

Prática dirigida

Configurando o cache de páginas e arquivos no serviço de proxy

Uma das configurações mais importantes com relação ao desempenho do proxy e à otimização do tráfego da rede é a configuração dos caches, onde o Squid guarda as páginas e arquivos já acessados de forma a fornecê-los rapidamente quando solicitados novamente. O Squid trabalha com dois tipos de cache:

- 1- Cache rápido, feito usando parte da memória RAM do servidor;
- 2- Cache um pouco mais lento porém maior, feito no HD.

O cache na memória RAM é ideal para armazenar arquivos pequenos, como páginas html e imagens, que serão entregues instantaneamente para os clientes. O cache no HD é usado para armazenar arquivos maiores, como downloads, arquivos do Windows Update e pacotes baixados via apt-get.

O cache na memória RAM é sempre relativamente pequeno, já que o volume de memória RAM no servidor é sempre muito menor do que o espaço em disco. O cache no HD pode ser mais generoso, afinal a idéia é que ele guarde todo tipo de arquivos, principalmente os downloads grandes, que demoram para serem baixados.

A configuração do cache é feita adicionando mais algumas linhas no arquivo de configuração:

1- A configuração da quantidade de memória RAM dedicada ao cache é feita adicionando a opção "cache_mem", que contém a quantidade de memória que será dedicada ao cache. Para reservar 64 MB, por exemplo, a linha ficaria:

cache_mem 64 MB

Como regra geral, você pode reservar 32 ou 64 MB para o cache em um servidor não dedicado, que atende a apenas alguns micros (como o servidor de uma pequena rede local) e até 1/3 da memória RAM total em um servidor proxy dedicado, que atende a um volume maior de usuários (veja mais detalhes a seguir). Em um servidor com 1 GB de RAM, por exemplo, você poderia reservar até 350 MB para o cache na memória.

2- Logo depois vai a opção "maximum_object_size_in_memory", que determina o tamanho máximo dos arquivos que serão guardados no cache feito na memória RAM (o resto vai para o cache feito no HD). O cache na memória é muito mais rápido, mas como a quantidade de RAM é muito limitada, é melhor deixá-la disponível para páginas web, figuras e arquivos pequenos em geral. Para que o cache na memória armazene arquivos de até 64 KB, por exemplo, adicione a linha:

maximum_object_size_in_memory 64 KB

3- Em seguida vem a configuração do cache em disco, que armazenará o grosso dos arquivos. Por default, o máximo são downloads de até 16 MB e o mínimo é zero, o que faz com que mesmo imagens e arquivos pequenos sejam armazenados no cache. Quase sempre é mais rápido ler a partir do cache do que baixar de novo da web, mesmo que o arquivo seja pequeno.

Se você faz download de arquivos grandes com frequência e deseja que eles fiquem armazenados no cache, aumente o valor da opção "maximum_object_size". Isso é especialmente útil para quem precisa baixar muitos arquivos através do apt-get ou Windows

Update em muitos micros da rede. Se você quiser que o cache armazene arquivos de até 512 MB, por exemplo, as linhas ficariam:

maximum_object_size 512 MB
minimum_object_size 0 KB

Você pode definir ainda a porcentagem de uso do cache que fará o Squid começar a descartar os arquivos mais antigos. Por padrão, sempre que o cache atingir 95% de uso, serão descartados arquivos antigos até que a porcentagem volte para um número abaixo de 90%:

cache_swap_low 90
cache_swap_high 95

4- Depois vem a opção "cache_dir", que é composta por quatro valores. O primeiro, (**/var/spool/squid**) indica a pasta onde o Squid armazena os arquivos do cache, enquanto o segundo (**2048**) indica a quantidade de espaço no HD (em MB) que será usada para o cache. Aumente o valor se você tem muito espaço no HD do servidor e quer que o Squid guarde os downloads por muito tempo.

Continuando, os números **16** e **256** indicam a quantidade de subpastas que serão criadas dentro do diretório. Por padrão, temos 16 pastas com 256 subpastas cada uma. O número "ideal" de pastas e subpastas para um melhor desempenho varia de acordo com o sistema de arquivos usado, mas esta configuração padrão é adequada para a maioria das situações. Combinando as quatro opções, a linha ficaria:

cache_dir ufs /var/spool/squid 2048 16 256

Assim como na maioria das opções do Squid, se a linha "cache_dir" for omitida, é usada a configuração default para a opção, que é usar o diretório "/var/spool/squid" e fazer um cache em disco de 100 MB.

5- Você pode definir ainda o arquivo onde são guardados os logs de acesso do Squid. Por padrão, o Squid guarda o log de acesso no arquivo "/var/log/squid/access.log". Este arquivo é usado pelo Sarg para gerar as páginas com as estatísticas de acesso:

cache_access_log /var/log/squid/access.log

Mais uma configuração que você pode querer alterar é o padrão de atualização do cache. Estas três linhas precisam sempre ser usadas em conjunto, ou seja, você pode alterá-las, mas sempre as três precisam estar presentes no arquivo. Eliminando uma, o Squid ignora as outras duas e usa o default.

Os números indicam o intervalo (em minutos) que o Squid irá aguardar antes de verificar se um item do cache (uma página, por exemplo) foi atualizado, para cada um dos três protocolos. O primeiro número (o 15) indica que o Squid verificará (a cada acesso) se as páginas e arquivos com mais de 15 minutos foram atualizados. Ele faz uma verificação rápida, checando o tamanho do arquivo e, se o arquivo não mudou, ele continua fornecendo aos clientes o arquivo que está no cache, economizando banda da conexão

O terceiro número (o 2280, equivalente a dois dias) indica o tempo máximo, depois do qual o objeto é sempre verificado. Além do http e ftp, o Squid suporta o protocolo gopher, que era muito usado nos primórdios da internet para localizar documentos de texto, mas perdeu a relevância hoje em dia:

```
refresh_pattern ^ftp: 15 20% 2280
refresh_pattern ^gopher: 15 0% 2280
refresh_pattern . 15 20% 2280
```

Depois de adicionar todas estas configurações, o nosso arquivo de configuração já ficará bem maior:

```
http_port 3128
visible_hostname gdh
```

cache_mem 64 MB

```
maximum_object_size_in_memory 64 KB
maximum_object_size 512 MB
minimum_object_size 0 KB
cache_swap_low 90
cache_swap_high 95
```

cache_dir ufs /var/spool/squid 2048 16 256

```
cache_access_log /var/log/squid/access.log
refresh_pattern ^ftp: 15 20% 2280
refresh_pattern ^gopher: 15 0% 2280
refresh_pattern . 15 20% 2280
```

```
acl all src 0.0.0.0/0.0.0.0
acl manager proto cache_object
acl localhost src 127.0.0.1/255.255.255.255
acl SSL_ports port 443 563
acl Safe_ports port 21 80 443 563 70 210 280 488 59 777 901 1025-65535
acl purge method PURGE
acl CONNECT method CONNECT
```

```
http_access allow manager localhost
http_access deny manager
http_access allow purge localhost
http_access deny purge
http_access deny !Safe_ports
http_access deny CONNECT !SSL_ports
```

```
acl redelocal src 192.168.1.0/24
http_access allow localhost
http_access allow redelocal
http_access deny all
```

Aqui já temos uma configuração mais completa, incluindo um conjunto de regras de segurança (para que o proxy seja usado apenas a partir da rede local) e a configuração do cache. Esta é uma configuração adequada para uso em uma rede doméstica ou em um pequeno escritório.

A acl "Safe_ports", alimentada com um conjunto de portas (80, 21 e outras) é usada para restringir as portas de saída do servidor proxy, evitando, por exemplo, que ele seja usado para enviar e-mails (porta 25). Isso evita um conjunto de abusos comuns e é uma configuração importante em qualquer servidor que precise ser configurado de forma minimamente segura. Naturalmente, você pode adicionar outras portas à lista, conforme necessário.

Em uma rede maior, você provavelmente iria querer adicionar algumas limitações de acesso, limitando o acesso a algumas páginas, criando um sistema de autenticação ou limitando o uso com base no horário, entre outras possibilidades.