FACULDADE SENAC GOIÁS

Laboratório de Redes de Computadores

Alunos: Lúcio Torres, Nayara Lopes, Silvanne Marinho

Logs

Um Servidor Web é o software que recebe uma requisição de acesso a uma página da web. Ele realiza algumas verificações de segurança na requisição HTTP e, em seguida, direciona para a página desejada. O Apache é o Servidor Web mais utilizado, sendo disponibilizado gratuitamente. Trata-se de um software de código aberto, rápido, confiável e seguro, que pode ser amplamente personalizado de acordo com as necessidades de diferentes ambientes, pelo uso de extensões e módulos. A maioria dos provedores de hospedagem WordPress utiliza o Apache.

Por sua vez, o log consiste em um processo de registo de eventos relacionados a um sistema operacional. Ele pode ser utilizado para restabelecer o estado original de um sistema ou para indicar ao administrador um evento já ocorrido. No caso dos logs de um servidor, trata-se de uma série de arquivos que são criados e mantidos por uma infraestrutura de software. Tais arquivos guardam a atividade do servidor e são utilizados para facilitar o diagnóstico de problemas ou para verificar o estado do ambiente em um determinado momento em um sistema computacional.

Logs são registros digitais que servem, inclusive como provas, no âmbito do Direito da Tecnologia da Informação, possibilitando a identificação em auditorias no ambiente virtual. Comumente, os sistemas operativos e programas incluem algum tipo de log de dados, livrando as aplicações do trabalho de mantê-los por si. Eles são de grande importância para todo tipo de serviço, devendo ter manutenção constante.

Existem diversos tipos de log, como o Error Log, Access Log, o Common Log Format, Piped Logs, Script Log, entre outros.

O Error Log é o log de erro do servidor, para onde o Apache enviará as informações de diagnóstico e gravará os erros encontrados nas solicitações de processamento. É o primeiro lugar acessado quando ocorre um problema com o início ou com alguma operação do servidor, pois ele pode ter detalhes de como ele surgiu e como corrigi-lo. O primeiro item na entrada do registro é a data e hora da mensagem. Em seguida, está o módulo que produz a mensagem e o nível de gravidade dessa mensagem.

Depois vem ID do processo e o endereço do cliente que fez o pedido. Por fim, existe a mensagem de erro detalhada.

O Access Log é um log de acesso ao servidor que registra todas as solicitações processadas pelo servidor. A localização e o conteúdo desse log são controlados pela instrução CustomLog, enquanto a instrução LogFormat pode ser utilizada para simplificar a seleção de conteúdo dos logs. Esta seção descreve como configurar o servidor para registrar informações no registro de acesso. O passo seguinte é analisar esta informação para produzir estatísticas úteis e o formato do log de acesso é altamente configurável.

O Common Log Format seria uma configuração típica para o log de acesso poder aparecer desta maneira:

```
LogFormat "%h %l %u %t \"%r\" %>s %b" common
CustomLog logs/access_log common
```

Isso define o apelido comum e o associa a uma String de formato de log específica. A cadeia de formato consiste em diretrizes percentuais, cada uma das quais diz ao servidor que registre uma determinada informação. Os caracteres literais também podem ser colocados na seqüência de formato e serão copiados diretamente para a saída de log. O caractere de citações (") deve ser escapado colocando uma barra invertida antes que ele evite que ele seja interpretado como o fim da seqüência de formato. A seqüência de formato também pode conter os caracteres de controle especiais " \ n " para nova linha e " \ t " para a guia.

Os Piped Logs são logs que aumentam a flexibilidade de um registro, por possibilitarem a capacidade de escrever erros e acessar arquivos de log por meio de um pipe ('|') para outro processo, em vez de diretamente em um arquivo. Para escrever logs em um pipe, basta substituir o nome do arquivo pelo caractere de pipe "|", seguido do nome do executável que deve aceitar entradas de log na sua entrada padrão. O servidor iniciará o processo de log de canais quando o servidor for iniciado e o reiniciará se ele falhar enquanto o servidor estiver executando.

E o Script Log serve para ajudar na depuração, a partir da instrução ScriptLog, que permite gravar a entrada e a saída dos scripts CGI, devendo ser usado apenas em testes.

Existem inúmeras opções para lidar com arquivos de log ao executar um servidor com muitos virtual hosts. Primeiramente, é possível usar logs exatamente como em um servidor de host único, bastando colocar as instruções de registro fora das seções <VirtualHost> no contexto do servidor principal, de forma que é possível registrar todas as solicitações no mesmo registro de acesso e log de erros.

Caso as diretivas CustomLog ou ErrorLog sejam colocadas dentro de uma seção <VirtualHost>, todos os pedidos ou erros para esse host virtual serão registrados apenas no arquivo especificado. Qualquer virtual host que não tenha diretrizes de registro ainda terá seus pedidos enviados para os logs do servidor principal.

Formatação de dados de log

Para a formatação dos dados de log, foi utilizado o terminal do CentOS 6.7.

Para a obtenção do arquivo de log a ser tratado, foi utilizado o sequinte comando:

```
[root@localhost ~]# wget http://pathalizer.sourceforge.net/wiki-access.log
```

Para a formatação final dos arquivos a serem utilizados pelo analisador de logs, foi utilizado:

```
[root@localhost ~]# cat wiki-access.log | awk {'print $2 $1 $4 $12'} | uniq > ip.txt
```

O comando 'cat' cria um arquivo, recebendo o texto digitado em seguida. 'wiki.access.log' é o arquivo que será tratado. O pipe é utilizado para a comunicação entre os processos, servindo como um certo tipo de encadeamento entre eles. O comando 'awk' serve para filtrar os conteúdos desejados no arquivo. O comando 'print', entre colchetes, delimita o conteúdo que será filtrado, no caso o endereço de IP, a data, o horário de acesso e o browser utilizado:

```
-222.64.146.118[19/Jun/2005:06:44:17"Mozilla/4.0
-218.84.191.50[19/Jun/2005:06:46:05"Mozilla/4.0
-202.201.245.20[19/Jun/2005:06:47:37"Mozilla/4.0
-138.243.201.10[19/Jun/2005:06:48:40"Mozilla/5.0
```

Por fim, o comando 'uniq' remove linhas duplicadas e, e em seguida, direciona o conteúdo gerado ao arquivo 'ip.txt'.