

Squid- Proxy/Cache



Prof. Dr. João Carlos Lopes Fernandes

E-mail: joao.fernandes1@docente.unip.br

Proxy / Cache

- Proxy
 - um agente que tem autorização para agir em nome de outro.
- Cache
 - local “disfarçado” para se preservar e esconder provisões (dados) que são inconvenientes para se transportar

O que é o *Squid* ?

- Agente que aceita solicitações de clientes (browsers) e as repassa aos servidores apropriados.
- Armazena uma cópia num cache de disco local
- Seu benefício só é sentido se o mesmo dado é requisitado várias vezes.

Squid

- É um programa proxy **HTTP** que faz cache (*caching proxy*), pois armazena os dados visitados
- Quais dados?
 - páginas HTML, sons, imagens (Objetos)
- Faz filtragem, mas não pode ser considerado como um sistema *firewall*.

Squid - Protocolos

- Desde que as requisições sejam enviadas por clientes via HTTP, Squid suporta (além de HTTP, obviamente):
 - FTP
 - *Gopher*
 - SSL (Secure Socket Layer)

Comunicação inter-cache

- Caches Squid “conversam” via UDP através dos protocolos:
 - HTTP: recuperar cópias de objetos de outros caches
 - ICP: *Internet Cache Protocol*. Tenta descobrir se um determinado objeto está em um determinado cache.
 - SNMP: Simple Network Management Protocol
 - *Cache Digest*. Recupera tabela de índices de objetos de outros caches.

Comunicação inter-cache

- Por que?
 - Base de usuário. Quanto maior a 'user base', maior a 'hit rate' (taxa de acerto). Grandes bases são possíveis se os caches cooperam
 - Redução de carga local
 - Espaço em disco. Cache usam muito (muito) disco...

Instalação

- Requisitos de hardware (em ordem decrescente de importância)
 - Disk Random Seek Time
 - Memória RAM
 - Throughput de disco sustentável
 - Poder de CPU

Quantidade de Disco

- Suponha um cache pessoal, com 1GB disco
- Navegação diária: 10 MB dados
 - 100 dias para encher o cache
- taxa de entrada influencia quantidade de disco a alocar
- Decida a quantidade de disco segundo a quantidade de dados que vai passar pelo cache por dia.

Quantidade de Disco

- Seja uma linha de 1 Mbs. -> 125000 bytes/s
- todos clientes acessando cache -> usaríamos disco na taxa de 125k por segundo -> 450 Mbytes por hora.
- Supondo maior parte do tráfego diário (8 hs)-> 3,6 GB por dia. Isso para 100% uso
- Vamos estimar, então, 2 GB Squid Cache, para cachear DIARIAMENTE os dados !

RAM

- Para otimizar busca, Squid utiliza memória para guardar tabelas do tipo Look-up.
- Cada objeto no disco consome cerca de 75 bytes de índice em RAM
- Média de um objeto na Internet 13 Kbytes. Supondo 1 GB cache -> 80000 objetos -> requerem 6 MB RAM.
- Um cache razoável (8 GB) requer 48 MB RAM !

Setup do Sistema

- Criar usuário e grupo squid
- Permissão de diretório
 -/squid/bin; ../squid/.....
- Executando 'squid -z' (../squid/cache). Devido as falhas isso pode falhar.
- Faça isso manualmente

```
mkdir /usr/local/squid/cache  
chown squid:squidadm ../cache  
chmod770 ../cache
```

Para o impacientes ...

```
Editar /etc/passwd: squid:x:1004:103:Squid Proxy:/:
```

```
Editar /etc/group: squid:x:103:
```

```
cd /usr/local/squid
```

```
chown root:squid -R etc/ ; chmod 775 -R etc/
```

```
chown squid:squid -R var/ ; chmod 770 -R var/
```

```
./sbin/squid -z (cria o diretorio de cache em  
./var/cache)
```

Configuração

- /usr/local/squid/etc/squid.conf
- porta http: **http_port 3128**
- local de cache: **cache_dir /usr/local/squid/cache/ 100 16 256**
 - 100 MB de cache; 16 diretórios, cada qual com 256 sub-diretórios

Configuração

- Usuários e grupos

`cache_effective_user squid * cache_effective_group squid *`

- Lista de controle de Acesso

- importante configurar
- Lembre-se: por *default*, tudo é negado !!!!

* Obs: pode usar UID e GID default: **nobody**

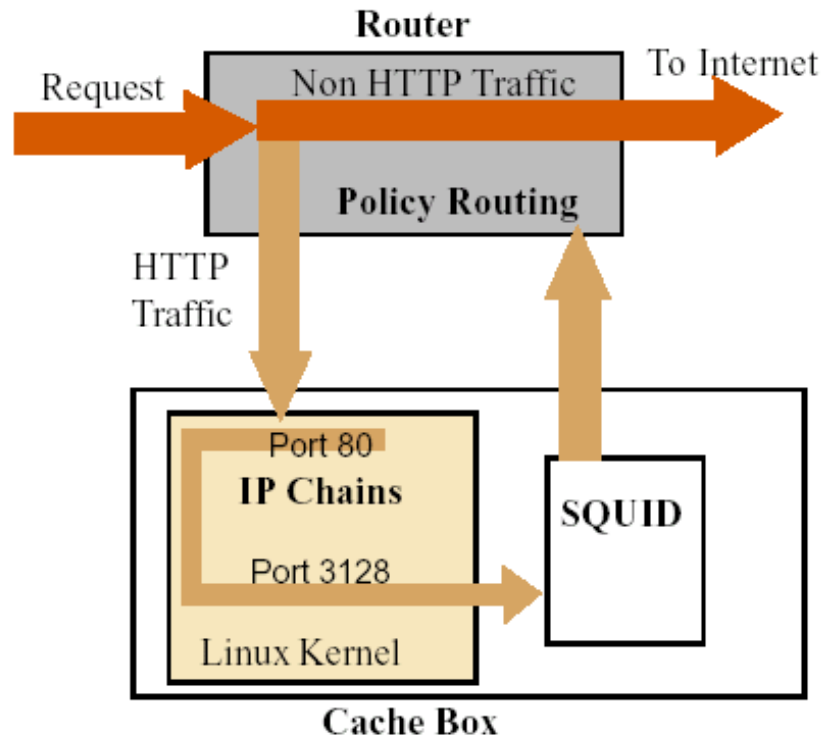
Proxy Transparente

- Clientes não precisam configurar o browser
- Squid se encarrega de interceptar os pacotes e colocá-los no cache.
- Setup é transparente, mas sua utilização não o é.
 - Há mascaramento de IPs. O IP origem do pacote será mudado !

Proxy Transparente

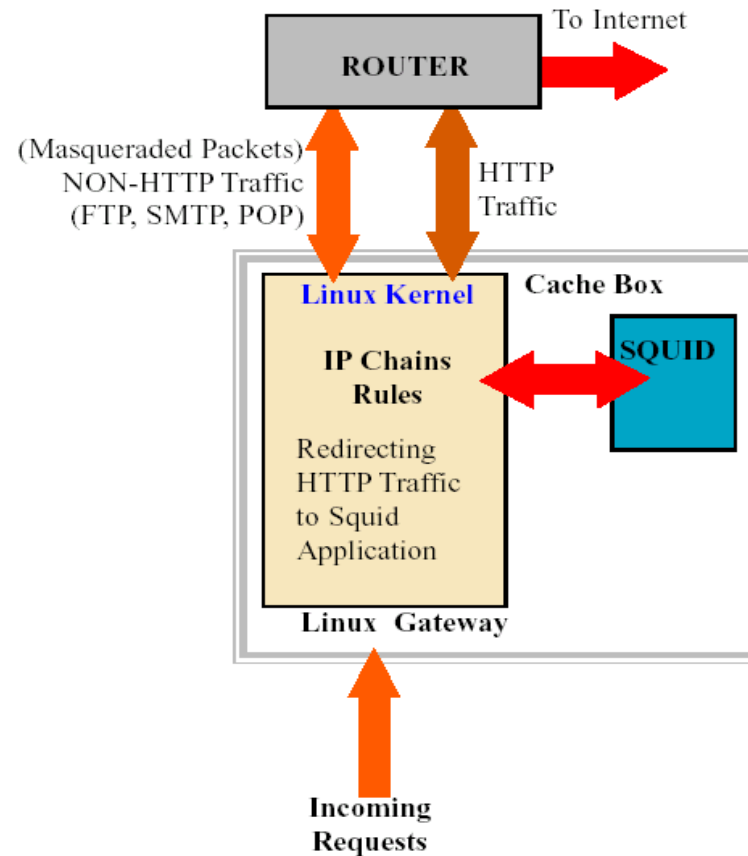
- Tipo de setup em que o squid roda na máquina que também é o gateway primário (uma possível configuração)
- É preciso:
 - configurar Kernel para transparência: redirecionar conexões porta 80 para squid
 - configurar squid (veja documentação)

Roteador que direciona tráfego IP para servidor proxy SQUID



Transparent Caching - Policy based routing

SQUID como um gateway



Configuração básica

- Redirecionamento da porta
 - iptables -t nat -A PREROUTING -p TCP --dport 80 -j REDIRECT --to-port 3128
- Modo transparente
 - httpd_accel_host **virtual**
 - httpd_accel_port **80**
 - httpd_accel_with_proxy **on**
 - httpd_accel_uses_host_header **on**

O Modo 'Acelerado'

Accelerator mode

- Squid rodar na porta 80 !
 - Formato de requisições proxy muito semelhantes às HTTPs
 - Web Server pode agir como web cache.
- Suponha que um sítio use IIS como solução Web (http e proxy ambos na porta 80), mas quer colocar squid para melhorar desempenho do serviço de proxy ???

O Modo 'Acelerado'

- Squid pode aceitar requisições no formato servidor web (trilha+nome do arquivo, não inclui o endereço do host) e repassá-las a outros servidores (apache, IIS).
- Squid aceita solicitações na porta 80 e repassa para o servidor Web na porta 81 (por exemplo)
- Isso é o que chamamos de modo acelerado.