Log Analyzer: Manual de Instalação e Uso

1. Introdução

O Log Analyzer é uma ferramenta de correlação de eventos, através do cruzamento de informações contidas nos arquivos de log gerados pelo Netfilter e pelo DNS. O cruzamento de informações dessas duas fontes de dados poderá ser útil para identificar atividades maliciosas geradas por um atacante a partir da rede interna, vírus que estejam instalados em alguns hosts da rede, ou mesmo identificar um evento de tráfego que tenha ocorrido na rede.

2. Instalação

O software é composto de dois módulos básicos: um servidor de consultas, desenvolvido em linguagem C, que é responsável pelo recebimento de consultas em formato XML e pelo processamento dos arquivos de log; e uma interface web, para a interação do usuário com a ferramenta.

2.1. Componentes do Software

O arquivo principal é o logquery-0.02.tar.gz. Ele contém o código fonte do software, binários, scripts em PHP, arquivos de configuração, DTDs, e arquivos de log para a demonstração de funcionamento. Os arquivos de log contidos referem-se ao dia 29 de novembro de 2005, no período de 12:00:00 às 12:59:59. Essa restrição se aplica devido ao tamanho dos arquivos de log. Portanto, as únicas consultas válidas, utilizando os arquivos enviados como exemplo, são as consultas referentes a esse intervalo de tempo.

Para descompactar o arquivo, utilize o GNU tar em conjunto com o GNU gzip. Supondo que o arquivo já se encontra no diretório pessoal do usuário que está utilizando o sistema no momento, execute a seqüência de comandos contidos na Tabela 1.

Tabela 1. Passos iniciais.

```
weverton@freebsd:~$ ls -l logquery-0.02.tar.gz
-rw-------    1  weverton  users  71056  2006-01-31  11:10  logquery-
0.02.tar.gz
weverton@freebsd:~$ tar xzf logquery-0.02.tar.gz
weverton@freebsd:~$ cd logquery-0.02/
weverton@freebsd:~/logquery-0.02$ ls -l
total 29
-rw-r--r--  1 weverton users   297 2006-01-31 10:15 CHANGELOG
-rw-r--r--  1 weverton users 18109 2001-06-26 20:33 COPYRIGHT
drwxr-xr-x  2 weverton users   160 2005-07-27 15:27 dtd/
drwxr-xr-x  4 weverton users   104 2006-01-31 10:16 log/
-rwxr-xr-x  1 weverton users  1559 2006-01-31 10:23 logquery.sh
drwxr-xr-x  2 weverton users   104 2005-07-27 16:42 query/
drwxr-xr-x  2 weverton users   560 2006-01-31 10:28 src/
drwxr-xr-x  2 weverton users   408 2006-01-31 10:25 www/
drwxr-xr-x  2 weverton users    96 2005-12-14 12:25 xml/
weverton@freebsd:~/logquery-0.02$
```

A Tabela 2 apresenta uma breve descrição sobre cada arquivo e diretório mostrado na última listagem.

Tabela 2. Descrição da Listagem.

`CHANGELOG`	Arquivo que informa as últimas alterações relevantes feitas ao software
`COPYRIGHT`	Licença sob a qual o software está sendo distribuído
`dtd/`	Contém os arquivos de definição dos documentos XML trocados entre cliente e servidor, bem como descrição dos documentos XML que serão utilizados para fins de configuração do software.
`log/`	Diretório que contém os arquivos de log que serão processados pela ferramenta. Atualmente contém dois diretórios, *bind* e *netfilter*, que contém, respectivamente, os arquivos de log do servidor DNS e do Netfilter. Os diretórios também podem ser links simbólicos para a suas respectivas localizações na máquina na qual o software deverá ser instalado.
`logquery.sh`	Script de inicialização do software. Através do mesmo todos os parâmetros de configuração e o PATH dos diretórios que contém os arquivos de log são especificados.
`query/`	Contém exemplos simples de documentos XML de consulta que são enviados do cliente ao servidor. Útil para consulta utilizando-se o comando *telnet.*
`src/`	Contém o código fonte do software. Os binários resultantes da compilação estão atualmente armazenados neste diretório.
`www/`	Contém as páginas que formam a interface com o usuário.
`xml/`	Contém arquivos de configuração do software. Atualmente, dois arquivos de configuração são suportados, o arquivo de configuração principal e a tabela de NAT.

2.1. Requisitos Operacionais

Para que o software possa ser utilizado, os requisitos são um computador com processador com 1 GHz de clock, memória RAM de 256 Mb, sistema operacional GNU/Linux e os pacotes descritos na Tabela 3.

Tabela 3. Descrição da Listagem.

`libxml2 (2.6.16 ou superior)`	Biblioteca XML do GNOME
`zlib1g (1.2.2 ou superior)`	Biblioteca de Compressão de Arquivos
`libc6 (2.3.2 ou superior)`	Biblioteca do GNU C
`Apache (1.3.33 ou superior)`	Servidor HTTP de alta performance
`php4 (4.3.10 ou superior)`	Linguagem para criação de scripts *server-side*

O software foi testado utilizando o sistema operacional Debian GNU/Linux 3.1, com as versões acima mencionadas. Não há garantias de que o software possa funcionar com versões anteriores dos pacotes mencionados, bem como hardware inferior.

2.2. Instalação dos Binários

Os arquivos binários do software funcionam independentemente de sua localização no disco. Portanto, uma vez que o software tenha sido descompactado, o mesmo já está pronto para uso.

Para compilar o software, é necessário ter o compilador *gcc*, o GNU *make* e as biliotecas *libc6-dev*, *zlib1g-dev* e *libxml2-dev* instaladas no sistema. Para instalá-las, utilize o gerenciador de pacotes do Debian *apt-get*. Assim que essas dependencias forem satisfeitas, basta entrar no diretório src e executar o comando *make all*.

```
weverton@freebsd:~/logquery-0.02$ sudo -s
Password:
root@freebsd:~/logquery-0.02# apt-get install libc6-dev
(...)
root@freebsd:~/logquery-0.02# apt-get install zlib1g-dev
(...)
root@freebsd:~/logquery-0.02# apt-get install libxml2-dev
(...)
root@freebsd:~/logquery-0.02# apt-get install gcc
(...)
root@freebsd:~/logquery-0.02# apt-get install make
(...)
root@freebsd:~/logquery-0.02# exit
weverton@freebsd:~/logquery-0.02$ cd src
weverton@freebsd:~/logquery-0.02/src$ make all
```

Uma correção que poderá ser necessária na sua instalação do pacote libxml2-dev é em relação ao local onde os arquivos de header dessa biblioteca estão localizados. Caso o seu sistema seja Debian, o seguinte comando precisará ser executado:

```
weverton@freebsd:~/logquery-0.02/src$ sudo -s
root@freebsd:~/logquery-0.02/src#   ln   -s   /usr/include/libxml2/libxml
/usr/include/libxml
root@freebsd:~/logquery-0.02/src# exit
weverton@freebsd:~/logquery-0.02/src$
```

Um script, *logquery.sh*, é forncedido para servir como uma interface simples para a inicialização do servidor. Tais como os *daemons* de inicialização contidos no diretório */etc/init.d/*, o script *logquery.sh* pode receber os argumentos conforme mostrado na Tabela 4.

Tabela 4. Argumentos para o script de inicialização *logquery.sh*.

Start	Inicia o servidor de consultas
Stop	Encerra o servidor de consultas
Restart	Reinicia o servidor de consultas

Uma observação sobre este script é que no mesmo é configurado todas as opções que deverão ser passadas ao servidor de consultas. A Tabela 5 mostra as principais variáveis do script e seus respectivos valores padrão. Entenda-se por valor padrão o valor definido previamente nas variáveis contidas no arquivo *logquery.sh*, e não o valor que seria assumido pela aplicação, caso o parâmetro não tivesse sido especificado.

Tabela 5. Variáveis internas do script *logquery.sh*.

ADDRESS	Endereço no qual o servidor de consultas deverá escutar por requisições. O valor padrão é *0.0.0.0*.
BACKGROUND	Indica se o servidor de consultas deverá funcionar em Segundo plano. O valor padrão é *true*.
CONFIG	Indica a localização do arquivo XML de configuração. O valor padrão é *xml/conf.xml*, o qual está contido no diretório que contém o script *logquery.sh*.
DNS	Indica a localização dos arquivos de log do servidor DNS. O valor padrão é o diretório *log/bind/*, o qual está contido no diretório que contém o script *logquery.sh*.
FIREWALL	Indica a localização dos arquivos de log do Netfilter. O valor padrão é o diretório *log/netfilter/*, o qual está contido no diretório que contém o script *logquery.sh*.
NAT	Indica a localização do arquivo XML que contém as regras de NAT implementadas pelo firewall. O valor padrão é *xml/nat.xml*, o qual está contido no diretório que contém o script *logquery.sh*.
PORT	Porta TCP na qual o servidor de consultas deverá escutar por requisições. O valor padrão é 65001.
USER	Informa o usuário que o software deverá utilizar para executar. Este está desabilitado, mas pode ser habilitado adicionando a opção "–u $USER" em OPT_START, e executando o script *logquery.sh* como superusuário (root).

Um problema que pode ocorrer durante a inicialização do servidor de consultas é o mostrado na Tabela 6.

Tabela 6. Erro típico de inicialização.

```
weverton@freebsd:~/logquery-0.02$ ./logquery.sh start
Starting logquery server: logquery.
logquery: error: (bind) Address already in use
weverton@freebsd:~/logquery-0.02$
```

O erro mostrado na Tabela 6 indica que o par endereço e porta escolhidos no arquivo *logquery.sh* já estão sendo utilizados por outro processo. Portanto, basta apenas alterar o valor ou da variável ADDRESS ou PORT, ou ambos, no arquivo *logquery*.sh.

Para os fins de avaliação do software, estão sendo fornecidos, juntamente com os componentes do software, exemplos de arquivos de log, um do servidor DNS e outro do Netfilter, ambos do dia xx de xx de 2006, e uma tabela de NAT (xml/nat.xml) que contém um resumo das regras de NAT aplicadas pelo *firewall* da rede. Esses arquivos servirão como uma base de consulta para a utilização do software. O procedimento para utilizar o software com os arquivos de log gerados pela própria máquina envolve modificações no *logrotate*. Basicamente, é necessário fazer com que os arquivos de log sejam rotacionados exatamente à meia noite, todos os dias, e que cada um seja mantido em um diretório específico: os arquivos do servidor DNS ficam em um diretório exclusivo, tais como os arquivos de log do Netfilter.

2.3. Instalação da Interface com o Usuário

A interface dom o usuário é formada por páginas HTML, escritas utilizando a linguagem PHP. Para a sua instalação, basta apenas copiar o conteúdo do diretório *www/* para o diretório no qual a página deverá ser publicada. É possível que a página seja servida por um computador que não o mesmo que rodará o servidor de consultas, desde que haja uma comunicação entre os mesmos, através do protocolo IP. Os procedimentos são mostrados conforme a Tabela 7, assumindo que o diretório raiz de documentos HTML configurado atualmente no servidor Apache é */var/www/*.

Tabela 7. Instalação da interface com o usuário.

```
weverton@freebsd :~/logquery-0.02$ ls -l www/
total 88
-rw-r--r--  1 weverton users  1079 2005-11-24 11:26 common.js
-rw-r--r--  1 weverton users  2853 2005-11-24 11:26 common.php
-rw-r--r--  1 weverton users   139 2005-11-30 12:56 config.php
-rw-r--r--  1 weverton users  4220 2005-11-24 11:26 css.css
-rw-r--r--  1 weverton users 11245 2005-12-01 16:06 dns.php
-rw-r--r--  1 weverton users 12632 2005-12-01 16:06 firewall.php
-rw-r--r--  1 weverton users 14140 2005-12-05 13:45 index.php
-rw-r--r--  1 weverton users    19 2005-11-24 11:26 Makefile
-rw-r--r--  1 weverton users  5851 2005-12-01 14:05 process.php
-rw-r--r--  1 weverton users   152 2005-11-24 11:26 reload.gif
-rw-r--r--  1 weverton users   104 2005-11-24 11:26 search.gif
-rw-r--r--  1 weverton users  3711 2005-11-24 11:26 style.css
weverton@freebsd:~/logquery-0.02$ su
Password:
freebsd:/home/weverton/logquery-0.02# mkdir /var/www/logquery
freebsd:/home/weverton/logquery-0.02# cp www/* /var/www/logquery
freebsd:/home/weverton/logquery-0.02# exit
weverton@freebsd:~/logquery-0.02$
```

A Tabela 8 mostra uma descrição de cada arquivo que compõe a interface com o usuário.

Tabela 8. Descrição dos arquivos que compõem a interface com o usuário.

`common.js`	Arquivos que contém funções em JavaScript
`common.php`	Arquivo que contém funções compartilhadas entre outros arquivos de scripts PHP
`config.php`	Arquivo de configuração da interface com o usuário
`css.css`	Arquivo CSS
`dns.php`	*Front-end* para a exibição dos resultados relacionados a consultas do log do servidor DNS
`firewall.php`	*Front-end* para a exibição dos resultados relacionados a consultas do log do Netfilter
`index.php`	Página inicial
`Makefile`	Script para tarefas automatizadas
`process.php`	Script que cria o documento XML de consulta e envia ao servidor
`reload.gif`	Arquivo GIF
`search.gif`	Arquivo GIF
`style.css`	Arquivo CSS

Uma atenção principal deve ser dada ao arquivo *config.php*. Atualmente ele suporta a configuração sobre qual o endereço IP e porta no qual o servidor de consultas está escutando por requisições. Portanto, suponhamos que o servidor de consultas está escutando no IP 200.129.1.1, porta 45328. O conteúdo do arquivo *config.php* ficaria como mostrado na Tabela 9.

Tabela 9. Arquivo *config.php*.

```
weverton@freebsd:~/logquery-0.02$ cat www/config.php
<?
  $ip_dst['firewall'] = '200.129.1.1';
  $ip_dst['dns'] = '200.129.1.1';

  $port_dst['firewall'] = 45328;
  $port_dst['dns'] = 45328;
?>
weverton@freebsd:~/logquery-0.02$
```

3. Manual de Uso

Uma vez que o servidor de consultas tenha sido iniciado (através do comando *./logquery.sh start*), o software está pronto para ser utilizado. A interação com o mesmo é feita utilizando-se apenas a interface web. Supondo que o endereço do servidor HTTP no qual a página está hospedada é 10.16.2.21, basta apenas acessar no navegador o endereço *http://10.16.2.21/logquery*.

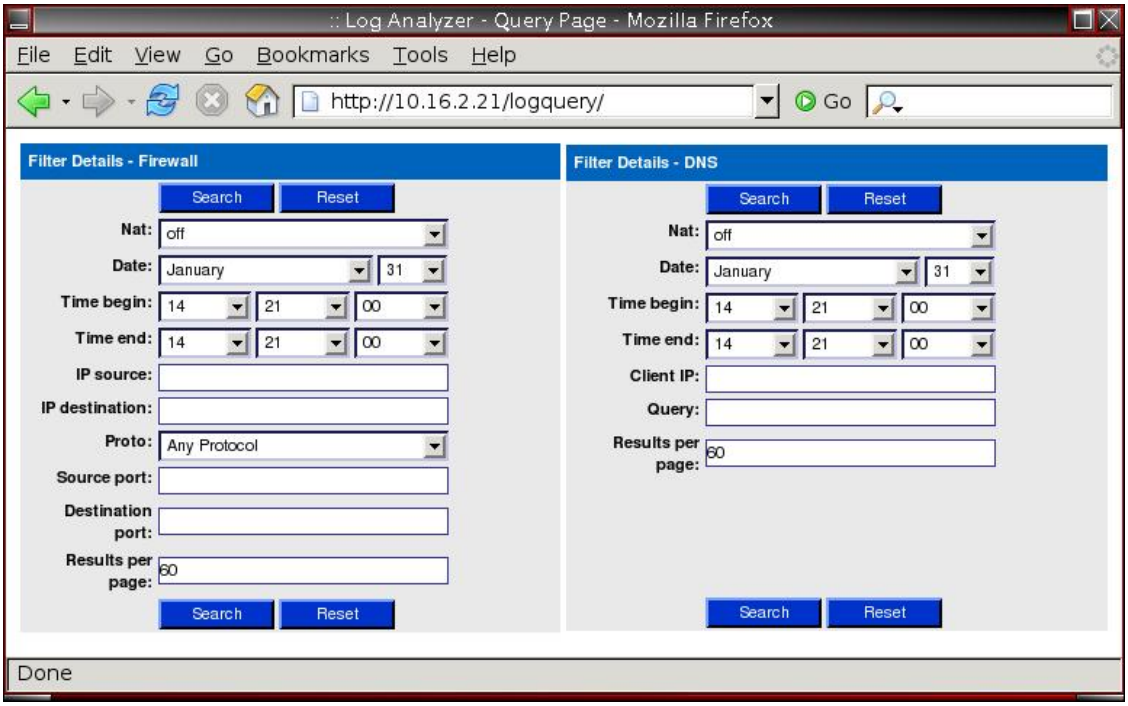

Figura 1. Interface Web do Software Log Analyzer.

A partir dessa interface é possível fazer qualquer consulta. Dois tipos de consultas podem ser efetuadas a partir desse ponto: consultas relacionadas a informações contidas em arquivos de log do Netfilter (em *Filter Details - Firewall*) e consultas relacionadas a informações contidas em arquivos de log do Bind (em *Filter Details - DNS*). As Tabelas

10 e 11 apresentam descrições sobre os campos de consulta, respectivamente, do *Filter Details – Firewall* e *Filter Details - DNS*.

Tabela 10. Descrição dos campos de consulta do Filter Details - Firewall.

Nat	Indica se a tabela de NAT sera utilizada ou não para traduzir o IP especificado em *IP source*
Date	Indica a data a que se refere a consulta
Time begin	Indica a partir de que momento desejamos consultar o arquivo de log do Netfilter
Time end	Indica até que momento desejamos consultar o arquivo de log do Netfilter
IP source	Indica o IP de origem reportado na entrada do arquivo de log do Netfilter que esta sendo processada
Ip destination	Indica o IP de destino reportado na entrada do arquivo de log do Netfilter que esta sendo processada
Proto	Indica o protocolo utilizado pelo pacote que foi registrado
Source port	Indica a porta de origem que foi utilizada no pacote registrado
Destination port	Indica a porta de destino que foi utilizada no pacote registrado
Results per page	Indica quantas entradas devem ser mostradas em uma única página. Apenas para questões de performance e visibilidade.

Tabela 11. Descrição dos campos de consulta do Filter Details - DNS.

Nat	Indica se a tabela de NAT sera utilizada ou não para traduzir o IP especificado em *IP source*
Date	Indica a data a que se refere a consulta
Time begin	Indica a partir de que momento desejamos consultar o arquivo de log do Netfilter
Time end	Indica até que momento desejamos consultar o arquivo de log do Netfilter
Client IP	Indica o endereço IP do cliente que fez a consulta naquele momento
Query	Indica parcialmente a query que o cliente do campo anterior fez ao servidor DNS
Results per page	Indica quantas entradas devem ser mostradas em uma única página. Apenas para questões de performance e visibilidade.

Levando em consideração os dados da Tabela 10, vamos fazer uma consulta, utilizando resolução de NAT, da seguinte situação: quero saber todas as entradas adicionadas ao arquivo de log do Netfilter no dia 29 de novembro, entre 12:30:00 e 12:31:00, que se referem ao host que efetuou NAT para o endereço 200.129.136.46, que tem como destino o IP 67.15.225.11, utilizando protocolo TCP, cujo pacote teve como destino a porta 80. O preenchimento do formulário ficaria como mostrado na Figura 2.

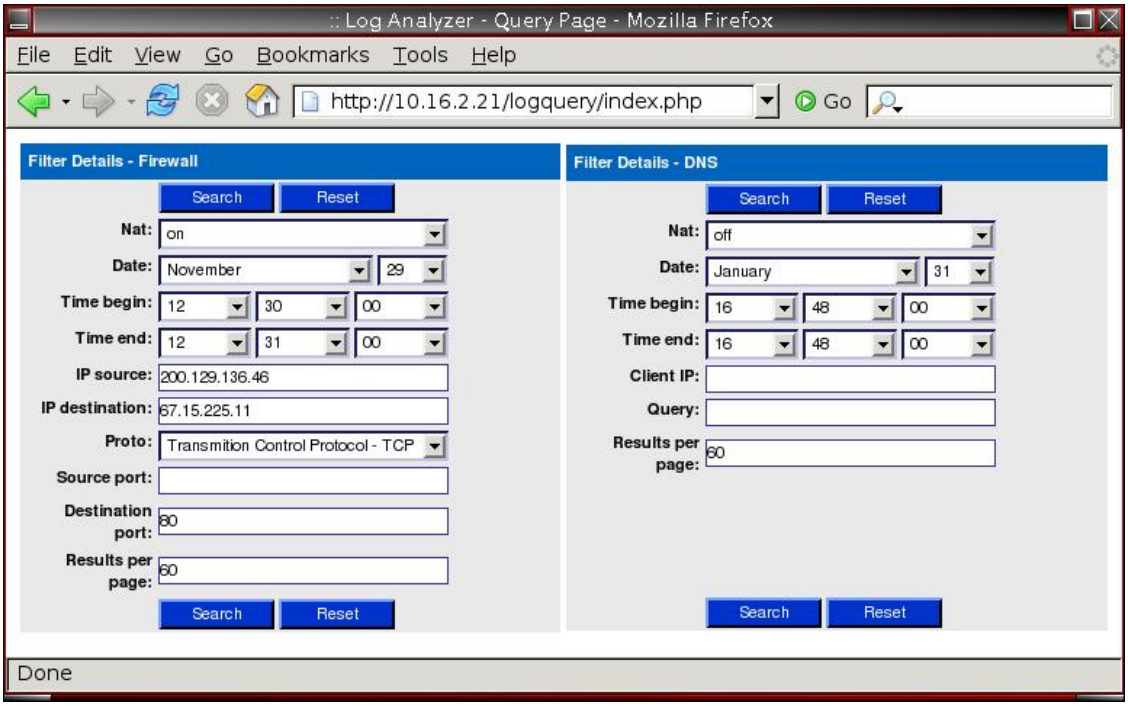

Figura 2. Exemplo de uma consulta ao log do Netfilter.

Os campos que são deixados em branco, no caso da Figura 2 o campo *Source port*, são interpretados pelo servidor de consulta como tendo qualquer valor aceitável. As restrições para campos que não podem ser omitidos são os campos *Date*, *Time begin* e *Time end*. Além disso, caso o campo *Nat* seja marcado como *on*, o campo *IP source* não poderá ser ignorado.

A consulta terá o seguinte efeito: todas as entradas registradas no dia 29 de novembro, adicionadas entre 12:30:00 e 12:31:00, serão adicionadas ao conjunto de resultados. Do conjunto de resultados, serão excluídas as entradas que contem IP de origem que não é traduzido, de acordo com a tabela de NAT, para o IP 200.129.136.46. No caso da tabela de NAT passada como parâmetro ao servidor de consulta (*xml/nat.xml*), todos os IPs que não fazem parte da rede 10.16.2.0/24 serão removidos do conjunto de resultados. Em seguida, serão excluídas do conjunto de resultados todas as entradas que não se referem ao IP de destino 67.15.225.11. Do conjunto resultante, serão excluídas as entradas que não se referem a um pacote que tenha sido destinado à porta 80, protocolo TCP. A tela de resposta seria como mostrada na Figura 3.

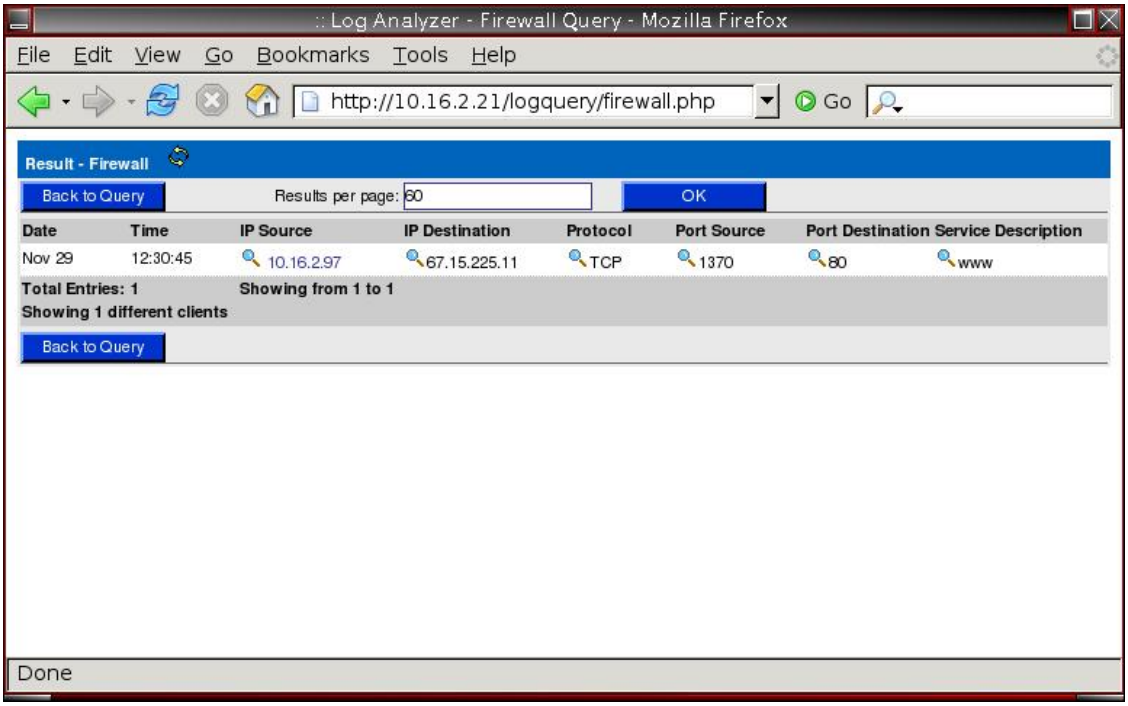

Figura 3. Resposta à consulta ilustrada na Figura 2.

Clicando no IP encontrado, 10.16.2.97, será mostrada uma nova tela, a qual informará todas as consultas ao servidor DNS que aquele cliente efetuou naquele intervalo de tempo. A tela resultante dessa nova consulta é como mostrada na Figura 4.

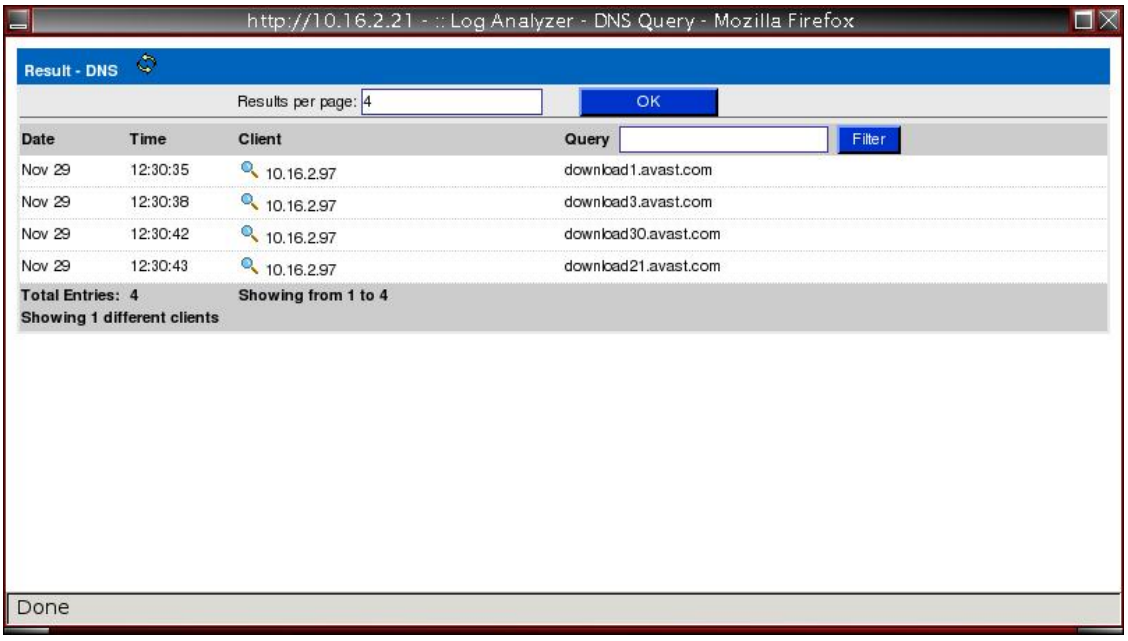

Figura 4. Correlacionamento com informações do arquivo de log do DNS.

O mesmo procedimento pode ser iniciado a partir de uma consulta ao arquivo de log do servidor DNS. Por exemplo, vamos saber quais clientes que fazem NAT para o IP 200.129.136.46 consultaram no servidor DNS, entre 12:35:00 e 12:40:00, algum

endereço relacionado ao site *orkut.com*. O formulário de consulta fica conforme mostrado na Figura 5.

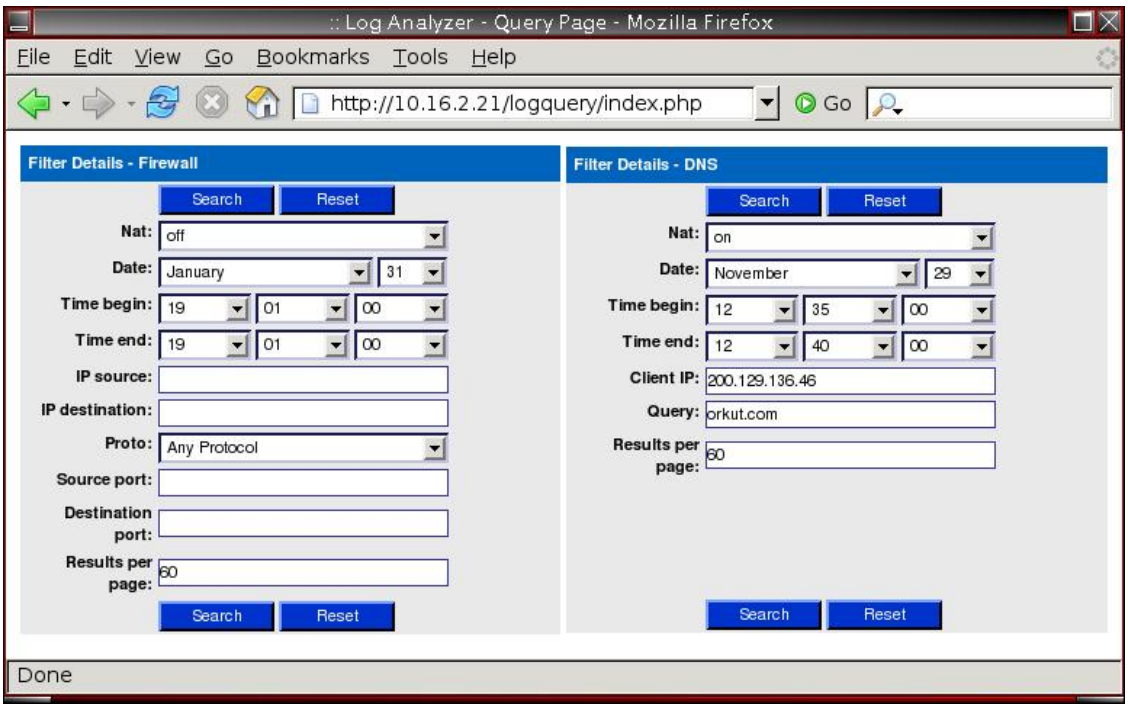

Figura 5. Exemplo de uma consulta ao log do DNS.

Acionando o botão *Search* associado à *Filter Details – DNS*, uma tela de resposta será mostrada conforme mostra a Figura 6.

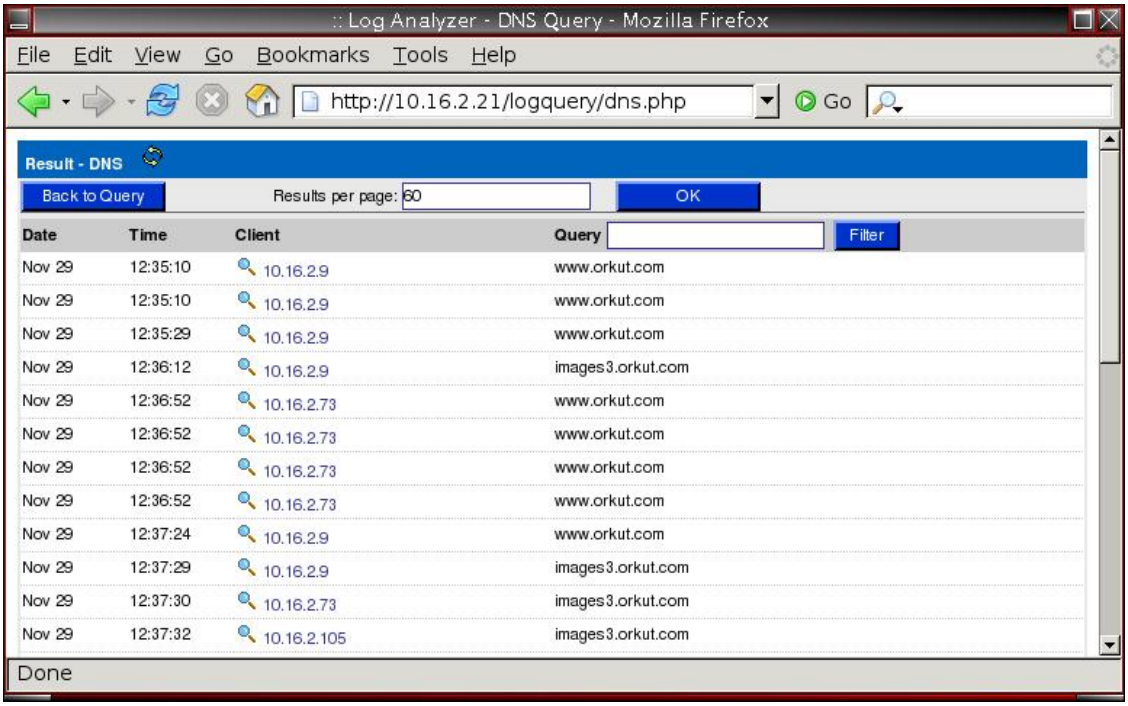

Figura 6. Resposta à consulta ilustrada na Figura 5.

Portanto, são mostrados todos as entradas que registram um cliente consultando alguma string que contenha a substring *orkut.com*. Podemos fazer um correlacionamento, cruzando a informação sobre um dos clientes encontrados e as informações processadas, no mesmo momento, pelo Netfilter. A tela resultante é ilustrada pela Figura 7.

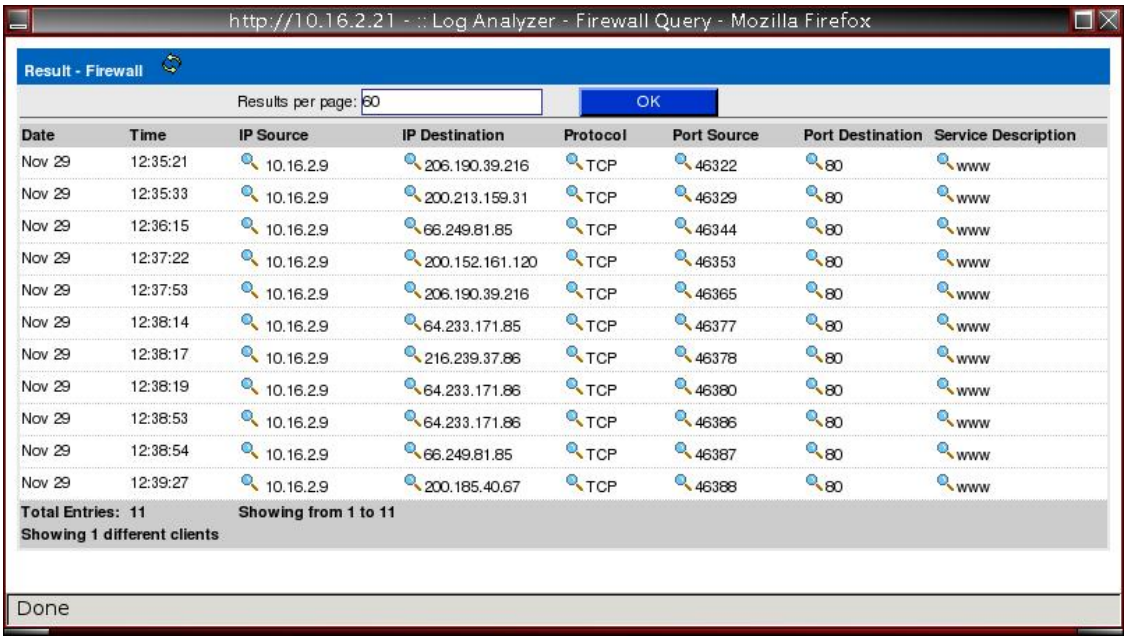

Figura 7. Correlacionamento com informações do arquivo de log do Netfilter.