解读前端性能优化之“配置ETag”

原创 2016年04月10日 18:39:28

**1、什么是ETag？**

实体标签(EntityTag)是唯一标识了一个组件的一个特定版本的字符串，是[web服务器](http://so.chinaz.com/cse/search?s=13215756937974215344&entry=1&q=web%E6%9C%8D%E5%8A%A1%E5%99%A8)用于确认缓存组件的有效性的一种机制，通常可以使用组件的某些属性来构造它。（总结：实体标签是web服务器和浏览器用于确认缓存组件的有效性的一种机制。）

**条件GET请求**

[浏览器](http://so.chinaz.com/cse/search?s=13215756937974215344&entry=1&q=%E6%B5%8F%E8%A7%88%E5%99%A8)下载组件的时候，会将它们存储到浏览器缓存中。如果需要再次获取相同的组件，浏览器将检查组件的缓存时间，假如已经过期，那么浏览器将发送一个条件GET请求到服务器，服务器判断缓存还有效，则发送一个304响应，告诉浏览器可以重用缓存组件。

那么服务器是根据什么判断缓存是否还有效呢?答案有两种方式，一种是前面提到的ETag，另一种是根据最新修改时间。先来看看最新修改时间。

**（1）最新修改时间**

原始服务器通过Last-Modified响应头来返回组件的最新修改时间。

以一个实际例子来说明，当我们不带缓存访问www.google.com.hk的时候，我们需要下载google的logo，这时会发送这样一个HTTP请求：

**Request：**

GET /logo.png HTTP 1.1

Host: www.google.com.hk

**Response:**

HTTP 1.1 200 OK

Last-Modified:Wed, 09 Oct 2013 01:35:39 GMT

当需要再次访问相同组件的时候，同时缓存已经过期，浏览器会发送如下条件GET请求：

**Request：**

GET /logo.png HTTP 1.1

If-Modified-Since:Wed, 09 Oct 2013 01:35:39 GMT

Host: www.google.com.hk

**Response:**

HTTP 1.1 304 Not Modified

**实体标签**

ETag提供了另外一种方式，用于检测浏览器缓存中的组件与原始服务器上的组件是否匹配。摘抄自书上的例子：

不带缓存的请求：

**Request：**

GET /i/[yahoo](http://so.chinaz.com/cse/search?s=13215756937974215344&entry=1&q=yahoo)/gif HTTP 1.1

Host: us.yimg.com

**Response:**

HTTP 1.1 200 OK

Last-Modified:Tue,12 Dec 200603:03:59 GMT

ETag:”10c24bc-4ab-457elc1f“

再次请求相同组件:

GET /i/yahoo/gif

**Request：**

HTTP 1.1

Host: us.yimg.com

If-Modified-Since:Tue,12 Dec 200603:03:59 GMT

If-None-Match:”10c24bc-4ab-457elc1f“

**Response:**

HTTP 1.1 304 Not Midified

当ETag和Modified-Time都出现了，则原始服务器禁止返回304除非请求中的条件头字段全部一致。

**2、为什么要引入ETag?**

**ETag主要是为了解决Last-Modified无法解决的一些问题：**

1. 一些文件也许会周期性的更改，但是他的内容并不改变(仅仅改变的修改时间)，这个时候我们并不希望客户端认为这个文件被修改了，而重新GET;

2. 某些文件修改非常频繁，比如在秒以下的时间内进行修改，(比方说1s内修改了N次)，If-Modified-Since能检查到的粒度是s级的，这种修改无法判断(或者说UNIX记录MTIME只能精确到秒);

3. 某些服务器不能精确的得到文件的最后修改时间。

**3、ETag带来的问题**

ETag的问题在于通常使用某些属性来构造它，有些属性对于特定的部署了网站的服务器来说是唯一的。当使用集群服务器的时候，浏览器从一台服务器上获取了原始组件，之后又向另外一台不同的服务器发起条件GET请求，ETag就会出现不匹配的状况。

**最佳实践**

1. 如果使用Last-Modified不会出现任何问题，可以直接移除ETag，google的搜索首页则没有使用ETag。

2. 确定要使用ETag，在配置ETag的值的时候，移除可能影响到组件[集群服务器](http://so.chinaz.com/cse/search?s=13215756937974215344&entry=1&q=%E9%9B%86%E7%BE%A4%E6%9C%8D%E5%8A%A1%E5%99%A8)验证的属性，例如只包含组件大小和[时间戳](http://so.chinaz.com/cse/search?s=13215756937974215344&entry=1&q=%E6%97%B6%E9%97%B4%E6%88%B3)。