

马  
哥  
教  
育

# 网站架构演化

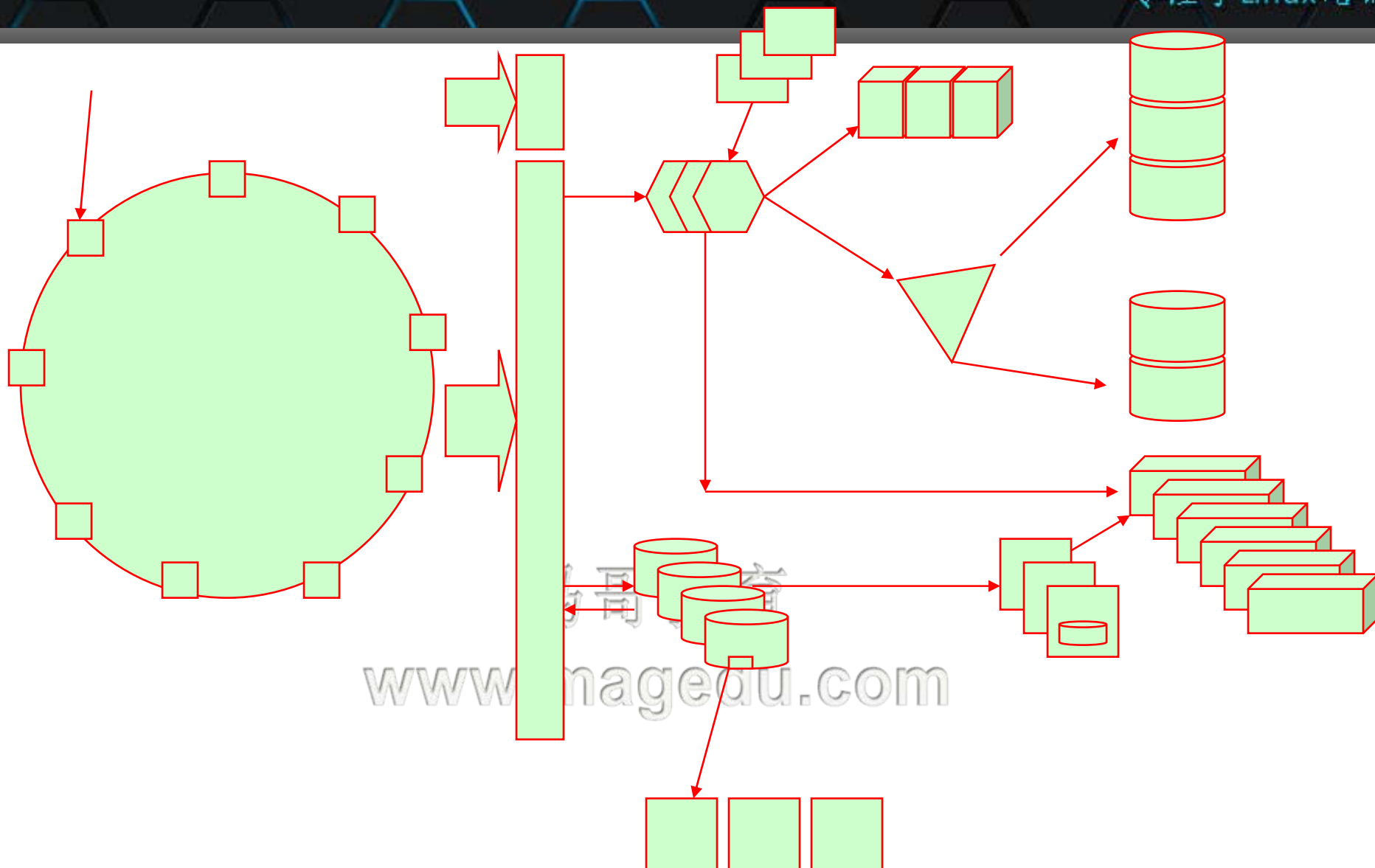
主讲：马永亮(马哥)

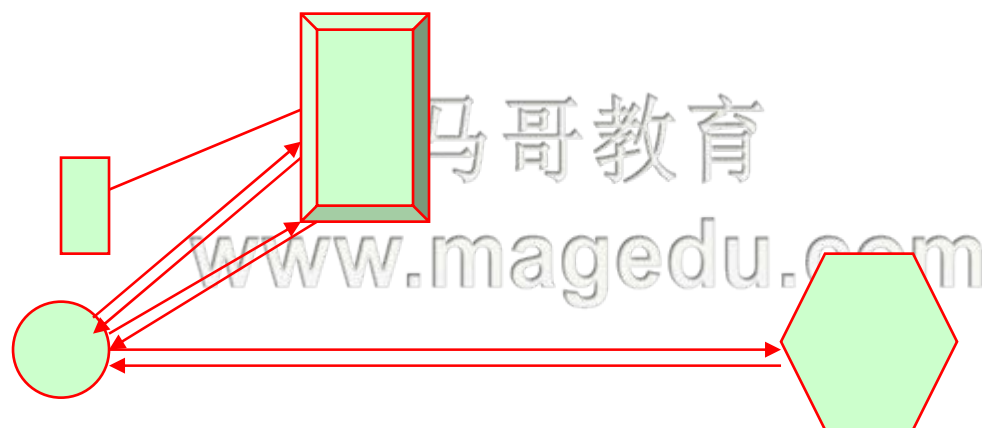
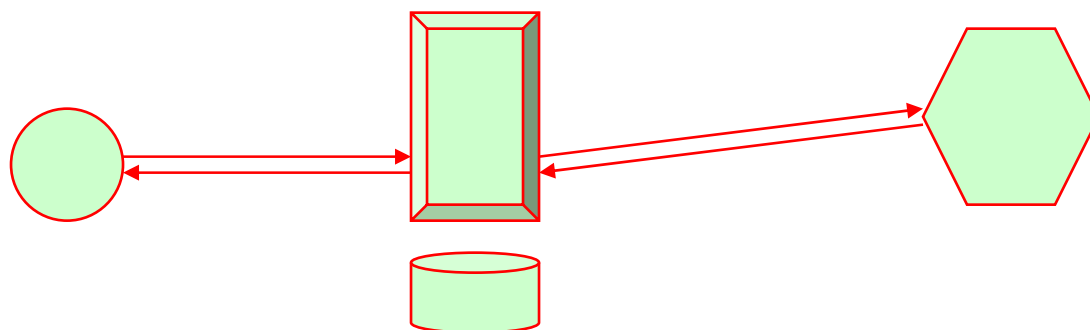
QQ群:169777636

客服QQ: 2813150558, 1661815153

<http://www.magedu.com>

<http://mageedu.blog.51cto.com>





马哥教育

www.magedu.com

马哥教育

# Web Cache

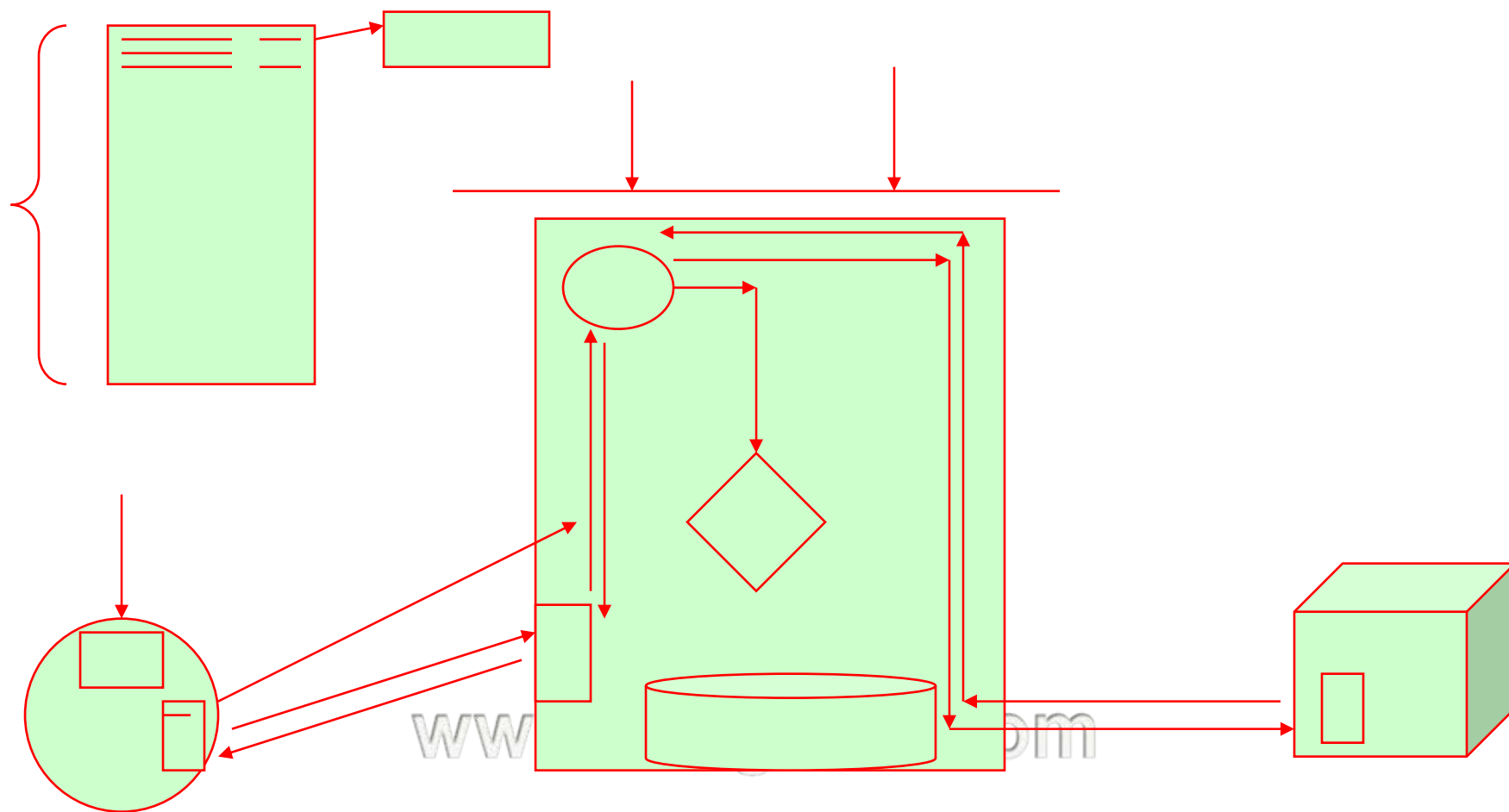
主讲：马永亮(马哥)

QQ群:169777636

客服QQ: 2813150558, 1661815153

<http://www.magedu.com>

<http://mageedu.blog.51cto.com>



马哥教育

www.magedu.com



马  
哥  
教  
育

# Squid

主讲：马永亮(马哥)

QQ群：169777636

客服QQ：2813150558, 1661815153

<http://www.magedu.com>

<http://mageedu.blog.51cto.com>



## ❖ Nginx

➡ Reverse proxy

➡ LNAMP = Nginx → Apache+php+MySQL

➡ Php: CPU

➡ MySQL: CPU

➡ Php → mysql.sock ← MySQL

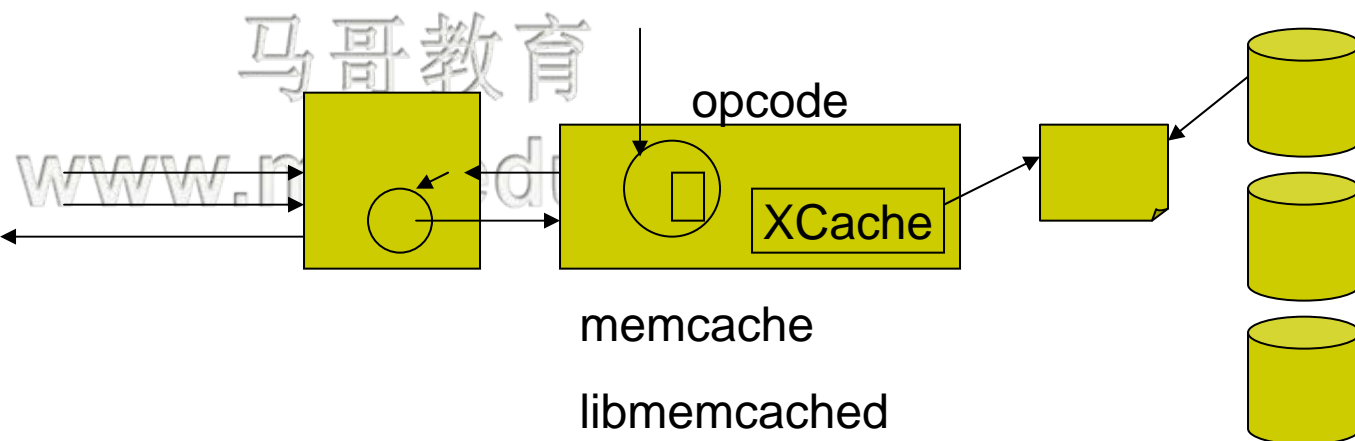
➡ Php → Memcached ← MySQL (query cache)

➡ LNAMP

• LNMP

• LAMP

▪ PHP



# What is Squid?



- ❖ Squid的诞生
- ❖ Squid功能
- ❖ An ever-growing number of companies use Squid to save on their internet web traffic, improve performance, deliver faster browsing to their end-clients and provide static, dynamic and streaming content to millions of internet users worldwide
- ❖ Squid 的特点
- ❖ offers a rich access control, authorization and logging environment to develop web proxy and content serving applications

## ❖ What are cachable objects?

➡ An Internet Object is a file, document or response to a query for an Internet service such as FTP, HTTP, or gopher

Not dns

➡ A client requests an Internet object from a caching proxy

➡ if the object is not already cached, the proxy server fetches the object (either from the host specified in the URL or from a parent or sibling cache) and delivers it to the client

www.magedu.com

## ❖ ICP protocol: Internet Cache Protocol

- ➡ Used for communication among squid caches
- ➡ ICP is primarily used within a cache hierarchy to locate specific objects in sibling caches
- ➡ ICP is currently implemented on top of 3130/UDP

马哥教育

www.magedu.com

❖ [www.squid-cache.org](http://www.squid-cache.org)

❖ squid软件包

- ➔ 软件包名: `squid-2.6.STABLE6`
- ➔ 服务名: `squid`
- ➔ 主程序: `/usr/sbin/squid`
- ➔ 配置目录: `/etc/squid/`
- ➔ 主配置文件: `/etc/squid/squid.conf`
- ➔ 默认监听端口: `TCP 3128`
- ➔ 默认访问日志文件: `/var/log/squid/access.log`

www.magedu.com

# 主配置文件squid.conf

## ❖ 常用配置项

- ➔ http\_port 3128
- ➔ cache\_mem 64 MB
- ➔ maximum\_object\_size 4096 KB
- ➔ reply\_body\_max\_size 10240000 allow all
- ➔ access\_log /var/log/squid/access.log squid
- ➔ visible\_hostname proxy.test.com

为缓存目录分配的  
磁盘空间（MB）

- ➔ cache\_dir **ufs** /var/spool/squid **100 16 256**

缓存数据的存  
储格式

缓存空间的一级  
子目录个数

缓存空间的二级  
子目录个数

❖ `export http_proxy=http://proxy_ip_address:port`

马哥教育

www.magedu.com



- ❖ `acl all 0.0.0.0/0.0.0.0`
- ❖ `http_access deny all`

马哥教育  
www.magedu.com

## ❖ ACL (Access Control List, 访问控制列表)

➡ 可以从客户机的**IP**地址、请求访问的**URL/域名/文件类型**、访问时间、并发请求数等各方面进行控制

## ❖ 应用访问控制的方式

➡ 定义**acl**列表

**acl** 列表名称 列表类型 列表内容 ...

➡ 针对**acl**列表进行限制

**http\_access allow或deny** 列表名.....

www.magedu.com

## ❖ 最基本的ACL访问控制示例

➡ 禁止任何客户机使用代理服务

```
acl all src 0.0.0.0/0.0.0.0
```

```
http_access deny all
```

马哥教育

www.magedu.com

❖ `reply_body_max_size 10240000 allow all`

马哥教育  
www.magedu.com

## ❖ 常用的acl列表类型

➔ src

➔ dst

➔ port

➔ srcdomain

➔ dstdomain

➔ time

➔ maxconn

➔ url\_regex

➔ urlpath\_regex

马哥教育

www.magedu.com

TAG NAME	http_access
Description	Using this, you can allow or deny the access lists defined by <a href="#">acl</a>
Build Option	Default
Usage	<b>http_access allow deny [!] aclname ...</b>
Default	http_access deny all

## Synopsis

This is used for filtering based on the acl matchings. If none of the "access" lines cause a match, the default is the opposite of the last line in the list. If the last line was deny, then the default is allow. Conversely, if the last line is allow, the default will be deny. For these reasons, it is a good idea to have an "deny all" or "allow all" entry at the end of your access lists to avoid potential confusion.

## Arguments

<i>allow deny</i>	Allow or deny on matching the acl
<i>aclname</i>	Access list to be allowed/denied on match

## Example(s)

[To allow http\\_access for only one machine with MAC Address 00:08:c7:9f:34:41](#)

[To restrict access to work hours \(9am - 5pm, Monday to Friday\) from IP 192.168.2/24](#)

[Can i use multitime access control list for different users for different timing](#)

[Rules are read from top to bottom](#)

## Note

The deny all line is very important. After all the http\_access rules, if access isn't denied, it's ALLOWED !! So, specifying a LOT of http\_access allow rules, and forget the deny all after them, is the same of NOTHING. If access isn't allowed by one of your rules, the default action ( ALLOW ) will be triggered. So, don't forget the deny all rule AFTER all the rules.

And, finally, don't forget rules are read from top to bottom. The first rule matched will be used. Other rules won't be applied.

- ❖ `acl worktime time 08:00-12:00`
- ❖ `acl worktime time 14:00-18:00`
- ❖ `http_access allow worktime mynetwork`

马哥教育  
www.magedu.com



TAG NAME	acl
Description	Used to define an Access List
Build Option	Default
Usage	<b>acl</b> aclname acltype string1 ... <b>acl</b> aclname acltype "file"
Default	Required minimum configuration for normal functioning

### Synopsis

The first field represents the name of the ACL, which is simply an assigned name, that can be just about anything the user chooses. The second field is the type of the ACL, which can be one of a number of choices, that indicates to Squid what part of a request should be matched against for this ACL. The possible types include the requesting clients address, the Web server address or host name, a regular expression matching the URL, and many more. The final field is the actual string to match. Depending on what the ACL type is, this may be an IP address, a series of IP addresses, a URL, a host name, etc.

When using "file", the file should contain one item per line.

### Arguments

<i>aclname</i>	Name of the acl
<i>acltype</i>	Type of acl
<i>string</i>	String to match
<i>file</i>	File that containing strings

www.magedu.com

ACL TYPE	arp
Description	Ethernet (MAC) address matching
Build Option	--enable-arp-acl
Usage	<b>acl</b> aclname <b>arp</b> mac-address

### Synopsis

This option only works for clients on the same local subnet, and only for certain platforms. Linux, Solaris, and some BSD variants are the supported operating systems for this type of ACL. This ACL can provide a somewhat secure method of access control, because MAC addresses are usually harder to spoof than IP addresses, and you can guarantee that your clients are on the local network (otherwise no ARP resolution can take place).

### Arguments

<i>aclname</i>	Access list name
<i>mac-address</i>	Physical address to be matched

### Example(s)

```
acl ACLARP arp 11:12:13:14:15:16
```

ACLARP refers MACADDRESS of the ethernet 11:12:13:14:15:16

### Note

Squid can only determine the MAC address for clients that are on the same subnet. If the client is on a different subnet, then Squid cannot find out its MAC address.

ACL TYPE	url_regex
Description	Matches using a regular expression on the complete URL
Build Option	Default
Usage	<b>acl</b> aclname <b>url_regex</b> [-i] pattern

### Synopsis

This ACL can be used to provide access control based on parts of the URL or a case insensitive match of the URL, and much more. The regular expressions used in Squid are provided by the GNU Regex library which is documented in the section 7 and 3 regex manpages. Regular expressions are also discussed briefly in a nice article by Guido Socher at LinuxFocus.

### Arguments

<i>aclname</i>	Access list name
<i>pattern</i>	Matching to be done for

### Example(s)

```
acl reject url_regex -i ^http://www.google.co.in/index.html
```

reject contains the entire url <http://www.google.co.in/index.html>. If in the [http\\_access](#) if you deny reject, it will deny only the url <http://www.google.co.in/index.html> but allows the url <http://www.google.co.in>

www.magedu.com

ACL TYPE	urlpath_regex
Description	This ACL matches on the URL path minus any protocol, port, and host name information
Build Option	Default
Usage	<b>acl</b> aclname <b>urlpath_regex</b> [-i] pattern

## Synopsis

This acl type expects for the particular pattern alone from the given URLs. If found the words, it will match it.

## Arguments

<i>aclname</i>	Access list name
<i>pattern</i>	Pattern which is expected for match

## Example(s)

```
acl reject url_regex -i index.html
```

reject contains the pattern index.html. If in the [http\\_access](#) if you deny reject, then for any url containing the pattern index.html will be denied.

www.magedu.com

ACL TYPE	port
Description	This ACL matches on the destination port for the request
Build Option	Default
Usage	<b>acl</b> aclname <b>port</b> port-number <b>acl</b> aclname <b>port</b> range

## Synopsis

Matching done on the destination port.

## Arguments

<i>aclname</i>	Access list name
<i>port-number</i>	Destination port number
<i>range</i>	Range of port numbers

## Example(s)

The following allows access only to the destination 172.16.1.115:80 from network 172.16.1.0

```
acl acceleratedhost dst 172.16.1.115/255.255.255.255
```

```
acl acceleratedport port 80
```

```
acl mynet src 172.16.1.0/255.255.255.0
```

```
http_access allow acceleratedhost acceleratedport mynet
```

```
http_access deny all
```

ACL TYPE	proto
Description	This ACL matches on the protocol of the request
Build Option	Default
Usage	<b>acl</b> aclname <b>proto</b> protocol

## Synopsis

Matching done based on protocol used in the request such as FTP, HTTP, ICP, etc.

## Arguments

<i>aclname</i>	Access list name
<i>protocol</i>	Protocol to be matched

## Example(s)

`acl myproto proto HTTP FTP` - This refer protocols HTTP and FTP



ACL TYPE	browser
Description	A regular expression that matches the clients browser type based on the user agent header
Build Option	Default
Usage	<b>acl</b> aclname <b>browser</b> [-i] regexp

### Synopsis

This allows for ACL's to operate based on the browser type in use, for example, using this ACL type, one could create an ACL for Netscape users and another for Internet Explorer users.

### Arguments

<i>aclname</i>	Access list name
<i>regexp</i>	Browser name

### Example(s)

acl aclname browser MOZILLA

This refers to the requests, which are coming from the browsers who have "MOZILLA" keyword in the user-agent header.

马哥教育

www.magedu.com



ACL TYPE	maxconn
Description	Matches when the client's IP address has more than the specified number of HTTP connections established
Build Option	Default
Usage	<b>acl</b> aclname <b>maxconn</b> number

### Synopsis

Matching is true when the defined number of HTTP connections has been established for a client.

### Arguments

<i>aclname</i>	Access list name
<i>number</i>	Maximum number of HTTP connections

### Example(s)

```
acl someuser src 1.2.3.4
```

```
acl twoconn maxconn 5
```

```
http_access deny someuser twoconn
```

```
http_access allow !twoconn
```

The above configuration will allow a maximum of 5 http connection to access.

www.magedu.com

ACL TYPE	src
Description	The IP address of the requesting client, or the clients IP Address
Build Option	Default
Usage	<b>acl</b> aclname <b>src</b> ip-address/netmask ... (clients IP address) <b>acl</b> aclname <b>src</b> addr1-addr2/netmask ... (range of addresses)

## Synopsis

Matching done based on clients ip address

## Arguments

<i>aclname</i>	Access list name
<i>ip-address</i>	Clients ip address
<i>netmask</i>	Network mask
<i>add1-addr2</i>	Range of ip address

## Example(s)

`acl network1 src 172.16.1.0/24` - refers to the whole Network with address 172.16.1.0

`acl source1 src 172.16.1.25/32` - refers to a single source

`acl srcrange src 172.16.1.25-172.16.1.35/32` - refers range of IP Addresses from 172.16.1.25-172.16.1.35

ACL TYPE	dst
Description	Same as src but looks for destination IP Address
Build Option	Default
Usage	<b>acl</b> aclname <b>dst</b> ip-address/netmask ... (URL host's IP address)

## Synopsis

First Squid will dns-lookup for IPAddress from the domain-name, which is in request header. Then this acl is interpreted.

## Arguments

<i>aclname</i>	Access list name
<i>ip-address</i>	ip-address of the origin server/machine
<i>netmask</i>	Network mask

## Example(s)

```
acl dest dst 63.194.81.67/32
```

www.magedu.com

ACL TYPE	myip
Description	The local IP address on which the client connection exists.
Build Option	Default
Usage	<b>acl</b> aclname <b>myip</b> ip-address/netmask ... (local socket IP address)

### Synopsis

Using this acl type, we can specify the local IP address on which the clients connections exists. This allows ACLs to be constructed that only match one physical network, if multiple interfaces are present on the proxy, among other things.

### Arguments

<i>aclname</i>	Access list name
<i>ip-address</i>	ip address of the local machine
<i>netmask</i>	Network mask

### Example(s)

```
acl myip1 myip 172.16.1.53/32
```

→ 可访问

www.magedu.com

ACL TYPE	srcdomain
Description	This acl type provides matches against the client domain name
Build Option	Default
Usage	<b>acl</b> aclname <b>srcdomain</b> .domain-name ...

## Synopsis

Matching can be either a single domain name, or a list of domain names, or the path to a file that contains a list of domain names. If a path to a file, it must be surrounded by parentheses. This ACL type can increase the latency, and decrease throughput significantly on a loaded cache, as it must perform an address-to-name lookup for each request, so it is usually preferable to use the Client IP Address type.

Note: Here "." is more important.

## Arguments

<i>aclname</i>	Access list name
<i>domain-name</i>	Clients domain name

## Example(s)

```
acl mydomain srcdomain .kovaiteam.com
```

www.magedu.com

ACL TYPE	dstdomain
Description	This refers to destination domain i.e. the source domain where the origin server is located
Build Option	Default
Usage	<b>acl</b> aclname <b>dstdomain</b> .domain-name ...

## Synopsis

When matching to be done based on the destination servers domain, you can use this acl type.

Note: Here "." is more important.

## Arguments

<i>aclname</i>	Access list name
<i>domain-name</i>	Destination domain

## Example(s)

```
acl domain1 dstdomain .kovaiteam.com
```

This looks for \*.kovaiteam.com from URL

www.magedu.com



ACL TYPE	srcdom_regex
Description	Matches against the client domain name
Build Option	Default
Usage	<b>acl</b> aclname <b>srcdom_regex</b> [-i] pattern ...

### Synopsis

Matching can be either a single domain name, or a list of domain names, or a path to a file that contains a list of domain names. If a path to a file is specified, it must be surrounded parentheses.

### Arguments

<i>aclname</i>	Access list name
<i>pattern</i>	Pattern to match

### Example(s)

acl aclname srcdom\_regex kovai - This looks for the word kovai from the client domain name

马哥教育

www.magedu.com



ACL TYPE	dstdom_regex
Description	Provides match against destination domain
Build Option	Default
Usage	<b>acl</b> aclname <b>dstdom_regex</b> [-i] pattern ...

## Synopsis

Matching done on destination domain based on regular expression.

## Arguments

<i>aclname</i>	Access list name
<i>pattern</i>	Pattern to match

## Example(s)

acl domain1 dstdom\_regex redhat - This looks for the word redhat from the client's request.

www.magedu.com

ACL TYPE	time
<b>Description</b>	This type is just what it sounds like, providing a means to create ACLs that are active during certain times of the day or certain days of the week
<b>Build Option</b>	Default
<b>Usage</b>	<b>acl</b> aclname <b>time</b> [day-abbrevs] [h1:m1-h2:m2]

## Synopsis

Provides timed filter settings. This feature is often used to block some types of content or some sections of the internet during business or class hours. Many companies block pornography, entertainment, sports, and other clearly non-work related sites during business hours, but then unblock them after hours. This might improve workplace efficiency in some situations (or it might just offend the employees). This ACL type allows you to enter days of the week and a time range, or select all hours of the selected days.

<b>Day-abbrevs:</b>	
S	Sunday
M	Monday
T	Tuesday
W	Wednesday
H	Thursday
F	Friday
A	Saturday
<b>Time:</b>	
h1:m1 must be less than h2:m2	

## Arguments

<i>aclname</i>	Access list name
<i>day-abbrevs</i>	Days list (as shown above)
<i>h1:m1-h2:m2</i>	from and to time

## Example(s)

acl acltime time M 9:00-17:00 - acltime refers day of Monday and time from 9:00 to 17:00.

- ❖ localnet,
- ❖ denytime: 18:00-23:59

马哥教育  
www.magedu.com

## ❖ ACL列表定义示例

- ➔ `acl LAN1 src 192.168.10.0/24`
- ➔ `acl PC1 src 192.168.10.12/32`
- ➔ `acl Blk_Domain dstdomain .qq.com`
- ➔ `acl Work_Hours time MTWHF 08:30-17:30`
- ➔ `acl Max20_Conn maxconn 20`
- ➔ `acl Blk_URL url_regex -i ^rtsp:// ^mms://`
- ➔ `acl Blk_Words urlpath_regex -i sex adult`
- ➔ `acl RealFile urlpath_regex -i \.rmvb$ \.rm$`

www.magedu.com

- ❖ 根据已经定义的部分**ACL**列表进行访问控制
  - ➔ `http_access deny LAN1 Blk_URL`
  - ➔ `http_access deny LAN1 Blk_Words`
  - ➔ `http_access deny PC1 RealFile`
  - ➔ `http_access deny PC1 Max20_Conn`
  - ➔ `http_access allow LAN1 Work_Hours`

马哥教育

www.magedu.com

## ❖ 访问控制规则的匹配顺序

### ➡ 没有设置任何规则时

—— 将拒绝所有客户端的访问请求

### ➡ 有规则但找不到相匹配的项时

—— 将采用与最后一条规则相反的权限，即如果最后一条规则是 **allow**，那么就拒绝客户端的请求，否则允许该请求

马哥教育

www.magedu.com

# 配置透明代理

## ❖ 实现透明代理的基本条件

### ➡ 前提：

- 客户机的**Web**访问数据要能经过防火墙
- 代理服务构建在网关（防火墙）主机中

### ➡ 配置要求：

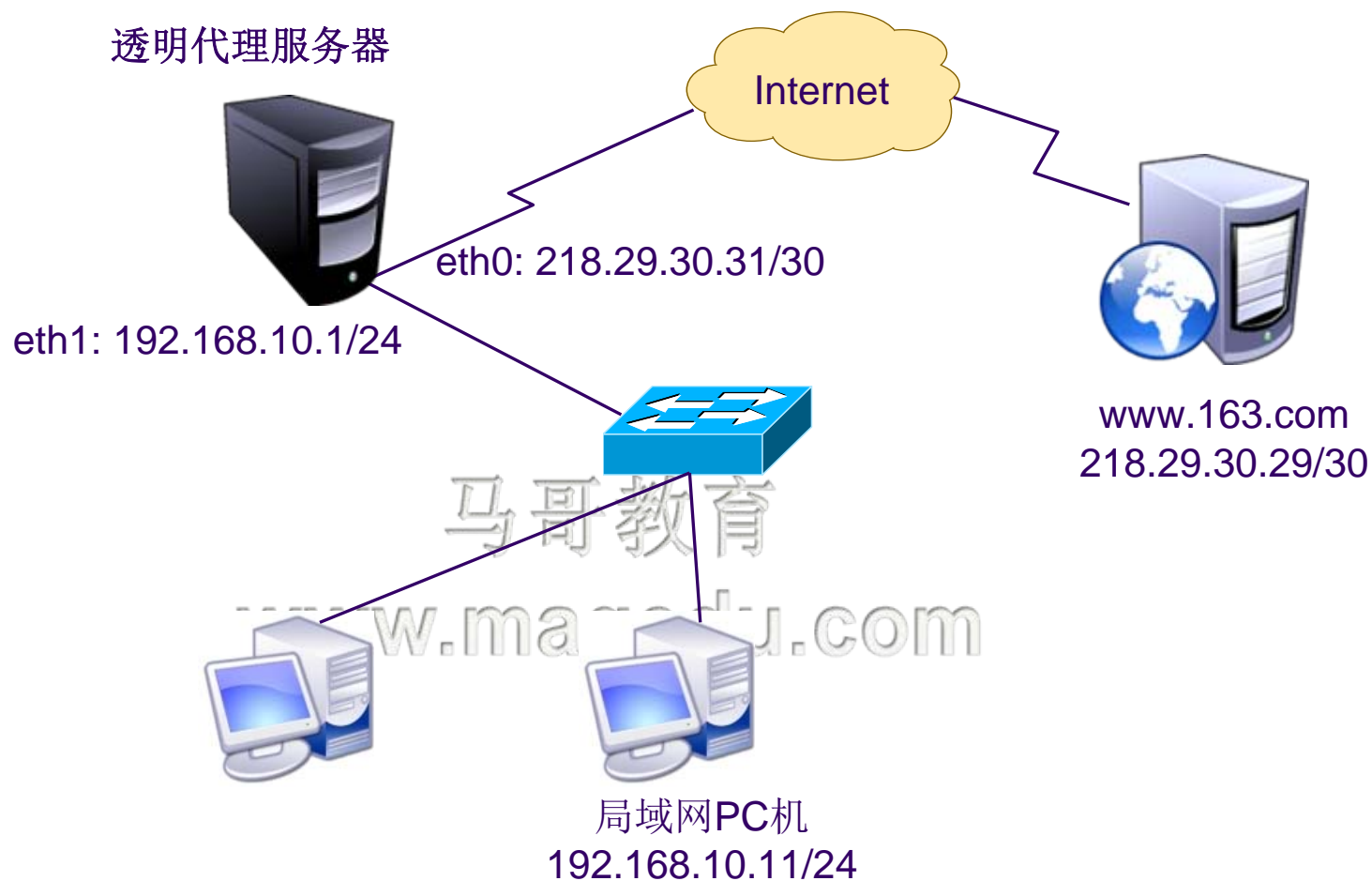
- 代理服务程序能够支持透明代理
- 设置防火墙规则，将客户机的**Web**访问数据自动重定向给代理服务程序处理

马哥教育

www.magedu.com



## ❖ 透明代理服务的典型应用环境



# 配置透明代理

## ❖ 基本实现步骤

➡ 修改**squid.conf**配置文件，并重新加载该配置

    ➤ **http\_port 192.168.10.1:8080 transparent**

➡ 添加**iptables**规则

    ➤ **iptables -t nat -I PREROUTING -i eth1 -s 192.168.10.0/24  
    -p tcp --dport 80 -j REDIRECT --to-ports 8080**

➡ 客户机浏览器

    ➤ 不需要在浏览器中指定代理服务器的地址、端口

➡ 验证透明代理的实施效果

马哥教育

www.magedu.com

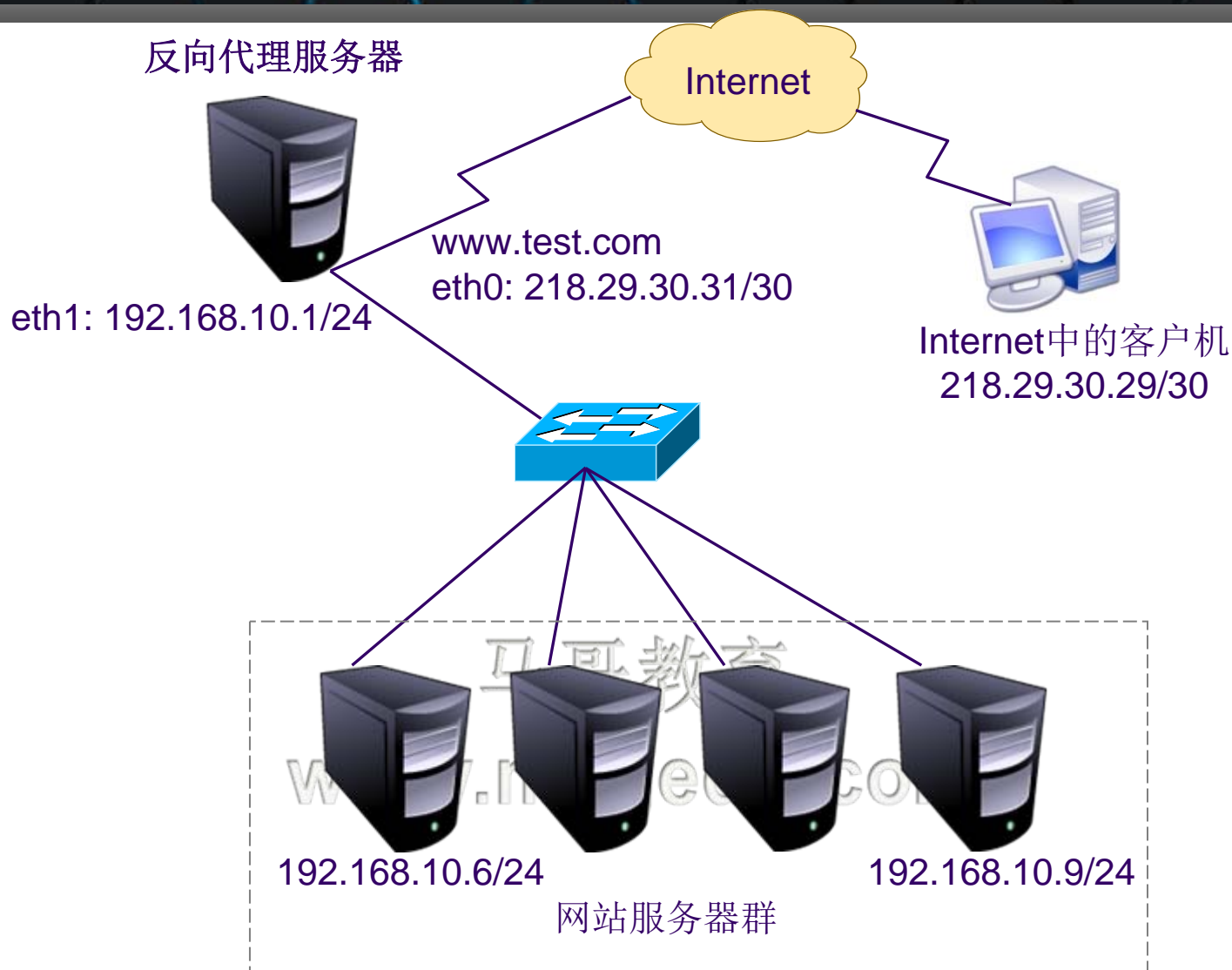
❖ `Iptables -t nat -A POSTROUTING -s 192.168.10.0/24 -j SNAT -to-source 192.168.0.34`

马哥教育

www.magedu.com

❖ http\_port 192.168.10.1:3128 transparent

马哥教育  
www.magedu.com



## ❖ accel

- ➡ Accelerator mode. Also needs at least one of vhost/vport/defaultsite

## ❖ defaultsite=domainname

- ➡ What to use for the Host: header if it is not present in a request. Determines what site (not origin server) accelerators should consider the default. Implies accel.

## ❖ vhost

- ➡ Accelerator mode using Host header for virtual domain support. Implies accel.

www.magedu.com

## ❖ cache\_peer

➡ To specify other caches in a hierarchy, use the format:

➤ `cache_peer hostname type http-port icp-port [options]`

➡ For example

➤ `cache_peer parent.foo.net parent 3128 3130 proxy-only default`

➤ `cache_peer sib1.foo.net sibling 3128 3130 proxy-only`

➤ `cache_peer sib2.foo.net sibling 3128 3130 proxy-only`

➡ **type:** either 'parent', 'sibling', or 'multicast'

➡ **proxy-port:** The port number where the cache listens for proxy requests

➡ **icp-port:** Used for querying neighbor caches about objects



## ❖ weight=n

- ➡ To affect the selection of a peer during any weighted peer-selection mechanisms
- ➡ The weight must be an integer; default is 1, larger weights are favored more

## ❖ max-conn=n

- ➡ To limit the amount of connections Squid may open to this peer

## ❖ originserver

- ➡ Causes this parent peer to be contacted as a origin server
- ➡ Meant to be used in accelerator setups

# 配置反向代理

## ❖ 基本实现步骤

➡ 修改**squid.conf**文件，并重新加载该配置

```
http_port 218.29.30.31:80 vhost
```

```
cache_peer 192.168.10.7 parent 80 0 originserver weight=5 max-conn=30
```

```
cache_peer 192.168.10.8 parent 80 0 originserver weight=5 max-conn=30
```

```
cache_peer 192.168.10.9 parent 80 0 originserver weight=5 max-conn=30
```

```
cache_peer 192.168.10.6 parent 80 0 originserver weight=1 max-conn=8
```

**cache\_peer** Web服务器地址 服务器类型 http端口 icp端口 [可选项]

www.magedu.com

马哥教育

# memcached

主讲：马永亮(马哥)

QQ:113228115

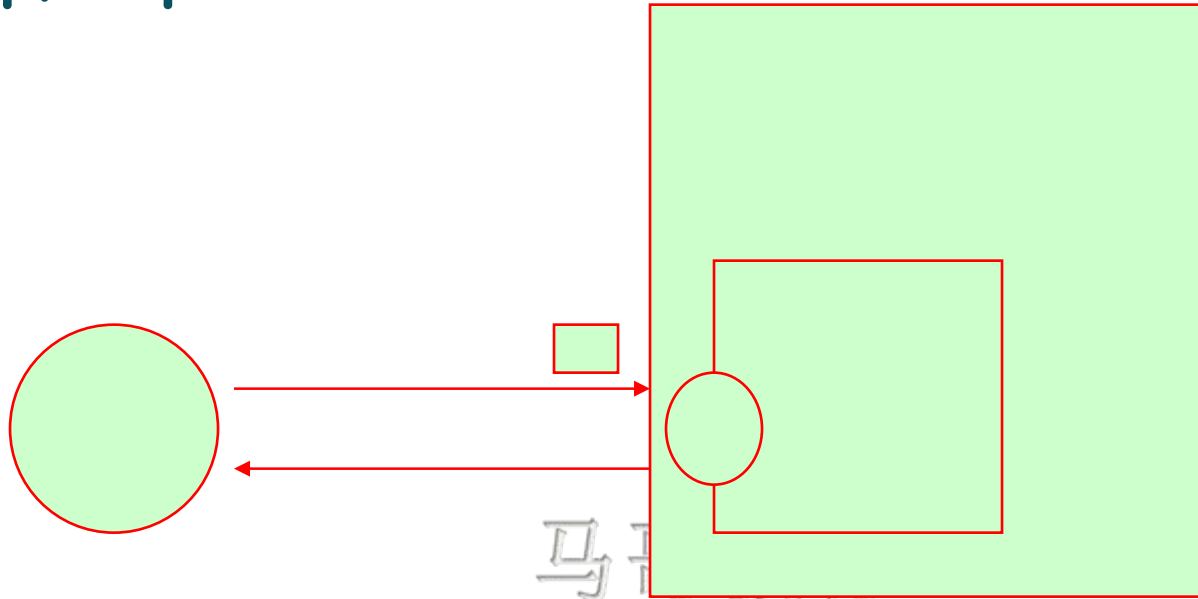
客服QQ: 1661815153

<http://www.magedu.com>

<http://mageedu.blog.51cto.com>

❖ ip

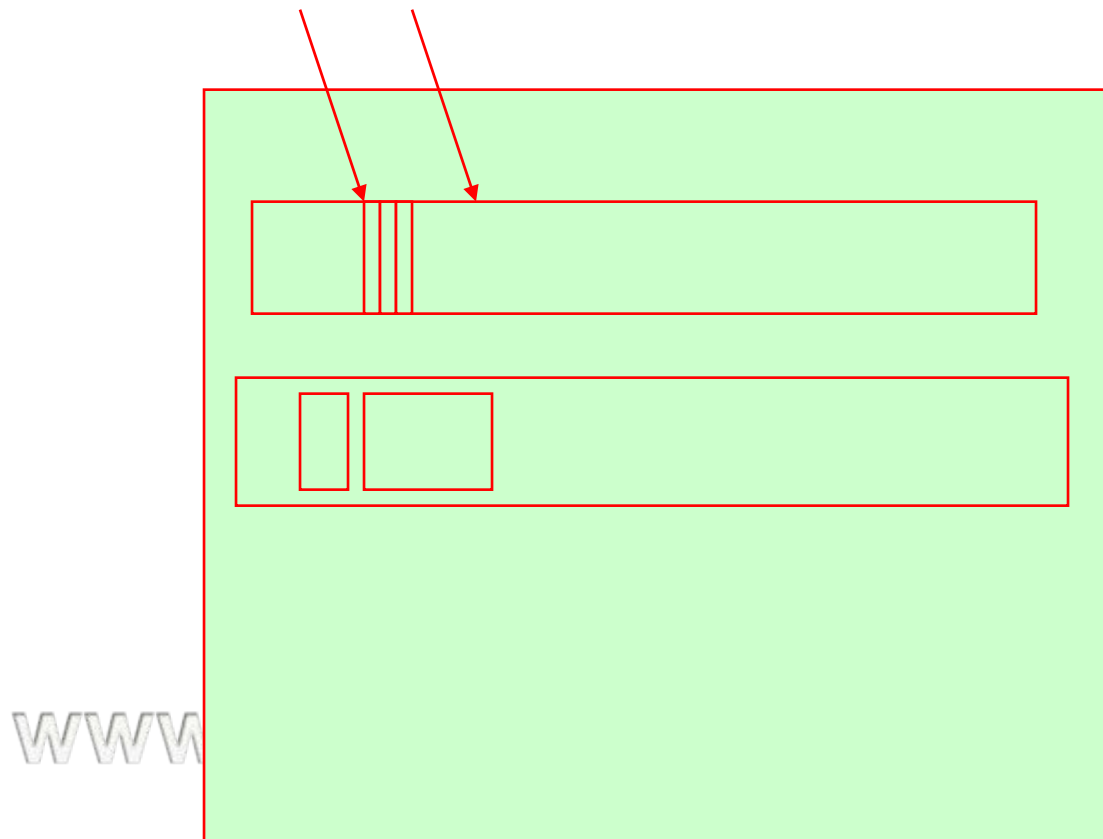
❖ tcp, udp



www.magedu.com

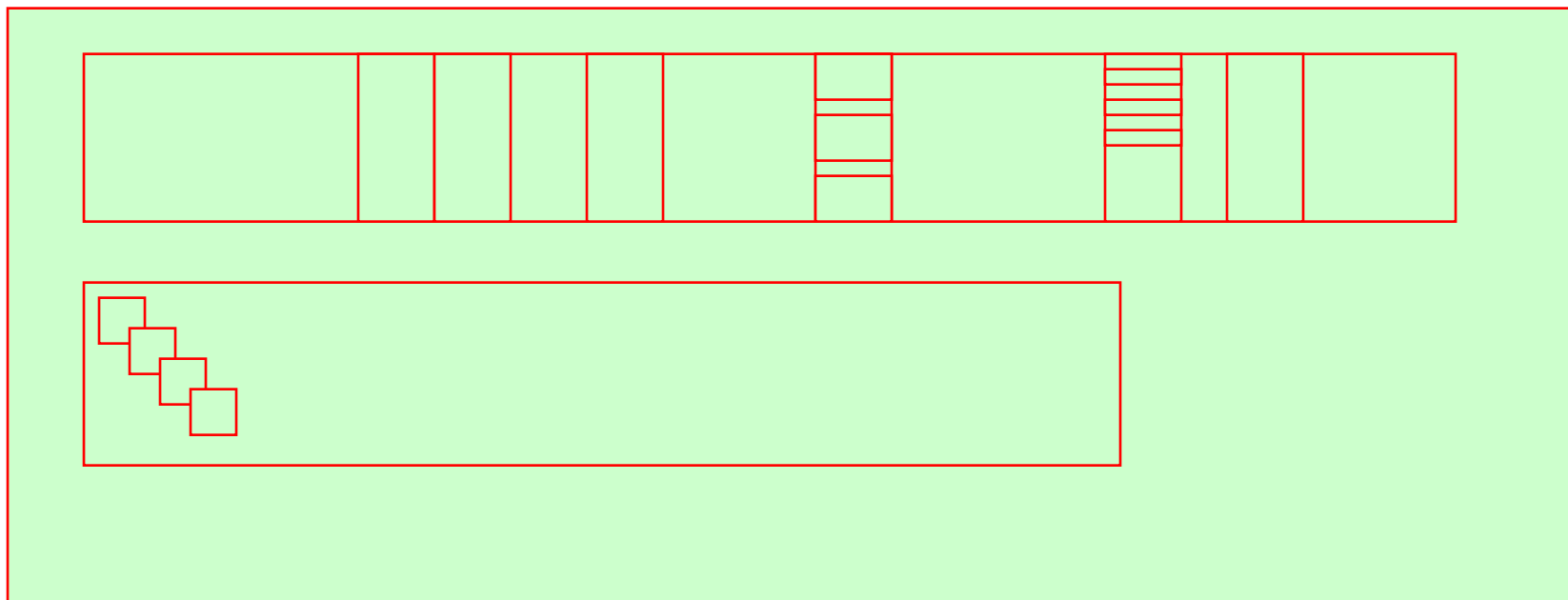
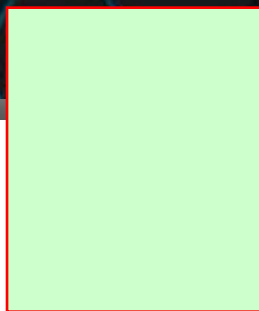
telnet ip 80

## ❖ 10k, 34K



❖ 10M

➡ 已用，空闲



www.magedu.com

❖ 48bytes

❖ 80bytes

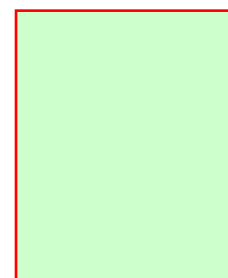
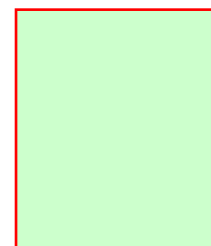
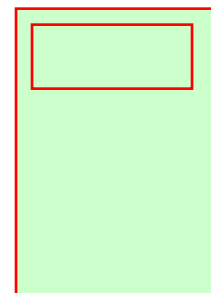
➡ 72

➡ 增长因子

➡ growth factor, 1.25

➡ 48bytes: slab class, slab chunk

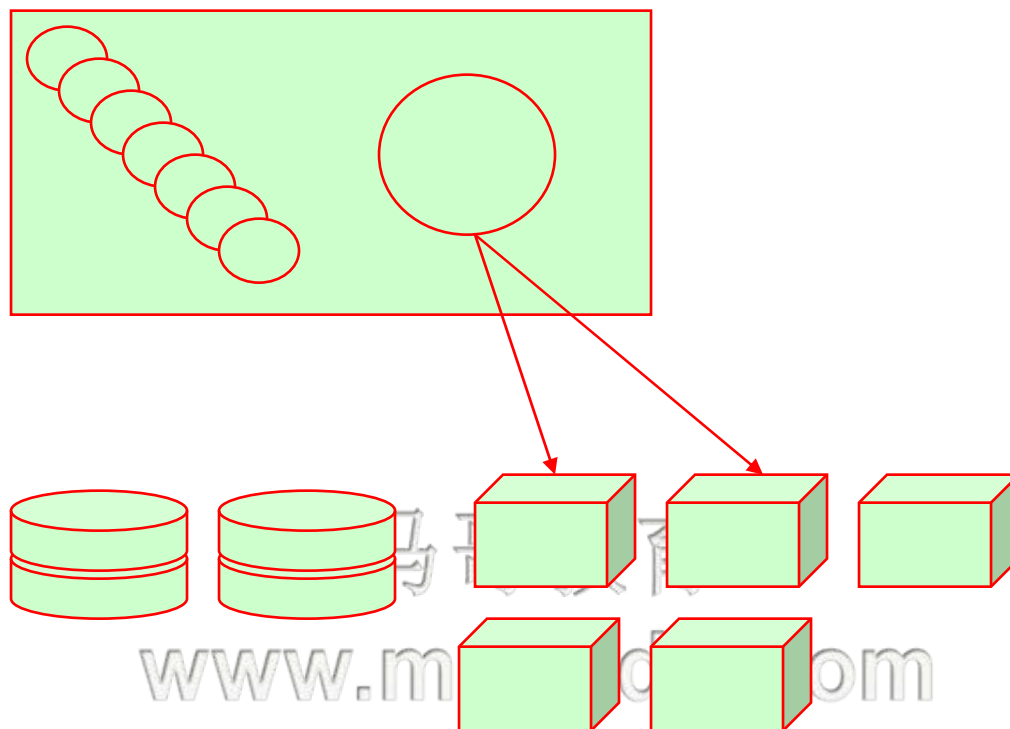
➡ 80bytes



马哥教育

www.magedu.com

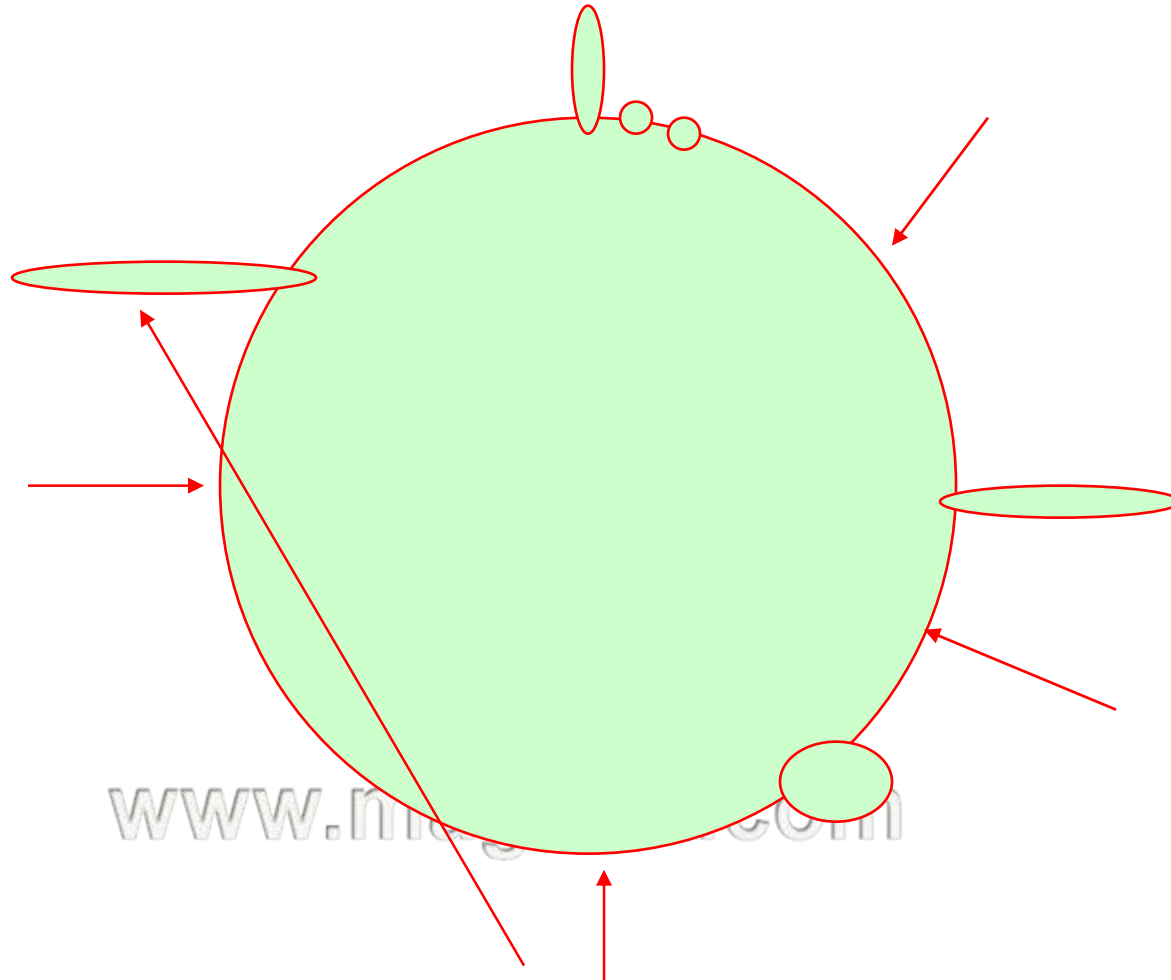




- ❖ perl module
  - ➡ cache::memcached
- ❖ php
  - ➡ memcache
  - ➡ memcached
- ❖ C/C++
  - ➡ libmemcached
    - 命令行工具
- ❖ memadmin

马哥教育

www.magedu.com



❖ Get URI

❖ PUT /etc/issue

❖ get, put, mget, mput

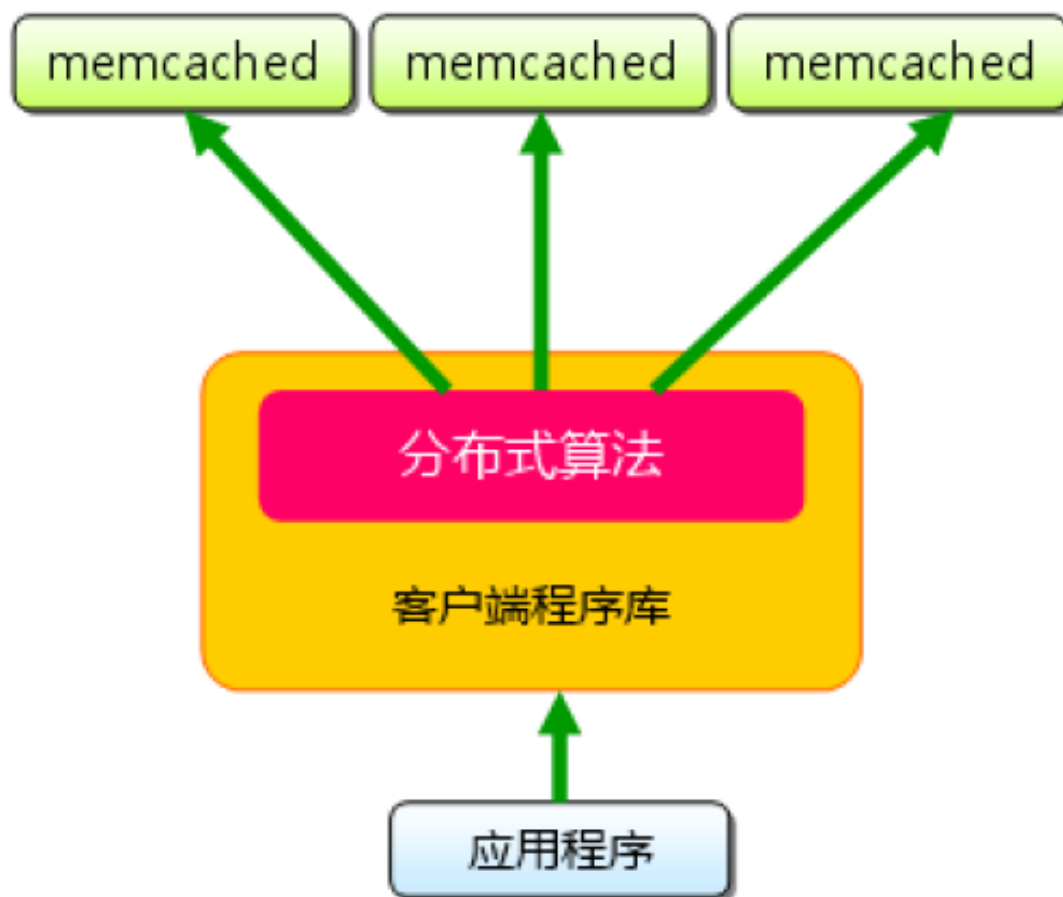
❖ simple protocol,

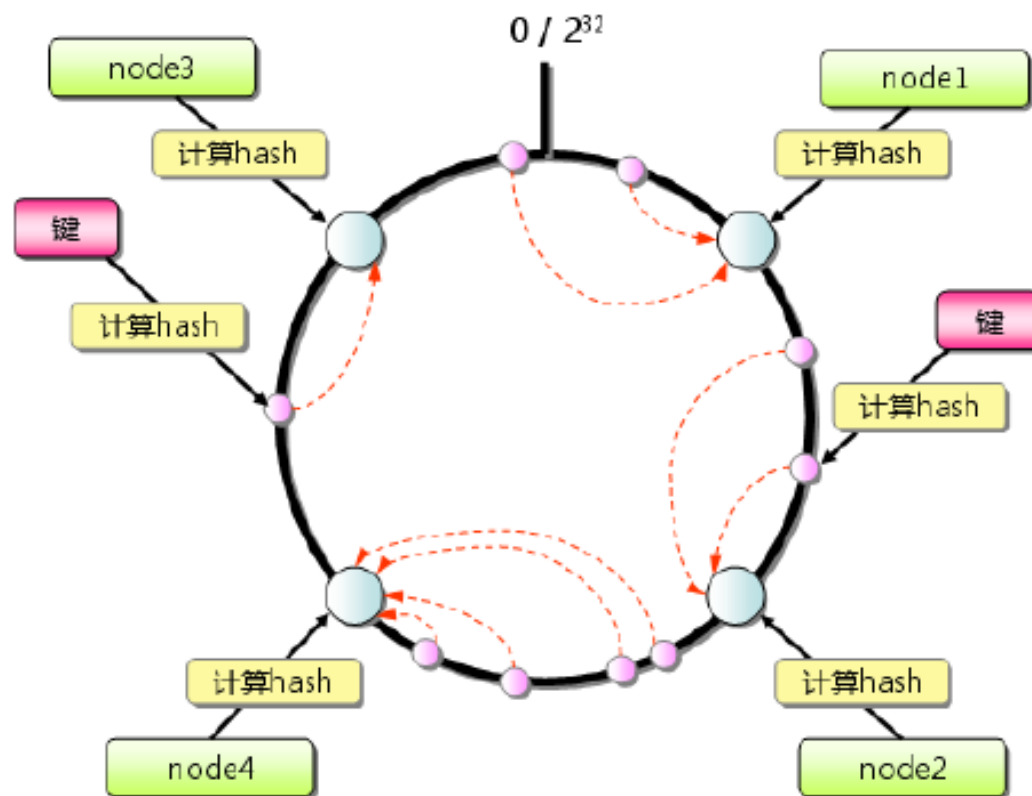
➡ http: text

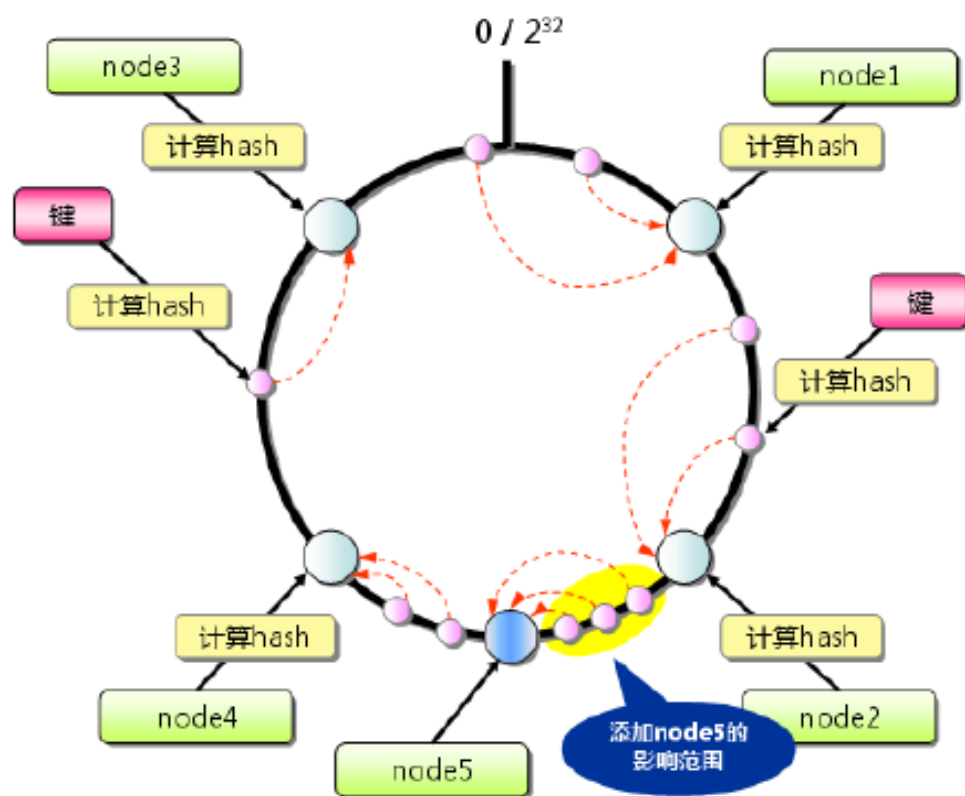
➡ ftp: text, binary

马哥教育

www.magedu.com









马  
哥  
教  
育

# Varnish

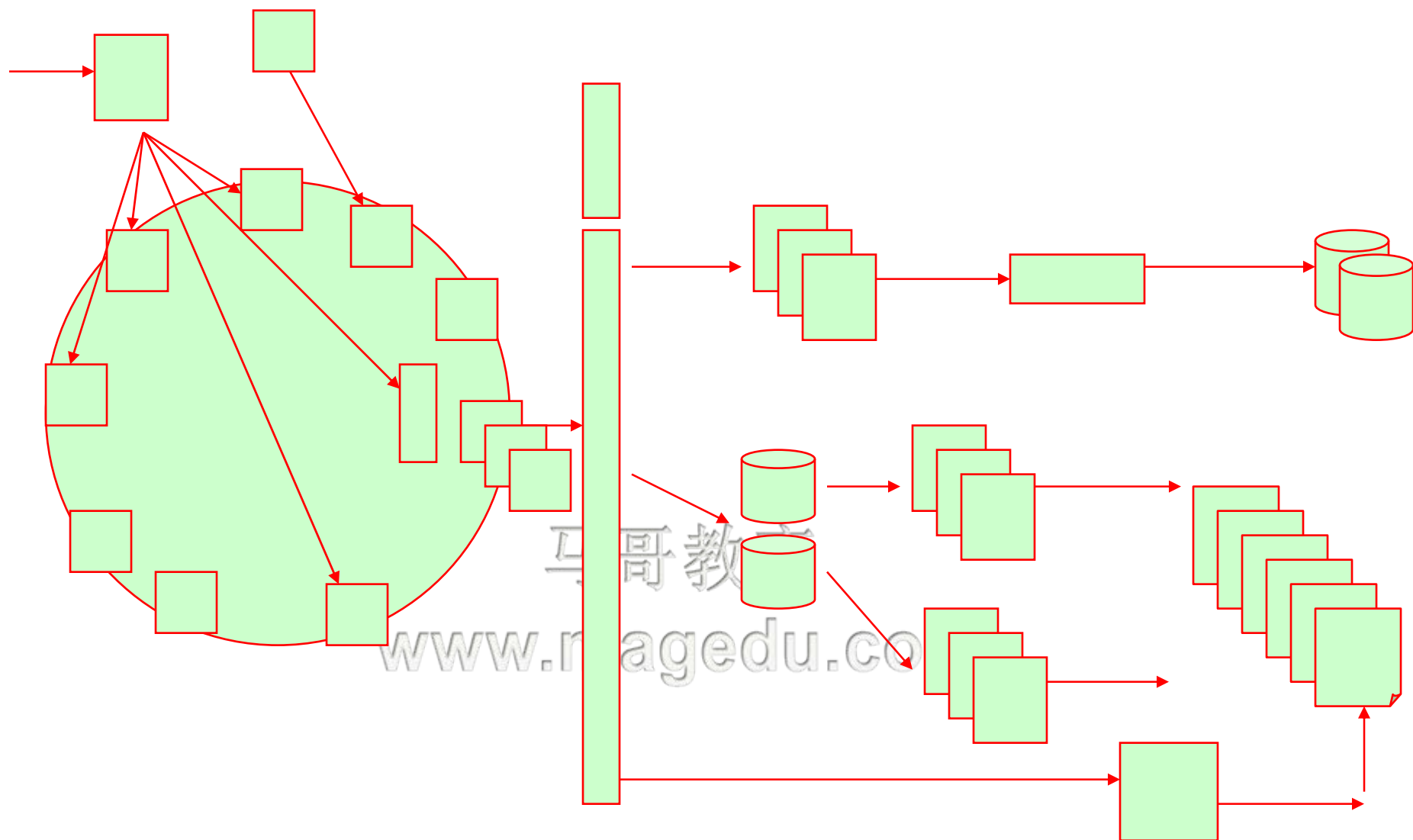
主讲：马永亮(马哥)

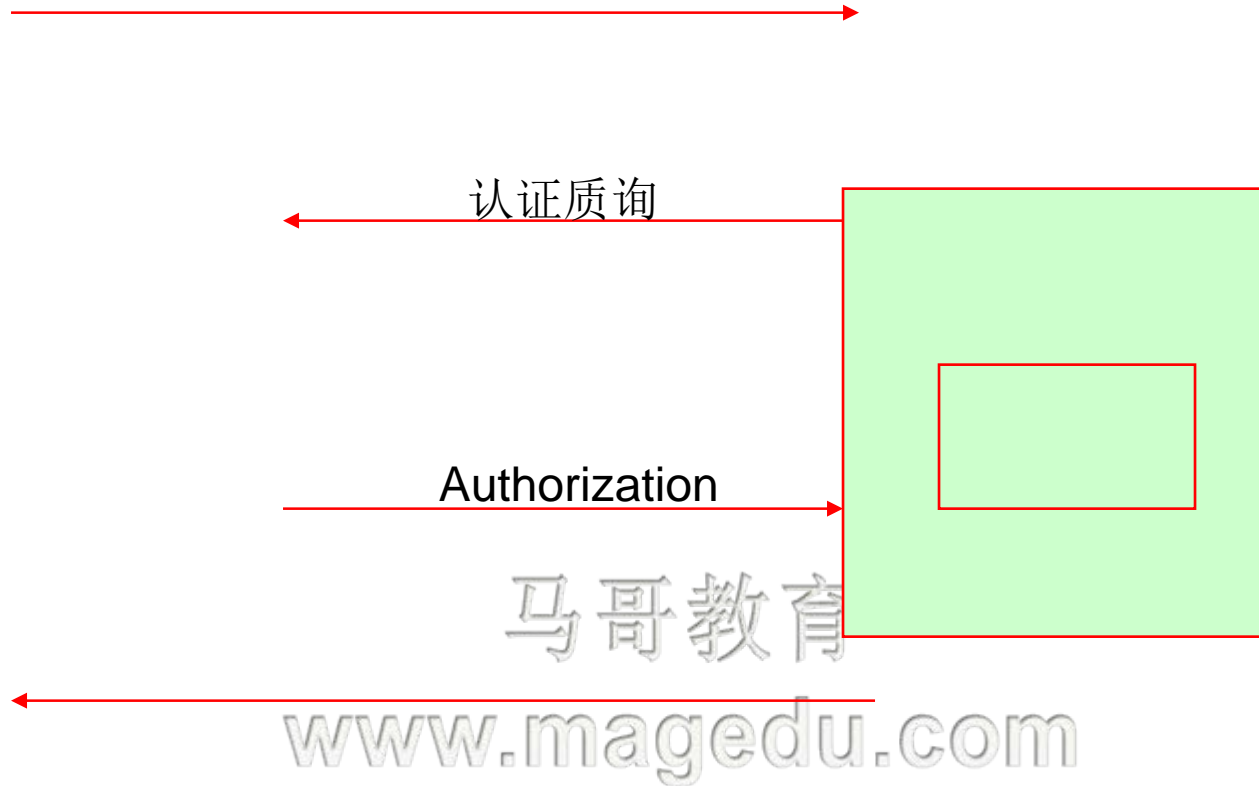
QQ群：169777636

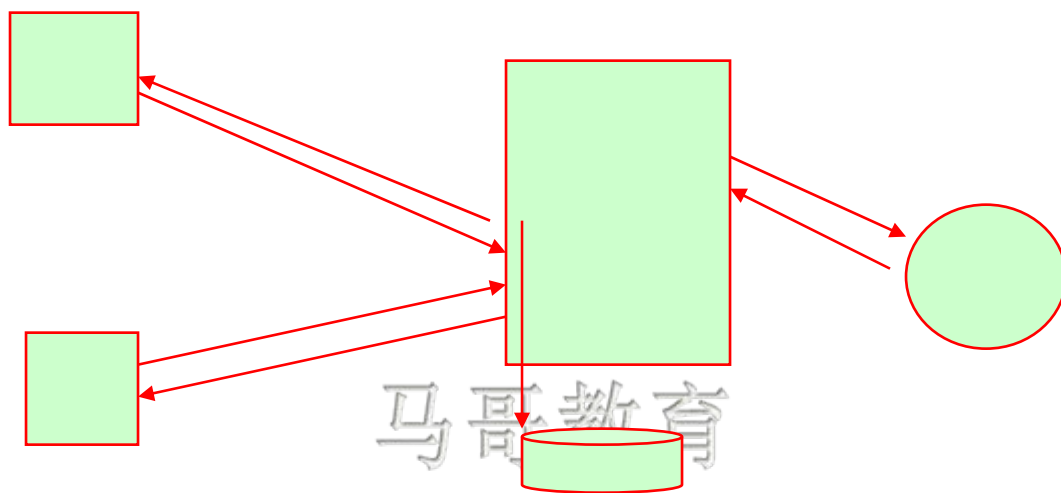
客服QQ：2813150558, 1661815153

<http://www.magedu.com>

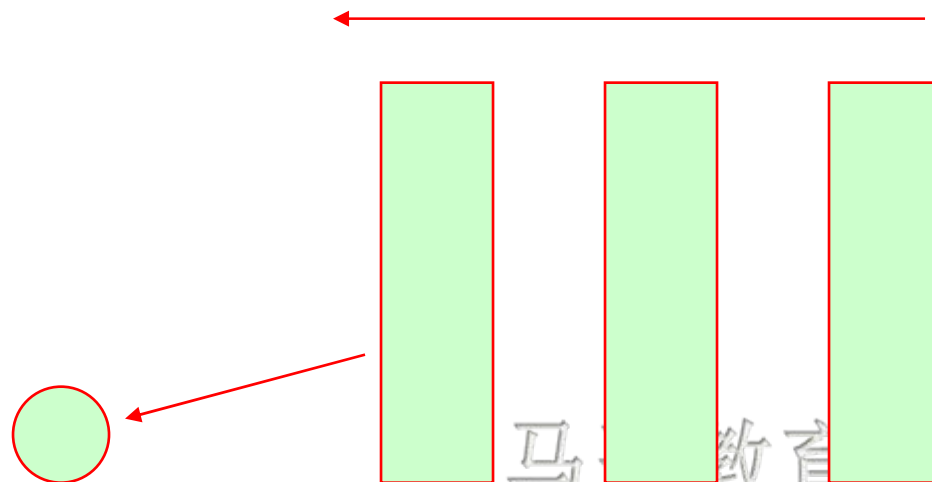
<http://mageedu.blog.51cto.com>





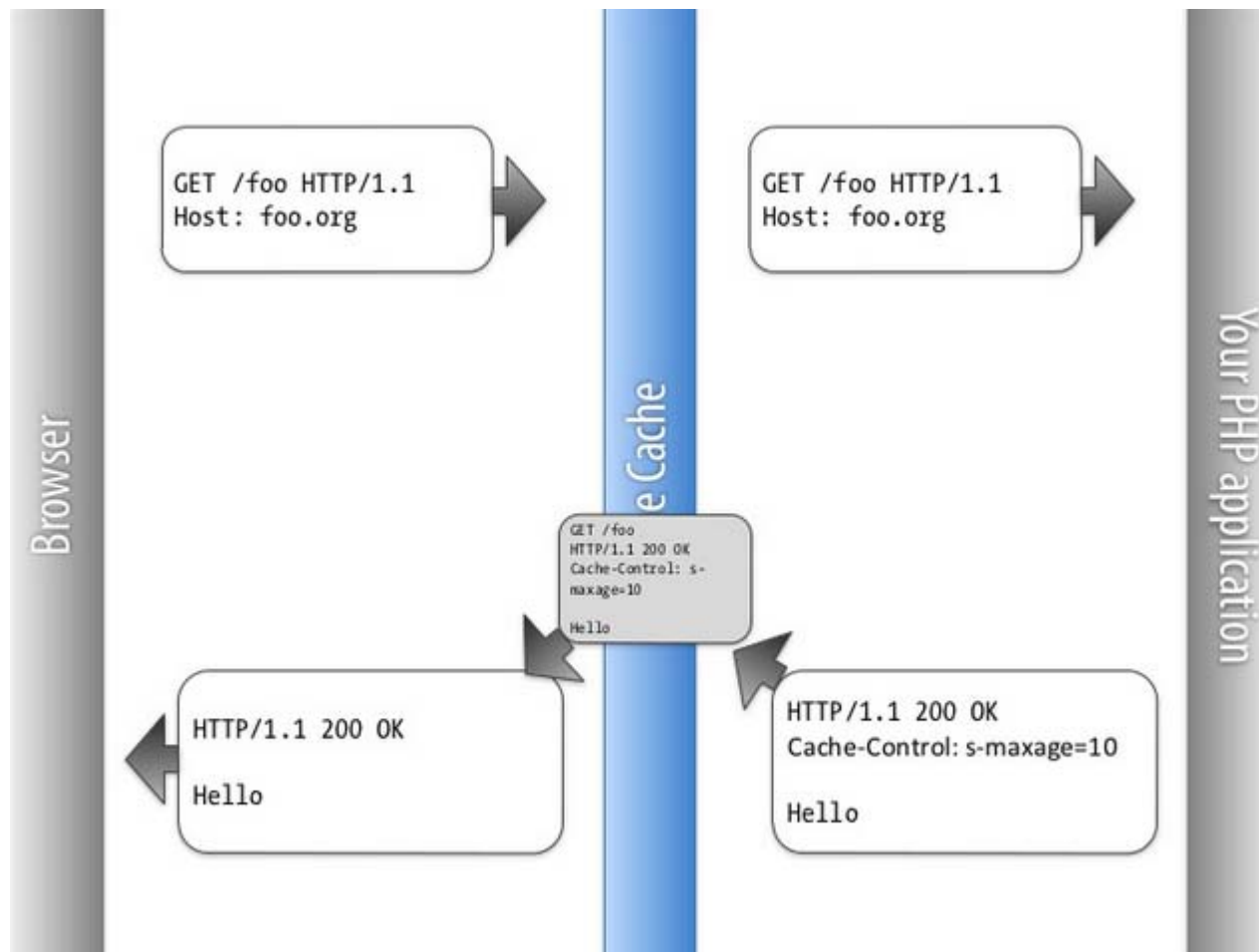


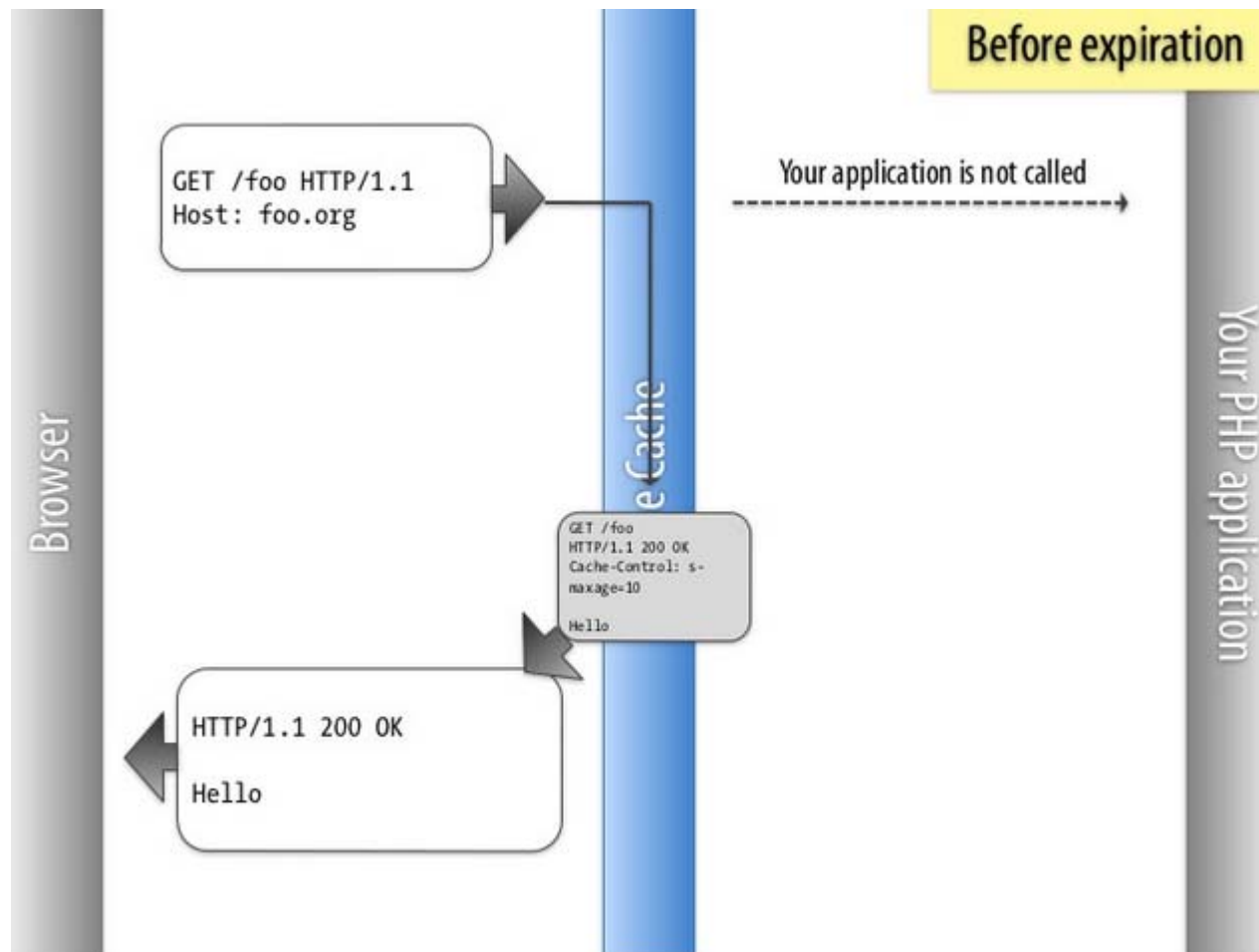
www.magedu.com



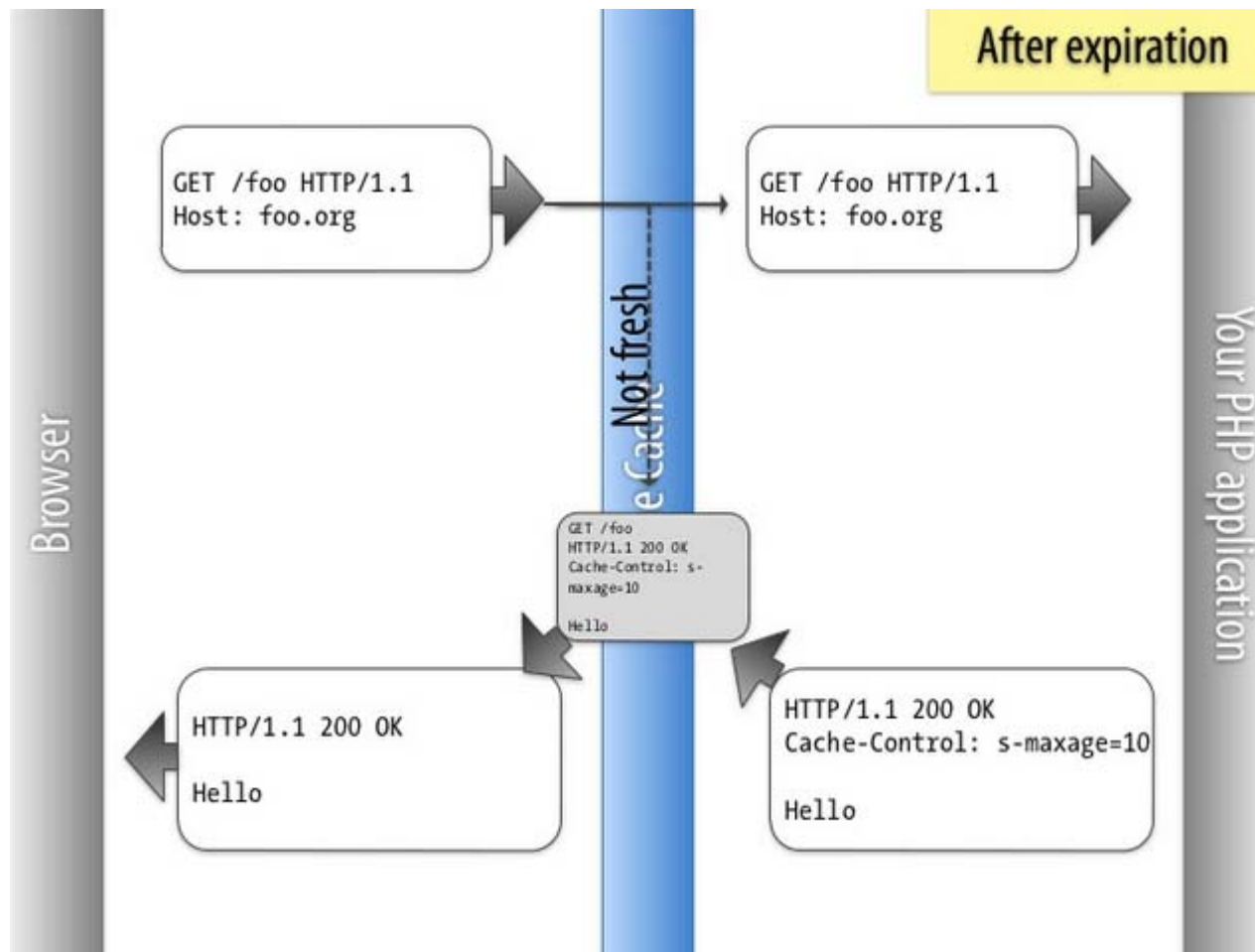
Via: b,  
a

# HTTP Expiration





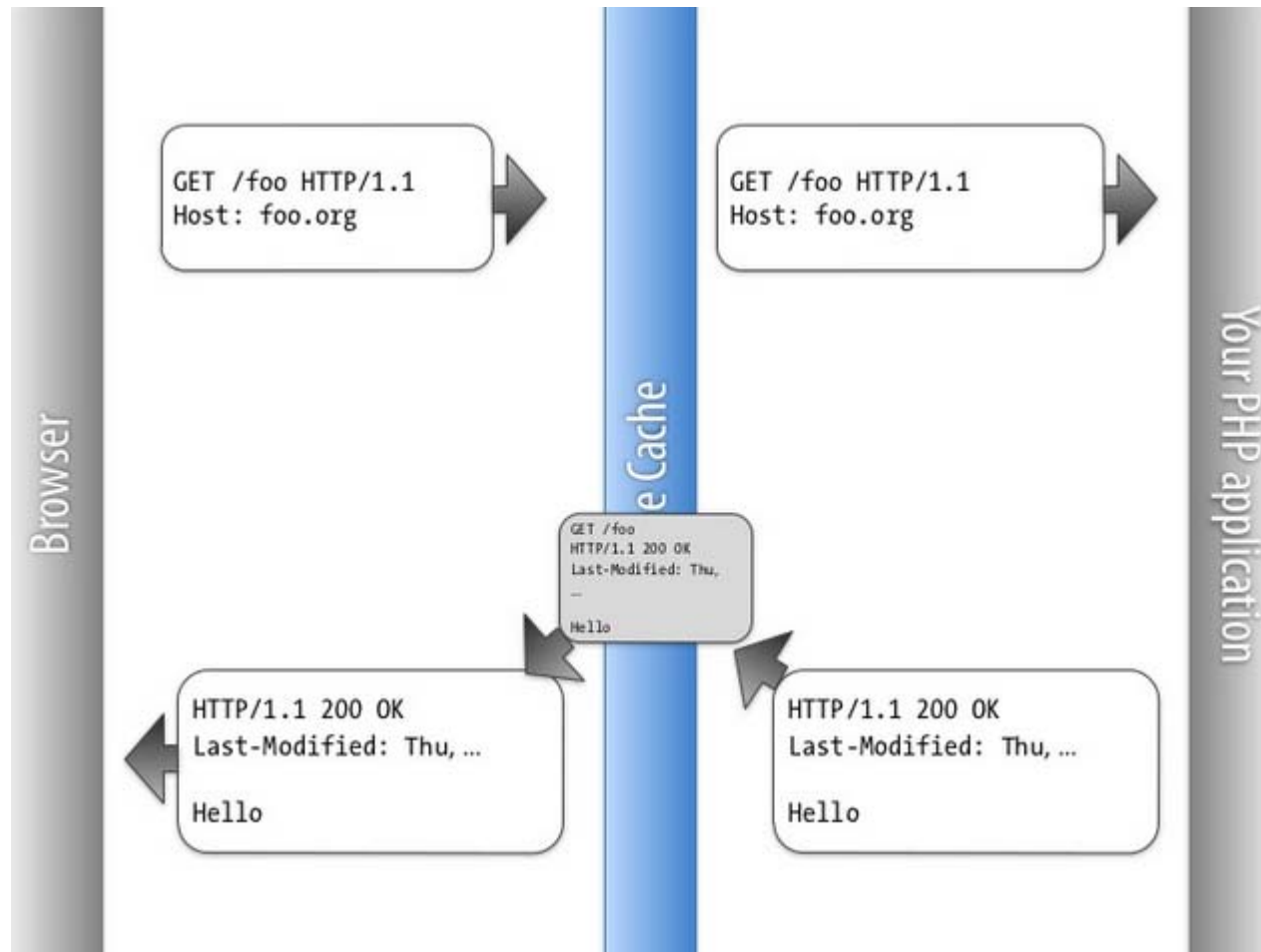


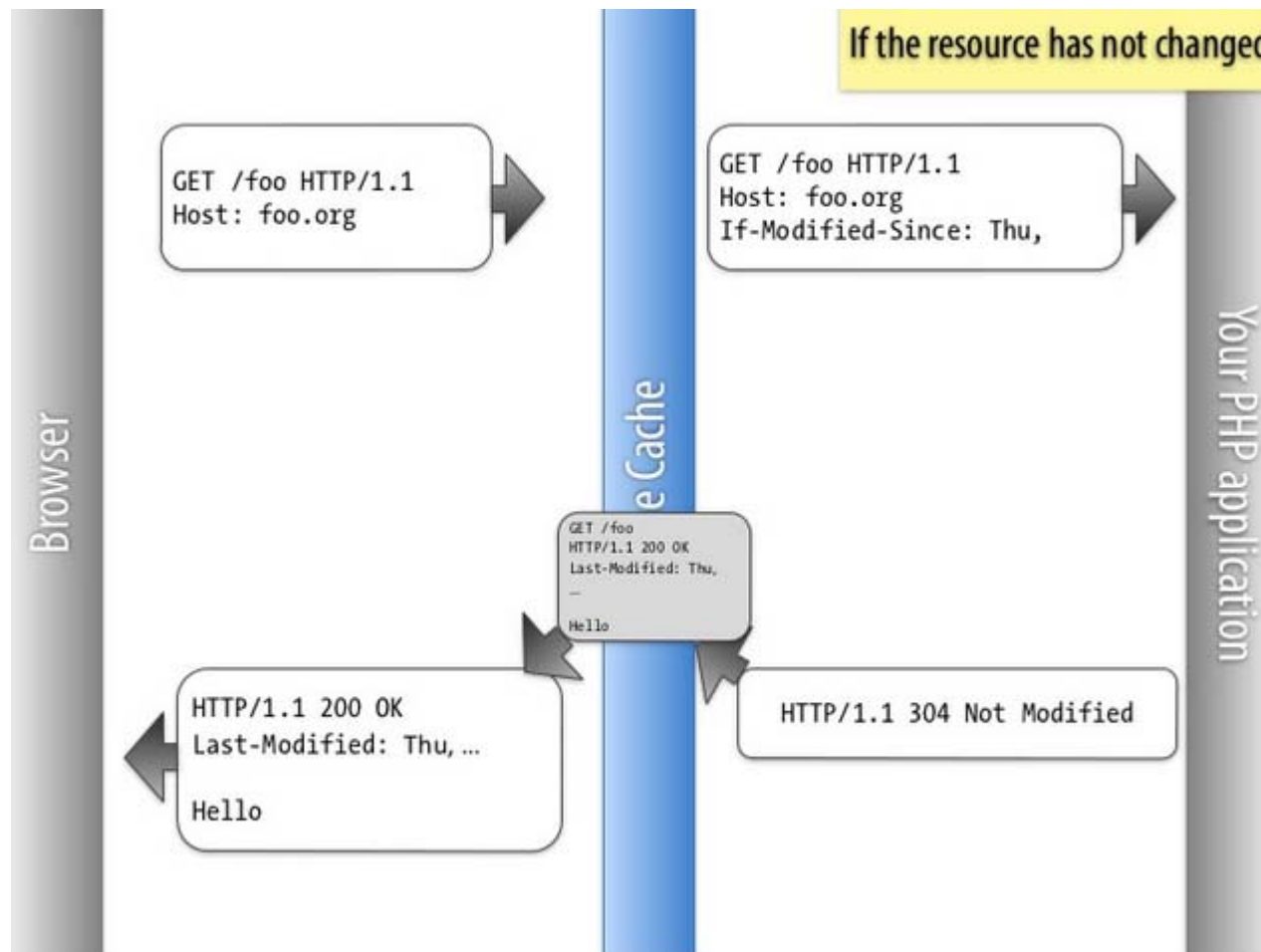


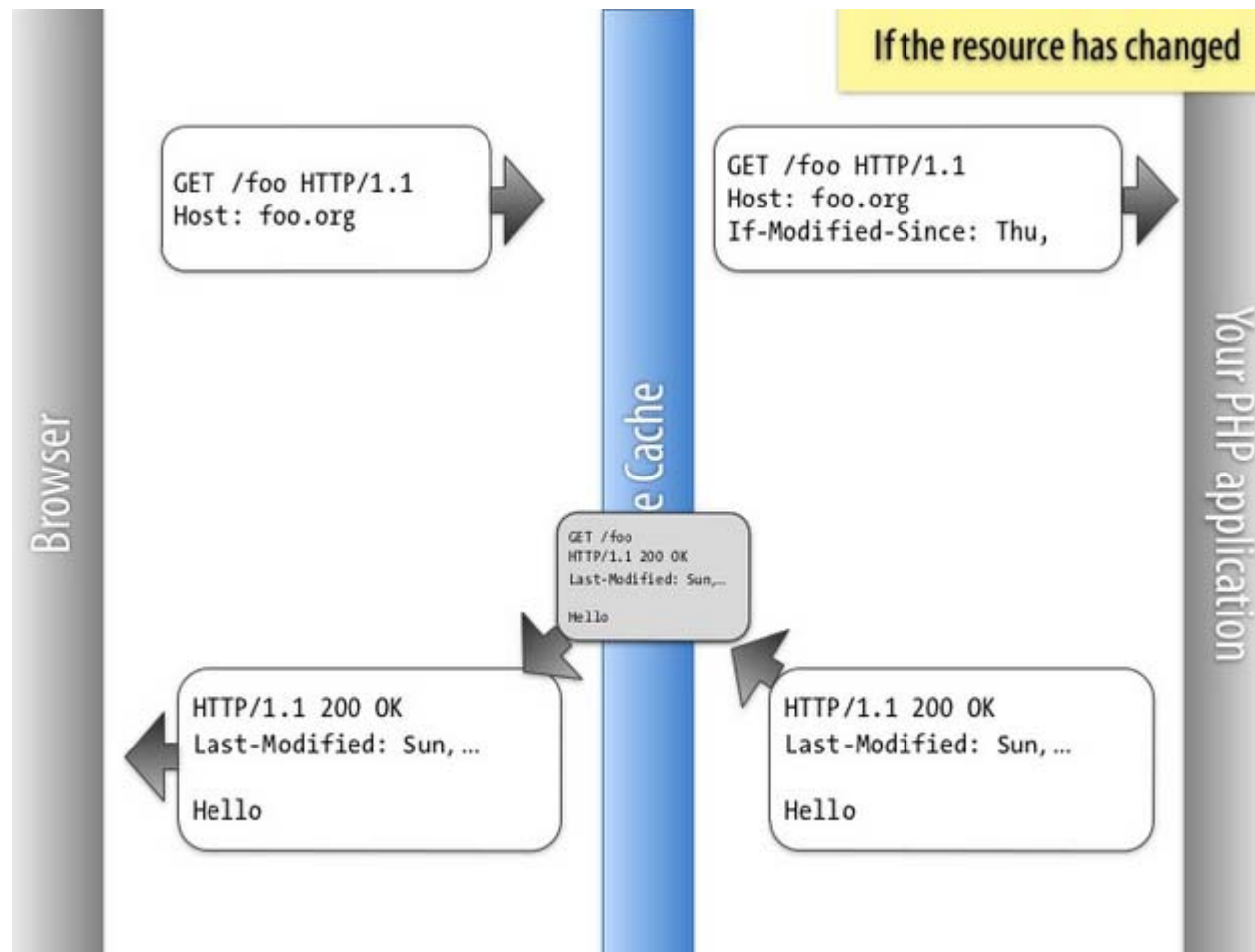
- ❖ Last Modified/If-Modified-Since
- ❖ Etag/If-None-Match

马哥教育

www.magedu.com



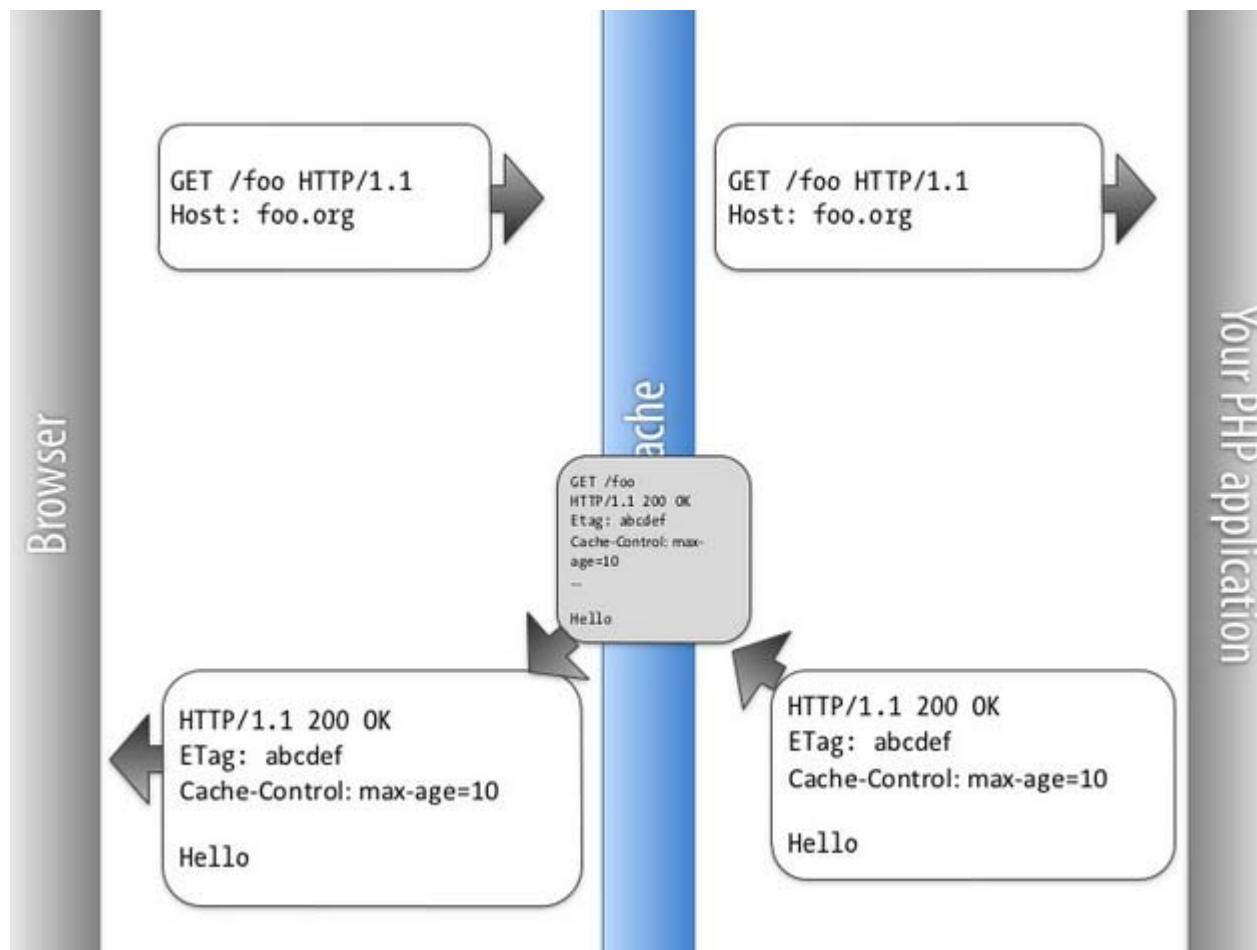




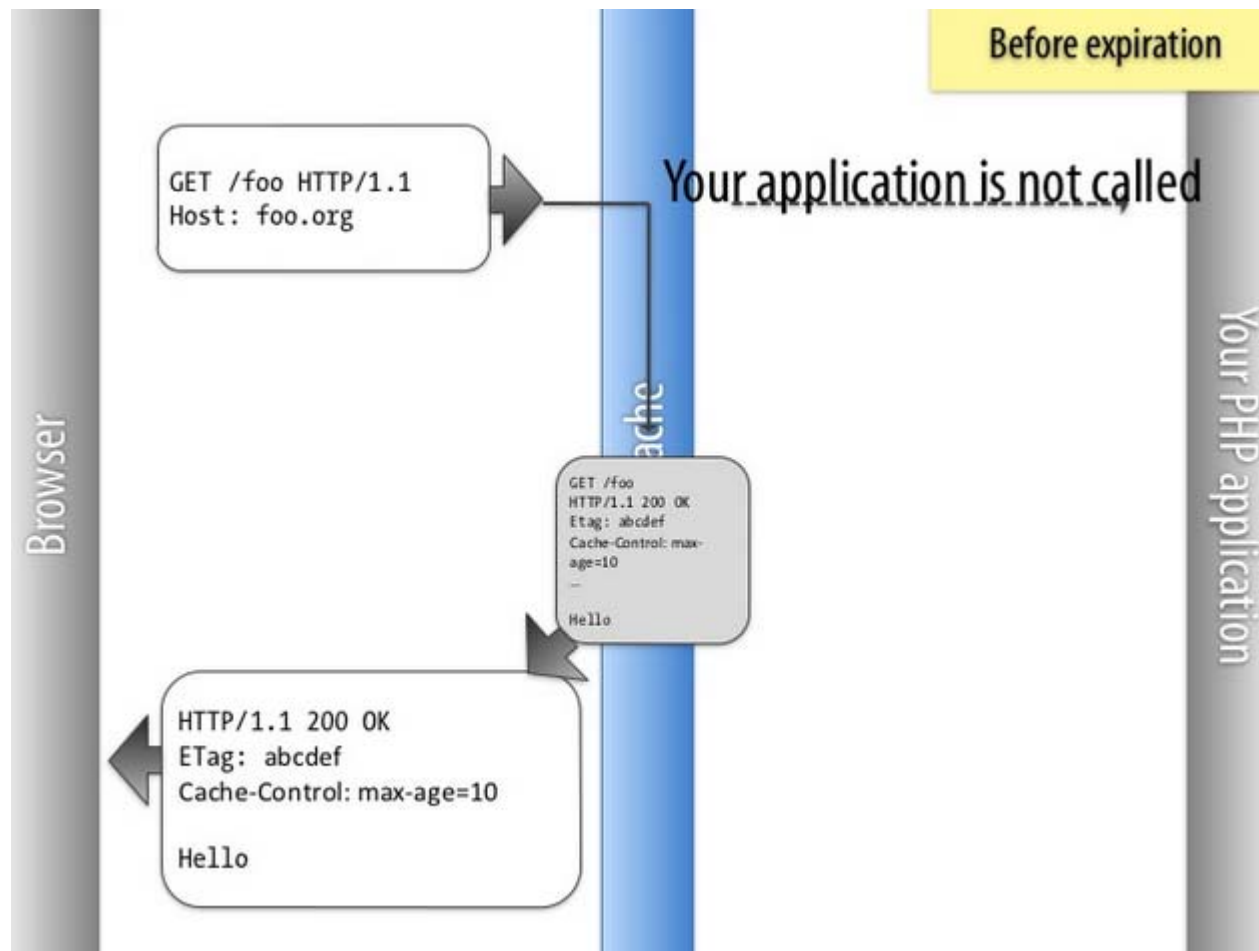
You can combine HTTP headers  
the way you want

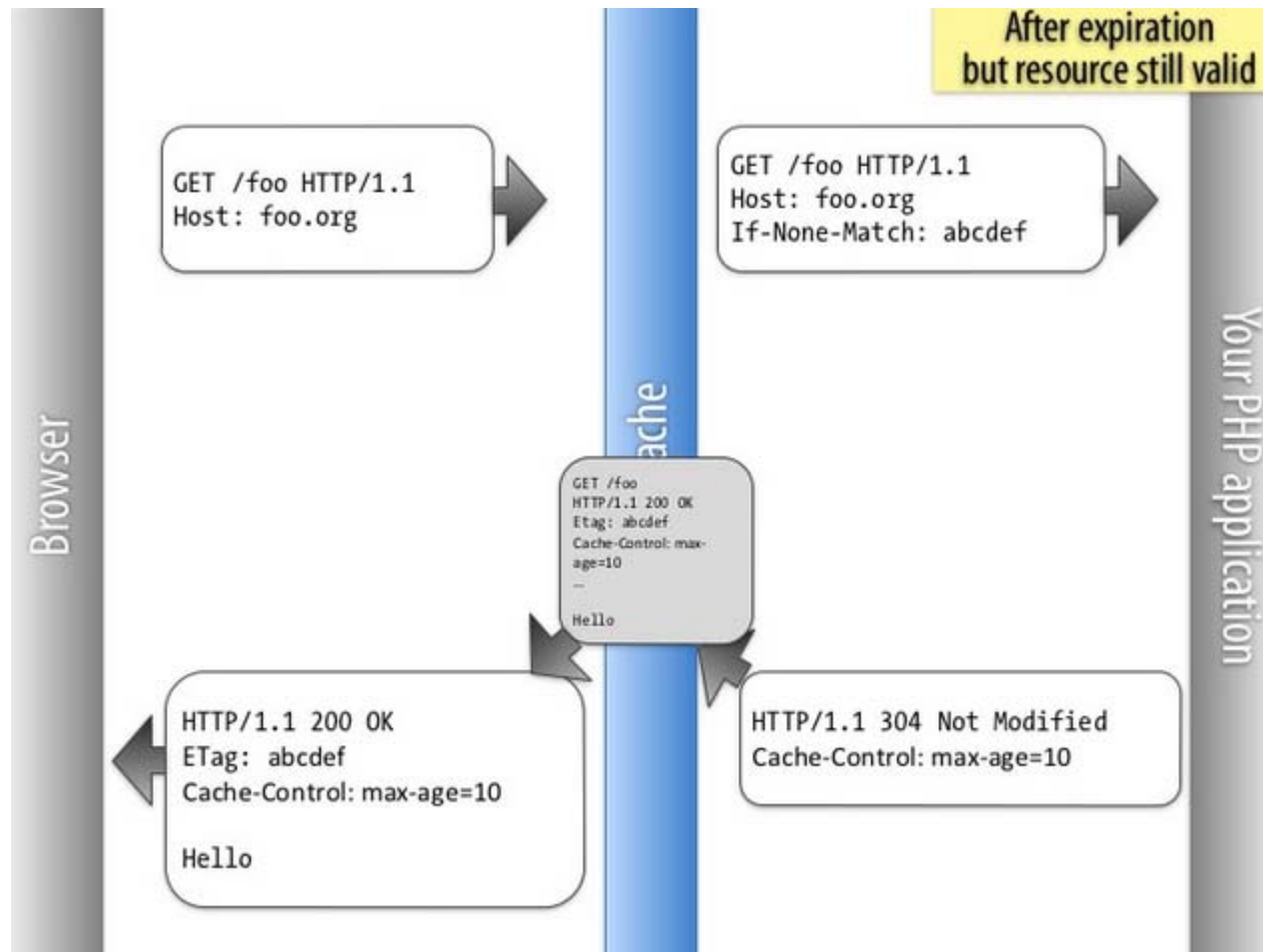
Expiration wins over Validation

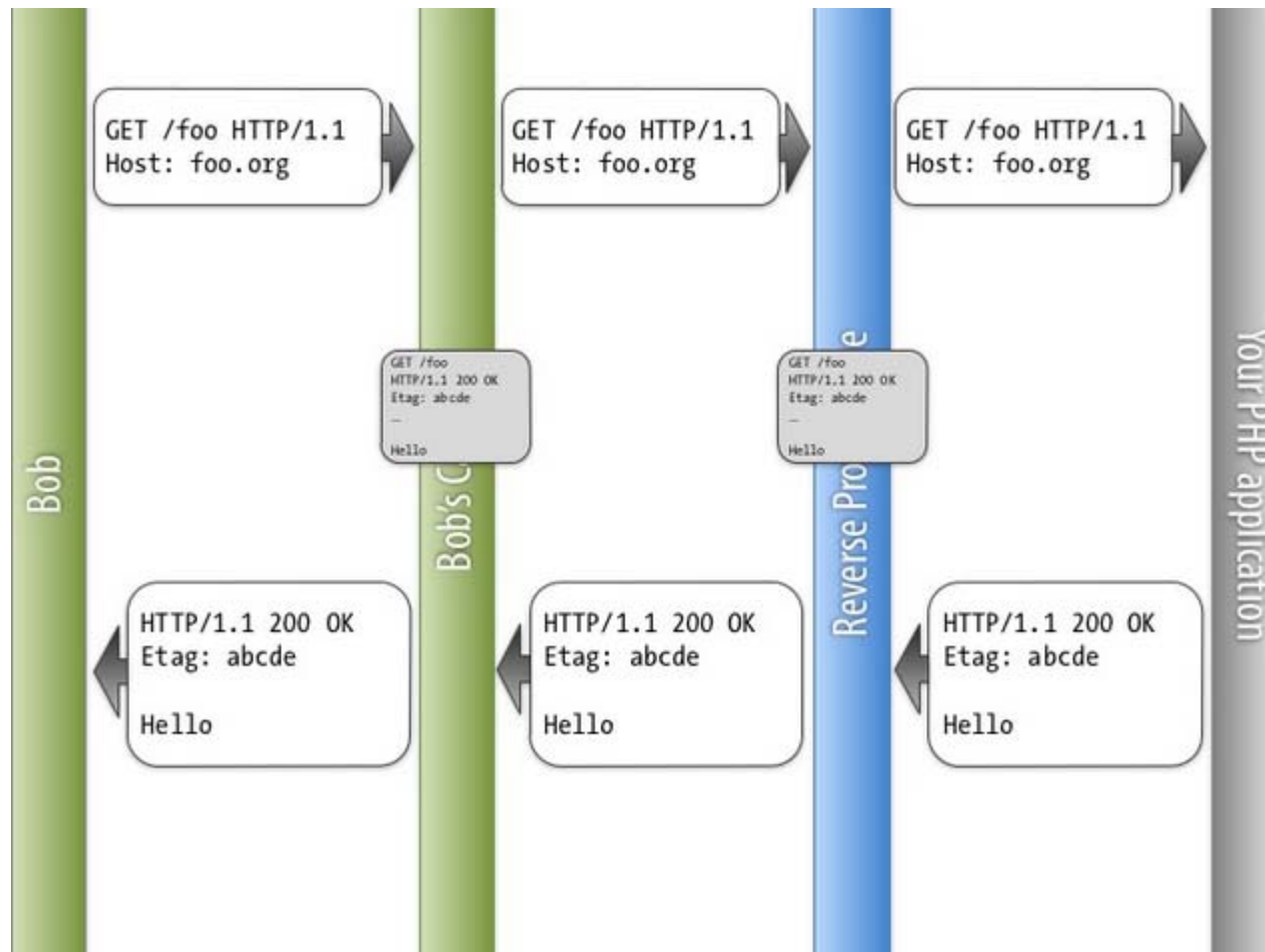
[www.magedu.com](http://www.magedu.com)

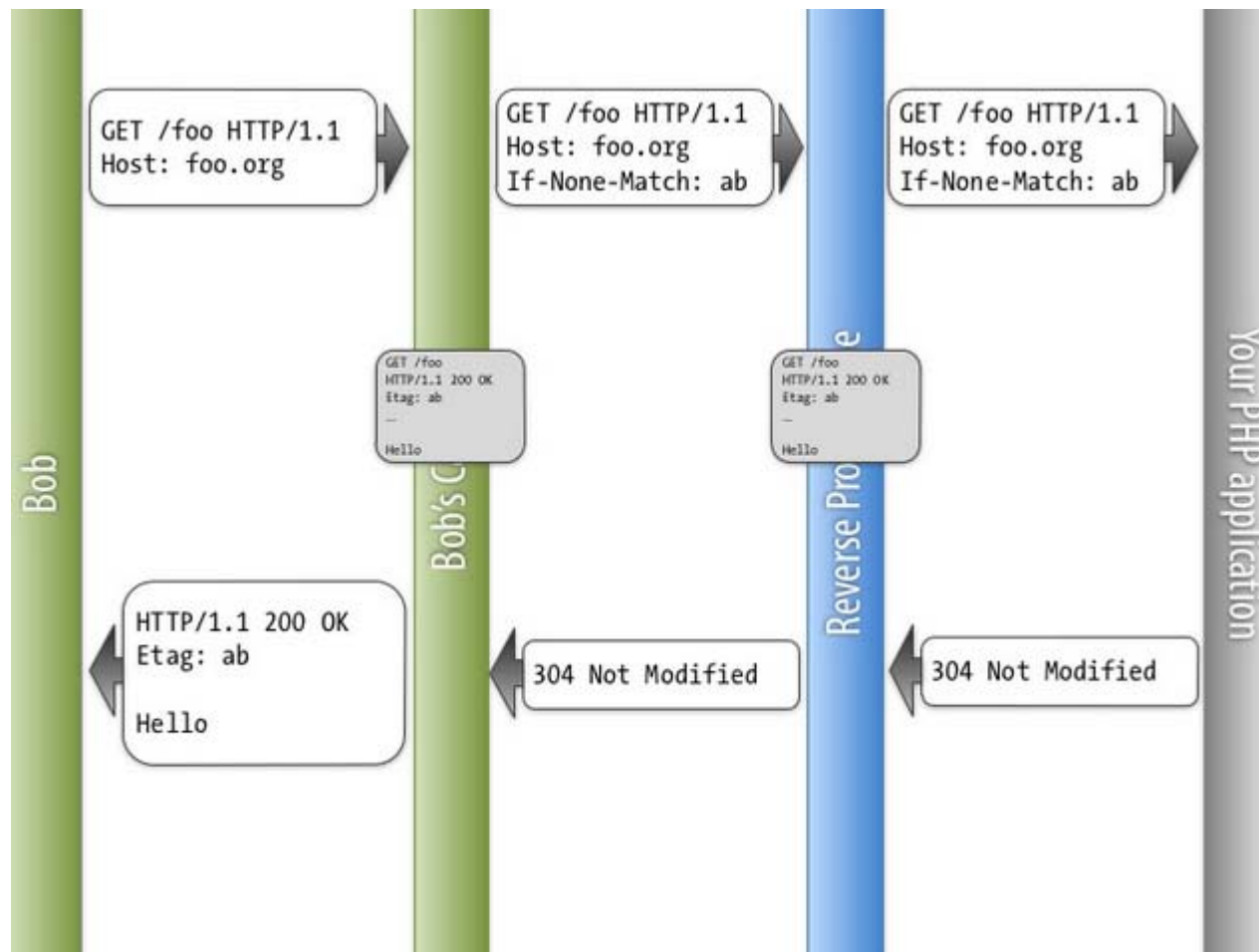


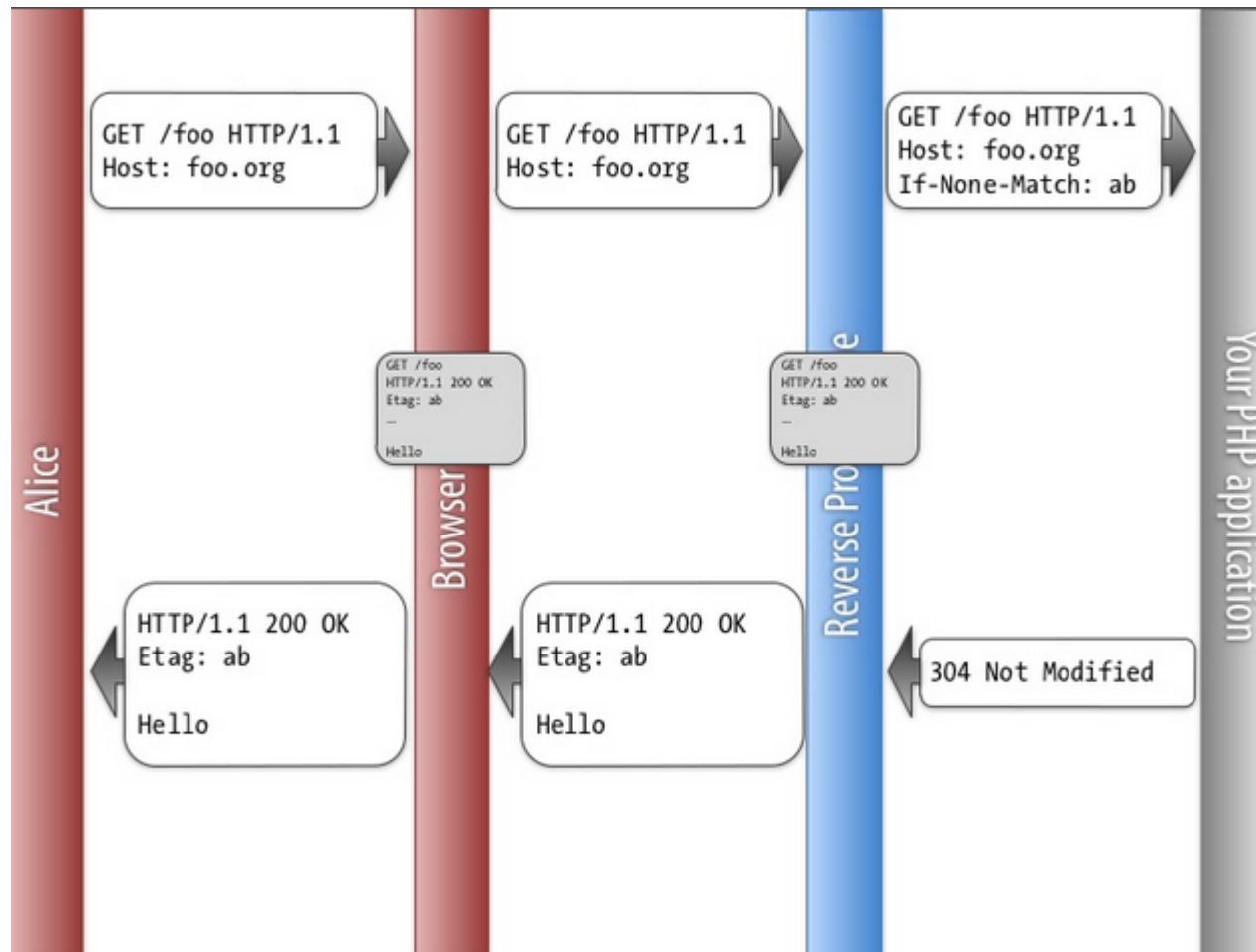


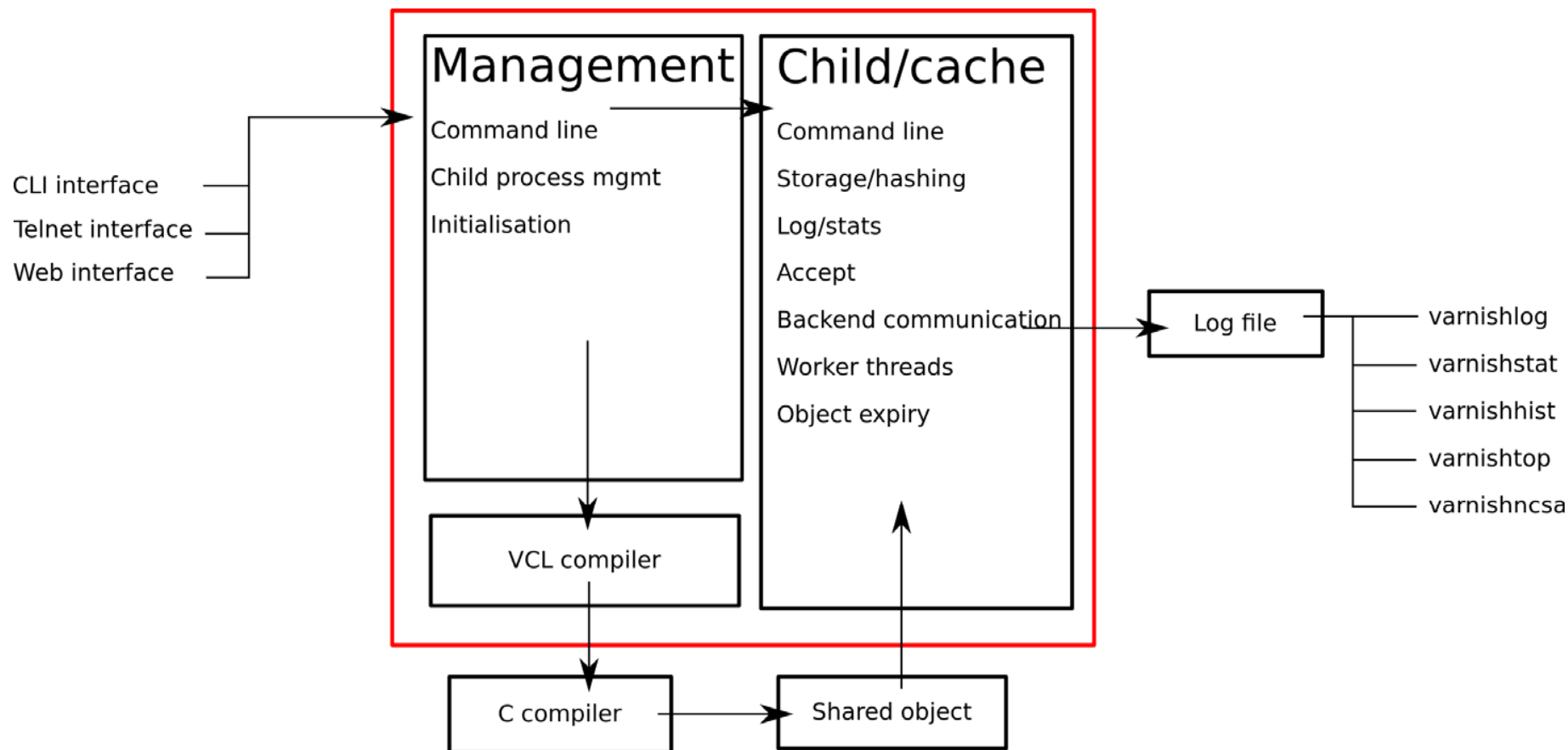


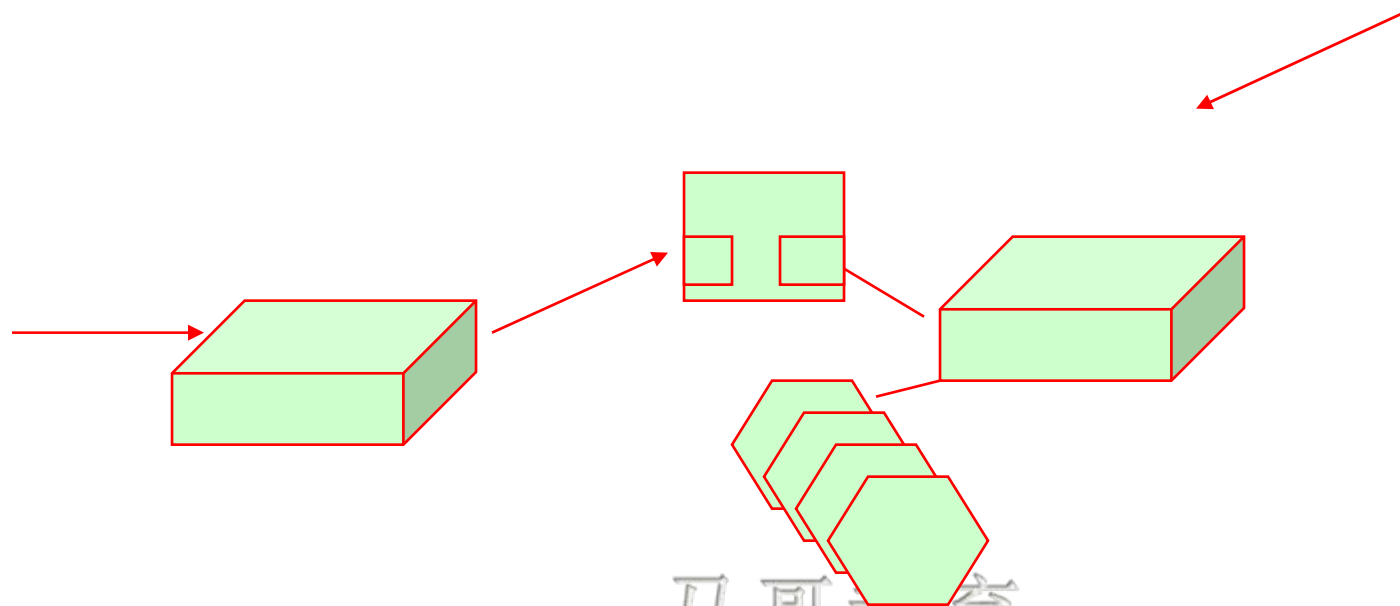








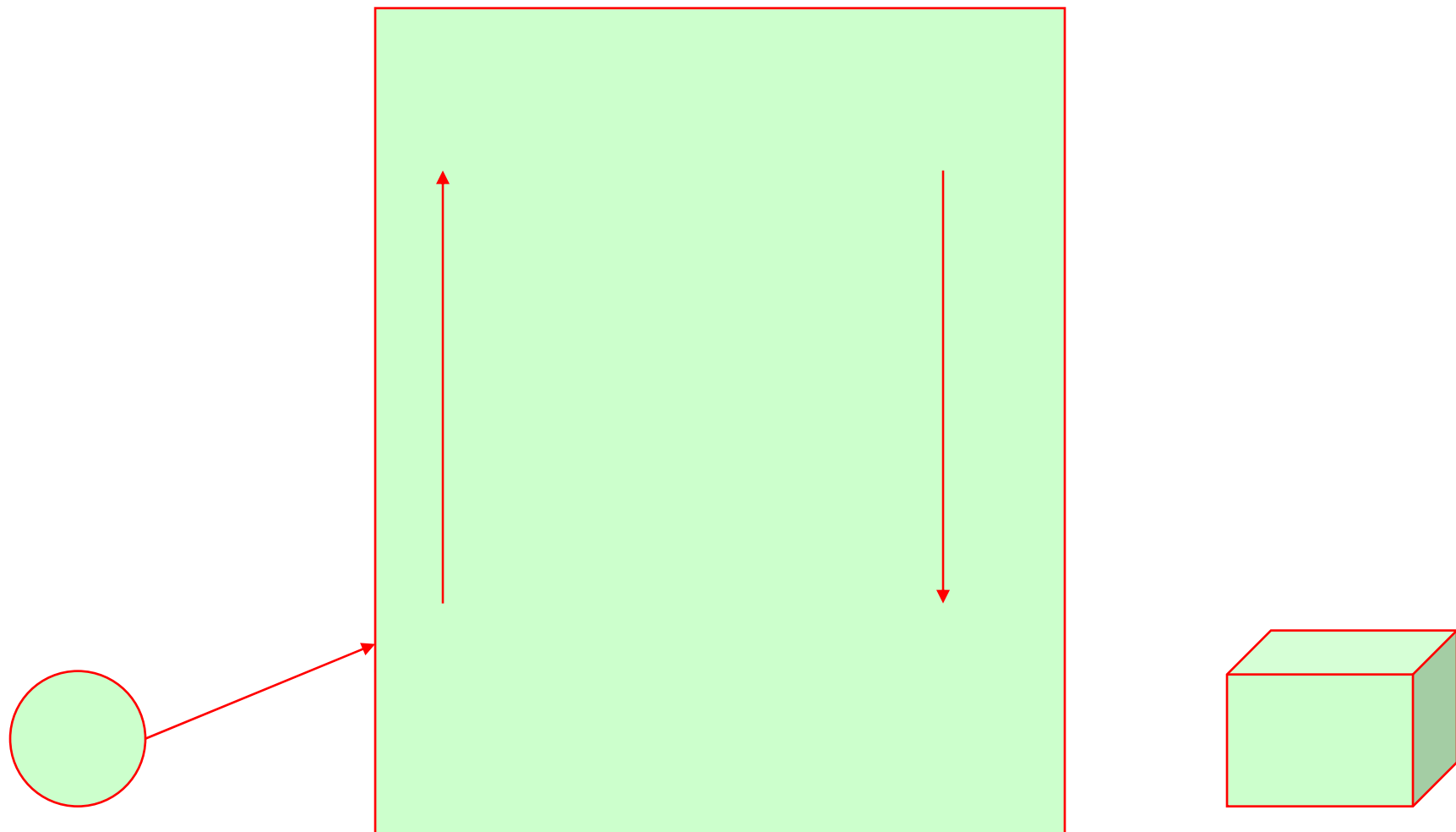




马哥教育

www.magedu.com





# VCL (Varnish)

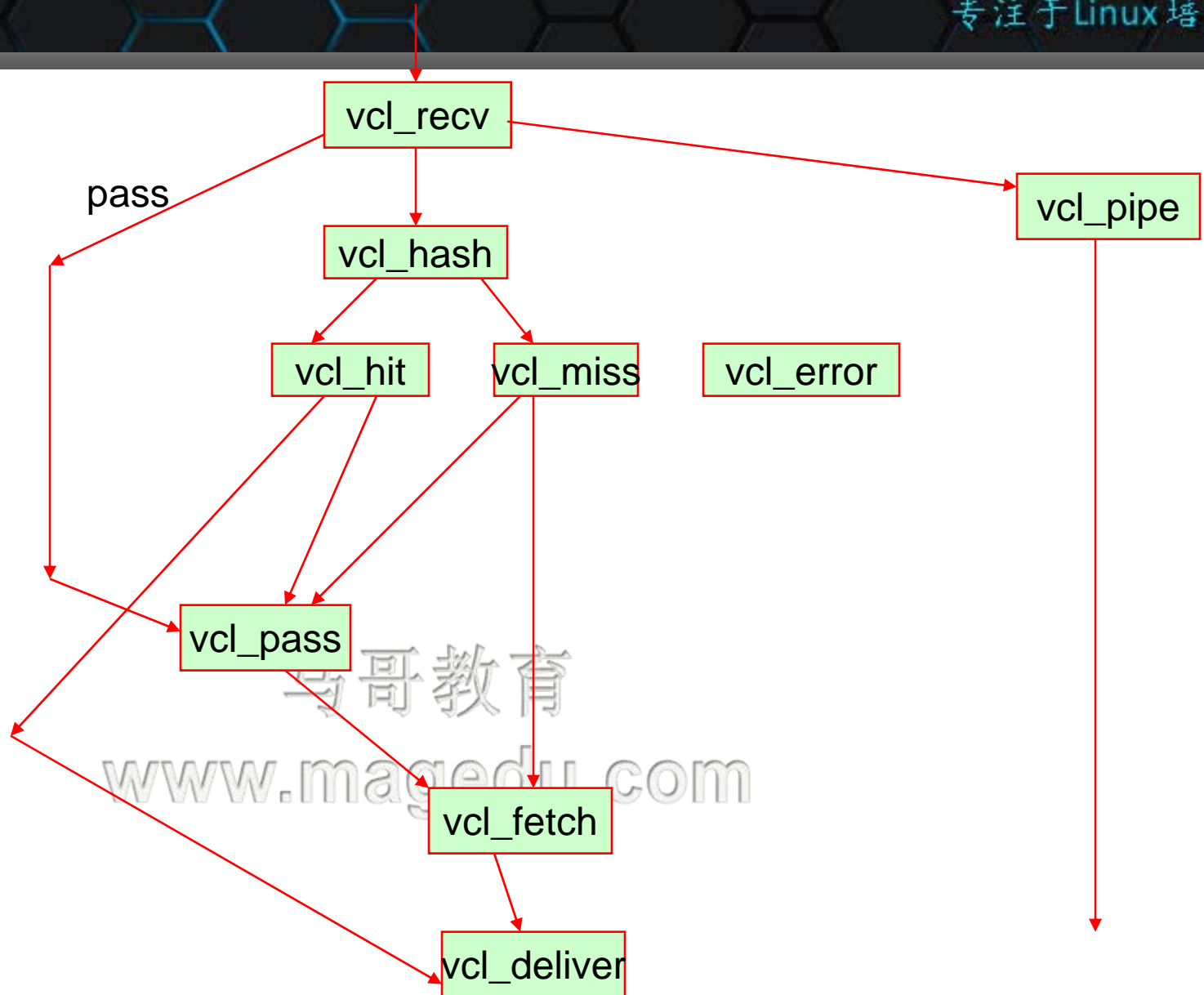
主讲：马永亮(马哥)

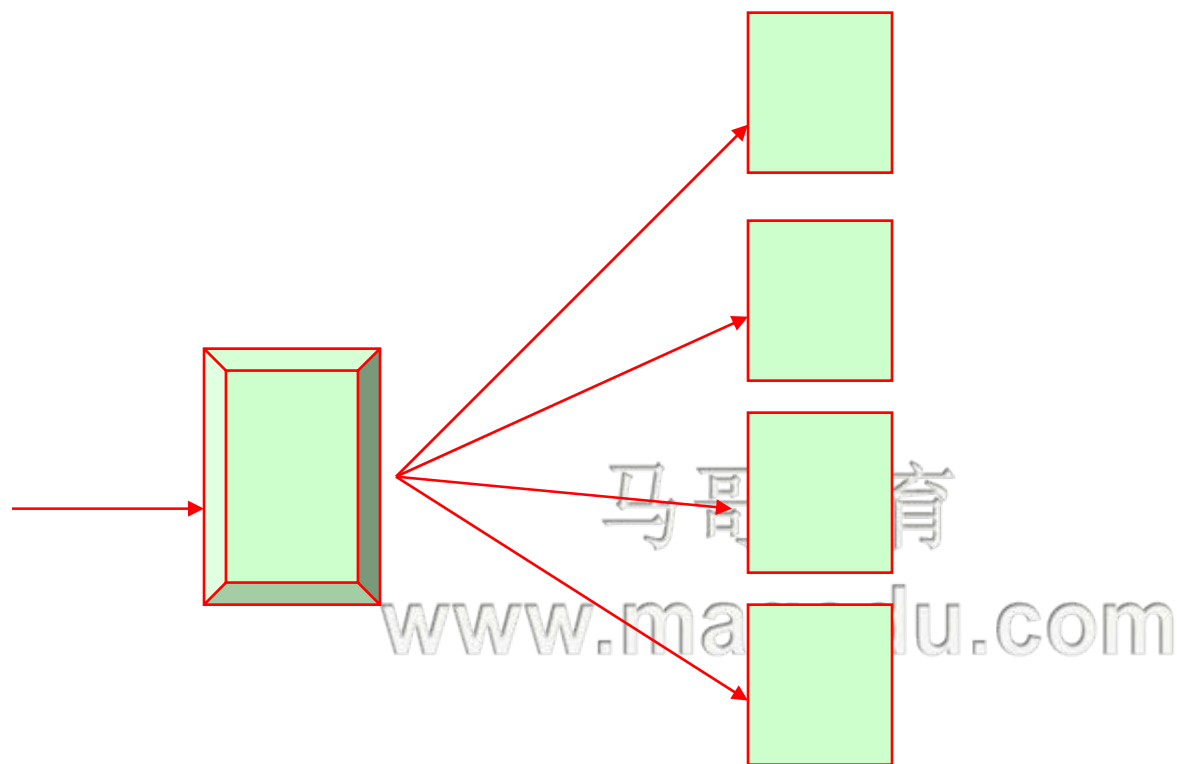
QQ群：169777636

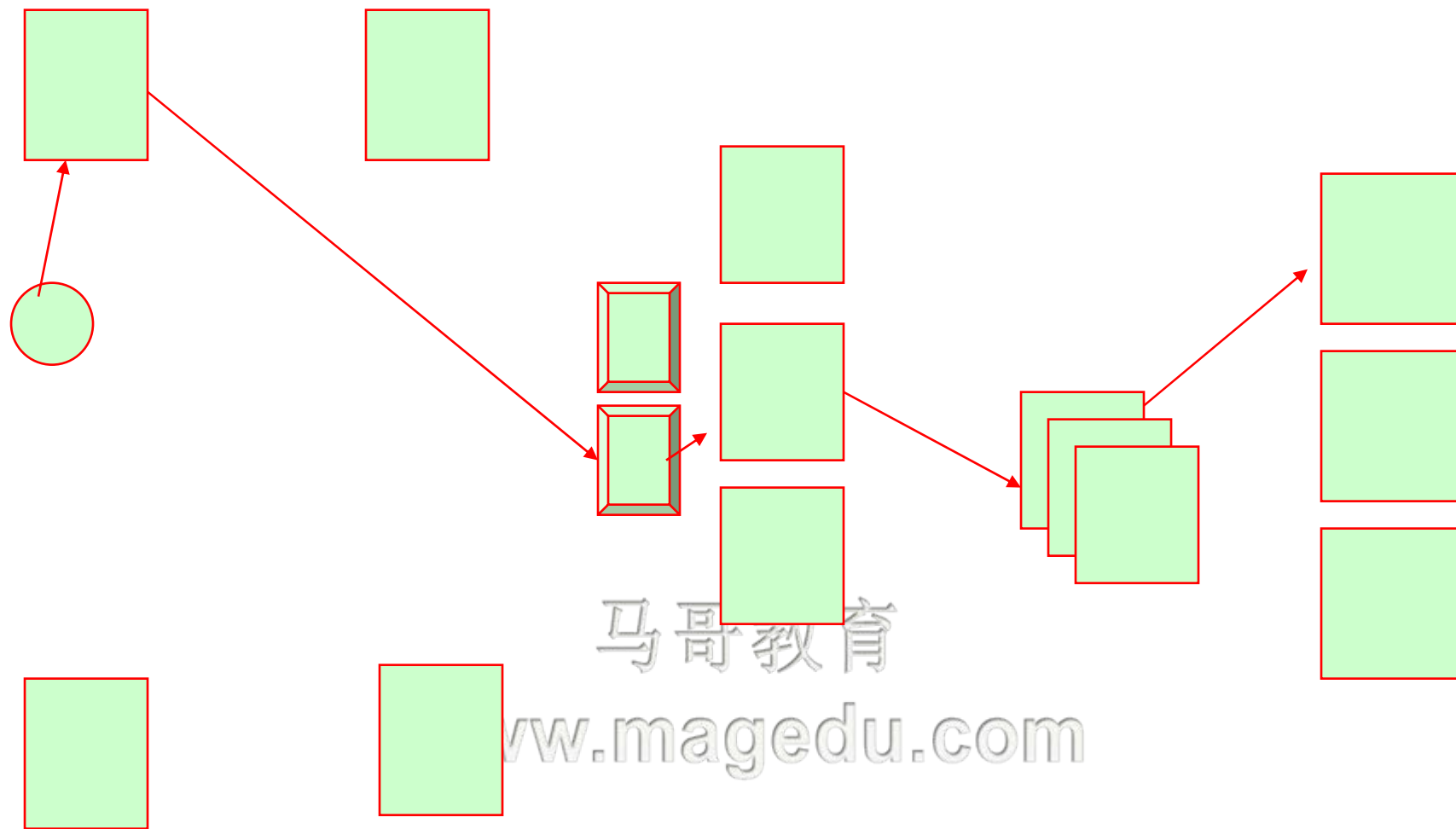
客服QQ：2813150558, 1661815153

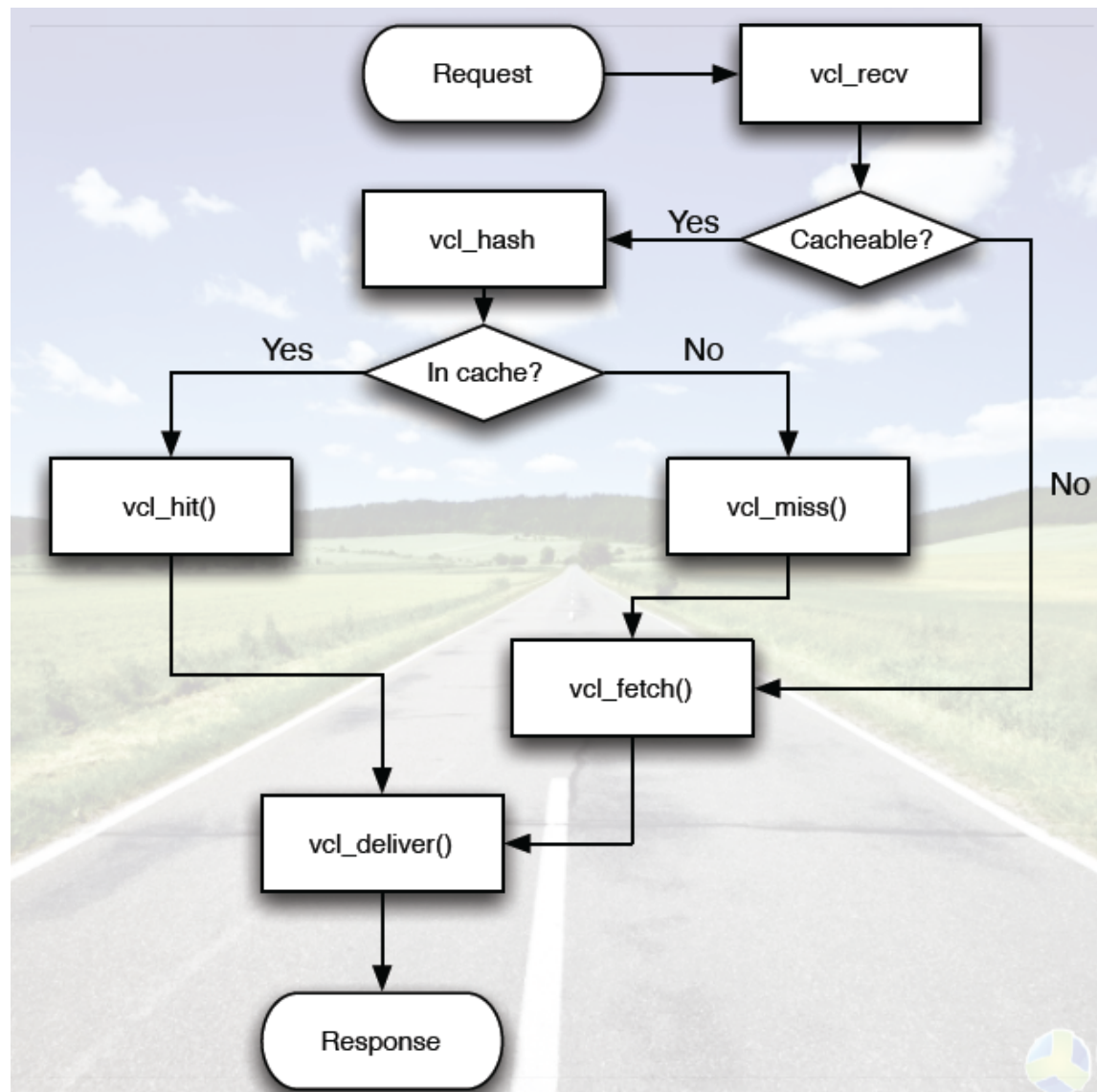
<http://www.magedu.com>

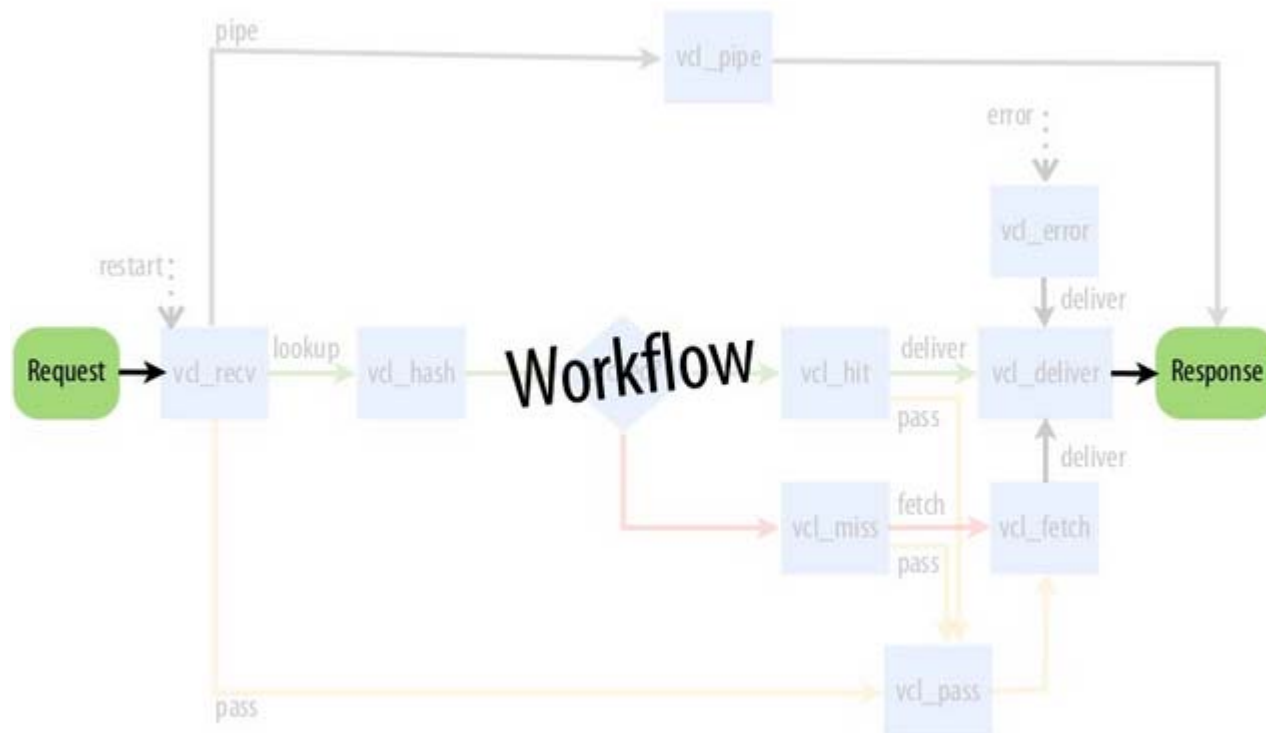
<http://mageedu.blog.51cto.com>



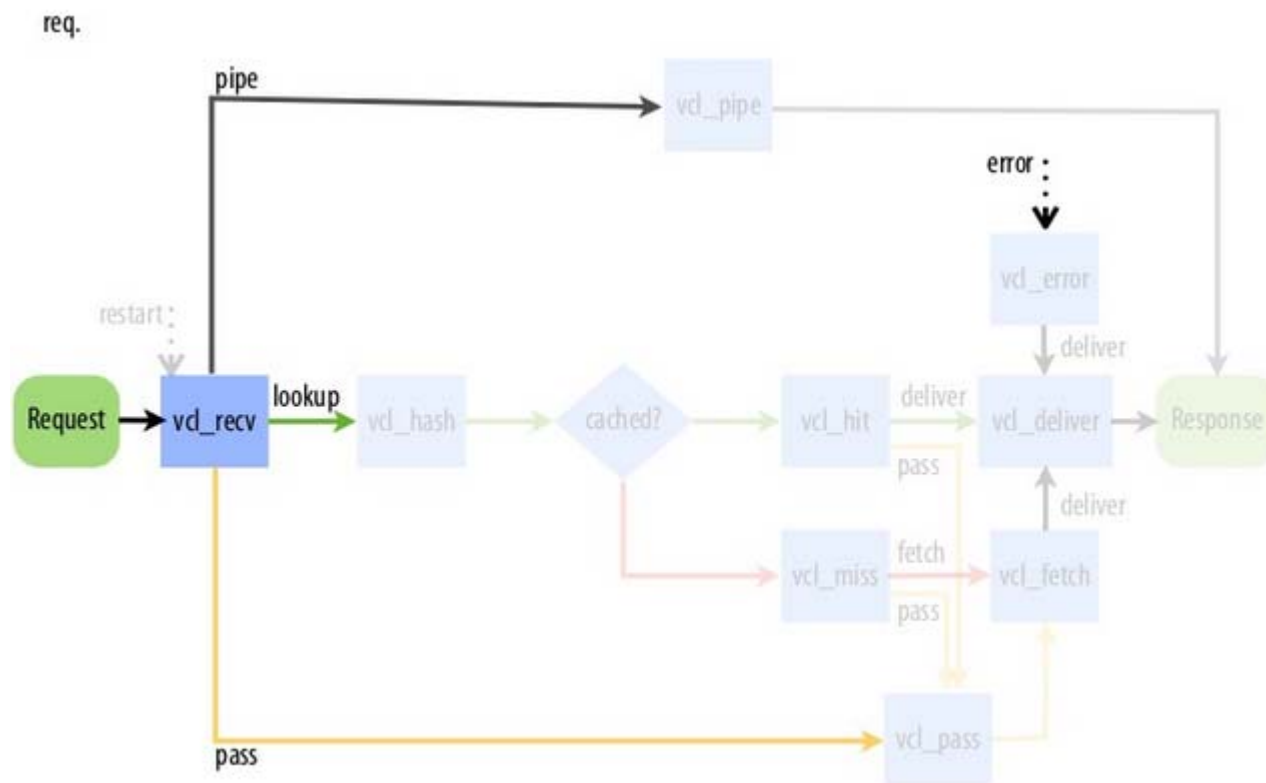




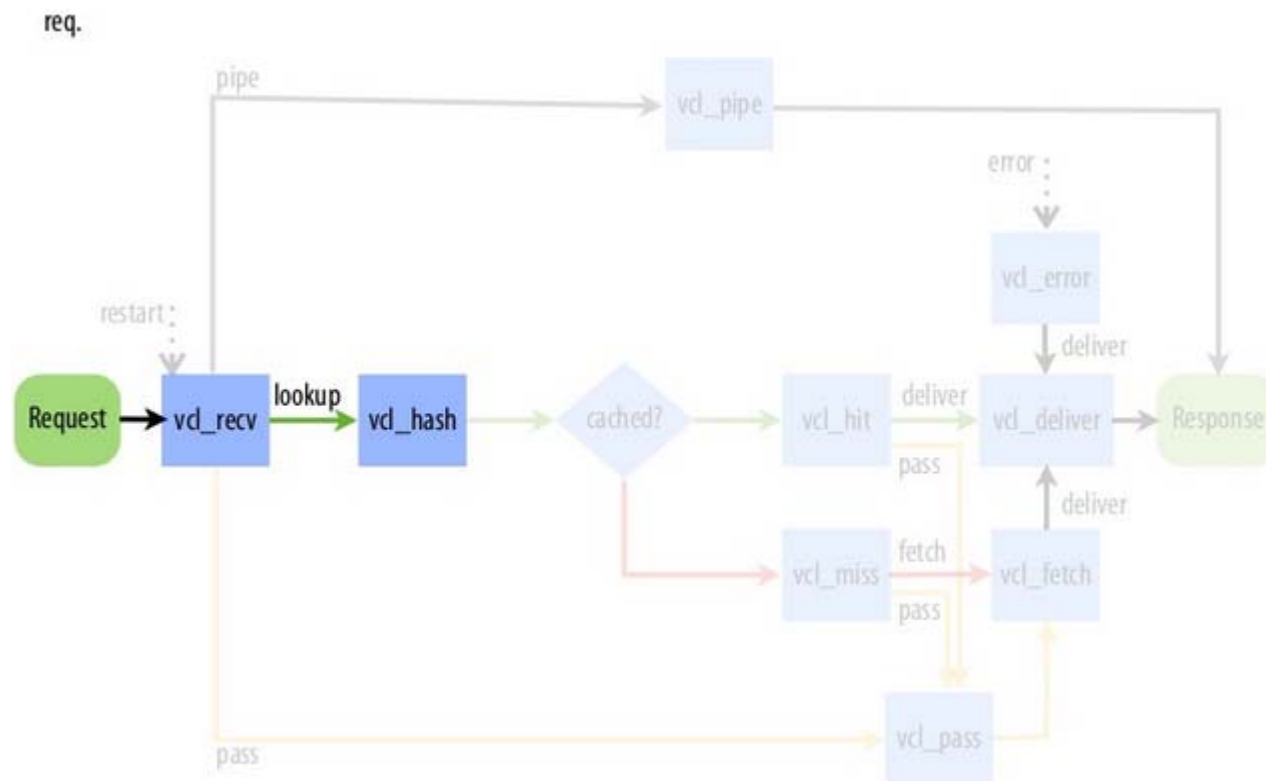




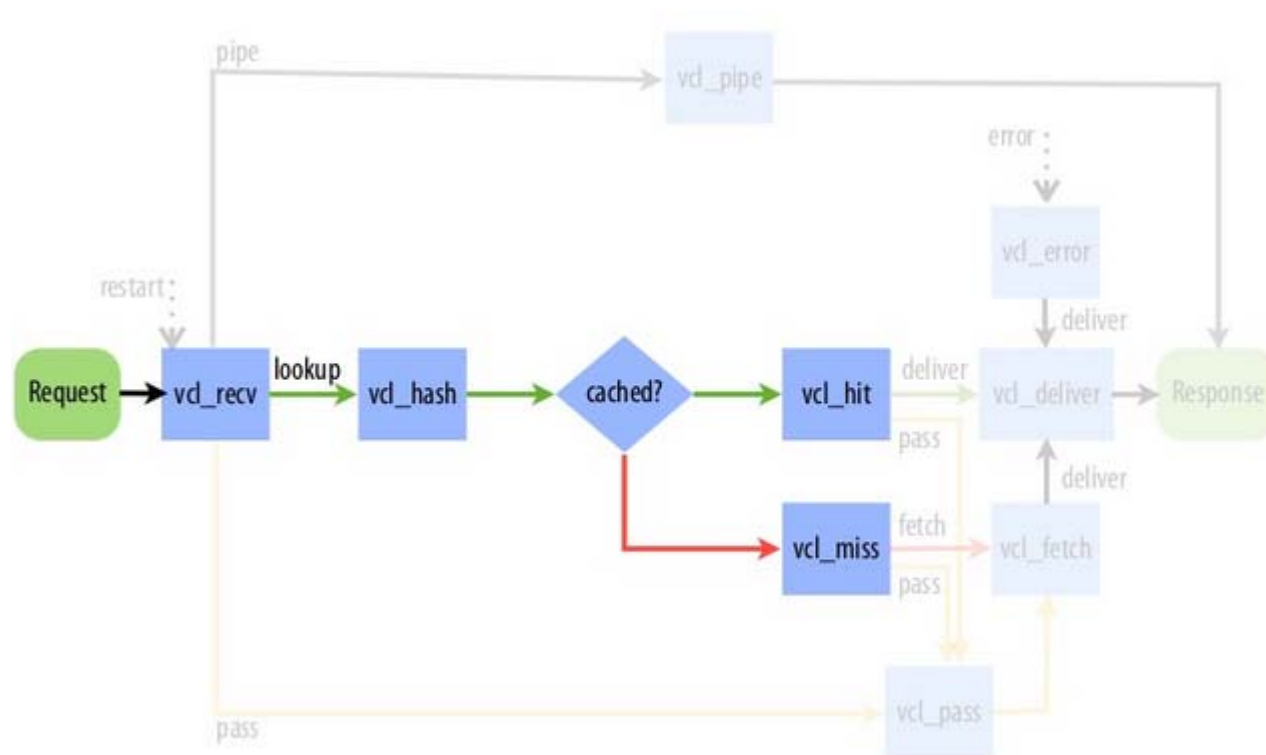


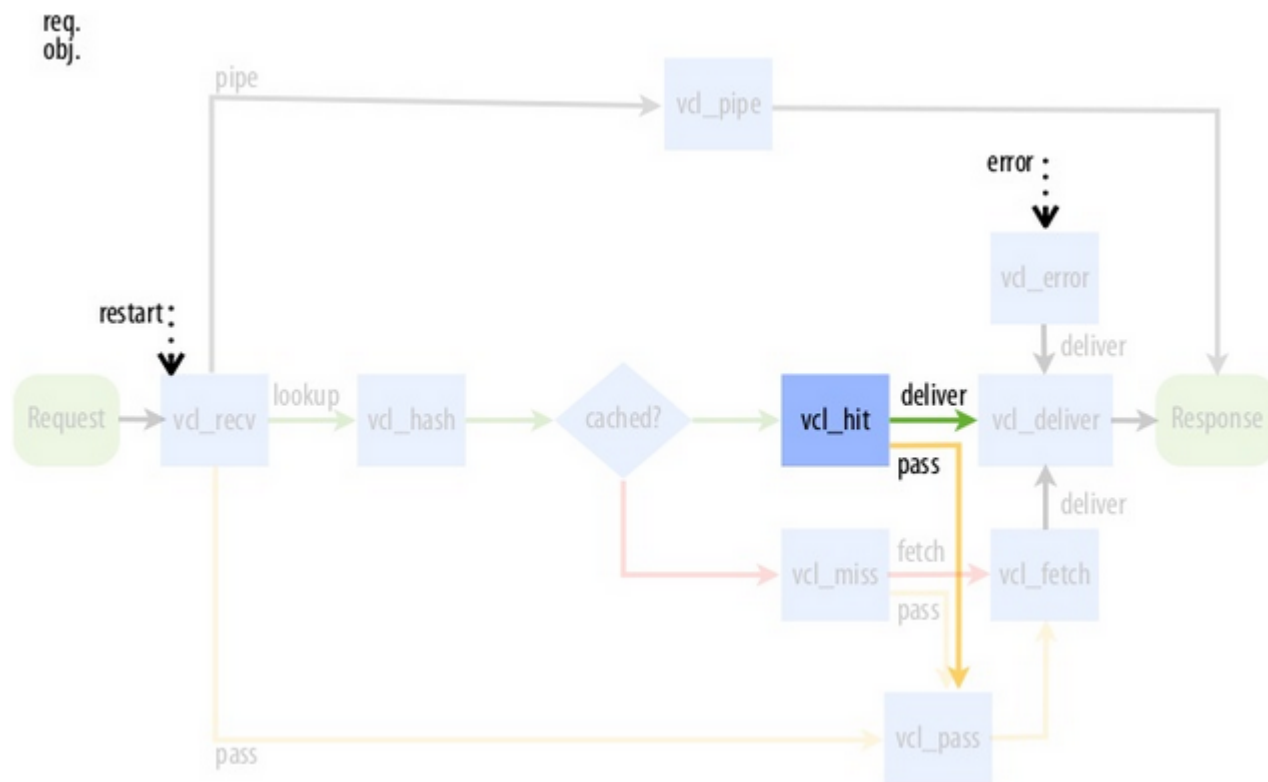


```
if (req.restarts == 0) {  
    if (req.http.x-forwarded-for) {  
        set req.http.X-Forwarded-For =  
            req.http.X-Forwarded-For + ", " + client.ip;  
    } else {  
        set req.http.X-Forwarded-For = client.ip;  
    }  
}  
if (req.request != "GET" && req.request != "HEAD" && req.request !=  
    "PUT" && req.request != "POST" && req.request != "TRACE" &&  
    req.request != "OPTIONS" && req.request != "DELETE") {  
    return (pipe);  
}  
if (req.request != "GET" && req.request != "HEAD") {  
    return (pass);  
}  
if (req.http.Authorization || req.http.Cookie) {  
    return (pass);  
}  
return (lookup);
```

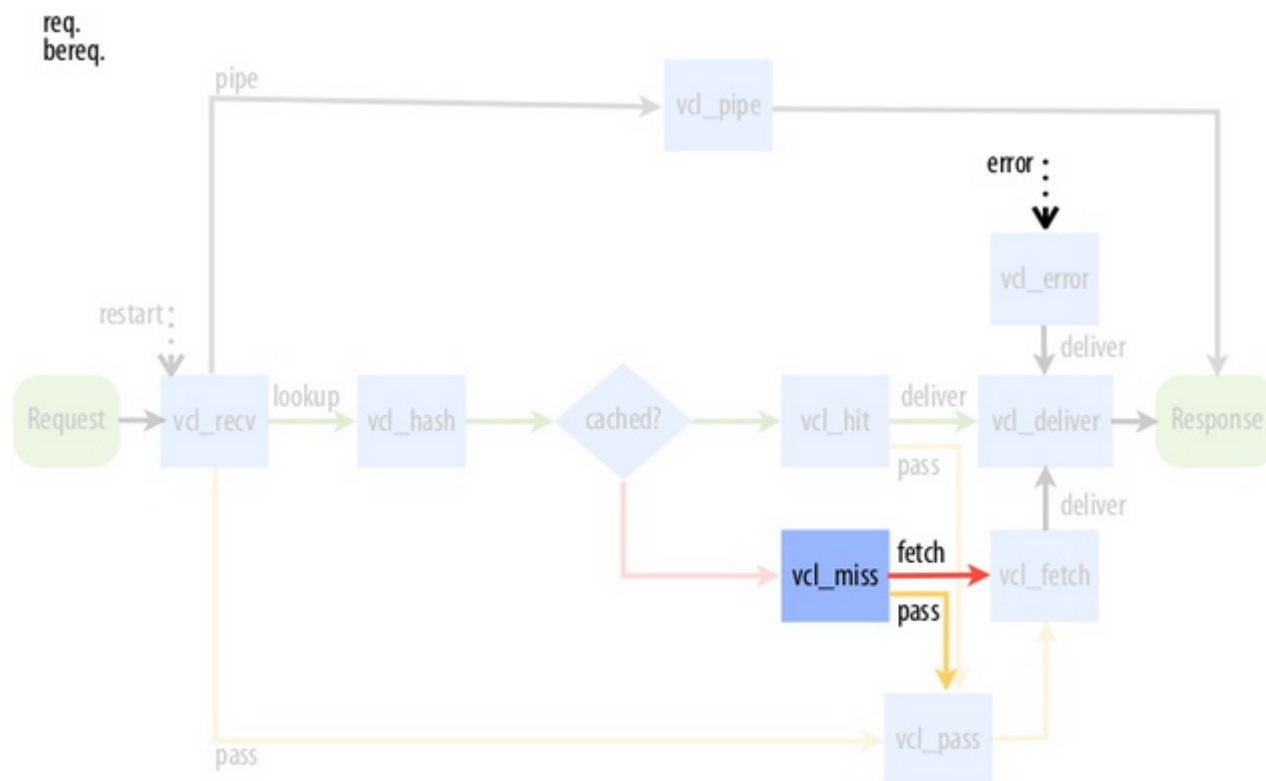


```
hash_data(req.url);  
if (req.http.host) {  
    hash_data(req.http.host);  
} else {  
    hash_data(server.ip);  
}  
return (hash);
```



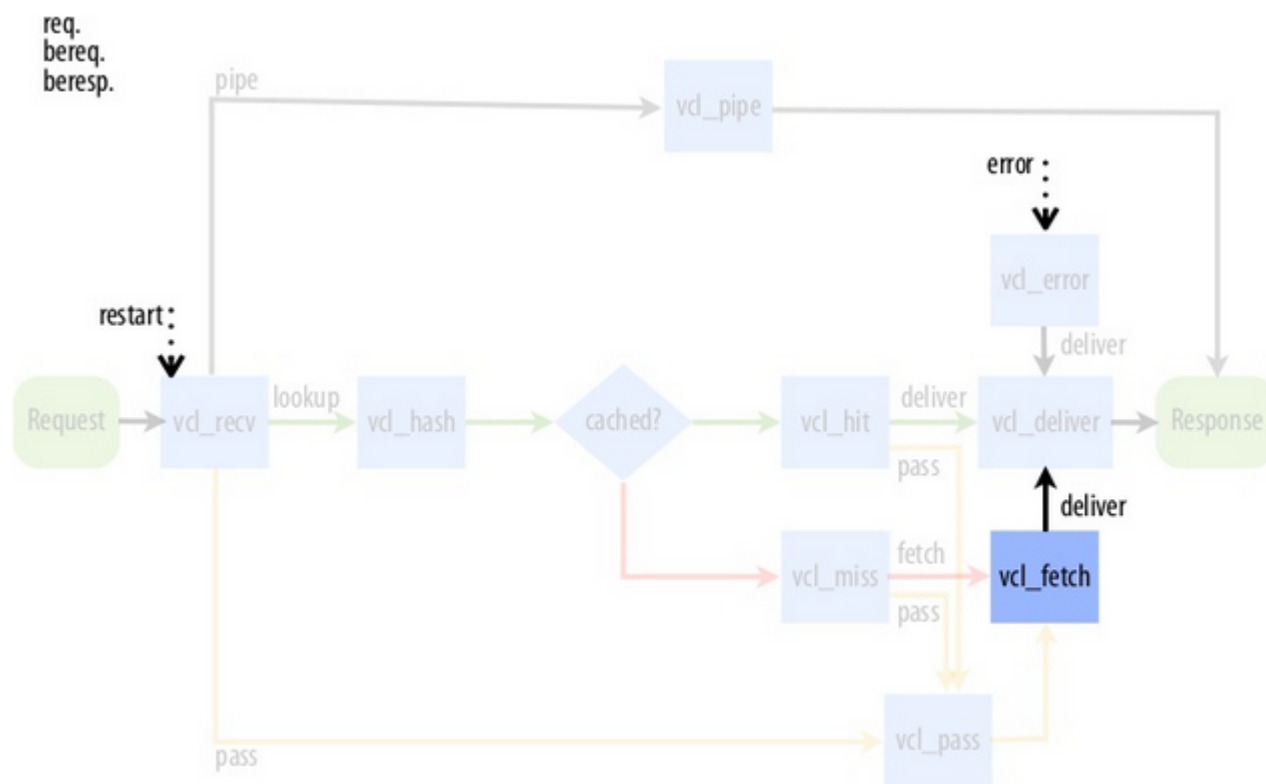


```
return (deliver);
```

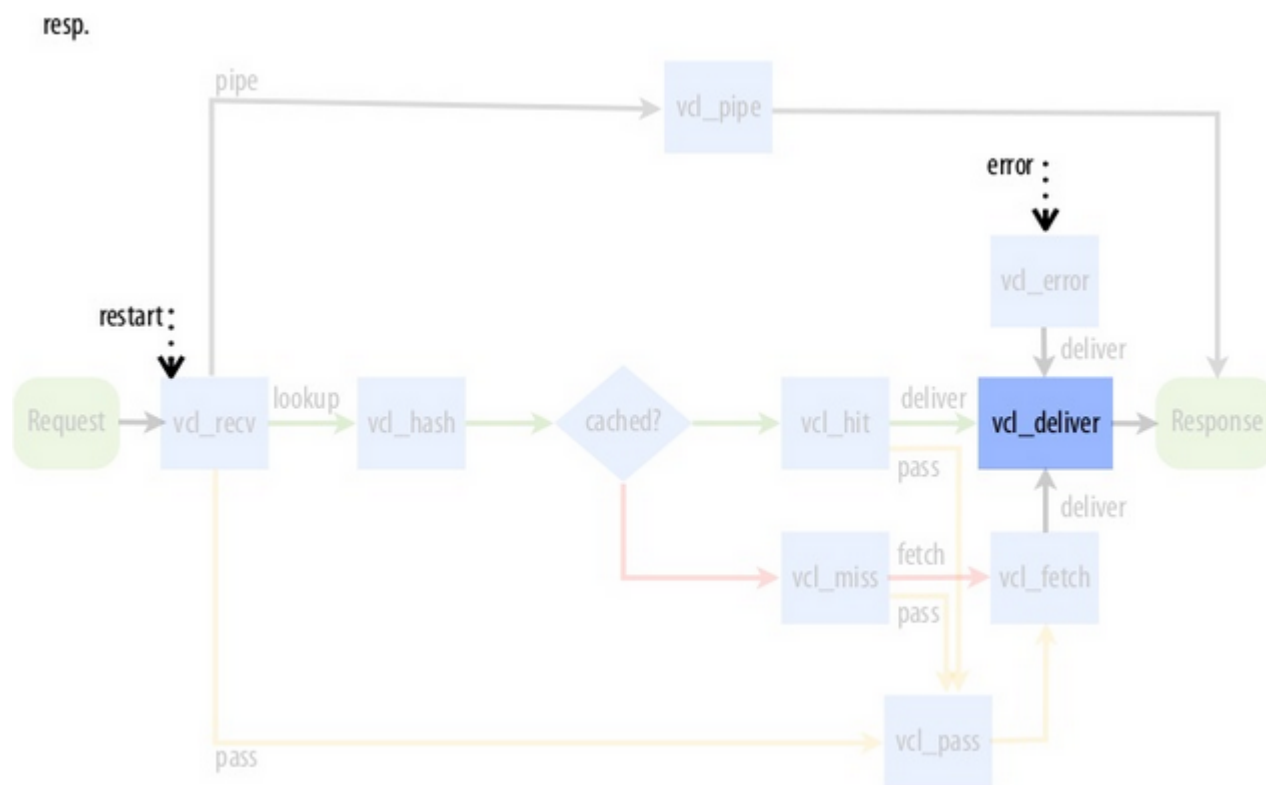




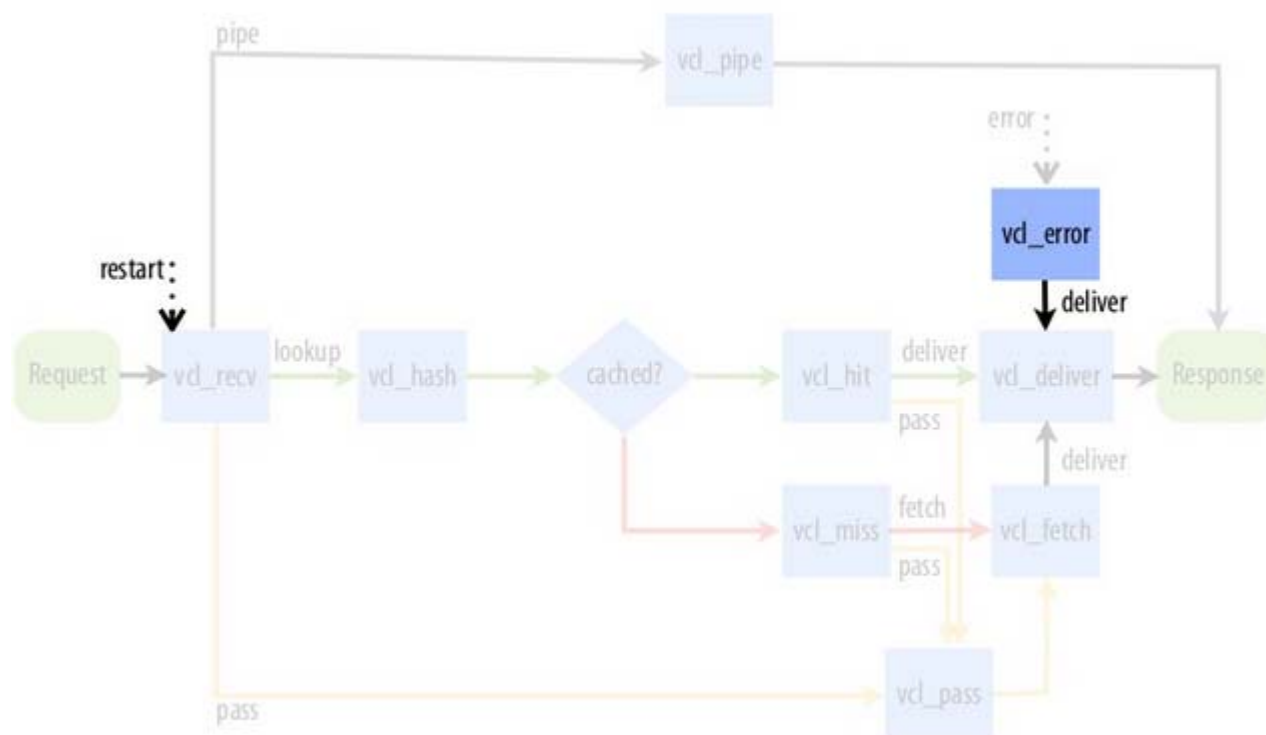
```
return (fetch);
```



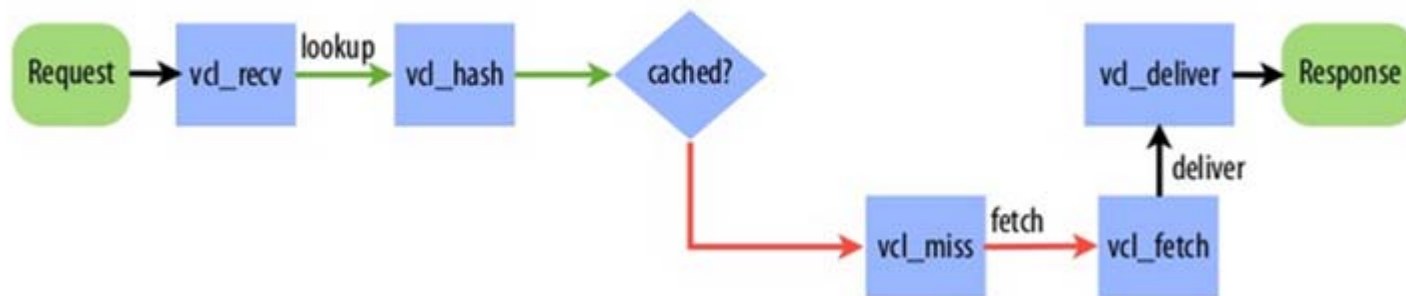
```
if (beresp.ttl <= 0s || beresp.http.Set-Cookie ||  
beresp.http.Vary == "*") {  
    set beresp.ttl = 120 s;  
    return (hit_for_pass);  
}  
return (deliver);
```



```
return (deliver);
```



```
set obj.http.Content-Type = "text/html; charset=utf-8";
set obj.http.Retry-After = "5";
synthetic {
<?xml version="1.0" encoding="utf-8"?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html>
<head>
<title>"} + obj.status + " " + obj.response + {"</title>
</head>
<body>
<h1>Error "} + obj.status + " " + obj.response + {"</h1>
<p>"} + obj.response + {"</p>
<h3>Guru Meditation:</h3>
<p>XID: "} + req.xid + {"</p>
<hr>
<p>Varnish cache server</p>
</body>
</html>
"};
return (deliver);
```

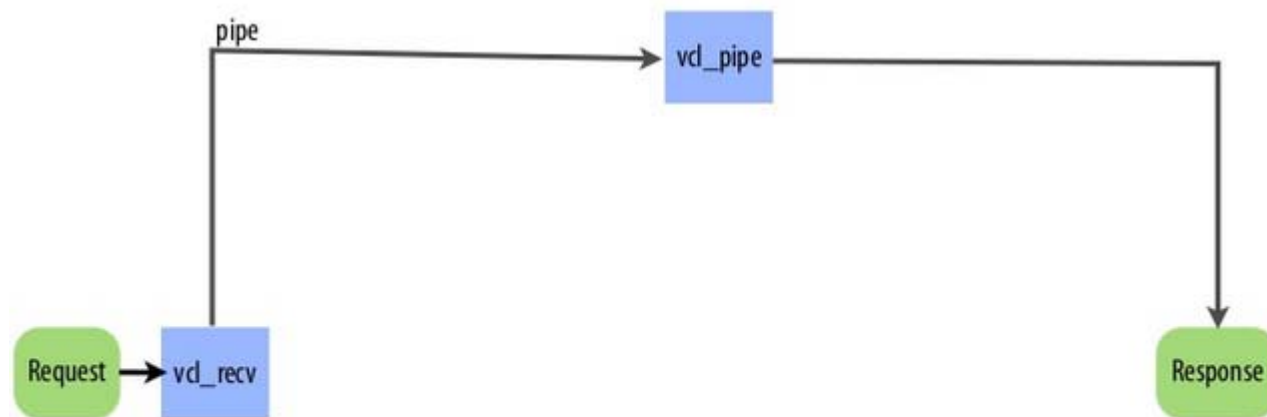




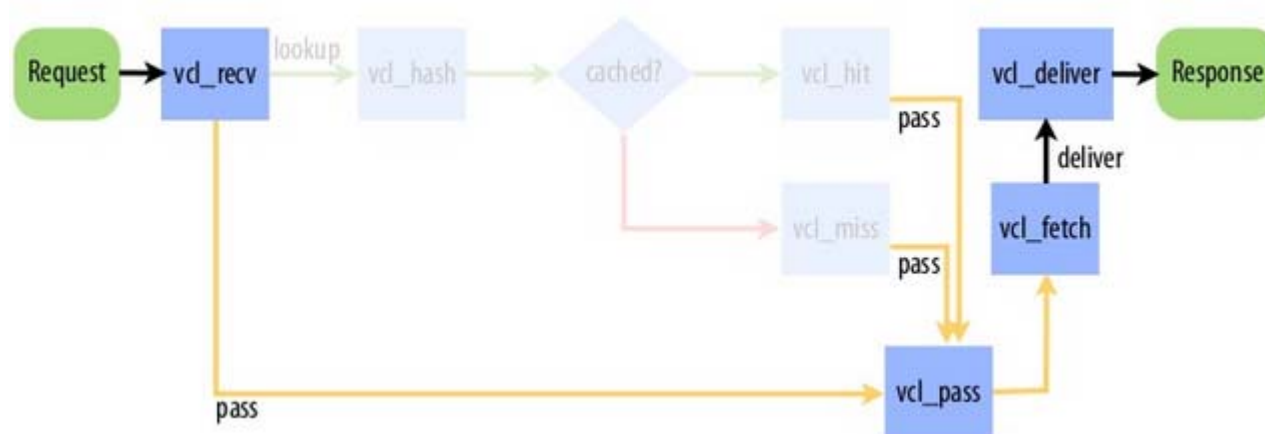


马哥教育

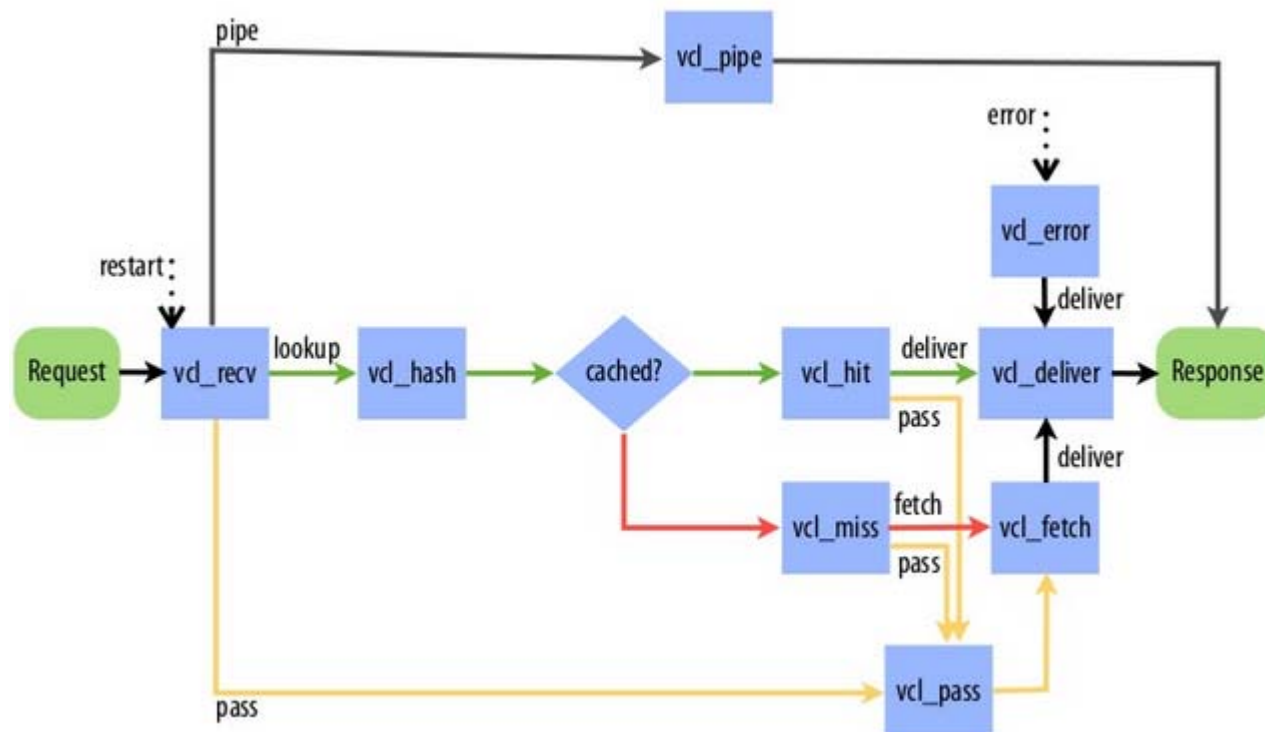
www.magedu.com



www.magedu.com



www.magedu.com

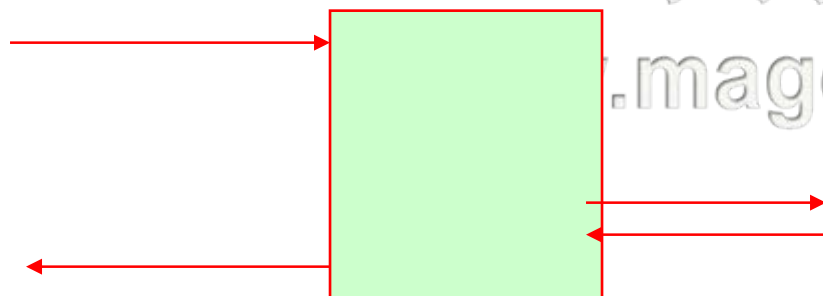


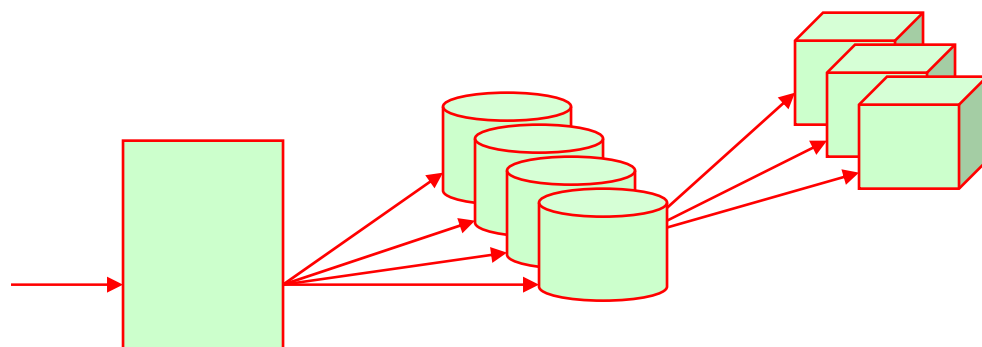
# Variable availability in VCL

Variable	recv	fetch	pass	miss	hit	error	deliver	pipe	hash
req.*	R/W	R/W	R/W	R/W	R/W	R/W	R/W	R/W	R/W
bereq.*		R/W	R/W	R/W				R/W	
obj.hits					R		R		
obj.ttl					R/W	R/W			
obj.grace					R/W				
obj.*					R	R/W			
beresp.*		R/W							
resp.*						R/W	R/W		

马哥教育

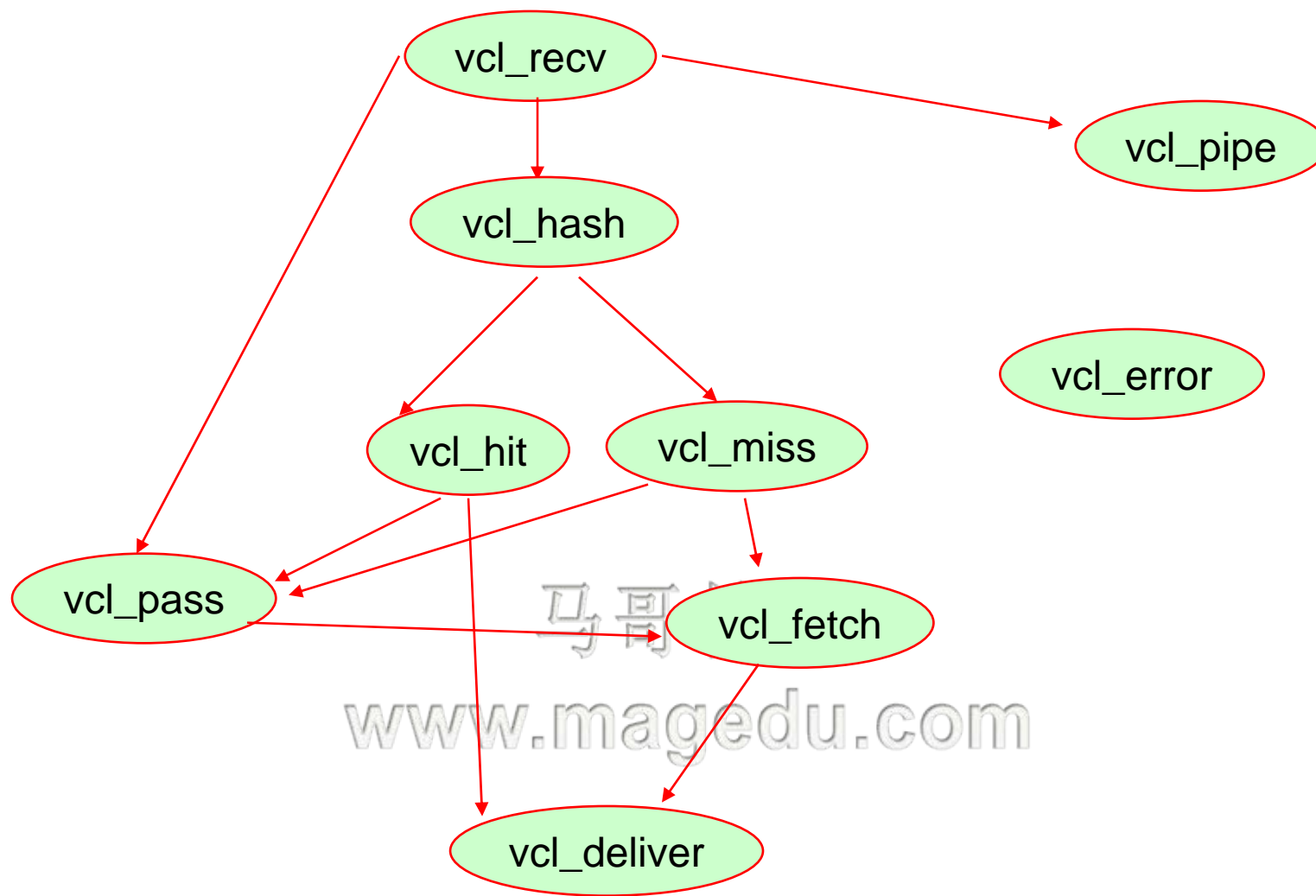
.magedu.com

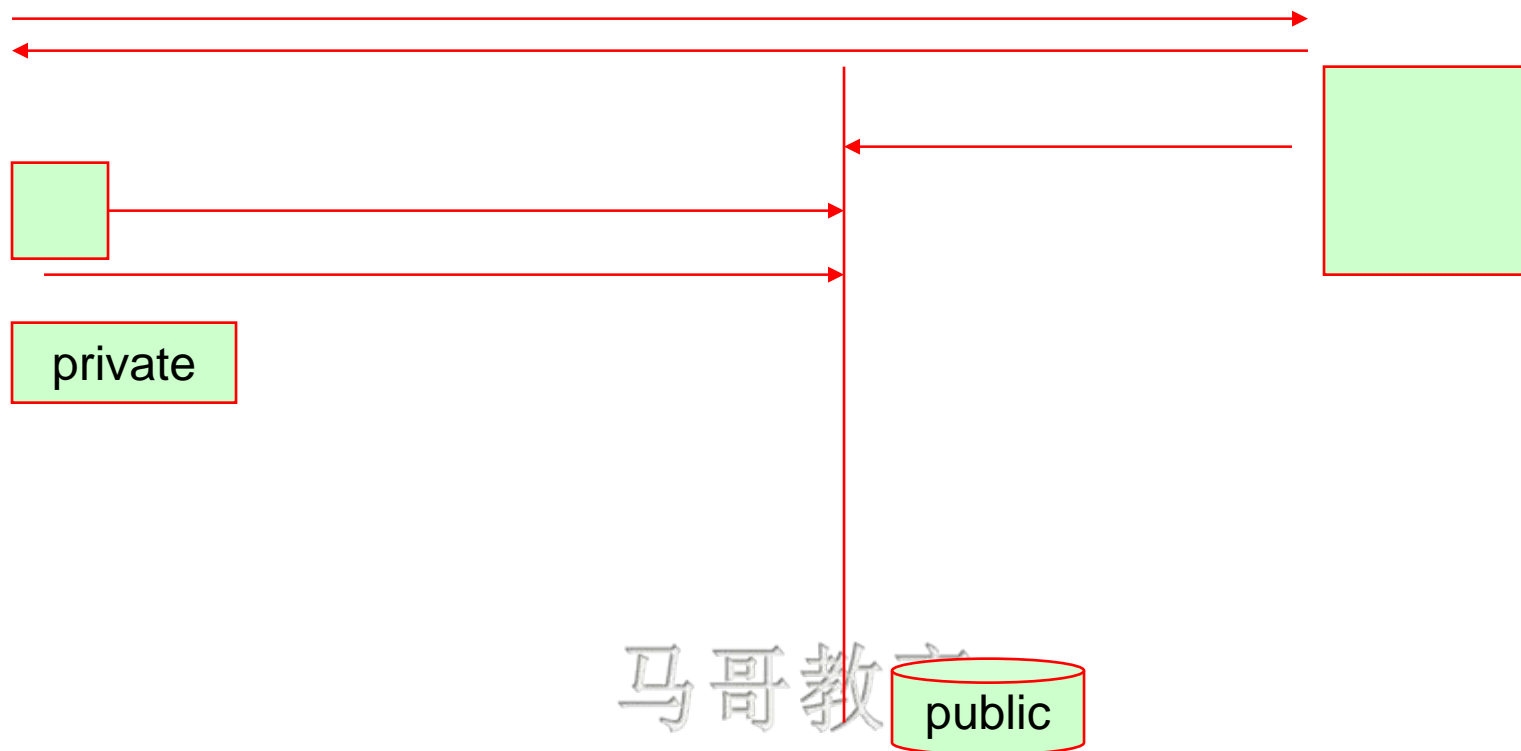




马哥教育

www.magedu.com

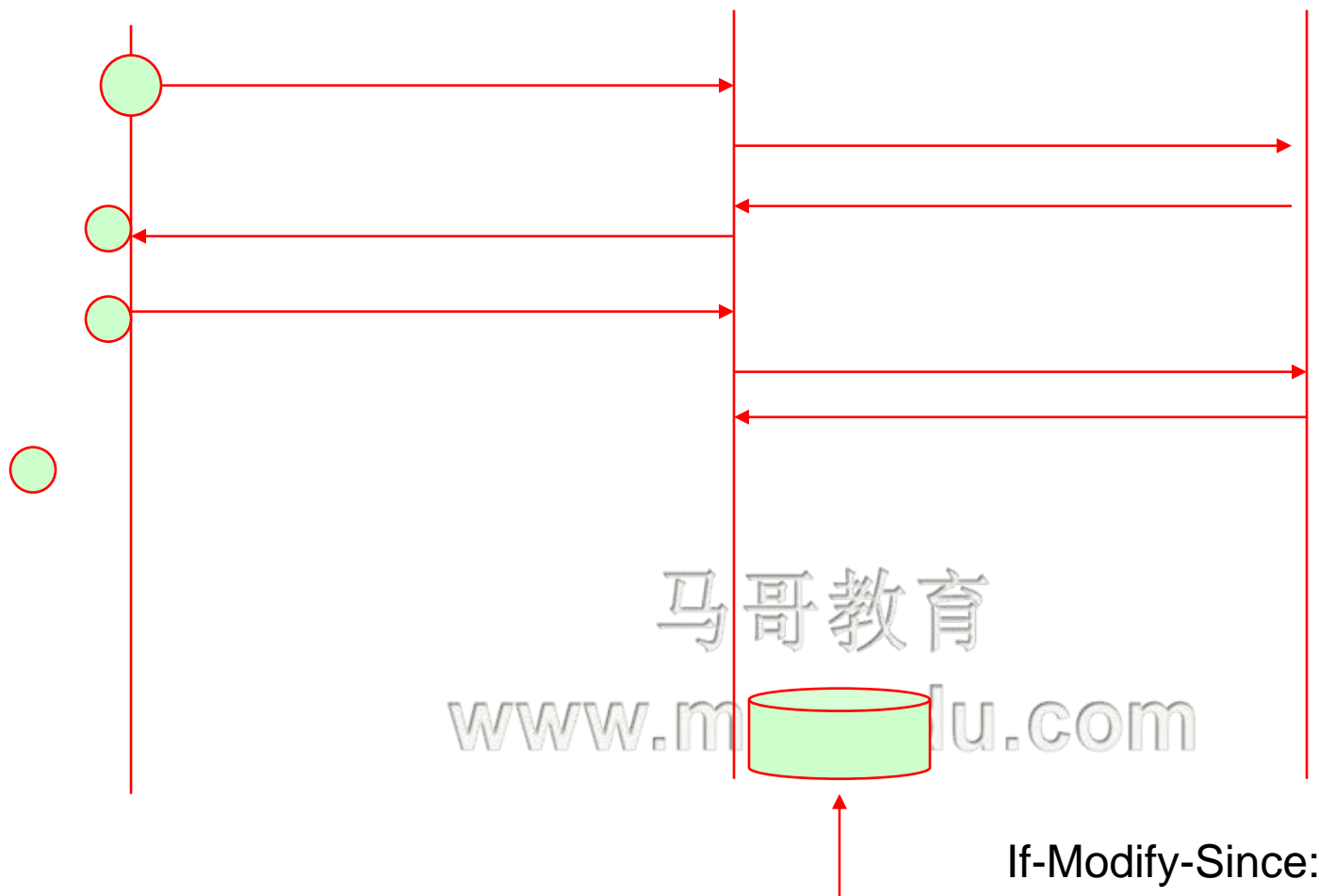




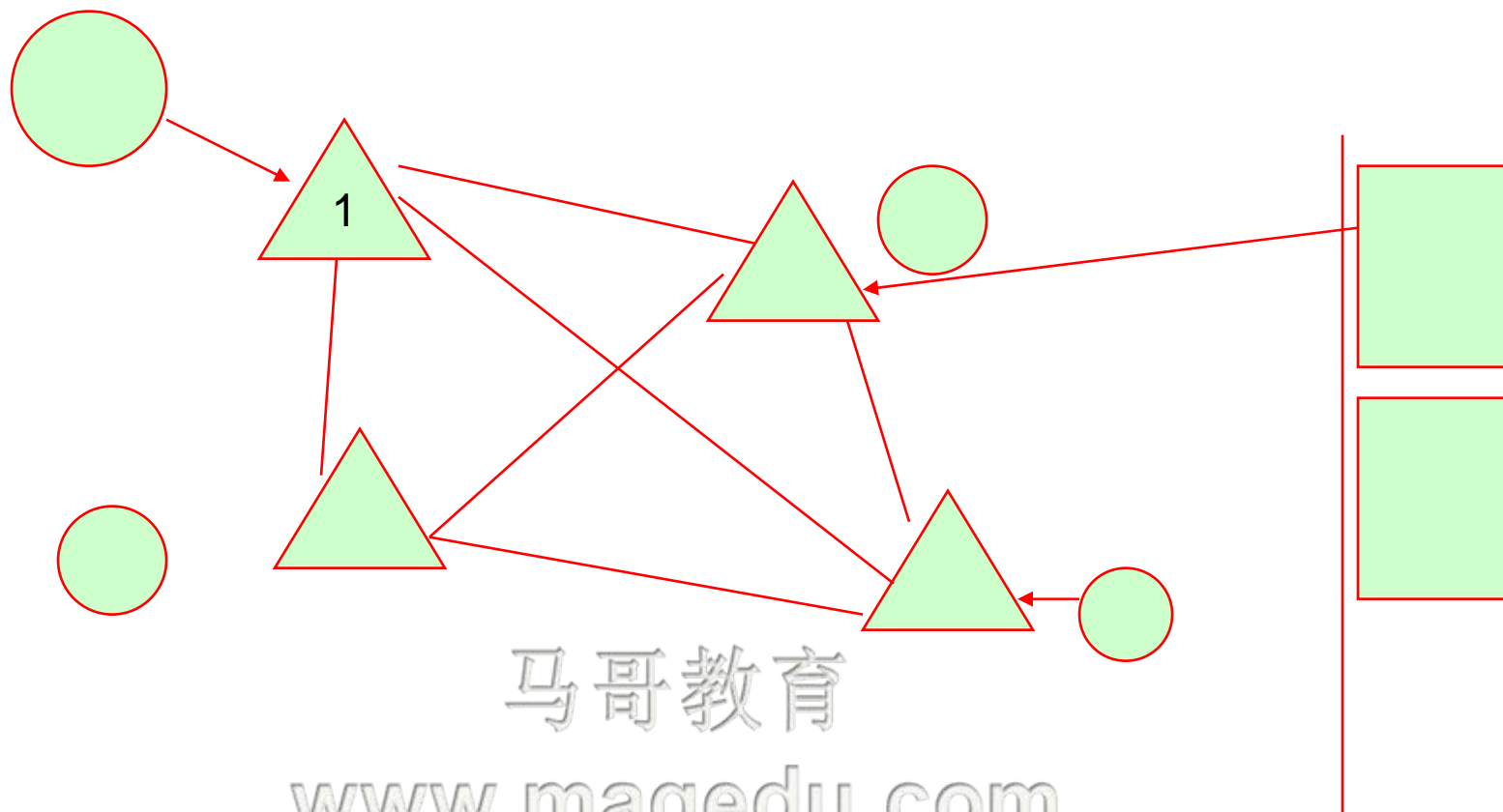
马哥教育  
www.magedu.com



# Cache-Control: max-age=600



# CDN: Content Delivery Network



www.magedu.com in CNAME cache1.cdn.com

www.magedu.com in CNAME cache4.cdn.com

## ❖ Bind, view

➡ Bind-dlz + MySQL

➡ Pgsq, Oracle, db4

➡ Squid: varnish

➡ Httpd: nginx

➡ Nginx + varnish

➡ Nginx + Squid

马哥教育

www.magedu.com

## ❖ 浏览器Cache

➡ 私有缓存

➡ 可以缓存“private”响应

## ❖ Cache Proxy

## ❖ Surrogates - 缓存代理

➡ Web加速

➡ 共享缓存

➡ 只能缓存“public”响应

马哥教育  
www.magedu.com



# 一些数据

- ❖ 多至**43%**的**Web**请求访问不可缓存的内容
- ❖ **Web**缓存命中率为**40%**左右是比较现实的

马哥教育

www.magedu.com

# 页面静态化

## ❖ 理想状态：

➡ 静态化所有页面——实际做不到

## ❖ 如何静态化？

➡ 由模板生成静态页面

➡ 定时或有更新时

## ❖ 困难：

➡ 成千上万个页面的静态化不现实

➡ 页面有变体，不适合静态化

➡ 多台服务器维护麻烦

➡ 页面既包含静态内容，又包含动态内容

# 页面缓存

- ❖ 按需静态化
- ❖ 前端**cache**存储动态页面内容
  - ➔ **mod\_cache**
  - ➔ **Squid**
  - ➔ **Varnish\***
- ❖ 遵循**HTTP**的**Cache**规范
- ❖ 可精细控制**Cache**
  - ➔ **Expires**
  - ➔ **Cache-control**
- ❖ 适合**GET**类**CGI**请求

马哥教育

www.magedu.com

- ❖ 要充分发挥**Cache**的效用，就必须了解**HTTP Cache**规范和机制
- ❖ 条件请求
- ❖ **Expiration**
- ❖ **Cache-control**
- ❖ **Cookie**

马哥教育

www.magedu.com



# 条件请求

## ❖ If-Modified-Since/Last-Modified

### ❖ 1. 服务器响应:

➤ 200 OK

➤ Last-Modified: ...

### ❖ 2. 浏览器请求

➤ GET ... HTTP/1.1

➤ If-Modified-Since: ...

### ❖ 3. 服务器响应

➤ 200 OK 或

➤ 304 Not Modified

马哥教育

www.magedu.com

## 条件请求(2)

### ❖ If-None-Match / ETag

#### ❖ 1. Server响应

- 200 OK
- ETag: abcdef

#### ❖ 2. 浏览器请求

- GET ... HTTP/1.1
- If-None-Match: abcdef

#### ❖ 3. Server响应

- 200 OK 或
- 304 Not Modified

马哥教育

www.magedu.com

## ❖ Expires

- ➡ 设置对象的绝对失效时间
- ➡ 是HTTP/1.0规范

## ❖ max-age

- ➡ 是Cache-control的一个指令
- ➡ 设置对象的年龄（秒数）
- ➡ 是HTTP/1.1规范

## ❖ max-age优先于Expires

马哥教育

www.magedu.com

```
cache-control:max-age=120,no-transform,proxy-revalidation
```

## ❖ no-cache=[Set-cookie]

- ➡ 浏览器和squad都不可以缓存
- ➡ 再次使用前需发送条件请求

## ❖ max-age=NNN

- ➡ 指定缓存有效时间（秒）

## ❖ public

- ➡ 浏览器和squad都可以缓存
- ➡ max-age指定缓存时间

## ❖ private

- ➡ 只有浏览器可以缓存

## ❖ HTTP请求携带Cookies:

➡ Cookie: name=profile, domain=www.bigmoney.com,  
path=mysecrets.html, cookie=day-trader

## ❖ HTTP响应设置Cookies:

➡ Set-cookie: name=profile, domain=www.bigmoney.com,  
path=mysecrets.html, cookie=day-trader

## ❖ Cached content

➡ 需要的时候才设置Cookies——太理想化

➡ Squid添加一个IMS头

www.magedu.com

# 一些典型的Cache场景

## ❖ 返回个人信息

➡ Cache-control: private, max-age=NNN

## ❖ CGI写操作

➡ Cache-control: no-store, no-cache

## ❖ CGI获取状态等信息

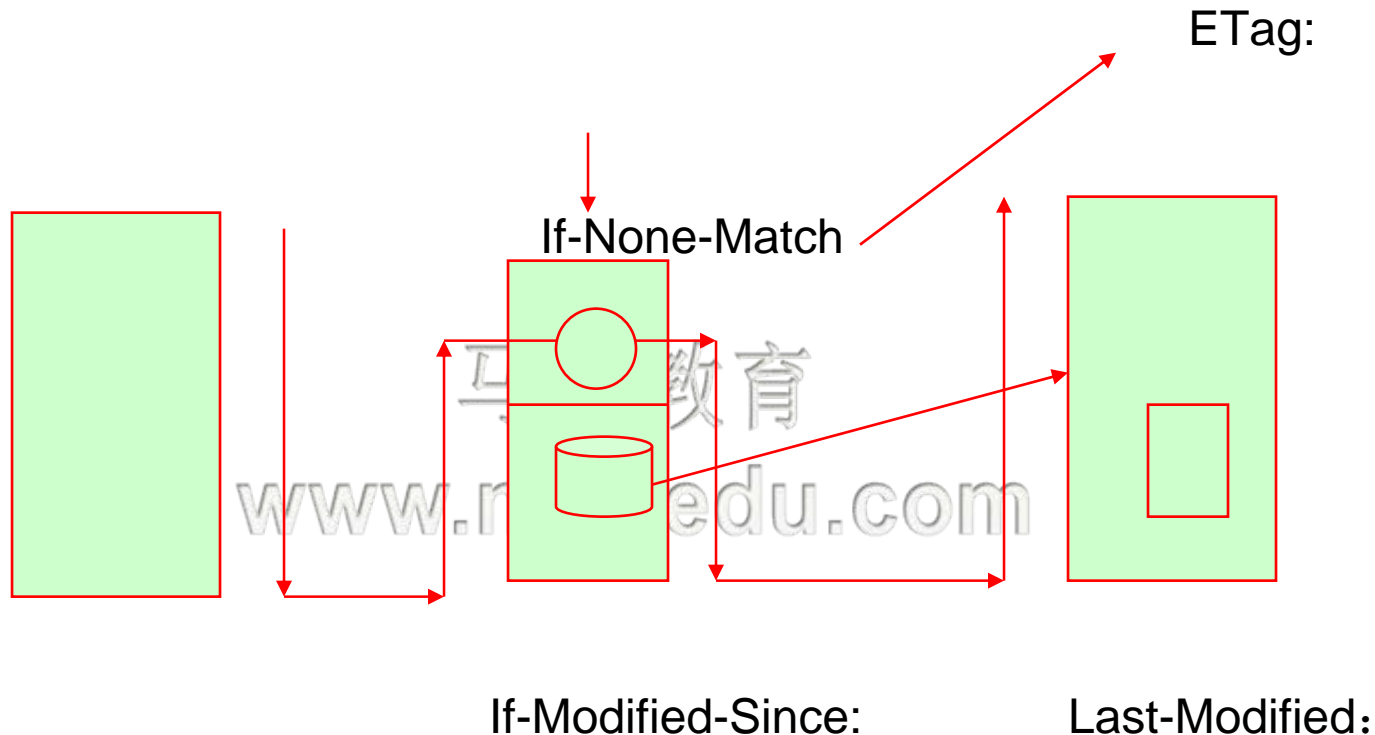
➡ Cache-control: public, max-age=NNN

马哥教育

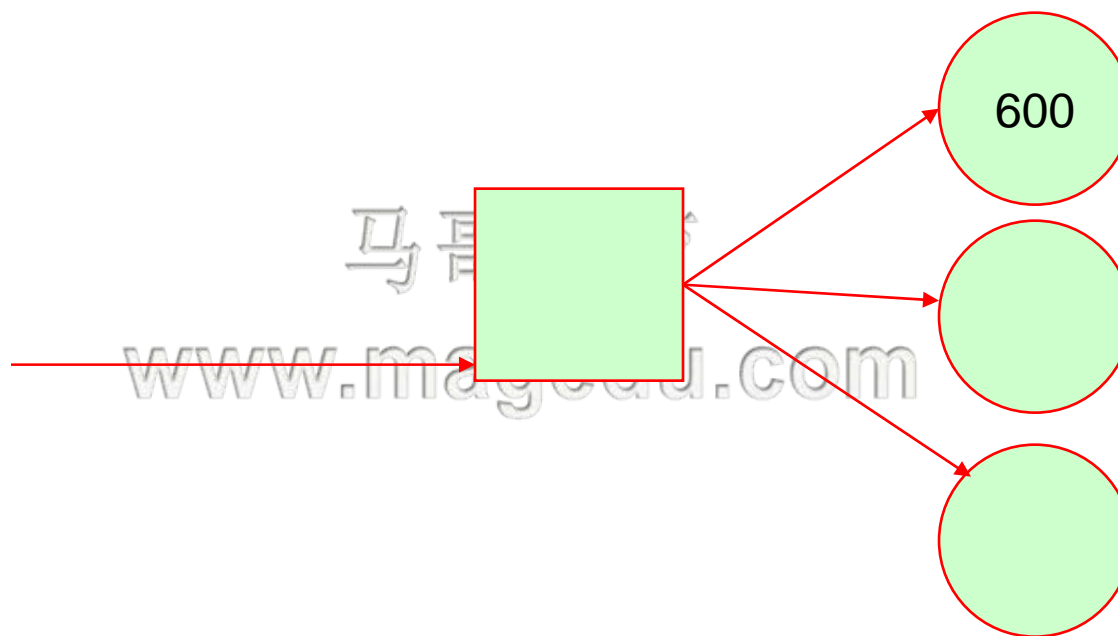
www.magedu.com

# GET, HEAD

- ❖ [www.magedu.com](http://www.magedu.com)
- ❖ Via: [www.myproxy.com](http://www.myproxy.com)
- ❖ Expire: 2012-12-21 00:00:00



- ❖ apache, mod\_proxy, cache (disk, memory)
- ❖ Nginx, cache (disk, )
- ❖ squid
- ❖ varnish (disk, memory)



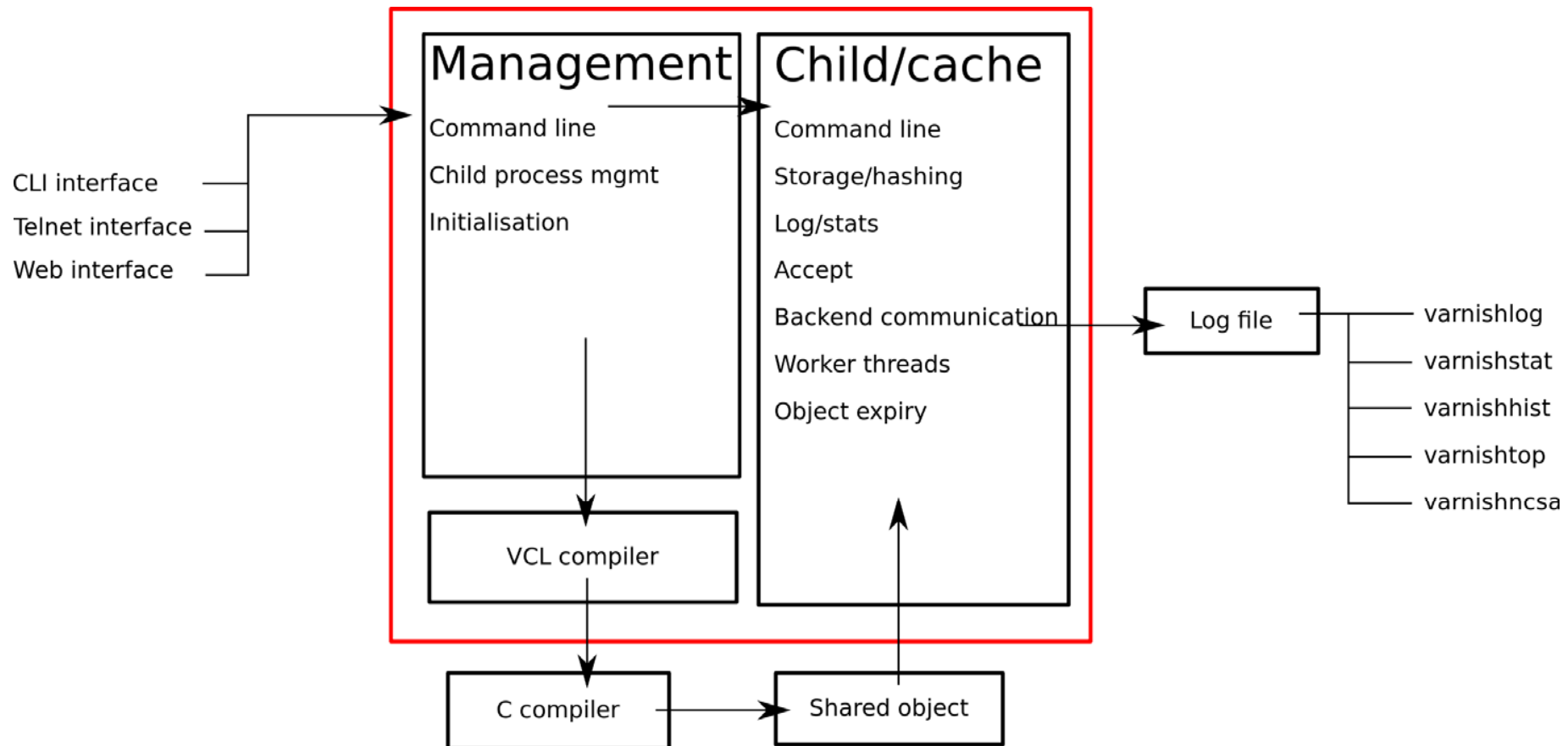


❖ Cache-Control: max-age=10

➡ s-maxage=10

马哥教育

www.magedu.com

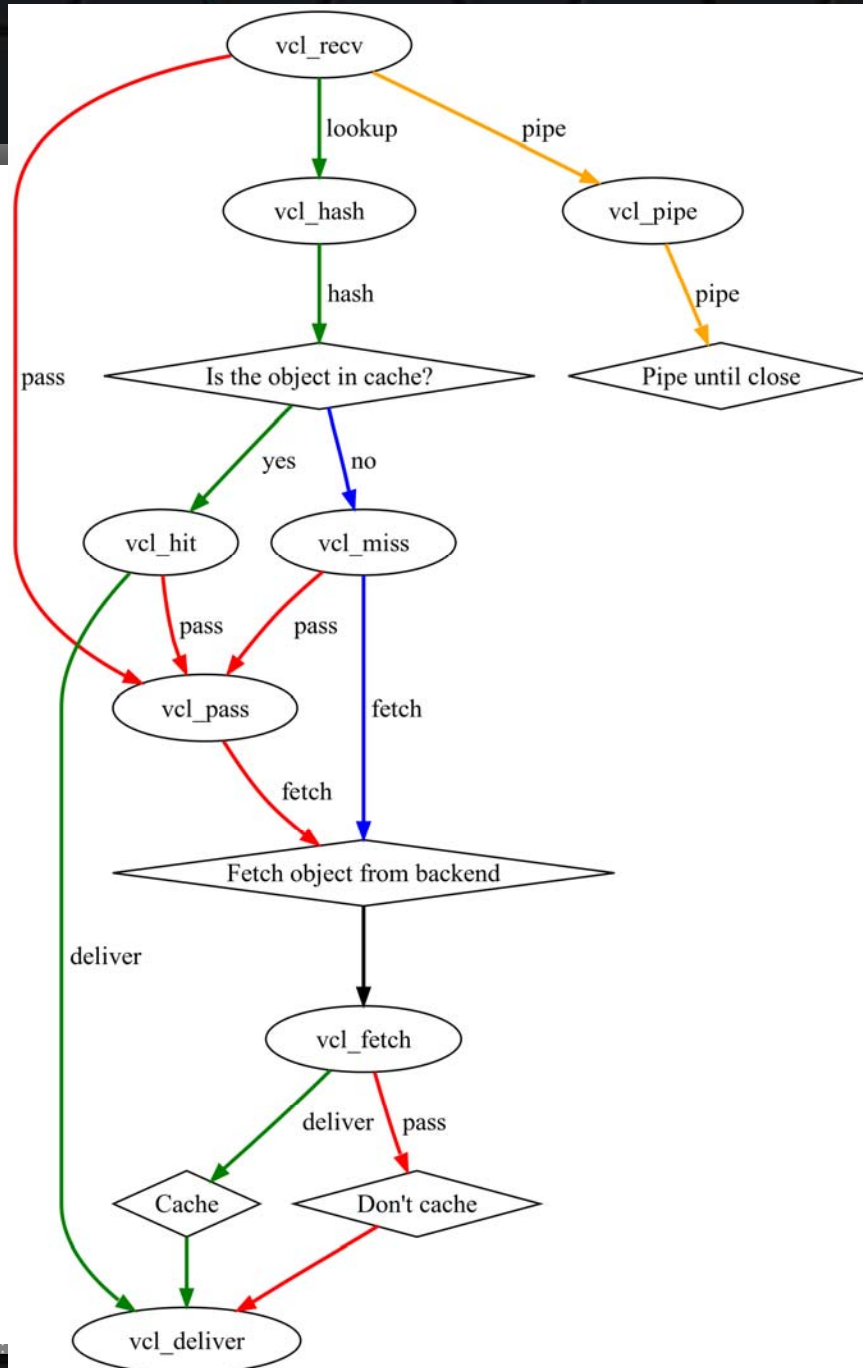


<i>stonith-action</i>	<i>reboot</i>	Action to send to STONITH device. Allowed values: <i>reboot</i> , <i>poweroff</i> .
<i>cluster-delay</i>	<i>60s</i>	Round trip delay over the network (excluding action execution). The "correct" value will depend on the speed and load of your network and cluster nodes.
<i>stop-orphan-resources</i>	<i>TRUE</i>	Should deleted resources be stopped
<i>stop-orphan-actions</i>	<i>TRUE</i>	Should deleted actions be cancelled
<i>start-failure-is-fatal</i>	<i>TRUE</i>	When set to <i>FALSE</i> , the cluster will instead use the resource's failcount and value for resource-failure-stickiness
<i>pe-error-series-max</i>	<i>-1 (all)</i>	The number of PE inputs resulting in <i>ERRORs</i> to save. Used when reporting problems.
<i>pe-warn-series-max</i>	<i>-1 (all)</i>	The number of PE inputs resulting in <i>WARNINGs</i> to save. Used when reporting problems.
<i>pe-input-series-max</i>	<i>-1 (all)</i>	The number of "normal" PE inputs to save. Used when reporting problems.

Parameter	Default value
thread_pool_add_delay	2 [milliseconds]
thread_pool_add_threshold	2 [requests]
thread_pool_fail_delay	200 [milliseconds]
thread_pool_max	500 [threads]
thread_pool_min	5 [threads]
thread_pool_purge_delay	1000 [milliseconds]
thread_pool_stack	65536 [bytes]
thread_pool_timeout	300 [seconds]
thread_pools	2 [pools]
thread_stats_rate	10 [requests]

马哥教育

www.magedu.com



```
backend default {
    .host = "127.0.0.1";
    .port = "8080";
}

sub vcl_recv
{
    if ((req.request == "GET" || req.request == "HEAD") && req.url ~ "\.(png|gif|jpg|swf|css|js)$") {
        unset req.http.cookie;

        return(lookup);
    }
}

sub vcl_fetch
{
    if (req.url ~ "\.(png|gif|jpg|swf|css|js)$") {
        remove beresp.http.set-cookie;

        set beresp.ttl = 2h;
    }
}
```

```
if (req.restarts == 0) {  
    if (req.http.x-forwarded-for) {  
        set req.http.X-Forwarded-For =  
            req.http.X-Forwarded-For + ", " + client.ip;  
    } else {  
        set req.http.X-Forwarded-For = client.ip;  
    }  
}  
if (req.request != "GET" && req.request != "HEAD" && req.request !=  
    "PUT" && req.request != "POST" && req.request != "TRACE" &&  
    req.request != "OPTIONS" && req.request != "DELETE") {  
    return (pipe);  
}  
if (req.request != "GET" && req.request != "HEAD") {  
    return (pass);  
}  
if (req.http.Authorization || req.http.Cookie) {  
    return (pass);  
}  
return (lookup);
```



```
director balance fallback
```

```
{  
    { .backend = web1; }  
    { .backend = web2; }  
}
```

```
director balance client
```

```
{  
    { .backend = web1; .weight = 2 }  
    { .backend = web2; .weight = 1 }  
}
```

```
director balance hash
```

```
{  
    { .backend = web1; .weight = 2 }  
    { .backend = web2; .weight = 1 }  
}
```



```
director balance round-robin
```

```
{  
    { .backend = web1; }  
    { .backend = web2; }  
}
```

```
director balance random
```

```
{  
    { .backend = web1; .weight = 2 }  
    { .backend = web2; .weight = 1 }  
}
```

马哥教育

www.magedu.com

```
probe health {  
    .url = "/test_health";  
    .expected_response = 200;  
    .interval = 5s;  
    .timeout = 1s;  
    .window = 5;  
    .threshold = 3;  
}
```

```
backend default  
{  
    .host = "127.0.0.1";  
    .port = "8080";  
    .probe = health;  
}
```

WV

```
acl purge_acl {  
    "localhost";  
    "192.168.55.0/24";  
}  
  
sub vcl_recv  
{  
    if (req.request == "PURGE") {  
        if (!client.ip ~ purge_acl) {  
            error 405 "Not allowed";  
        }  
  
        purge_url(req.url);  
        error 200 "Purged";  
    }  
}
```

- ❖ 博客: <http://magedu.blog.51cto.com>
- ❖ 主页: <http://www.magedu.com>
- ❖ QQ: 2813150558, 1661815153, 113228115
- ❖ QQ群: 203585050, 279599283



马哥教育

Thank You!