Proxy/Cache-Squid

ACHRAF FAYAD

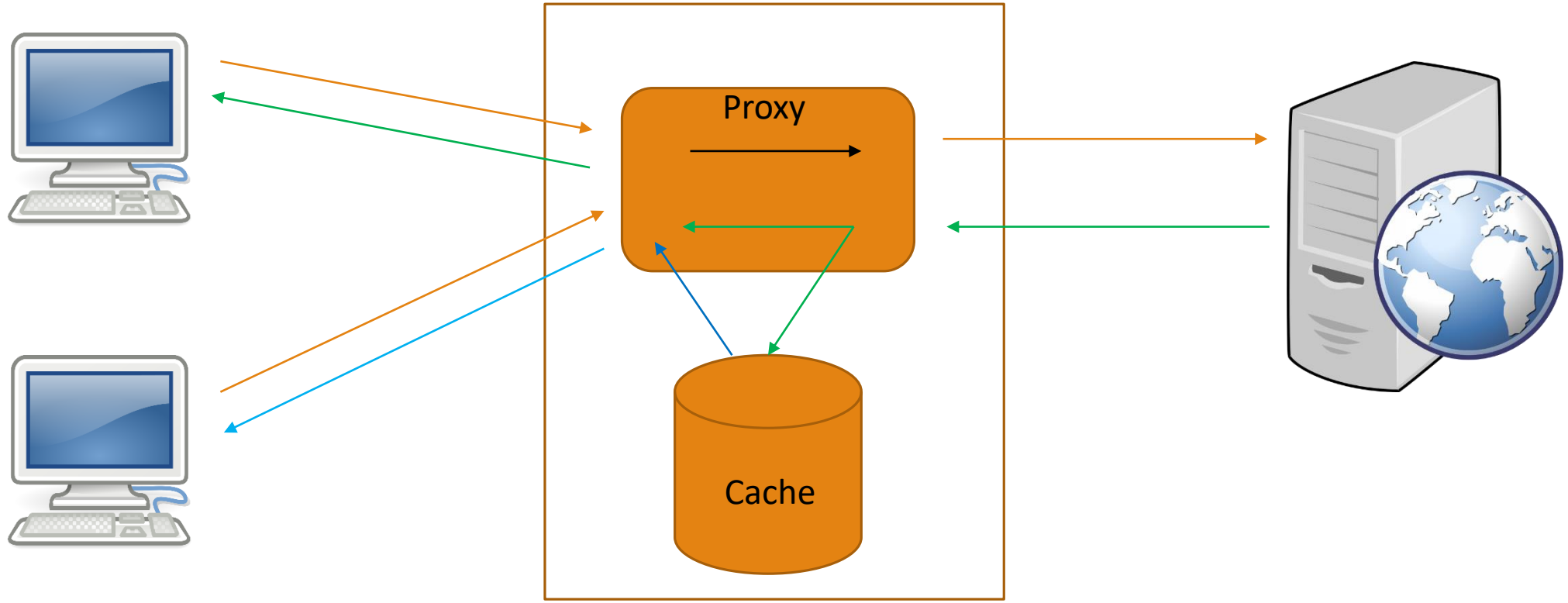
ECE

2018/2019

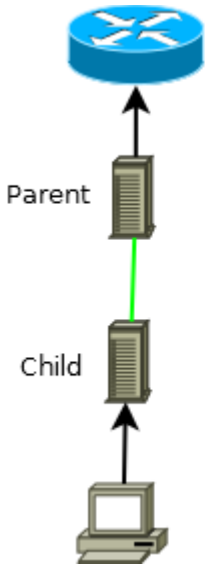
Why Proxy-Cache ?

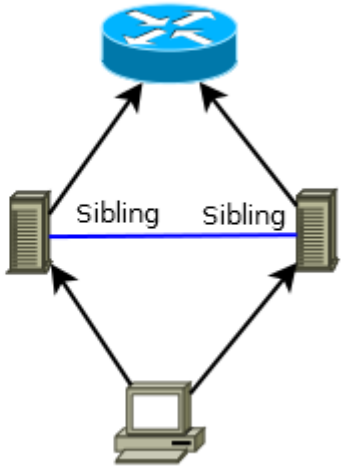
- Provides a nearby cache of Web pages and files available on remote Web servers, allowing local network clients to access them more quickly or reliably.
- Filters the content of Web pages served. Some censorware applications - which attempt to block offensive Web content - are implemented as Web proxies.
- Reformats web pages for a specific purpose or audience.
- Network operators can also deploy proxies to intercept computer viruses and other hostile content served from remote Web pages.

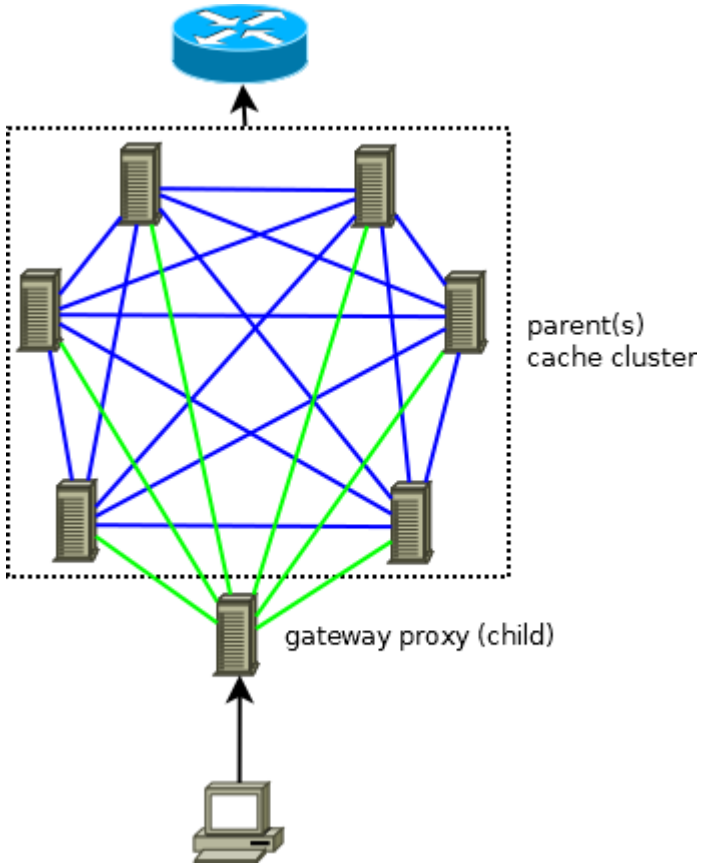
Architecture



Cache Hierarchy







Type of Proxies

- Transparent Proxy

  This type of proxy server identifies itself as a proxy server and also makes the original IP address available through the http headers. These are generally used for their ability to cache websites and do not effectively provide any anonymity to those who use them. However, the use of a transparent proxy will get you around simple IP bans. They are transparent in the terms that your IP address is exposed, not transparent in the terms that you do not know that you are using.

- Anonymous Proxy

  This type of proxy server identifies itself as a proxy server, but does not make the original IP address available. This type of proxy server is detectable, but provides reasonable anonymity for most users.

- Distorting Proxy

  This type of proxy server identifies itself as a proxy server, but make an incorrect original IP address available through the http headers.

- High Anonymity Proxy

  This type of proxy server does not identify itself as a proxy server and does not make available the original IP address.

Squid

Squid is a caching proxy for the Web supporting HTTP, HTTPS, FTP, and more. It reduces bandwidth and improves response times by caching and reusing frequently-requested web pages. Squid has extensive access controls and makes a great server accelerator. It runs on most available operating systems, including Windows and is licensed under the GNU GPL (http://www.squid-cache.org/)

Squid basic files

- **squid.conf** (/etc/squid3/)
  - file defines the configuration for squid.

  the port default of Squid is 3128/tcp

- **Squid Log File (/var/log/squid3/)**
  - cache.log
    - The cache.log file contains the debug and error messages that Squid generates
  - access.log
    - Squid result codes (separated by underscore characters)

1458500174.848 1 192.168.113.1 **TCP_MISS**/404 590 **GET** http://192.168.11.2/favicon.ico - HIER_DIRECT/192.168.11.2 text/html

1458588772.610 24 192.168.113.1 **TCP_MEM_HIT**/200 3510 **GET** http://192.168.11.2/ - HIER_NONE/- text/html

1458588991.132 167 192.168.113.1 **TCP_MISS**/200 1940 **CONNECT** 192.168.11.2:443 - HIER_DIRECT/192.168.11.2 -

http://wiki.squid-cache.org/SquidFaq/SquidLogs#Squid_Log_Files

Squid Configuration

- Configuration Cache

cache_dir ufs Directory-Name Mbytes L1 L2 [options]

'Mbytes' is the amount of disk space (MB) to use under this directory.

'L1' is the number of first-level subdirectories which will be created under the 'Directory'. The default is 16.

'L2' is the number of second-level subdirectories which will be created under each first-level directory. The default is 256.

Default configuration:

cache_dir ufs /var/spool/squid3 100 16 256

Squid Configuration

- ACL

Define criteria for a access list :

acl aclname acltype argument ...

acl aclname acltype "file" ...

- acl aclname src ip-address/mask ... # clients IP address [fast]
- acl aclname src addr1-addr2/mask ... # range of addresses [fast]
- acl aclname dst [-n] ip-address/mask ... # URL host's IP address [slow]
- acl aclname srcdomain .foo.com ... # reverse lookup, from client IP [slow]
- acl aclname dstdomain [-n] .foo.com ...
- acl aclname time [day-abbrevs] [h1:m1-h2:m2]

- More details : http://www.squid-cache.org/Doc/config/acl/

Squid Configuration

- http_access

Defining the ACLs alone does not actually block anything – it's just a definition. ACLs can be used in various places of your squid.conf. The most useful feature is the http_access statement. It works similar to the way a firewall would handle rules.

http_access (allow|deny) aclname

- Combining ACLs (AND)

http_access allow aclname1 aclname2

- Combining ACLs (OR)

http_access allow aclname1

http_access allow aclname2

Squid Configuration

Authentication in Squid :

- Users will be authenticated if squid is configured to use ***proxy_auth*** ACLs
- The Squid source code bundles with a few authentication backends ("helpers") for authentication. These include:
  - **DB:** Uses a SQL database.
  - **getpwam:** Uses the old-fashioned Unix password file.
  - **LDAP:** Uses the Lightweight Directory Access Protocol.
  - **MSNT:** Uses a Windows NT authentication domain.
  - **MSNT-multi-domain:** Allows login to one of multiple Windows NT domains.
  - **NCSA:** Uses an NCSA-style username and password file.
  - **NIS :** Uses the NIS database
  - **PAM:** Uses the Unix Pluggable Authentication Modules scheme.
  - **POP3:** Uses an email server to validate credentials. Useful for single-signon to proxy and email.
  - **RADIUS:** Uses a RADIUS server for login validation.
  - **SASL:** Uses SASL libraries.
  - **SMB:** Uses a SMB server like Windows NT or Samba.
  - **SSPI:** Windows native authenticator

HTTP Headers using Squid

Client -> Proxy

```
Internet Protocol Version 4, Src: 192.168.113.1, Dst: 192.168.113.129
Transmission Control Protocol, Src Port: 63658 (63658), Dst Port: 3128 (3128), Seq: 1, Ack: 1, Len: 368
Hypertext Transfer Protocol
  GET http://192.168.11.2/ HTTP/1.1\r\n
  Host: 192.168.11.2\r\n
  User-Agent: Mozilla/5.0 (Windows NT 10.0; WOW64; rv:39.0) Gecko/20100101 Firefox/39.0\r\n
  Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8\r\n
  Accept-Language: fr,fr-FR;q=0.8,en-US;q=0.5,en;q=0.3\r\n
  Accept-Encoding: gzip, deflate\r\n
  Proxy-Authorization: Basic dGVzdDp0ZXN0\r\n
     Credentials: test:test
  Connection: keep-alive\r\n
  \r\n
```

Proxy -> HTTP Server

```
GET / HTTP/1.1\r\n
Host: 192.168.11.2\r\n
User-Agent: Mozilla/5.0 (Windows NT 10.0; WOW64; rv:39.0) Gecko/20100101 Firefox/39.0\r\n
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8\r\n
Accept-Language: fr,fr-FR;q=0.8,en-US;q=0.5,en;q=0.3\r\n
Accept-Encoding: gzip, deflate\r\n
Via: 1.1 proxy.lab.com (squid/3.4.8)\r\n
X-Forwarded-For: 192.168.113.1\r\n
Cache-Control: max-age=259200\r\n
Connection: keep-alive\r\n
\r\n
```

Proxy Auto-Configuration

A Proxy Auto-Configuration **(PAC)** file contains a set of rules coded in JavaScript which allows a web browser to determine whether to send web traffic direct to the Internet or be sent via a proxy server.