http协议

Hypertext Transfer Protocal是无状态的应用层协议，用于分布式的、协作的、富文本信息系统。

Hypertext Transfer Protocal是无状态的应用层请求、响应协议，其利用可扩展的语义和自描述的消息载荷来与基于网络的超文本信息系统进行灵活的交互。

HTTP在可靠的传输层或会话层连接之上，通过交换消息来通信。HTTP客户端与HTTP服务器建立连接，发送一个或更多的HTTP请求。HTTP服务程序接受连接，通过发送HTTP响应来服务服务HTTP请求。

经常出现的『user agent』指发起HTTP请求的各种各样的客户端程序之一，包括但不限于浏览器、爬虫、命令行程序、自定义的应用和移动应用等。

HTTP依赖于URI标准来指示目标资源和资源之间的关系。消息传输应用了类似Internet mail 和MIME的格式

大部分的HTTP通信由一个检索资源的GET请求构成，资源由URI给出。最简单的情况，可以通过单独的双线连接来完成用户代理和原始服务器之间HTTP通信。

request >

UA ======================================= O

< response

客户端通过请求消息给服务器发送HTTP请求，请求消息由请求行开始，包括请求方法、URI和协议版本，接下来是头部，头部由一个空行来结束，最后是消息体。

服务器通过一次或多次HTTP响应消息来回应客户端的请求。每个响应消息也包括三部分，开始行是状态行包括协议版本、成功或失败状态码，之后是用于说明状态码的文本，接着是响应头部（包括服务器信息、资源元数据和展示元数据），空行代表头部结束，之后是响应消息体。

列子

Client request:

GET /hello.txt HTTP/1.1

User-Agent: curl/7.16.3 libcurl/7.16.3 OpenSSL/0.9.7l zlib/1.2.3

Host: www.example.com

Accept-Language: en, mi

Server response:

HTTP/1.1 200 OK

Date: Mon, 27 Jul 2009 12:28:53 GMT

Server: Apache

Last-Modified: Wed, 22 Jul 2009 19:15:56 GMT

ETag: "34aa387-d-1568eb00"

Accept-Ranges: bytes

Content-Length: 51

Vary: Accept-Encoding

Content-Type: text/plain

Hello World! My payload includes a trailing CRLF.

HTTP的中间人

Proxy、gateway和tunnel

A "gateway" (a.k.a. "reverse proxy") is an intermediary that acts as

an origin server for the outbound connection but translates received

requests and forwards them inbound to another server or servers.

Gateways are often used to encapsulate legacy or untrusted

information services, to improve server performance through

"accelerator" caching, and to enable partitioning or load balancing

of HTTP services across multiple machines.

All HTTP requirements applicable to an origin server also apply to

the outbound communication of a gateway. A gateway communicates with

inbound servers using any protocol that it desires, including private

extensions to HTTP that are outside the scope of this specification.

However, an HTTP-to-HTTP gateway that wishes to interoperate with

third-party HTTP servers ought to conform to user agent requirements

on the gateway's inbound connection.

用gateway可以做负载均衡

缓存

缓存是对先前的响应消息进行本地存储系统和控制消息存储、重新检索和删除的子系统的总和。缓存系统存储那些可以被缓存的响应，将来遇到相同请求，可以减少响应时间，减少带宽的消耗