barra de filtros – Situada logo abaixo da barra de ferramentas. Fornece um modo rápido de filtrar as informações capturadas;

4. Lista de pacotes – Mostra um resumo de cada pacote capturado de acordo com os seguintes atributos: número, tempo, origem, destino, protocolo, tamanho e info. Cada linha representa um pacote capturado. Ao clicar nos pacotes aqui, você irá controlar o que será mostrado em outros dois painéis da tela;

5. Detalhes dos pacotes – Mostra informações detalhadas dos pacotes selecionados;

6. Bytes dos pacotes – Mostra os dados dos pacotes selecionados em formato hexadecimal;

7. Barra de status – Mostra informações detalhadas sobre os dados capturados e o status da captura.

Iniciando a captura

Para iniciar uma captura, vá até o menu *Capture* e clique em *Interfaces*. Feito isso, uma tela com a lista de interfaces irá aparecer. Selecione a(s) interface(s) desejada(s) e clique em *Start*.

- DICA: a coluna "Packets" na janela de seleção de interfaces de rede mostra uma contagem em tempo real de pacotes em cada interface. Isso pode ser um critério importante na hora de escolher a interface para a captura dos pacotes.

Após a realização desses passos, se houver qualquer movimentação de pacotes na interface de rede selecionada, os pacotes irão aparecer na lista de pacotes. Para parar uma captura, vá ate o menu *Capture* e clique em *Stop*. Ou, simplesmente clique no ícone *Stop* na barra de ferramentas.

Como nenhum filtro foi definido, todos os pacotes trafegando pela rede foram capturados pelo Wireshark. Esse tipo de captura de pacotes, em que nenhum tipo de filtro é definido, geralmente exige muito espaço em disco, além de tempo para processar e analisar os pacotes. Em uma rede com muito tráfego, 5 minutos de captura pode representar vários gigabytes de espaço em disco e um grande trabalho para a análise dos dados.

Um tipo de filtro que auxilia a diminuir o número de pacotes apresentados para a visualização é o filtro por dispositivo ou *host*. Tais filtros são inseridos antes de iniciar a captura dos pacotes e podem ser muito úteis no cotidiano dos administradores de rede. Por exemplo, deseja-se analisar os pacotes que trafegam por uma determinada interface de rede, mas somente destinados para o próprio computador. Ou ainda, analisar os pacotes que trafegam em uma determinada interface de rede em um computador específico, por exemplo, um servidor que apresenta problemas de lentidão.

Este filtro é construído da seguinte forma: vá até o menu *Capture* e clique em *Interfaces*. Em seguida, uma tela com a lista de interfaces irá aparecer. Selecione a(s) interface(s) desejada(s) e clique em *Options*, quando uma tela semelhante à **Figura 4** será exibida. Toda a customização da captura dos pacotes pode ser encontrada aqui. Um arquivo externo, por exemplo, simulando uma captura, pode ser carregado e analisado. Outra opção interessante que pode ser definida é o critério de parada da captura. Os critérios incluem número de pacotes, tamanho e tempo. Além disso, opções de visualizações como a atualização dos pacotes em tempo real e o recurso da barra de rolagem acompanhar a captura automaticamente também são apresentados.

Para configurar os filtros, dê um clique duplo na interface escolhida e uma nova janela será apresentada.

Nessa nova janela existe um campo chamado *Capture Filter*. Escreva *host x.x.x.x*, onde x.x.x.x é o número IP que você deseja filtrar. Ao iniciar a captura, somente serão mostrados os pacotes enviados por esse IP ou com destino a esse IP.

Um experimento prático pode ser realizado utilizando esse filtro. Sendo assim, crie um filtro direcionado para o seu próprio computador, feche as suas conexões de rede (navegadores, chats e etc.) e inicie a captura. Depois disso, abra o navegador e acesse *www.google.com.br*. O resultado pode ser visto na **Figura 5**. Toda a sequência dos pacotes até a exibição da página requisitada pode ser visualizada no Wireshark.

O pacote número 4 corresponde a um pedido do meu computador (192.168.1.102) ao servidor DNS (201.82.0.61) para a resolução do nome *www.google.com.br*. O pacote número 5 é a resposta do servidor DNS endereçada para o meu computador.

Os próximos três pacotes (6, 7 e 8) são relacionados ao *handshake* de três vias do TCP. Primeiramente o pedido de conexão (SYN) de minha máquina, seguido da resposta do servidor do Google (SYN ACK) e por fim o ACK de minha máquina, endereçado para o servidor. Somente após esses passos é que a página HTTP é descarregada em meu computador. Esse processo é descrito nos pacotes 9 e 11 com a solicitação de recurso (GET) e a resposta do servidor, devolvendo o conteúdo do site *www.google.com.br*.

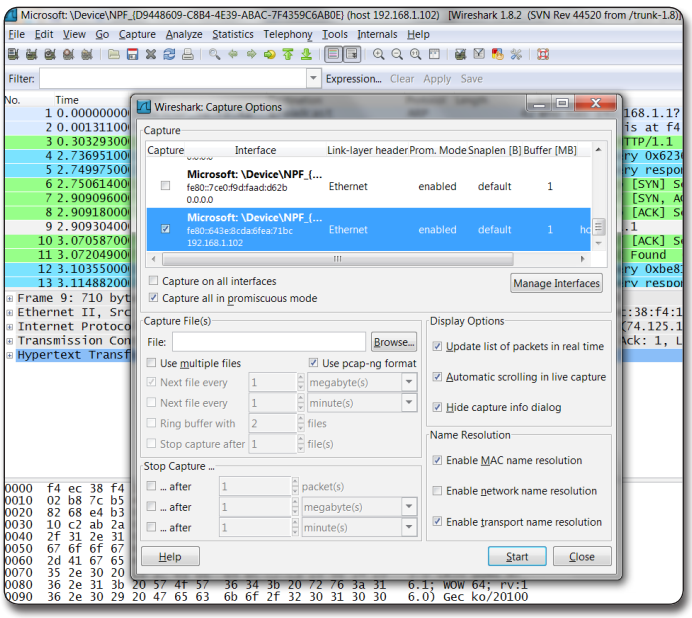


Figura 4. Wireshark – Opções



Figura 5. Wireshark – Captura utilizando filtros

Esse pequeno exemplo mostra como o Wireshark, quando utilizado corretamente, pode ser útil para a administração de redes. Em nosso caso, a página foi carregada com sucesso, porém um erro no decorrer do processo seria capturado pelo Wireshark e as informações fornecidas pela ferramenta poderiam ser valiosas para o tratamento de tal erro. Com o Wireshark seria possível encontrar em qual protocolo o erro aconteceu. O problema foi no servidor DNS? No estabelecimento da conexão TCP ou na obtenção da página HTTP?

Aplicações práticas

Conforme discutido anteriormente, além da monitoração local, o Wireshark pode ser utilizado para a detecção de problemas de

rede. Nesta seção serão mostradas algumas aplicações práticas da ferramenta estudada.

Problemas de rede

Problemas de desempenho são recorrentes para quem trabalha com administração de redes. Porém, descobrir a causa-raiz de tais problemas geralmente demanda muito tempo e esforço. O Wireshark pode ser utilizado para auxiliar a investigação de problemas desse tipo.

A grande vantagem do Wireshark nesse caso reside no fato de seu processamento mostrar a monitoração em tempo real e em diferentes protocolos. Qualquer comportamento anômalo nos pacotes pode ser notado na visualização proporcionada pela ferramenta. As estatísticas embutidas no programa também são úteis para desvendar possíveis gargalos na rede. Estatísticas como o resumo das trocas de mensagens (*Statistics -> Conversations*), gráfico do tempo total de ida e volta dos pacotes (*Statistics - TCP Stream Graph -> Round Trip Time Graph*) e a distribuição de carga dos servidores HTTP (*Statistics -> HTTP -> Load distribution*) ajudam a reconstruir o cenário da rede para a realização de análises mais profundas.

Situações como equipamentos mal configurados, monitoração de serviços de telefonia pela Internet (VoIP), configuração de DNS e resposta de servidores de e-mails são casos clássicos e bem documentados da aplicação do Wireshark.

Análise de segurança

Uma preocupação recorrente nas redes de computadores é a segurança da informação. O uso de firewalls, definição de políticas, implementação de controles de acesso, criptografia e utilização de túneis seguros (redes privadas virtuais) nem sempre significa que a organização está segura e ficará ausente de vulnerabilidades e ataques em seu perímetro.

O Wireshark pode ser utilizado para auxiliar nas análises de segurança em dispositivos da rede. Um exemplo de análise que pode ser feita é a busca por senhas fracas ou identificação de servidores sem autenticação segura. O Wireshark possui uma funcionalidade chamada Follow TCP Stream (botão direito no pacote -> Follow TCP Stream), que mostra o conteúdo completo de uma troca de mensagens TCP. Usando esse recurso, o conteúdo dos pacotes pode ser lido e eventualmente senhas em claro (não cifradas) podem ser descobertas. O Wireshark também pode capturar pacotes cifrados, porém, tais pacotes serão inteligíveis para o usuário, mesmo com o recurso Follow TCP Stream.

É muito importante deixar claro que a captura de pacotes destinados a outras máquinas ou outras pessoas é uma forma de invasão de privacidade, o que em muitos casos pode ser algo ilegal. Busque sempre a aprovação antes de realizar esse tipo de análise. Um profissional ético deve sempre conhecer os limites entre a análise de rede e a violação da privacidade dos usuários envolvidos.

CURSOS ONLINE

A Revista Java Magazine oferece aos seus assinantes uma série de Cursos Online de alto padrão de qualidade.



CONHEÇA ALGUNS DOS CURSOS:

- Java Básico: Aplicações Desktop
- JSF com Primefaces
- Conhecendo o Apache Struts
- Introdução ao JSP com HTML 5: Intranet Escolar
- Bluetooth: Controle de Aplicação desktop via celular

Para mais informações :
www.devmedia.com.br/curso/javamagazine
(21) 3382-5038

Outro exemplo da aplicação do Wireshark na análise de segurança é a detecção de códigos maliciosos. Antivírus são os responsáveis pela detecção e remoção de arquivos maliciosos, porém em muitas vezes o antivírus não está atualizado ou não consegue, por algum motivo, detectar o arquivo que está causando problemas. Um sintoma clássico é a lentidão da máquina alvo.

Se o arquivo contendo o código malicioso tenta se comunicar usando a rede, o Wireshark pode ser utilizado para identificar essa comunicação. Por exemplo, a máquina alvo pode estar enviando pacotes para alvos suspeitos e usando portas suspeitas. Nesse caso é importante olhar para os pacotes referentes aos protocolos DNS, HTTP (existe algum pedido GET para servidores não conhecidos?) e também usar o TCP Follow Stream para remontar a troca de mensagens TCP. Um exemplo de troca de mensagens entre um alvo infectado com um bot e um servidor de controle pode ser visto na **Figura 6**.

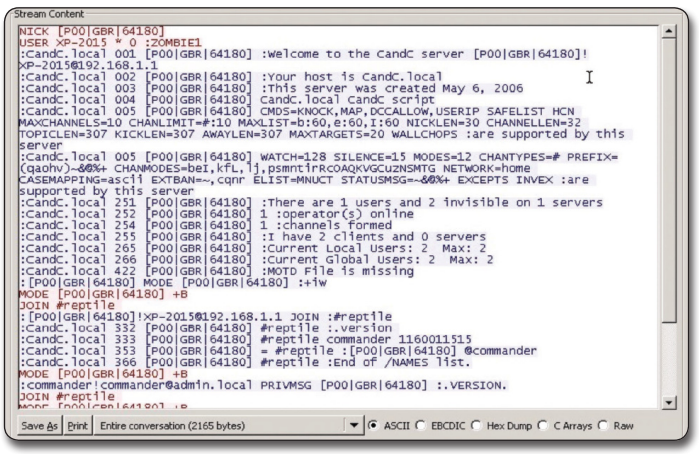


Figura 6. Wireshark – Conteúdo da troca de mensagens entre uma máquina infectada e um servidor BOT. Extraído de "Security Analysis with Wireshark"

Conclusões

O Wireshark é uma ferramenta valiosa para a gerência de redes, seja na monitoração do ambiente ou na análise minuciosa de problemas. Além disso, a observação da sequência de mensagens trocadas entre os envolvidos e os detalhes da operação dos protocolos também contribui para melhorar a compreensão dos protocolos de redes. Muitas vezes nos esquecemos dos fundamentos relacionados aos elementos que compõem uma rede. A resposta para os problemas pode residir na compreensão correta desses fundamentos e o Wireshark, a partir da captura e visualização dos pacotes, fornece subsídios para isso.

O artigo mostrou somente uma pequena parte de uma área que está em plena expansão. A fundação Wireshark University foi criada com o intuito de orientar os profissionais de TI sobre a importância da ferramenta. Além disso, a fundação também gerencia a distribuição de certificações sobre o software, chamadas de Wireshark Certified Network Analyst (WCNA), e publica livros e artigos sobre o assunto. Outras referências sobre o assunto desvinculadas a Wireshark University também podem ser encontradas, seja por meio de livros ou artigos.

De uma forma geral, a busca pela gerência de redes eficiente e segura exige mecanismos que envolvem o conhecimento dos protocolos de rede, do ambiente e ferramentas para auxiliar esses processos. Neste contexto, o Wireshark pode trazer inúmeros benefícios para seus usuários.

Rodrigo Sanches Miani
rodrigomiani@gmail.com



Graduado em Matemática pela Universidade Federal de São Carlos (2005) e mestrado em Engenharia Elétrica pela Universidade Estadual de Campinas (2009), onde atualmente cursa doutorado em Engenharia Elétrica. É também pesquisador visitante da University of Maryland, College Park, EUA. Possui experiência de sete anos em pesquisa e desenvolvimento de soluções para governo eletrônico e segurança de redes. É também professor das disciplinas relacionadas à Segurança da Informação na Faculdade Campo Limpo Paulista - São Paulo (FACCAMP). Seus interesses de pesquisa incluem modelos de quantificação de segurança da informação, métricas de segurança, e estudos empíricos em segurança e criptografia.

Site do Wireshark
http://www.wireshark.org/

Wireshark User's Guide
http://www.wireshark.org/docs/wsug_html_chunked/

Artigo "Security analysis with Wireshark", escrito por Russ McRee
http://holisticinfosec.org/toolsmith/docs/november2006.pdf

Artigo "Evaluation of the Capabilities of WireShark as a tool for Intrusion Detection" escrito por Usha Banerjee, Ashutosh Vashishtha e Mukul Saxena
http://www.ijcaonline.org/volume6/number7/pxc3871427.pdf

Artigo "Usos práticos do Wireshark" escrito por Armijn Hemel
http://www.linuxnewmedia.com.br/images/uploads/mags/lm/articles/LM32_wireshark.pdf

Dissertação de mestrado "Wireshark para sistemas distribuídos" escrito por Nuno Miguel Galego Farruca
http://hdl.handle.net/10362/2288

Livro "Redes de Computadores" escrito por Andrew Tanenbaum
Editora Elsevier, 4.ª edição, 2003

Splunk: Monitorando o ambiente de TI – Parte 1

Com o Splunk você monitora a TI e os negócios da sua empresa e ainda toma decisão em tempo real

Resumo DevMan

De que se trata o artigo:

Este artigo apresentará, de maneira geral, a ferramenta de coleta, indexação e análise de informação vista como uma plataforma universal de dados de máquina, ou seja, todo e qualquer tipo de informação que seja possível ler a olhos nus, independentemente de onde estiver e que se encontre de maneira estruturada – caso contrário não poderá ser coletada e disponibilizada para o usuário final. O grande diferencial do Splunk em relação a outras ferramentas do gênero é que ele possibilita, de forma fácil e rápida, a indexação de informações de quaisquer fontes de dados, e as disponibiliza para análise e desenvolvimento de dashboards por meio de poucos cliques. Assim, possibilita a correlação de tais dados com as necessidades dos negócios.

Em que situação o tema é útil:

Muitos dos cases de sucesso do mercado nacional e/ou internacional ainda têm uma postura reativa em relação à TI para suporte aos negócios (produtos e serviços) que tais empresas entregam. O Splunk é um software que permite que a organização passe a ter uma visão mais proativa com base nas informações que elas nem sabem que possuem (Big Data) e, na maioria das vezes, informações que podem trazer, além de melhor suporte à decisão, mais inteligência competitiva para manter os negócios de maneira mais confiável por parte da TI. Assim, tem-se maior visibilidade operacional em relação à capacidade instalada, além da informação necessária para manter as estratégias de crescimento da empresa em tempo real.

Splunk: Monitorando o ambiente de TI – Parte 1:

O Splunk é uma ferramenta que permite coletar, indexar e analisar a informação gerada pelo negócio, também conhecida pelo nome de dados de máquina. A partir da coleta e da indexação de toda a informação de syslog de uma máquina, por exemplo, é possível determinar vários problemas que a TI não fazia nem ideia de que ocorriam com o ambiente. Com isso, várias análises poderão ser realizadas, desde a implementação de um sistema de alertas com base em uma quantidade x de erros ou ocorrência de determinado evento, até a apresentação de gráficos que refletem indicadores de performance do negócio, em tempo real. Tudo isso de maneira muito fácil é rápida, o que possibilita que a organização, ou mesmo os departamentos envolvidos, possam tomar determinadas decisões baseadas em fatos, sem o uso da intuição.

São muitos os controles e métricas que normalmente são criados para monitorar os elementos críticos que compõem o ambiente de TI de pequenas, médias e grandes empresas. Tais elementos fazem parte de um grande stack de suporte às operações de negócio: sistemas OLTP, bancos de dados, sistemas de entrada e saída de dados e aqueles de apoio e suporte à decisão.

Obviamente, esses são elementos básicos, mas ainda podemos abrir o leque ou mesmo fazer um *drill-down* e abordar os servidores de aplicação como o WebSphere, servidores web como o Apache e/ou Tomcat, ou ainda bancos de dados como Oracle, MySQL, SQL Server e sistemas desenvolvidos em Java, .NET e qualquer outra linguagem. Outros elementos compõem o que é tido como serviço networking, que é composto por vários outros pedaços como firewall, roteadores e switches, além dos vários servidores que são controladores de domínio, DNS, servidores de impressão e vários outros que precisam ser monitorados em tempo real para, por exemplo, saber se estão online e servindo aos funcionários.

A grande questão é que todos estes pontos têm os seus controles e, na maioria das vezes, as práticas não permitem uma análise real do que está acontecendo por debaixo de todos aqueles aparatos, uma análise completa no que diz respeito à saúde dos serviços de TI que suportam o negócio.

Na maioria das vezes, quando um problema acontece, ele passa por todos os profissionais do departamento de TI e acaba no conhecido jogo de empurra entre as áreas – isso realmente acontece. Infelizmente não existe uma ferramenta gráfica que colete a informação para que se possa fazer uma análise mais detalhada sem que essa situação aconteça.

Por causa disso, geralmente acontece o seguinte: alguém nota alguma instabilidade no ambiente e a comunica. É quando o gerente de área toma conhecimento da situação e aciona, além de pessoal interno, uma consultoria ou mesmo uma auditoria de sistemas para dar conta do problema. Nestes casos e em muitos outros, o tempo dos profissionais envolvidos na resolução reativa dos problemas toma grande parte dos recursos que a empresa poderia empregar para tornar a TI mais proativa e menos reativa. Reagir a um problema é trabalhar para colocar um sistema de volta às operações normais após um crash de algum dos elementos de TI já citados; ou seja, o problema já aconteceu e nesse momento a empresa parou de oferecer determinado serviço para o seu cliente.

Rotineiramente, o profissional precisa analisar vários logs utilizando um editor de textos, ou mesmo via terminal em tela preta, e tomar nota dos vários registros após consultar vários arquivos. Este processo de investigação de informação ou busca pela ocorrência de um evento em vários arquivos de log pode consumir muito tempo e pessoal dedicado a realizar esse tipo de trabalho, que acaba por obter um resultado mais por inferência do que algo que realmente seja uma verdade com fatos a serem exibidos. Tais arquivos de log fazem parte de um conceito que é muito discutido nos dias atuais, que é o Big Data, dados que as empresas só descobrem que possuem quando precisam acessá-los, na maioria das vezes, em momentos de downtime dos sistemas (*o conceito de Big Data será discutido em uma próxima oportunidade*).

No entanto, hoje as empresas já estão mais atentas em relação a estes tipos de dados, procurando utilizá-los em prol da criação ou da extração de inteligência operacional e, assim, saltar em competitividade na frente de seus concorrentes. Dessa forma, alcançam mais visibilidade nos negócios e obtêm serviços com mais *uptime* e com equipes mais resilientes, podendo escalar menos os problemas que ocorrem no dia a dia.

Considerando o citado cenário, vamos apresentar neste artigo uma ferramenta que está revolucionando a maneira de monitorar em tempo real a TI e os negócios e ainda relacionar os dados de máquinas (ver **Nota do DevMan 1**) gerados por *devices* (*routers* e *switches*), sistemas operacionais, serviços de network e vários outros pontos com dados obtidos dos negócios.

O Splunk é uma Plataforma Universal de Dados de Máquina desenvolvida por uma empresa de nome Splunk, Inc., localizada na cidade de São Francisco, no estado da Califórnia, EUA, que fez o IPO (veja **Nota do DevMan 2**) há pouco tempo (Abril de 2012) e que se mantém em um crescimento enorme a cada ano. O conceito do produto é fácil de entender: basta que você adicione todas as fontes de dados – como logs de aplicações/softwares, logs de segurança, de sistemas operacionais, logs de operações de billing, logs de acesso a servidor web, logs de bancos de dados, logs gerados por roteadores e/ou switches e várias outras fontes de dados, não importa onde elas estejam –, e inicie a coleta e a indexação de tal informação.



Nota do DevMan 1

O que são Dados de Máquina?

Dados de máquina é toda a informação que é gerada por serviços que são disponibilizados para usuários finais e funcionários de uma empresa para que os seus sistemas e dispositivos funcionem. Entenda dados de máquina como todos os dados que são gerados por roteadores, switches, bancos de dados, servidores de aplicação, etc., para que estes possam registrar tudo o que estão fazendo. Podemos pensar nesses dados como logs, mas eles são muito mais que isso. Podem ser dados de configuração, dados de APIs e filas de mensagens, eventos de mudança, saída dos comandos de diagnóstico, informação sobre integridade de arquivos e muito mais. Além disso, há uma variedade muito mais ampla de dados de log com os quais os sistemas de gerenciamento de eventos e informações de segurança (SIEMs) e gerenciamento de logs voltados para segurança de rede e conformidade foram criados para lidar.



Nota do DevMan 2

IPO

O processo usualmente referido como IPO (do inglês Initial Public Offering) é o evento que marca a primeira venda de ações de uma empresa no mercado de ações. Seu principal propósito para empresas novas/pequenas é levantar capital pela sociedade para utilizar como investimento para expansão, porém também ocorre em empresas/corporações maiores por motivos de alavancagem. Após este processo, a empresa que incorre se torna uma companhia de capital aberto e passa a ser reconhecida como Sociedade Anônima ou S.A..

Com as fontes de dados configuradas, os dados serão indexados de forma contínua, e sobre estes dados o administrador de sistemas, DBA, analista de sistemas e/ou gestor de área poderão facilmente criar alertas e gráficos para monitoramento em tempo real dos elementos do ambiente de TI. A granularidade do monitoramento vai da necessidade dos profissionais envolvidos, pois o Splunk é capaz desde meramente monitorar um arquivo de log de forma contínua, até mesmo pesquisar toda a movimentação de um usuário por todos os elementos da TI, correlacionando os vários eventos por vários arquivos de log.

A ideia é justamente obter em minutos o que é feito normalmente em horas, baseando-se em pesquisa textual, centralização da informação e dados de máquina, sendo que ainda é possível contar com uma arquitetura de escala linear, o que veremos no próximo tópico.

Conhecendo o Splunk

O Splunk é o mecanismo de coleta, indexação e reconhecimento automático de padrões de informação considerada como dado de máquina. Dado de máquina é tudo aquilo que é gerado para manter os serviços que são oferecidos pela TI de uma organização funcionando satisfatoriamente. Qualquer arquivo de log, seja ele um log gerado por um firewall IronPort, um BlueCoat, ou mesmo um log de alerta do Oracle ou o log de

erro do MySQL, ou ainda os logs de acesso e erros do Apache, todos estes também contêm dados de máquina que podem ser coletados e indexados de forma contínua pelo Splunk para fornecer análise do ambiente em tempo real. A **Figura 1** mostra alguns exemplos de fontes de dados.

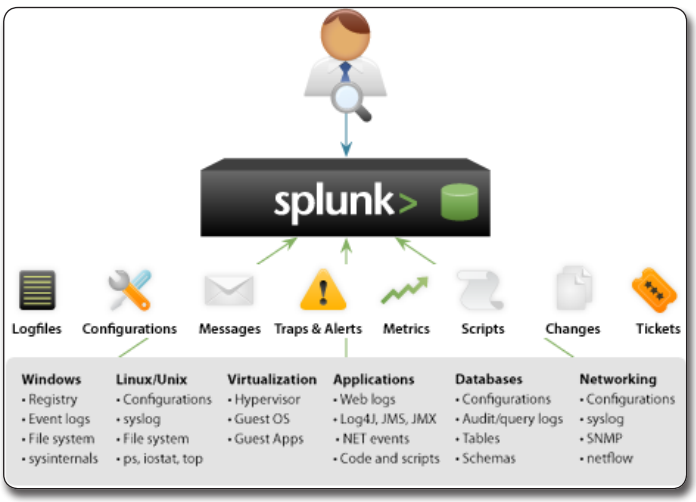


Figura 1. Fontes de dados

O Splunk não utiliza conectores, não utiliza qualquer tipo de bancos de dados por trás e nem impõe limite em relação ao volume de coleta e indexação de informação, já que estamos falando de uma ferramenta de Big Data. Você não precisa se preocupar em desenvolver conectores para coletar e indexar determinados tipos de informação ou mesmo desenvolver um novo pedaço de software sempre que um novo padrão de informação chegar ao mercado. O Splunk reconhece os padrões de um arquivo de log automaticamente, faz a sua indexação e armazena a informação organizada em disco com 50% de taxa de compressão. Os dados são armazenados no formato *key/value* e são gerenciados pelo próprio Splunk, que desde a versão 4 tem uma versão otimizada do MapReduce (ver **Nota do DevMan 3**) com várias melhorias em relação ao release do Google de 2004.



Nota do DevMan 3

Splunk e sua implementação de MapReduce

O MapReduce é um modelo de processamento de dados de maneira distribuída, primeiramente escrito pelo Google em 2004. Seu conceito fundamental recai sobre a divisão de problemas em partes como mapear (Map) os processos de acordo com suas fontes para então reduzir (Reduce) os processos a um só e fornecer a resposta para quem o solicitou. Uma chamada a uma função Map poderá ser executada paralelamente em vários nós de um cluster ou mesmo em um único cluster (single node), e uma vez que os dados são recuperados para uma área denominada Temporal MapReduce, a função de redução é aplicada aos dados e o resultado é entregue. O modelo em si ainda requer muito desenvolvimento para que possa ser utilizado, mas o Splunk introduziu várias melhorias ainda na versão 4.2 que automatizam os processos de chamadas a mapeamento e redução, o que tornam o MapReduce 100% transparente ao usuário, diferenciando-o de qualquer outro produto que utilize o mesmo framework. Mais à frente neste artigo, quando falarmos de buscas por meio da Search App, chamaremos a atenção para consultas que utilizam o modelo MapReduce.

Toda informação é armazenada em buckets, que são a menor porção de dados dentro da estrutura de armazenamento de informação. Tais buckets são então amarrados a índices, que por sua vez são gerenciados e também mantidos pelo Splunk em disco. O mais interessante é que você poderá criar índices sempre que um novo tipo de informação for adicionado para ser coletado e indexado automaticamente pelo Splunk.

Os buckets podem ser ainda dos tipos *hot*, *warm*, *cold* e *frozen* (veja o subtópico a seguir). Os dados vão se movendo internamente nessa estrutura de acordo com a quantidade e a idade da informação armazenada. Tal estrutura é própria do Splunk e por isso o usuário não precisa se preocupar em ter um banco de dados para armazenar os dados coletados pela ferramenta, visto que muitas ferramentas no mercado ainda precisam ter um Oracle, SQL Server ou MySQL. No Splunk os dados vão sendo coletados e sua arquitetura acontece em tempo real, de forma que o esquema de dados é desenhado de acordo com a identificação dos padrões nos eventos dos arquivos configurados.

O Splunk coleta e indexa a informação que o usuário deseja e, assim, reconhece os padrões dentro dos eventos armazenados em arquivos de logs ou até mesmo, gerados por devices como roteadores e switches. Essa informação, considerada como dados de máquina, é coletada, indexada e armazenada pelo próprio Splunk em um conjunto de arquivos que contém essencialmente dois tipos:

- Arquivos contendo dados de máquina em forma comprimida, conhecidos também como "rawdata";
- Arquivos de índices que apontam para os arquivos de dados – rawdata – e arquivos de metadados.

Tais arquivos juntos formam os índices no Splunk, que referenciam um determinado tipo de dado de máquina coletado/indexado e que pode ser utilizado na Search App para busca e análise. Todos os dados indexados são organizados em buckets, e estes, são classificados em cinco formas diferentes, com base na idade dos dados. São elas:

- **Hot:** Aqui são armazenados os eventos mais recentes. Estes buckets estão abertos para escrita e os novos dados são sempre escritos neles. Para sistemas 32 bits, hot buckets podem ter até 1GB de tamanho, ao passo que para sistemas 64 bits, podem ter até 10GB de tamanho antes que virem Warm buckets. Por padrão, os hot buckets podem ser encontrados em *$SPLUNK_HOME/var/lib/splunk/defaultdb/db/**. Tal localização pode ser configurada para que cada tipo de bucket seja armazenado em um disco separado, o que é considerada uma boa prática. Caso um Hot bucket não atinja 1/10GB, em 90 dias ele será migrado para Warm com base na sua idade;
- **Warm:** Dentro deste tipo de informação estão todos os dados que pertencem a buckets maiores que 1/10GB que passaram do estágio de Hot para Warm, processo reconhecido como "rolled data". A cada bucket que é migrado do estágio Hot para o Warm, um novo diretório é criado (em uma instalação padrão) em *$SPLUNK_HOME/var/lib/splunk/defaultdb/db/**, seguindo uma convenção especial de nomenclatura. Uma instalação padrão do

Splunk poderá contar com até 300 Warm buckets. Quando for criado o Warm bucket número 301, o mais antigo bucket deste estágio será migrado para Cold;

• **Cold:** Armazena os buckets que foram migrados de Warm, sendo que podem chegar até 500 GB antes de serem migrados para Frozen. O aspecto da idade dos dados também está presente neste estágio, sendo que, quando determinados buckets atingirem seis anos, eles serão migrados para Frozen;

• **Frozen:** por padrão, o Splunk exclui permanentemente todos os buckets de dados que são migrados de Cold para Frozen; os dados são removidos do índice nesse momento. Caso seja necessário arquivar os dados que são movidos para este estágio, edite o arquivo *indexes.conf* localizado em *$SPLUNK_HOME/etc/system/default* e configure o parâmetro *coldToFrozenDir* para arquivar os novos Frozen buckets, como exibido na **Listagem 1**.

Listagem 1. Editando o arquivo inputs.conf para configurar o coldToFrozen.

```
# edição do arquivo $SPLUNK_HOME/etc/system/default/inputs.conf
[<nome_índice>]
coldToFrozenDir = "/frozenPartition"
```

Também no arquivo *indexes.conf*, utilize o parâmetro *coldToFrozenScript* para executar um script que normalmente é chamado para fazer o arquivamento dos dados e alguma outra tarefa que pode ser executada por meio de scripts, o que dá a possibilidade de utilizar a criatividade de forma livre. Veja como fazer esta configuração na **Listagem 2**.

Listagem 2. Adicionando a parâmetro coldToFrozenScript ao arquivo inputs.conf.

```
# edição do arquivo $SPLUNK_HOME/etc/system/default/inputs.conf
[<nome_índice>]
coldToFrozenScript = [$SPLUNK_HOME/bin/] "coldToFrozenExample.py"
```

Por padrão, o script especificado na **Listagem 2**, chamado *coldToFrozenExample.py*, está pronto para ser utilizado e está localizado em *$SPLUNK_HOME/bin*. O parâmetro *coldToFrozenDir* tem precedência sobre o *coldToFrozenScript*, portanto cuidado ao colocar as duas funcionalidades ao mesmo tempo;

• **Thawed:** Se o usuário traduzir para o Português a palavra *Thawed*, vai perceber algo até engraçado e que, nesse contexto, faz todo sentido. Literalmente, este estágio trata do descongelamento dos dados ou buckets que foram congelados, ou seja, antes de serem arquivados, eles estavam em um formato de Frozen buckets e agora serão posicionados no diretório de Thawed buckets, localizado em *$SPLUNK_HOME/var/lib/splunk/defaultdb/thaweddb*. Após colocar os Frozen buckets como "descon-

DÊ UM SALTO
EM CONHECIMENTO!



Acesse o maior portal para desenvolvedores da América Latina!

20 mil posts

430 mil cadastrados

10 milhões de page-views por mês

gelados" no diretório citado, eles podem ficar lá o tempo que for necessário, e quando não os quiser mais, basta excluí-los. Uma boa estratégia é criar um script para mover os Frozen buckets para Thawed buckets e eliminá-los de acordo com uma política defendida em um script no nível do SO.

Por padrão, o Splunk realizará buscas somente nos Hot/Warm/Cold buckets, uma vez que os Frozen buckets são excluídos por padrão. O usuário poderá ainda arquivar os Frozen buckets em um disco para fazer o seu restore, caso seja necessário. O comportamento ou política para que buckets sejam considerados antigos e assim passem de um estágio para outro, ou mesmo, as práticas de arquivamento de Frozen buckets, poderão ser reconfigurados de acordo com os interesses da organização por meio do arquivo *indexes.conf*. Neste arquivo, o administrador do Splunk especificará variáveis para configuração de políticas de retenção de eventos indexados. Veja mais sobre o assunto em *http://tinyurl.com/aatf2gr*.

Cada evento é indexado a partir do seu *timestamp* e as partes que compõem o evento são identificadas e adicionadas abaixo de estruturas denominadas *fields*. Os *fields* são estruturas que nomeiam a informação identificada em meio aos eventos e que podem ser utilizados para efetuar buscas por meio do aplicativo padrão do Splunk, a Search App. Neste aplicativo, ou no aplicativo de busca, é onde o usuário poderá escrever suas próprias consultas para extrair a informação mais valiosa para sua organização, aquela que responde a perguntas estratégicas.

A Search App é considerada como o Google do lado interno da sua organização, pois com ela você pode buscar por termos como "error" ou "fail" e tem a possibilidade de percorrer todos os logs adicionados ao Splunk e saber exatamente o que está acontecendo com a sua TI. Além disso, o Splunk possui uma linguagem própria de consulta aos dados de máquina indexados denominada SPL, ou *Search Processing Language*.

Além de armazenar os dados sem precisar de um banco de dados, não necessitar de conectores, lidar com qualquer volume de dados e ainda possuir um aplicativo que possibilita manipular dados com uma linguagem própria, o Splunk possibilita que o usuário crie os seus próprios relatórios baseados em vários tipos de gráficos (*column*, *line*, *area*, *bar*, *pie*, *scatter*, *radial*, *filter* e *marker*), que podem ser criados a partir da Search App e assim serem salvos em um painel que poderá conter vários gráficos relacionados com um departamento específico. Além dos gráficos, alertas sobre consulta a dados também podem ser criados e dentro destes alertas, o usuário poderá definir o envio de e-mails e a execução de scripts para realizar ilimitadas ações de correção para um problema identificado.

Com isso, o Splunk entrega o que tem defendido como sua missão, que é "tornar os dados de máquina acessíveis, utilizáveis e valorizados por todos". Com apenas quatro passos, você conseguirá extrair informação inteligente de um emaranhado de dados, nos quais, a olhos nus, o usuário ou uma equipe não teria tanta facilidade em fazer.

Os quatros passos são:

1. Configurar a fonte de dados;
2. Consultar a informação com a *Search App*;
3. Criar relatórios com vários tipos de gráficos;
4. Criar alertas sobre a informação retornada pelas consultas.

O Splunk ainda oferece a facilidade de poder facilmente ser integrado com vários outros sistemas. Um exemplo é a integração que pode ser feita com sistemas de controle de SLA (*Service Level Agreement* ou Acordo de Nível de Serviço), para determinar quais tickets estão para vencer, quais já venceram e quão efetivo é o departamento em não deixar SLAs vencerem. Integrar o Splunk com portais de e-commerce também é uma tarefa simples. O administrador de sistemas poderia facilmente saber quais produtos estão sendo abandonados no carrinho de compras, ou até mesmo o tempo de entrega de uma requisição de usuário, correlacionando logs de Apache, WebSphere e MySQL.

Os sistemas operacionais suportados são os seguintes:

- Solaris 9, 10 (x86, SPARC);
- Linux Kernel versão 2.6.x e superiores (x86: 32 e 64-bit);
- FreeBSD 6.1 (x86: 32-bit), 6.2, 7.x, 8.x (x86: 32 e 64-bit);
- Windows Server 2003/2003 R2;
- Windows Server 2008/2008 R2;
- Windows XP (32-bit);
- Windows Vista (32-bit, 64-bit);
- Windows 7 (32-bit, 64-bit);
- Mac OS X 10.5 e 10.6;
- AIX 5.2, 5.3, e 6.1;
- HP-UX.

Além disso, o Splunk tem como característica ser escalável desde o seu laptop até máquinas maiores em um Data Center. Sendo assim, o administrador Splunk poderá iniciar os testes em laptops e depois migrar aquele ambiente para máquinas maiores. Uma ressalva interessante é que a empresa não precisará adotar máquinas muito grandes para rodar o Splunk, já que o produto tem escalabilidade horizontal ou linear, ou seja, pode operar bem com várias máquinas pequenas (*hardware commodity*). E é aí que entram alguns novos conceitos: as três funções básicas do Splunk, que é a de coletar, indexar e consultar dados de máquina, funções estas que podem ser facilmente divididas entre várias máquinas.

A máquina que servirá para busca de informação (utilizando a *Search App*) passará a ser chamada de *Search Head*; as máquinas que recebem as informações e as armazenam em disco passam a ser chamadas de *Indexers*; e aquelas que enviam as informações para os *Indexers* passam a contar com um pequeno agente, que é um software Splunk denominado *Universal Forwarder*. Este *Universal Forwarder*, ou somente *Forwarder*, coletará a informação de interesse nas máquinas onde rodam, por exemplo, o Apache, o Linux e o Oracle, e enviará aquelas informações para os *Indexers* para serem indexadas e, assim, estarem disponíveis para os *Search Heads*, onde o usuário entra com as consultas (ufa!). Para vermos melhor como funciona essa arquitetura, observe a **Figura 2**.

Para fazer a instalação do Splunk é necessário acessar o site splunk.com e fazer um pequeno cadastro. Assim, além de fazer o download de acordo com a sua plataforma, o usuário já contará com um perfil para acessar o fórum oficial de usuários Splunk, conhecido por *Splunk Answers*, e também baixar aplicativos no *Splunkbase*, onde a própria Splunk disponibiliza vários aplicativos desenvolvidos por eles e também aplicativos de terceiros (uma boa oportunidade para quem deseja vender aplicativos para Splunk).

Tais aplicativos podem ser baixados de dentro da sua instalação do Splunk e são como inteligência empacotada para resolver determinado problema de monitoramento. Com um aplicativo, ou somente "app", como são chamados, o usuário conseguirá monitorar um IronPort, BigIP e servidores Unix/Linux e Windows em 10 minutos. Saiba mais detalhes sobre a documentação do Splunk na **Nota do DevMan 4**.

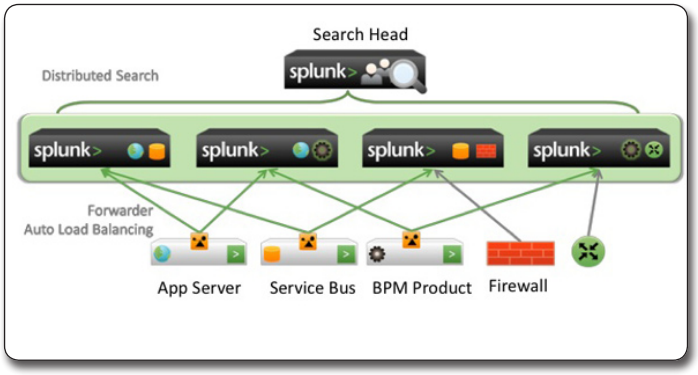


Figura 2. Splunk em três camadas, dividindo funções



Nota do DevMan 4

Documentação do Splunk

A documentação do Splunk está disponível online. Caso este seja seu primeiro contato com o produto, recomendamos que faça uma leitura rápida do Splunk Tutorial para entender os principais conceitos e já dar início à sua experiência com o produto. Após tal leitura, a documentação geral é bem indicada, pois trata dos tópicos com muito mais detalhes. O link para ambos é apresentado na seção **Links**.

Conclusão

Neste artigo vimos algumas características do funcionamento do Splunk, que é uma plataforma universal de dados de máquina que possibilita que a organização não só aprenda mais com os dados que já possui, como com os dados que produz no dia a dia de suas operações. Uma vez coletados os dados, o Splunk os armazenará em disco, em buckets, como vimos, e utilizará mecanismos de indexação destes dados para que informações sejam recuperadas com muita rapidez e facilidade. A partir de tais dados armazenados, será possível utilizar a Search App para então consolidar relatórios e gráficos para que os membros de cada time dos vários departamentos de uma empresa possam acompanhar a operação em tempo real, o que é o mais interessante.

No próximo artigo, você terá a oportunidade de aprofundar um pouco mais no Splunk, ao realizarmos uma exploração mais prática passando pelos principais tópicos como coleta de informação, análise, consulta e implementação de relatórios e gráficos.



Wagner Bianchi a.k.a. 007
wbianchi@splunk.com

Com MBA em Bancos de Dados Oracle pela Uniara (SP), Pós-graduação em Administração de Empresas e Gestão de Negócios pela Fundação Getúlio Vargas (SP/MG) e Nível Superior em Gerenciamento de Bancos de Dados pela Faculdade Infórium de Tecnologia (MG), Bianchi vem atuando como Consultor de TI com foco em servidores de bancos de dados há mais de 10 anos, tendo passado por empresas como IBM, Oracle e Percona. Atualmente é Sr Sales Engineer na empresa norte-americana Splunk, Inc., com sede em São Francisco na Califórnia, onde mantém seu foco em Big Data e Inteligência Operacional, alinhando a TI aos Negócios para obtenção de melhores resultados.

Exploring Splunk
http://www.splunk.com/goto/book

Documentação do Splunk
http://docs.splunk.com/Documentation

Download do arquivo de logs (Sampledata.zip)
http://tinyurl.com/bgmkod3

Site oficial Splunk em português
http://pt.splunk.com

Somos tão apaixonados por tecnologia que o nome da empresa

diz tudo.

Porta 80 é o melhor que a Internet pode oferecer para sua empresa.

Já completamos 8 anos e estamos a caminho dos 80, junto com nossos clientes.

Adoramos tecnologia.
Somos uma equipe composta de gente que entende e gosta do que faz,

assim como você.





Estrutura

100% NACIONAL. Servidores de primeira linha, links de alta capacidade.



Suporte diferenciado

Treinamos nossa equipe para fazer mais e melhor. Muito além do esperado.



Serviços

Oferecemos a tecnologia mais moderna, serviços diferenciados e antenados com as suas necessidades.



1-to-1

Conhecemos nossos clientes. Atendemos cada necessidade de forma única.
Conheça!

Porta 80
WEB HOSTING

Hospedagem | Cloud Computing | Dedicados | VoIP | Ecommerce | Aplicações | Streaming | Email corporativo

porta80.com.br | comercial@porta80.com.br | twitter.com/porta80

SP 4063-8616 | RJ 4063-5092 | MG 4063-8120 | DF 4063-7486