Divya Natesan & Shrikant Adhikarla

24 July 2012

Rules 3-9

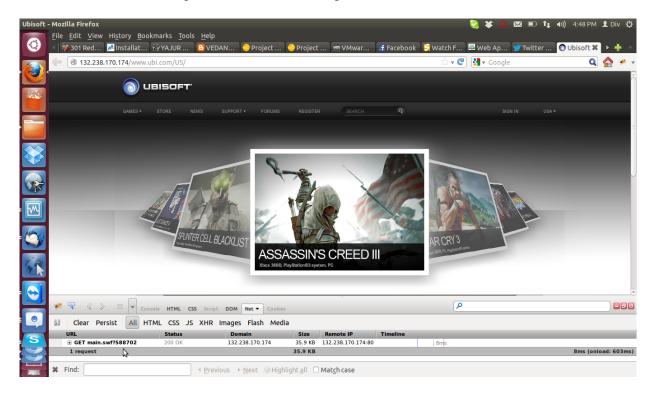
3. Add An Expires Header

We enabled the expires module in the apache configuration. A 'config' file was included for this module where caching was enabled for the resources of the website. The resources of this website were grouped into three categories – those that are very dynamic in nature, those that change but not often and those that hardly change. The html page was segregated under the first category. The stylesheets, javascripts, images of the format gif, jpg and png were classified under the second type while the ico image was put under the static category. The xml resources are also put under bucket2 since the rss-xml feeds that are collated by the webpage have different file names each time they come in. A revision for these file names would be maintained as the changes come in (during future iterations). The values assigned to each of these buckets are 1 week, 1 month and 3 months respectively. In addition, any resource of other types are also assigned to a default value equivalent to 1 month (like aspx pages). The ico image though might not change for a very long time (say, 6 months) – should be noted that cannot be changed in name. Hence, a nominal value was assigned for the cache expiry of this image. All these resources are set for expiry since access. Though 'M' is typically used for pages that serve like weekly notices (like the html page that changes often), it adds a constraint that the resource downloaded should exist in the disk of the client. Setting the expiry date with respect to the access date gives this added advantage. In any case, each time a resource is modified on the server, the filename would be revved anyways to take care of cache inconsistencies. Hence, the access date was chosen as a standard to define the expiry duration. We also noted that the cache-control directive was specified in each html request to the web server because of which when the page is loaded

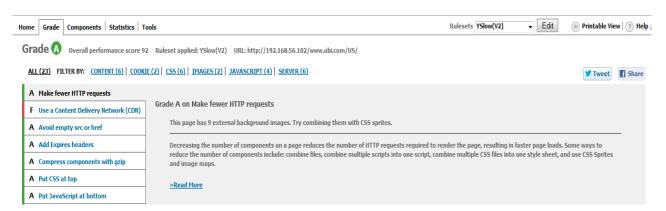
Divya Natesan & Shrikant Adhikarla

24 July 2012

again (not, refresh since If-Not-Modified directive would be invoked in this case) – the resources stored in the cache are retrieved rather than launching http requests to the server. The cache mode is set to public since this webpage does not require any authentication and the same page can be served to different clients. Given below is a screenshot of the firebug screen after the expires module was configured and activated.



Please see screenshot below for YSlow score.



4. Gzip Components

The deflate module was enabled and configured. We also made sure that all resources that have already been compressed were excluded from being compressed again on request for encoded webpages from the client. We also included a vary header that enables the proxies to validate this header directive and accordingly serve the pages to the clients – so that browsers that do not have the capability to decode this version are not served with such format. While analysing the different approaches in which the deflate rules must be configured, we were faced with the decision to choose between white-listing and blacklisting. We chose to white-list – meaning, specify the formats that should be gzipped. Such a decision was encouraged by the urge to ensure that no unknown resource is gzipped because sometimes compressed files if further gzipped can result in an increase in the size of the object. There are problems that are faced with Netscape version 4.0 to compress all formats except html documents. In a similar fashion, there are problems with Netscape versions 4.06-4.08 for support to gzip at all. Internet Explorer masquerades as Netscape but gzip is supported in all latest versions of these browsers. Internet Explorer 6 would not be supported by our webpage since we do not have those explicit checks (that rule out gzip) in place. This is because, it is sometimes noted that internet explorer 7 returns a user agent string with version 6.0's details and we do not want to retract gzip support to the latest version by this feature. A dont-vary flag is added specifically to all the image format files (present in the resource) so that the header does not contain the vary flag information (this decreases the size of the header) and this does not necessitate the proxy server to check for non-gzipped versions explicitly – the need for such a scenario does not even exist. Testing

Divya Natesan & Shrikant Adhikarla

24 July 2012

was performed through external websites apart from a command line uitlity for which the screenshots are given below. The command lwp-request was used to send a http request (can be customized) and receive a response as well.

Before:

```
div@ubuntu:~$ lwp-request -e http://132.238.170.53/www.ubi.com/US/
Cache-Control: max-age=2592000
Connection: close
Date: Wed, 25 Jul 2012 05:12:46 GMT
Accept-Ranges: bytes
ETag: "729-313c-4c5777da06bce"
Server: Apache/2.2.22 (Debian)
Vary: Accept-Encoding, User-Agent
Vary: Accept-encoding, oser-agent
Content-Length: 12604
Content-Type: text/html
Expires: Fri, 24 Aug 2012 05:12:46 GMT
Last-Modified: Mon, 23 Jul 2012 04:15:06 GMT
Client-Date: Wed, 25 Jul 2012 05:12:47 GMT
Client-Peer: 132.238.170.53:80
Client-Response-Num: 1
Link: <.../resources/min.css>; rel="stylesheet"; type="text/css"
Title: Ubisoft
X-Meta-Content-Language: en-US
X-Meta-Description: Discover Ubisoft's best selling games like Assassin's Creed, Just Dance and Ghost Recon. Watch trailers, play free games and
 X-Meta-Google-Site-Verification: hHAnwrfrv3xVFkRVwErSIim05Wh3vyTp4rT3_LyDdQA
<!doctype html>
<html dir="ltr" lang="en-US">
 <head>
   <title>Ubisoft</title>

         content="Discover Ubisoft's best selling games like Assassin's Creed, Just Dance and Ghost Recon. Watch trailers, play free games and jo
    <meta name="google-site-verification" content="hHAnwrfrv3xVFkRVwErSIim05Wh3vyTp4rT3_LyDdQA">
 </head>
 <body>
   <div id="header-top-bar">
      <div class="content-area">
  <div id="logo" onClick="document.location.href = '/www.ubi.com/US'" onMouseOver=""><a href="#"></a></div>
```

After:

Divya Natesan & Shrikant Adhikarla

24 July 2012

Please find in line item 3 the relevant screenshot for YSlow.

For rules 3 and 4, we considered different possible MIME types (the current MIME types ere referred from the IANA portal) that would make the expires and deflate features compatible with all browsers and their previous versions. In fact, different resources can be created with several MIME types that are not current standards. All such problems have been solved by the above mentioned approach.

5. Put Stylesheets At The Top (This was done in the previous iteration)

All stylesheets were combined into one .css and placed at the top of the html so that the pages appear to load faster to the user (especially, for end-users with slow connection or pages with a lot of content). Such a measure also helps the webpage to self-serve as a progress indicator. Some browsers like Internet Explorer do not render elements until the stylesheets are parsed because it might have to redraw the elements if otherwise. There is also a chance that the elements are unstyled while presented to the user during progressive load, which might not be appreciated by the user. Due to the two above stated reasons, we decided to put the stylesheets at the top instead of towards the bottom end of the page. The relevant screenshot from YSlow can be referred from line item 3.

6. Put Scripts At The Bottom (This was done in the previous iteration)

Scripts block parallel downloads (even those served from different host names). This mechanism allows the user to view the elements on the page progressively without pausing for a long time in order that the javascript loads. It was also taken care that these javascript functions were not referenced in index.html before the script is included to avoid scoping issues. The relevant screenshot from YSlow can be referred from line item 3.

7. Avoid CSS Expressions

Divya Natesan & Shrikant Adhikarla

24 July 2012

An expression is used to define the minimum width of the header div. Latest versions of browsers support the minwidth tag for elements in stylesheets – hence, the expression can be removed without consequences (there is no need to add an event handler). Expressions are removed since they are evaluated more frequently than when the element is initially drawn – in fact, for every event that occurs for the element. Please find below the screenshot from YSlow.



8. Make JS And CSS External

One inline script was found to exist within index.html that assigns name – value pair for a footer menu bar items. Since, this is a common element that might be repeated in other webpages, it should be cached. It is hence, collated with the existing javascript file. Please note that the number of http requests made remains unchanged.

9. Reduce DNS Lookups

The apache configuration file was studied and a parameter HostnameLookups was determined to be an attribute that is set at the server. This option allows the server to log the names of the clients, if turned on. We ensured that this feature was turned off to prevent the

Divya Natesan & Shrikant Adhikarla

24 July 2012

lookup request that is resulted each time a client requests a page from the webserver. Please find below screenshot from YSlow.



In addition to the above rules, we also categorized the image components in sprites based on their height and width dimensions. Accordingly, the sprite image previously created was split up into multiple sprite images to ensure that the image size is not huge – we also decreased the padding between image components in the sprite image. The resoultion of the promo images that were not included as part of the sprites (as reasoned in the previous document) was reduced by hundred times rather than scaling these images in the html – which is a costly process.