

TECHNOLOGY BRIEF

HTTP

What is HTTP?

HTTP, the Hypertext Transfer Protocol, is the application-level protocol that is used to transfer data on the Web. HTTP comprises the rules by which Web browsers and servers exchange information. Although most people think of HTTP only in the context of the World-Wide Web, it can be, and is, used for other purposes, such as distributed object management systems.

How Does HTTP Work?

HTTP Is a request-response protocol. For example, a Web browser initiates a request to a server, typically by opening a TCP/IP connection. The request itself comprises

- o a request line,
- o a set of request headers, and
- o an entity.

The server sends a response that comprises

- o a status line,
- o a set of response headers, and
- o an entity.

The entity in the request or response can be thought of simply as the payload, which may be binary data. The other items are readable ASCII characters. When the response has been completed, either the browser or the server may terminate the TCP/IP connection, or the browser can send another request.

An Example

As an illustration of HTTP, here is an example exchange between a Web browser and the Silicon Press server, www.silicon-press.com. In response to a user request to go to the URL

http://www.silicon.press.com

the browser sends the following HTTP request to www.silicon-press.com:

```
GET / HTTP/1.1
Connection: Keep-Alive
User-Agent: Mozilla/5.0 (compatible; Konqueror/2.2-11; Linux)
```



```
Accept: text/*, image/jpeg, image/png, image/*, */*
Accept-Encoding: x-gzip, gzip, identity
Accept-Charset: Any, utf-8, *
Accept-Language: en, en_US
Host: www.silicon-press.com
-- blank line -
```

A brief explanation:

- o The first line is the request line that comprises three fields:
 - 1. a method: The GET method indicates that the server is supposed to return an entity.
 - 2. a request-URI (Universal Resource Identifier). The / indicates the root of the document system on the server, and
 - 3. HTTP protocol version: 1.1 in this case.
- The second line is the optional Connection header informs the server that the browser would like to leave the connection open after the response.
- The third line is the optional User-Agent header that identifies the kind of browser that is sending the request, its version, and its operating system.
- The Accept headers specify the type, language, and encoding for the returned entity that the browser would prefer to receive from the server.

Responding to the browser, the www.silicon-press.com server sends the following response:

```
HTTP/1.1 200 OK
Date: Thu, 24 Jan 2002 17:33:52 GMT
Server: Apache/1.3.14
Last-Modified: Mon, 21 Jan 2002 22:08:33 GMT
Etag: "47bc6-25e0-3c4c9161"
Accept-Ranges: bytes
Content-Length: 9696
Connection: close
Content-Type: text/html
-- blank line--
-- HTML entity --
```

A brief explanation:

- o The first line is the status line consisting of three fields:
 - 1. HTTP protocol version of the response: 1.1 in this case,
 - 2. a three-digit numeric status code, and
 - 3. a short description of the status code.
- o The Content-Length, Content-Type, Etag, and Last-Modified header lines describe the entity returned.



HTTPS (Secure HTTP)

HTTPS denotes the use of HTTP with SSL (Secure Socket Layer) protocol or its successor protocol Transport Layer Security (TLS), a transport-layer protocol. Either of these protocols, which use encryption, can be used to create a secure connection between two machines. The browser uses SSL or TLS when connecting to a secure part of a website indicated by an HTTPS URL, that is, a URL with the prefix https://. The browser then uses HTTP to send and receive requests over this secure connection.

Where Can I Find More Information?

- o <u>HTTP/1.1 Specification</u> (http://www.rfc-editor.org/rfc/rfc2616.txt)
- o <u>TLS Protocol</u> (http://www.rfc-editor.org/rfc/rfc2246.txt)

Author: David M. Kristol, info@silicon-press.com

