

# 我国科学数据共享平台网站首屏时间分析

□ 姜吉栋 彭洁 赵辉 / 中国科学技术信息研究所 北京 100038  
刘润达 / 北京市科学技术委员会 北京 100195  
/ 中国科学技术信息研究所 北京 100038  
马海银 / 山东大学软件学院 济南 250100

**摘要:** 首屏时间会影响用户对网站的体验和使用。文章对我国84个科学数据共享网站的首屏时间进行了测试, 揭示了此类网站首屏时间的总体现状, 并使用Page Speed找出影响首屏时间的主要因素, 提出应在使用浏览器缓存、将图片组合为CSS贴图定位、优化图片、启用文本压缩四个方面改进网站性能, 缩短网站首屏时间。

**关键词:** 科学数据网站, 首屏时间, 网站性能优化

DOI: 10.3772/j.issn.1673—2286.2013.10.010

## 1 研究背景

科学数据是一种宝贵的科技资源, 是国家科技创新和可持续发展的重要战略资源。共享是充分发挥科学数据的科学价值、社会价值和经济价值的重要手段和方式之一<sup>[1]</sup>。近年来, 一些国际组织和发达国家在信息技术发展与广泛应用的基础上, 不断对加强科学数据的共享管理。随着互联网的日益成熟和发展, 网络用户信息素养的提高, 网络已成为科学数据共享和用户寻求所需科学数据的重要途径。在这种情况下, 我国也在各个层级开展了科学数据共享工作, 其主要内容之一是构建科学数据共享平台网站。该类网站是数据存储或发布的平台, 作为公益性网站, 服务性是其最主要的目的, 用户可以浏览、检索、下载相关科学数据, 这使得科学数据网站的性能显得十分重要。在拥有丰富资源的基础上, 如何让用户方便快捷地看到以及得到这些资源, 很大程度上依靠网站的性能建设。

网站首页作为一个网站的入口网页, 是抓住用户的关键因素, 首页性能的好坏直接影响到用户的使用, 其中首屏载入时间是用户体验的重要指标。首屏的定义以800×600像素尺寸为标准。从开始监测开始时计时, 到浏览器页面显示高度达到600像素, 且此区域有内容显示之后的时间就是首屏时间。相关研究发现, 对用户来说, 当页面充满800×600像素的区域时就可以看到内容并可以点击, 而非等待页面全部下载完成。如果首

屏时间太长, 用户就会离开转向其他网站。作为项目形式构建的科学数据共享网站一开始也往往不太注意首屏时间的优化, 从而造成了用户体验的下降。解决这一问题需要找出影响首屏时间的主要因素, 通过在网站设计技术上优化网站性能, 缩短首屏时间。

## 2 研究方法

### 2.1 研究对象的确定

目前我国的科学数据共享平台项目主要来源于三个层面: 一是由科技部主管的科学数据共享工程, 二是科学数据库项目, 三是地方政府建设的省、市级科学数据库。以各大科学数据共享计划网站的首页为出发点, 可以整理出主要的科学数据共享网站。在以上获取的网站基础上, 通过网络调研的方法, 在Google搜索引擎中, 检索关键词“科学数据”和“科学数据共享”获取其他我国主要的科学数据共享平台。对两方面获取的数据进行合并和去重, 共获取98个主要的科学数据共享网站。

98个科学数据网站中, 有7个由于各种原因无法访问, 还有7个是明显的综合导航性的网站, 网站本身并没有科学数据及其他服务, 本研究不对其进行评价。最终得到84个待评价网站, 本研究以这84个科学数据共享网站作为研究对象。

## 2.2 工具选择与测试过程

本文首先对研究对象的整体首屏时间现状作了分析,然后利用插件对每一个研究对象的首屏时间进行测试,找出影响首屏时间的关键因素,提出改善建议,并应用其他软件模拟改善后的效果,验证建议的有效性。

本次测试时间范围是2012年08月13日00:00 - 2012年11月12日00:00。各网站的首屏时间数据来自专业应用性能管理及用户体验评价公司的监测。此专业公司的监测原理是通过在地理上广泛分布的客户端(全国330多个城市的多个国家及地方的运营商网络节点)上安装监测软件,监测和收集相关网站的各性能指标。在首屏时间影响因素分析和缩短首屏时间分析方面,本文采用的工具是Page Speed插件。Page Speed最初是Google内部使用的改进网页设计的工具,它整合在Firefox的著名插件Firebug中。它能够用于评估、诊断和优化网页性能,为开发人员带来很大的辅助作用<sup>[2]</sup>。当用户运行Page Speed,可以通过一系列操作获得如何改进网页载入速度的建议。插件使用方法简单,生成的图形一目了然,因此很受欢迎。

以海洋科学数据共享中心网站为例说明测试过程,经专业公司监测,该网站的首屏时间是2.63秒,打开该网站,用Page Speed插件查看导致页面载入速度情况的各项指标,会按照高中低的优先级给出缩短载

入时间的建议,其中红色指标是高优先级,表示如果完成这些开发工作,可极大地提升载入速度。黄色表示中优先级,表示这些建议在性能方面的提升空间较小。蓝色表示低优先级,这些建议表示即使完成此类开发工作,在性能方面的提升空间也很小。绿色表示已完成,说明在这些方面该网页已采用了最佳做法。由此可以看出,高优先的建议正是网站没有做好的方面,而这些方面则是影响首屏时间的主要因素。Page Speed的测试页面如图1所示。

## 3 测试结果

### 3.1 首屏时间整体测试

专业公司的监测数据显示,84个网站中,首屏时间低于5s的有37个,处于5s和10s之间的有30个,超过10s的则有17个。

通常一个网站,如果首屏时间在5秒以内是比较优秀的,10秒以内是可以接受的,10秒以上就不可容忍了。用户会选择刷新页面或立刻离开。随着网络技术的发展 and 用户体验要求的提高,这几个数字只会变得更小。如腾讯网、凤凰网、网易网等主流门户网站的首屏时间早已经进入3s时代<sup>[3]</sup>。从数据上看,科学数据网站的整体首屏时间还不够理想,只有约44%的网站在5s以内,其他一多半的网站超过了5s,这必然会流失用户,需要作出调整和改进。首屏时间较短的网站如表1所示。

可以看出,排名在前十位的网站首屏时间都进入了两秒以内,由于科学数据网站的首页内容和设计程度要



图1 海洋科学数据共享中心网站  
的Page Speed测试分析

表1 首屏时间排名前十网站

排名	名称	首屏时间 (单位: 秒)
1	国家地震科学数据中心	0.448
2	渔业科学数据分中心	1.005
3	国际科学数据服务平台	1.243
4	海洋科学数据共享平台	1.412
5	渔业科学数据库	1.58
6	中国天文数据中心	1.595
7	中国科学院化合物参考数据库	1.642
8	海洋科学数据共享中心	1.752
9	喀斯特科学数据中心	1.757
10	作物科学数据分中心	1.771

比门户网站简单得多,因此并不能因为这些网站的首屏时间比门户网站短,就简单地认为它们在网站性能优化上已经接近完美。虽然这些网站的首屏时间在同类网站中做得比较好,但是它们同样存在很多问题,需要进一步测试,找出导致首屏时间不理想的主要因素。

### 3.2 首屏时间详细测试

本文使用Page Speed插件,对每个研究对象进行测试,以找出影响网站首屏时间的主要因素。Page Speed按照高、中、低优先级以及完成项目,可以给出共27项建议。其中,高优先级可以视作影响首屏时间的主要因素。测试结果如表2所示。

测试结果表明,Page Speed给出的高优先级改进建议主要有8项,分别是使用浏览器缓存,将图片组合为CSS贴图定位,启用压缩,优化图片,提供压缩后的图片,内嵌小型JavaScript,启用Keep-Alive,压缩JavaScript。而得到前六项建议的网站比例较高。这说明大部分网站并没有采取这六项措施,也意味着实施这六项措施可以大幅缩短网站的首屏时间。其中Page Speed对61个网站提出了使用浏览器缓存的高优先级建议,说明这是绝大部分网站需要重视的问题。

## 4 对策与优化结果

分析6项高优先级建议,使用浏览器缓存可以有效解决网络带宽慢所造成的诸如网络速度慢、网络带宽低等不利影响,而其它五项如“启用压缩”等,其本质上都是通过减少网页大小来缩小下载时间。因此,网络带宽和网页大小对减少网站首屏显示时间至关重要。对于网站建设方来说,应当从这两方面对网站性能进行优化。

### 4.1 避开网络带宽影响

科学数据共享平台网站大多分布在教育部负责的中国教育和科研网(Cernet)和中科院负责的中国科学技术网(Cstnet),而中国用户基本都分布在电信集团负责的中国公用Internet骨干网(Chinanet)。这三个网络之间互联速率比在其内部流通速率低很多,所以普通用户在访问科学数据共享平台网站时会有较大延迟。解决这一问题的最好方法就是将页面内容暂存在

表2 Page Speed测试结果

高优先级改进建议	需优化的网站数目	占总网站数目比例
使用浏览器缓存	61	72.6%
将图片组合为CSS贴图定位	40	47.6%
启用压缩	36	42.9%
优化图片	30	35.7%
提供压缩后的图片	18	21.4%
内嵌小型JavaScript	11	13.1%
启用Keep-Alive	4	4.8%
压缩JavaScript	2	2.4%

用户端,避开几大网络之间较低的接口速率,实现用户端的瞬间“访问”。主要措施便是使用浏览器缓存。

浏览器缓存是最靠近用户的缓存,如果启用该缓存,用户在访问同一个页面时,将不再从服务器下载页面,而是从本机的缓存目录中读取页面,然后在浏览器中展现这个页面。

网站上的静态内容,诸如图片、CSS、JS、flash等,变动一般很小,如一些固定版块中的图片,甚至几个月都不会有所更改,这些体积庞大的静态内容占据了大部分网页加载时间。因此,降低静态内容加载时间,减少二次加载静态内容,可以有效地缩短页面载入时间。通过在服务器端添加Expires和Cache-Control HTTP头信息都可以控制缓存的保质期。

#### (1) 方法一: 添加Expires头信息

向apache服务中的.htaccess文件添加以下代码:

```
#Expire Header
ExpiresActive on //开启缓存
ExpiresByType text/CSS "access 1 month"//CSS文件缓存1个月
ExpiresByType text/plain "access 2 days"//纯文本内容缓存2天
ExpiresByType text/html "access 2 days"//html文件缓存2天
ExpiresByType application/javascript "access 1 month"//JS文件缓存1个月
ExpiresByType image/jpeg "access 1 month"//jpeg图片缓存1个月
ExpiresByType image/x-icon "access 1 month"//icon缓存1个月
ExpiresByType image/gif "access 1 month"//gif图
```

片缓存一个月

ExpiresByType image/png "access 1 month"//png  
图片缓存一个月

ExpiresByType image/ico "access 1 month"//ico缓  
存1个月

ExpiresByType application/pdf "access 1 month"//  
pdf文件缓存一个月

ExpiresByType application/x-shockwave-flash  
"access 1 month"//flash缓存一个月

ExpiresDefault "access 1 month"// 默认(未提及  
的可缓存文件) 缓存1个月

(2) 方法二: 添加Cache-Control头信息

向apache服务中的.htaccess文件添加以下代码:

```
<FilesMatch".(ico|pdf|flv|jpg|jpeg|png|gif|js|CSS|  
swf)$">
```

```
Header set Cache-Control "max-age=5184000,  
public"
```

```
</FilesMatch>
```

然后, 就可以在页面加载后看到Http文件头中的  
cache-control:public, max-age=5184000

通过Yahoo!发布的一款基于FireFox的插件YSlow  
来测试, 可以发现在启用浏览器缓存后, 浏览器对远程  
服务器的访问量减少90%。

## 4.2 减少网页大小

科学数据共享平台网站包含大量数据信息, 为了引  
导用户找到相应的资源, 会添加很多导航图片或者检  
索数据等动态网页元素。通过测试发现, 科学数据共  
享平台网站主要的组成元素有html文件、js文件、CSS  
文件、图片文件和视频文件, 可以通过压缩、合并、清除  
冗余的方法减少占用空间。根据Page Speed给出的建  
议, 具体措施包括:

(1) 将图片组合为CSS贴图定位

网站的小图片耗费了很多http请求, 即使在改进的  
http1.1规范中, 连续的http请求也都要经过请求和响应  
的过程, 对于这种情况的网站, 我们可以将大量同系列  
的小图片组合为CSS贴图定位(即CSS Sprite技术), 目  
的就是通过整合图片, 减少对服务器的请求数量, 从而  
加快页面加载速度。CSS Sprite技术将许多图片组合到  
一个网格上, 然后通过程序将每个网格的内容定位到  
画面上。

例如在访问国家人口与健康科学数据共享平台网  
站(<http://www.ncmi.cn/yysjgx/index.jsp>) 首页的过程  
中会加载大量引导用户找寻资料的小图片, 图片总大  
小为320k, 共35个小图, 每个图片在加载的过程中都对  
服务器发送一次http请求, 所以, 图片越多请求次数越  
多, 造成延迟的可能性也就越大。35个小图需要发送请  
求35次, 如果将它们整合成一张图片, 只需要发送一次  
请求, 图片总大小也缩减为80k。加载完成后, 再利用  
CSS的“background-image”, “background-repeat”,  
“background-position”三个参数的组合进行背景定  
位, background-position可以用数字精确地定位出背景  
图片的位置, 从而使整合后的图片分部到页面相应位  
置上。

(2) 优化图片

主要包括优化图片编码和优化图片尺寸。

① 优化图片压缩编码

压缩能够大大减少各种文件的尺寸, 而从HTTP/1.1  
开始, Web服务端和客户端都开始支持HTTP请求中的  
压缩格式<sup>[4]</sup>。对于图片文件, 大都采用jpeg、png和gif格  
式, 虽然其本身就是压缩格式, 但是可以通过专业工具  
进一步压缩图片。

如中国科学院化合物参考数据库网站([www.chemcpd.csdb.cn/cmref/default.asp](http://www.chemcpd.csdb.cn/cmref/default.asp)), 该网站图片  
占到网站内容的90%以上, 下载时间也占到总下载时  
间的70%以上, 仅JPEG格式的图片下载时间就达到了  
68.38%, 如果对图片进行压缩, 可以有效缩小网页体  
积, 加快下载速度。

对于JPEG文件, 可以使用jpegtran或jpegoptim两  
个软件进行压缩。对于PNG文件, 可以使用OPTIPNG  
或PNGOUT软件进行压缩。通过剥去图片的描述数据  
(metadata), 优化图片文件的霍夫曼编码层, 图片占  
用存储的空间缩小范围在60%-90%之间, 而压缩后的  
图片质量损失用肉眼几乎看不出来。

② 优化图片尺寸

如果浏览器加载的实际图片尺寸大于页面中显示  
的图片尺寸, 也会导致过多不必要信息的加载, 从而浪  
费带宽, 影响网页下载速度。

如中国科学院化合物参考物参考数据库首  
页的某张图片在CSS中固定显示为height:32px,  
width:45px, 但是在服务器端该图片的实际大小却是  
920px\*1294px, 两者的不一致不仅占用下载时间而且  
还使图片显示变形。如果在图片处理工具中将该图像进



行裁剪处理并将大小调整为45px\*32px时,图像占用存储容量由之前的3.41M减为4.22k,效果显著。

此外,还可以通过去除不必要的空间、减少颜色数、去除注释<sup>[5]</sup>等方法,在不损失用户视觉体验的同时显著加快首页图片比例较高网站的载入速度。

### (3) 启用文本压缩

对于html、js、CSS文件,可以使用gzip软件进行压缩,进而减少通过网络发送的字节数。

例如大气科学数据管理与共享平台(<http://www.atmosphere.csdb.cn/>),利用测试网站<http://loadimpact.com>可以得到,该网站的html、js、CSS文件分别占到了网站资源总量的5.26%、63.16%和10.53%,在加载时间中,以上文件也分别占到总加载时间的25.59%、12.85%和11.73%。可以通过gzip软件压缩html文件。通过压缩,主页文件大小减少84.1%,js文件大小减少82.0%,CSS文件大小减少89.1%。

此外,调整脚本和样式顺序也能缩短网页下载时

间。所有浏览器在下载js的时候,会阻止一切其他活动,直到js下载、解析和执行完毕后才开始继续并行下载其他资源<sup>[6]</sup>。尽可能将脚本文件放在页面底部,可以使页面更快地显示更多内容。

## 5 总结

科学数据共享网站的主要目的是为用户服务,用户往往会通过首页浏览或查询所需的科学数据,或者直接从网站下载数据。因此,网站的性能建设不容忽视。首屏时间作为用户体验的重要指标,更应当引起网站建设方的重视。通过本次测试,发现很多科学数据共享网站的首页载入速度还有很大的提升空间。网站方可以使用相关工具,自行分析影响网站首屏时间的因素,得到相关改进意见,应当定期对首页性能进行监测,及时发现问题并加以改进。

## 参考文献

- [1] 孙九林.科学数据资源与共享[J].中国基础科学,2003(1):30-33.
- [2] 聂应高.基于Page Speed的网站前端性能优化分析[J].现代图书情报技术,2009,185(11):69-73.
- [3] 门户网站用户体验进入3s时代[EB/OL].(2010-8) [2013-03-02]. [http://www.networkbench.com/news/company\\_news100824.html](http://www.networkbench.com/news/company_news100824.html).
- [4] 包韶平.面向前端的网站性能优化方法与工具研究[J].科技风,2009,123(9):127.
- [5] 胡朝明.数字图书馆门户网站测试与分析:关注Web标准与页面性能[J].数字图书馆,2012,28(2):23-27.
- [6] 聂应高.高校图书馆网站前端性能调查与优化对策:以我国20所“211工程”院校为例[J].2011,55(1):114-117.

## 作者简介

姜吉栋 (1988-), 男, 硕士, 研究方向: 科学数据共享, 科技资源评价等。E-mail: [sdzfjdd@126.com](mailto:sdzfjdd@126.com)  
彭洁 (1965-), 女, 硕士, 研究员, 研究方向: 信息资源管理, 科技资源管理等。  
赵辉 (1971-), 女, 硕士, 副研究馆员, 研究方向: 信息资源管理、科技项目管理等。  
刘润达 (1981-), 男, 博士, 副研究员, 研究方向: 科学数据共享、科技资源管理等。  
马海银 (1988-), 男, 硕士, 研究方向: 分布式计算。

## The Analysis of the First Screen Time for Chinese Scientific Data Sharing Websites

Jiang Jidong, Peng Jie, Zhao Hui / Institute of Scientific and Technical Information of China, Beijing, 100038  
Liu Runda / Beijing Municipal Commission of Science and Technology, Beijing, 100195  
/ Institute of Scientific and Technical Information of China, Beijing, 100038  
Ma Haiyin / School of Software, Shandong University, Jinan, 250100

Abstract: The first screen time will affect the user experience and the use of the site. In this paper, 84 Chinese scientific data sharing websites' first screen time were tested. The authors show the overall status of the first screen time about this kind of sites and use Page Speed to identify the main factors that affect the time of the first screen. Finally, the authors suggest that the websites designer should take specific measures to improve site performance and shorten website first screen time in four areas: using the browser cache, picture combination as CSS sprites, optimizing images and using text compression.

Keywords: Scientific data website, First screen time, Web performance optimization

(收稿日期: 2013-07-07)