Cosa e Quali sono gli Header HTTP?

l vostro browser, ogni volta che si collega ad un web server per aprire una pagina, invia specifiche richieste ed il server gli ritorna delle risposte. Questi messaggi chiamati **header http**, contengono delle informazioni, alcune necessarie ed altre puramente informative.

Queste informazioni possono essere:

**Di richiesta**, in uscita (**out**) ed inviate dal browser al web server: cookies, il software che utilizzate, il tipo di browser, referrer, lingua utilizzata, etc.

Vediamo un esempio di **HTTP request**, utilizzando la richiesta effettuata tramite un browser Internet Explorer, all’indirizzo www.quickgo.it:

*GET / HTTP/1.1*  
*Accept: image/gif, image/x-xbitmap, image/jpeg, image/pjpeg, application/x-shockwave-flash, application/msword, application/vnd.ms-excel, \*/\**  
*Accept-Language: it*  
*Accept-Encoding: gzip, deflate*  
*User-Agent: Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1)*  
*Host: www.blogbyte.it*  
*Connection: Keep-Alive*  
*Cookie: \_\_utma=151439968.604684092.1165355506.1165355506.1165357918.2; \_\_utmz=151439968.1165357918.2.2.utmccn=(organic)|utmcsr=google|utmctr=%2Beclipse+%2B%22web+editor%22|utmcmd=organic*

La prima riga della request (request line) contiene il metodo utilizzato per la richiesta (nel nostro esempio GET), la URL associata alla richiesta stessa e la versione del protocollo HTTP utilizzata (nel nostro esempio: HTTP/1.1)

Dopo la request line, vediamo una serie di request header che forniscono informazioni sul browser e sulla richiesta specifica.

Nel nostro esempio, sono contenute sette request header anche se il numero di header, in generale, può essere variabile. Ogni request header inizia sempre con il nome della request header stessa (per es.: User-Agent), seguito dal carattere “:”, da uno spazio e, infine, dal valore associato all’header.

Le **request header** **http** più usate sono le seguenti:

* **Accept**: specifica l’ordine preferito dal browser relativamente ai MIME types supportati. La stringa “\*/\*” indica che il browser è in grado di gestire qualunque tipologia di MIME type.
* **Accept-Encoding**: specifica le tipologie di compressioni accettate dal browser
* **Accept-Charset**: specifica l’insieme di caratteri che il browser può accettare.
* **Accept-Language**: serve a specificare i codici dei linguaggi standard che il browser preferisce ricevere.
* **Authorization**: identifica il livello di autorizzazione per il browser. Quando si utilizza la sicurezza gestita dal container, il servlet container si occupa di impostare questo header in modo automatico.
* **Connection**: indica il tipo di connessione che viene utilizzata dal browser. Nella versione 1.1 del protocollo HTTP tale valore è valorizzato con la stringa “Keep-Alive”.
* **Cookie**: specifica i valori di eventuali cookie inviati dal server nelle precedenti comunicazioni con il browser.
* **Host**: contiene informazioni sull’host e la porta
* **Referer**: la url della pagina web cui si fa riferimento. Si noti, per curiosità, l’errore ortografico (la parola corretta sarebbe Referrer) commesso da uno degli autori del protocollo HTTP che si decise di non correggere nelle versioni successive.
* **User-agent**: indica il tipo di browser. Sia Internet Explorer che Netscape Navigator si identificano come “Mozilla” ma il primo include anche la stringa “MSIE”

**Di risposta**, in ingresso (**in**) e provenienti dal web server: cookies, data di modifica della pagina, etc.

HTTP/1.0 200 OK  
Date: Sun, 11 Mar 2007 15:38:49 GMT  
Server: Apache  
Cache-Control: no-store, no-cache, must-revalidate  
Expires: Fri, 15 Aug 1976 18:15:00 GMT  
X-Powered-By: PHP/4.3.10-18  
Pragma: no-cache  
Last-Modified: Sun, 11 Mar 2007 15:38:49 GMT  
Cache-Control: post-check=0, pre-check=0  
Content-Encoding: gzip  
Connection: close  
Content-Type: text/html; charset=iso-8859-1

La prima riga, di una response HTTP, viene chiamata status line. In essa vengono specificate:

* La versione del protocollo HTTP che si sta utilizzando
* Un codice definito status code (nel nostro caso valorizzato con 200)
* Un messaggio associato allo status code (“OK”).

Dopo la status line, sono presenti le response header, contenenti delle informazioni sul server e sulla risposta che si sta inviando al client.

Ogni response header inizia con il nome dell’header seguito dal carattere “:”, uno spazio e il valore associato all’header.

L’HTTP response, al termine dell’ultima response header, ha una riga vuota seguita dal cosiddetto *response body* che, tipicamente, è il contenuto HTML del documento.

Le **response header** più comunemente utilizzate sono le seguenti:

* **cache-control**: controlla la modalità di caching per una pagina web
* **content-disposition**: può essere utilizzato per specificare un file binario in allegato alla response
* **content-length**: specifica la lunghezza del body della response, espresso in byte
* **content-type**: specifica il MIME type del documento di risposta
* **content-encoding**: specifica il tipo di codifica utilizzato nella risposta. Spesso, utilizzare una codifica di tipo GZIP incrementa notevolmente le performance
* **expires**: specifica la durata della cache
* **last-modified**: indica, in termini temporali, l’ultimo aggiornamento effettuato su una pagina
* **pragma**: disabilita la cache sui vecchi browser attraverso l’utilizzo della stringa “no-cache”
* **refresh**: specifica il numero di secondi trascorsi i quali il browser deve richiedere un aggiornamento della pagina