

红心地瓜的专栏

专注于WebKit内核和浏览器相关技术(t.sina.com.cn/chenyuzhi)

目录视图

摘要视图

[RSS](#)

订阅

个人资料



红心地瓜

访问：243968次
积分：3011分
排名：第2422名

原创：69篇 转载：12篇
译文：10篇 评论：137条

文章搜索

文章分类

- [android \(12\)](#)
- [C/C++语言 \(1\)](#)
- [html5 \(2\)](#)
- [WebKit内核源代码分析 \(17\)](#)
- [五味瓶 \(1\)](#)
- [协议研读 \(15\)](#)
- [杂七杂八 \(3\)](#)
- [浏览器技术 \(6\)](#)
- [问答 \(5\)](#)
- [linux \(1\)](#)

文章存档

- [2013年09月 \(1\)](#)
- [2013年04月 \(1\)](#)
- [2012年04月 \(1\)](#)
- [2012年03月 \(1\)](#)
- [2011年12月 \(1\)](#)

[展开](#)

阅读排行

- [\[转\]Android中的Intent详解 \(25391\)](#)
- [WebKit内核源码分析 \(一\) \(18479\)](#)
- [WebKit 内核源代码分析 \(16436\)](#)

2014年1月微软MVP申请开始啦！

2013年10月微软MVP当选名单揭晓！

Tag功能介绍—我们为什么打Tag

说说家乡的互联网主题有奖征文

WebKit 内核源代码分析（四）

分类：[WebKit内核源代码分析](#) [浏览器技术](#)

2011-04-26 10:08 16436人阅读 [评论\(28\)](#) [收藏](#) [举报](#)

[代码分析](#) [webkit](#) [network](#) [cache](#) [hashmap](#) [callback](#)

[目录\(?\)](#) [\[+\]](#)

WebKit 内核源代码分析（四）

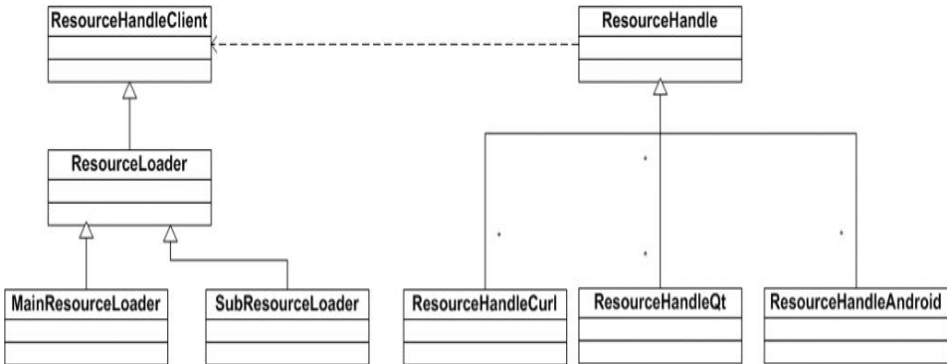
红心地瓜 (tomorrow.cyz@gmail.com)

摘要：本文介绍 WebCore 中 Loader 模块是如何加载资源的，分主资源和派生资源分析 loader 模块的类关系。

关键词：WebKit, Loader, Network, ResourceLoader, SubresourceLoader

一、类结构及接口

Loader 模块是 Network 模块的客户。Network 模块提供指定资源的获取和上传功能，获取的资源可能来自网络、本地文件或者缓存。对不同 HTTP 实现的适配会在 Network 层完成，所以 Loader 接触到的基本上是同 OS 和 HTTP 实现无关的 Network 层接口。



如上 Loader 和 NetWork 之间的类关系图，ResourceHandleClient 是 ResourceHandle 的客户，它定义一系列虚函数，这些虚函数是 ResourceHandle 的回调，继承类实现这些接口。

ResourceHandleClient 的接口同网络传输过程息息相关，一般为某一个网络事件对应的回调。下面是其中的一些接口。

// 一般情况下，在发起网络请求前调用，可以设置特定的 Http

WebKit内核源代码分析	(13481)
WebKit内核源代码分析	(12159)
WebKit内核源码分析 (二)	(8532)
WebKit加载网页的流程(c	(6321)
WebKit中的html词法解析	(6317)
RFC791中文翻译一(dlmu	(6116)
如何使用gdb调试android	(5385)

评论排行	
WebKit 内核源代码分析	(28)
[转]Android中的Intent详	(19)
WebKit内核源码分析 (一	(10)
WebKit内核源代码分析	(10)
WebKit内核源代码分析	(8)
MMS规范翻译 (MMS Ar	(5)
WebKit中的Chrome和Ch	(5)
WebKit内核源码分析 (二	(5)
How WebKit's loader wo	(4)
在ubuntu上编译chrome	(4)

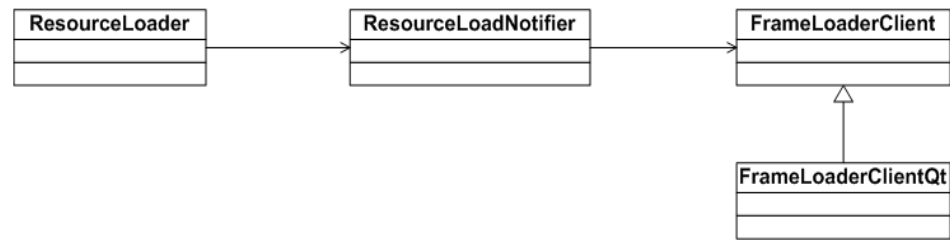
推荐文章	
* Windbg调试命令详解 (3)	
* 如何建立基于CocoaPods的ReactiveCocoa工程	
* Android 下拉刷新框架实现	
* linux内核奇遇记之md源代码解读之raid5数据流之同步数据流程	
* Android 类似电影倒计时效果 (两种随你选)	
* Linux系统NAT实现机制的升级改进	

最新评论	
换常用邮箱了	asd333333: 地瓜兄,我有些浏览器相关的问题想请教您,已经发邮件到您的163邮箱,希望您能抽空指点一二。谢谢地瓜兄...
五种利用strace查故障的简单方法	asd333333: 地瓜兄,我有些浏览器相关的问题想请教您,已经发邮件到您的163邮箱,希望您能抽空指点一二。谢谢地瓜兄...
五种利用strace查故障的简单方法	那时花儿: 顶。strace -p 15427这个对调试死锁,系统卡住非常有帮助,哈哈
WebKit内核源码分析 (一) ---dlmu	popo004: 请教楼主,哪里能下到webkit的源码,我正准备学习浏览器。
RFC793中文翻译一(dlmu2001)	sanbailiushiliuye: @dlmu2001:我是做Linux c的,想找点tcp底层的协议,就看到你的帖子了 :)
RFC793中文翻译一(dlmu2001)	红心地瓜: @sanbailiushiliuye:hehe,你做什么的?

```
头部,比如 user agent 等,在重定向请求的时候,也会自动调用
void willSendRequest(ResourceHandle*, ResourceRequest&, const
                        ResourceResponse&)
// 上传数据的时候,在 TCP wrtie 事件的时候,向对方发送数据的时候调用, loader 可以根据这个回调显示上传进度。
void didSendData(ResourceHandle*, unsigned long long
                  /*bytesSent*/, unsigned long long /*totalBytesToBeSent*/)
// 收到第一个响应包,此时至少 http 的部分头部已经解析 (如 status code ), loader 根据响应的头部信息判断请求是否成功等。
void didReceiveResponse(ResourceHandle*,
                        const ResourceResponse&)
// 收到 HTTP 响应数据,类似 tcp 的 read 事件,来 http 响应数据了, Network 的设计机制是来一段数据上传一段数据。
void didReceiveData(ResourceHandle*, const char*, int,
                    int /*lengthReceived*/)

// 加载完成,数据来齐。
void didFinishLoading(ResourceHandle*, double /*finishTime*/)
// 加载失败
void didFail(ResourceHandle*, const ResourceError&)
// 要求用户鉴权
void didReceiveAuthenticationChallenge(ResourceHandle*,
                                       const AuthenticationChallenge&)

WebCore 把要加载的资源分成两类,一类是主资源,比如 HTML 页面,或者下载项,一类是派生资源,比如 HTML 页面中内嵌的图片或者脚本链接。这两类资源对于回调的处理有很大的不同,比如,同样是下载失败,主资源可能需要向用户报错,派生资源比如页面中的一张图下载失败,可能就是图不显示或者显示代替说明文字而已,不向用户报错。因此有了 MainResourceLoader 和 SubresourceLoader 之分。它们的公共基类 ResourceLoader 则完成一些两种资源下载都需要完成的操作,比如通过回调将加载进程告知上层应用。
ResourceLoader 通过 ResourceNotifier 类将回调传导到 FrameLoaderClient 类。
```



主资源的加载是立刻发起的,而派生资源则可能会为了优化网络,在队列中等待 (这里的立刻发起是 loader 层面的,不是 Network 层面的)。ResourceScheduler 这个类就是用来管理资源加载的调度。主要调度对象就是派生资源,会根据 host 来影响资源加载的先后顺序。

主资源和派生资源的加载还有一个区别,主资源目前是没有缓存的,而派生资源是有缓存机制的。这里的缓存指的是 Resouce Cache ,用于保存原始数据 (比如 CSS

RFC793中文翻译一(dlmu2001)
sanbailiushiliuye: 校友呀，呵呵

WebKit 内核源代码分析（四）
u010353426: 您好，我是猎头Luna，想和您认识一下，我有些webkit高端职位，想跟您推荐。方便的话，加我QQ:...

浏览器背景知识(dlmu2001)
whb_18718: 你好，网页中的视频地址是浏览器解析还是flash 插件来解析？或者说播放网页视频的时候浏览器和fl...

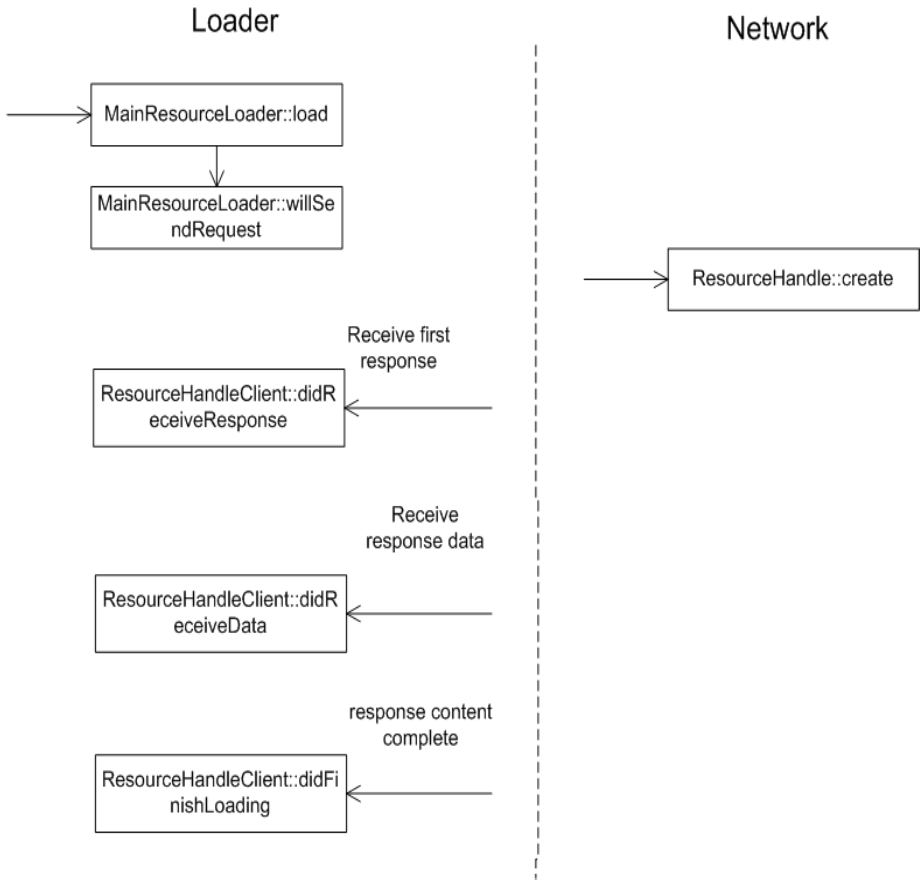
[转]Android中的Intent详细讲解
pimkle: 懂了一点

- 技术链接
- 云风
 - 王建硕
 - 阿里的aulik5同学
 - 李先静
 - tinyfool
 - 冯大辉
 - 伐木丁丁
 - 常高伟
 - 冷冰
 - 蜗牛
 - 陈罡
 - 流媒体技术工作组

， JS 等），以及解码过的图片数据，通过 Resource Cache 可以节省网络请求和图片解码的时候。不同于 Page Cache ， Page Cache 存的是 DOM 树和 Render 树的数据结构，用来在前进后退的时候快速显示页面。

二、加载流程

下图是加载 html 页面时，一个正常的加载流程。



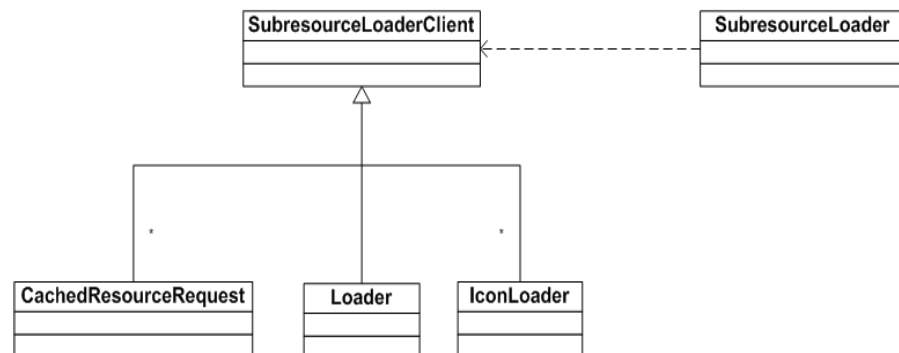
三、主资源加载过程

1. DocumentLoader 调用 MainResourceLoader::load 向 loader 发起请求
2. 调用 MainResourceLoader::loadNow
3. 调用 MainResourceLoader::willSendRequest
4. 调用 ResourceLoader::willSendRequest, 将 callback 通过 ResourceNotifier 传导给 FrameLoaderClient 。 Client 可以在回调中操作 ResourceRequest ，比如设置请求头部。
5. 调用 PolicyChecker::checkNavigationPolicy 过滤掉重复请求等
6. loader 调用 ResourceHandle::create 向 Network 发起加载请求
7. 收到第一个 HTTP 响应数据包 ,Network 回调 MainResourceLoader::didReceiveResponse ，主要处理 HTTP 头部。
8. 调用 PolicyChecker:: checkContentPolicy, 并最终通过 FrameLoaderClient 的 dispatchDecidePolicyForMIMEType 判断是否为下载请求（存在 "Content-Disposition"http 头部）

9. 调用 `MainResourceLoader::continueAfterContentPolicy` , 根据 `ResourceResponse` 检测是否发生错误。
10. 调用 `ResourceLoader::didReceiveResponse` , 将 `callback` 通过 `ResourceNotifier` 传导给 `FrameLoaderClient` 。
11. 收到 HTTP 体部数据, 调用 `MainResourceLoader::didReceiveData`
12. 调用 `ResourceLoader::didReceiveData` , 将 `callback` 通过 `ResourceNotifier` 传导给 `FrameLoaderClient`
13. 调用 `MainResourceLoader::addData`
14. 调用 `DocumentLoader::receivedData`
15. 调用 `DocumentLoader::commitLoad`
16. 调用 `FrameLoader::commitProvisionalLoad` , `FrameLoader` 从 `provisional` 状态跃迁到 `Committed` 状态
17. 调用 `FrameLoaderClientQt::committedLoad`
18. 调用 `DocumentLoader::commitData` , 启动 `Writer` 对象来处理数据 (`DocumentWriter::setEncoding` , `DocumentWriter::addData`)
19. 调用 `DocumentWriter::addData`
20. 调用 `DocumentParser::appendByte`
21. 调用 `DecodedDataDocumentParser::appendBytes` 对文本编码进行解码
22. 调用 `HTMLDocumentParser::append` , 进行 HTML 解析
23. 数据来齐, 调用 `MainResourceLoader::didFinishLoading`
24. 调用 `FrameLoader::finishedLoading`
25. 调用 `DocumentLoader::finishedLoading`
26. 调用 `FrameLoader::finishedLoadingDocument` , 启动 `writer` 对象接收剩余数据, 重复 19-22 进行解析
27. 调用 `DocumentWriter::end` 结束接收数据 (调用 `Document::finishParsing`)
28. 调用 `HTMLDocumentParser::finish`

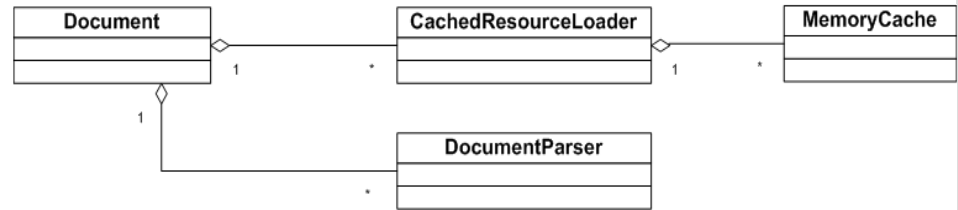
四、派生资源加载流程

在派生资源的加载中, `SubresourceLoader` 更多起到的是一个转发的作用, 通过它的 `client` (`SubresourceLoaderClient` 类) 来完成操作。



各个加载阶段的处理在 `SubresourceLoaderClient` 的派生类 `CachedResourceRequest`, `Loader`, `IconLoader` 中完成。Client 会创建 `SubresourceLoader` 。

请求发起阶段，ResourceLoadScheduler 负责对 SubresourceLoader 进行调度。



Document 类会创建 CachedResourceLoader 类的对象 m_cachedResourceLoader，这个类（对象）提供了对 Document 的派生资源的访问接口 requestImage，requestCSSStyleSheet，requestUserCSSStyleSheet，requestScript，requestFont，requestXSLStyleSheet，requestLinkPrefetch。为了实现这些接口，CachedResourceLoader 需要创建 CachedResourceRequest 对象来发起请求。

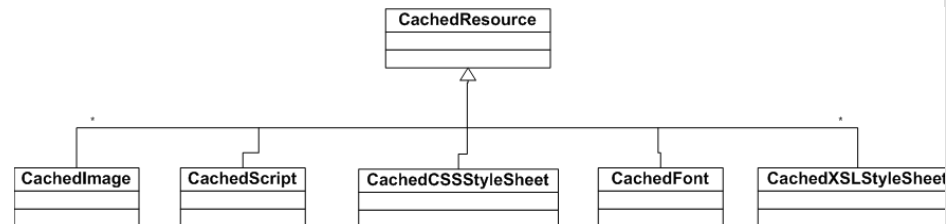
一般情况下，一个 Document 拥有一个 CachedResourceLoader 类实例。

MemoryCache 类则对提供缓存条目的管理，可以方便地进行 add，remove，缓存淘汰等。具体的缓存条目则是通过 CachedResource 类存储，MemoryCache 类维护了一个 HashMap 存储所有缓存条目。

```
HashMap <String,CachedResource> m_resources;
```

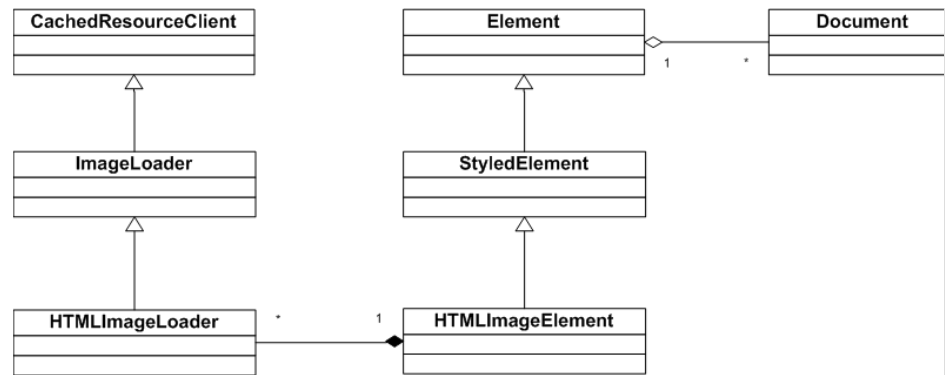
CachedResourceRequest 依赖于 CachedResource，在 CacheResourceRequest 的构造函数中，会传入 CachedResource 对象作为参数。CachedResource 既存储响应体部，也存储同 cache 相关的头部。在发起请求前，会检查是否有 cache 的 validator，在收到响应的时候，则需要更新对应的头部。CachedResource 类实现了 RFC2616 中的缓存一节。实际上 CachedResource 类真正完成了同网络的通信。

CachedResource 类根据申请的资源类型派生出不同的子类。



CachedResource 类的使用者必须是 CachedResourceClient，在这个类中维护了 CachedResourceClient 类的集合 m_clients。每一个 Client 通过 addClient 和 removeClient 将自己加入到该类的 Client 集合中。CachedResourceClientWalker 则提供了 CachedResourceClient 的一个遍历接口。当数据来齐的时候，CachedResource 类会通过 CachedResourceClient::notifyFinished 接口通知使用者。

下图是 Image 元素对应的几个类关系。



下面以 image 为例分析其加载过程

1. 解析 html 页面的时候，解析到 img 标签，调用 HTMLImageElement::create 创建 HTMLImageElement 对象，该对象包含 HTMLImageLoader 对象 m_imageLoader
2. 解析到 img 的 href 属性，调用 ImageLoader::updateFromElementIgnoringPreviousError
3. 调用 ImageLoader::updateFromElement
4. 调用 CachedResourceLoader::requestImage
5. 调用 CachedResourceLoader::requestResource (根据缓存的情况确定是否可以从缓存获取，或者需要 revalidate，或者需要直接从网络获取)
6. 调用 CachedResourceLoader::loadResource
7. 根据 Resource 的类型调用 createResource 创建对应的 CachedResource
8. 调用 MemoryCache::add 在 cache 中查找是否有对应的 cache 条目，如果没有创建之
9. 调用 CachedImage::load
10. 调用 CachedResource::load
11. 调用 CachedResourceLoader::load
12. 调用 CachedResourceRequest::load
13. 创建 CachedResourceRequest 对象，它将作为 SubresourceLoader 的 client
14. 调用 ResourceLoaderScheduler::scheduleSubresourceLoad
15. 调用 SubresourceLoader::create
16. ResourceLoadScheduler::requestTimerFired
17. 调用 ResourceLoader::start
18. 调用 ResourceHandle::create 发起请求
19. 收到 HTTP 响应头部，调用 ResourceLoader::didReceiveResponse
20. 调用 SubresourceLoader::didReceiveResponse
21. 调用 CachedResourceRequest::didReceiveResponse 处理响应头部，特别是同缓存相关的头部，比如 304 的 status code
22. 调用 ResourceLoader::didReceiveResponse
23. 收到体部数据，调用 ResourceLoader::didReceiveData
24. 调用 SubresourceLoader::didReceiveData
25. 调用 ResourceLoader::didReceiveData
26. 调用 ResourceLoader::addData 将数据存储到 SharedBuffer 里面

27. 调用 `CachedResourceRequest::didReceiveData`
28. 数据来齐，调用 `ResourceLoader::didFinishLoading`
29. 调用 `SubresourceLoader::didFinishLoading`
30. 调用 `CachedResourceRequest::didFinishLoading`
31. 调用 `CachedResource::finish`
32. 调用 `CachedResourceLoader::loadDone`
33. 调用 `CachedImage::data` ，创建对应的 `Image` 对象，解码

分享到：

上一篇：[linux下编译和使用7-zip](#)
下一篇：[htc magic 刷机](#)

NIKE中国
官方商城

24HOUR
专业运动装备库



¥ 1599



¥ 839



¥ 1199



¥ 1699



¥ 1249

查看评论

20楼 [u010353426](#) 2013-04-18 15:38发表



您好，我是猎头Luna,想和您认识一下，我有些webkit高端职位，想跟您推荐。方便的话，加我QQ:1012456121~或者发送最新的联系方式给我，我们详细聊~
谢谢~

19楼 [zshshr001](#) 2013-02-07 15:51发表



楼主写的非常好，请教一下：我如何实现各种子资源的调度，比如说:我打开网页的时候，先不下载图片资源和js脚本资源，让他们在后台自动下载，先把所有html资源下载下来？

18楼 [ckw20xx](#) 2012-03-13 10:36发表



请问大大，Document Frame::m_doc成员是在什么地方初始化的？

17楼 [hsupoyang](#) 2012-03-02 11:36发表



可以請問一下有關使用didReceiveAuthenticationChallenge()這個method的方法嗎？謝謝。

16楼 [anhongwan](#) 2011-12-07 16:19发表



楼主问一下，在webkit支持wml后，打开移动测试专用网址后不能在用户名处输入字母时。。

Re: [红心地瓜](#) 2011-12-11 19:16发表



回复anhongwan：什么意思，没听明白

15楼 [wj321666](#) 2011-12-05 13:28发表



问下楼主 怎么才能让android的webkit支持html5的<video>标签（不使用adobe flash）？

Re: [红心地瓜](#) 2011-12-07 10:48发表



回复wj32166：不知道你使用什么版本的android，2.3以下用的webcore的版本是比较老的版本。最新的webcore里面应该已经有video的支持了，但是各种平台具体实现方式可能有区别，比如有的实现自己解码，有的则利用平台的MediaPlayer的组件。很抱歉我没有做过深入的研究，你可以通过HTMLVideoElement找找看有没有相关的线索。

14楼 [lwjtutu_1985](#) 2011-12-02 14:47发表



楼主写到：ImageLoader::updateFromElementIgnoringPreviousError
加载image的资源是从这个调用开始的，我在跟从解析到href属性到上面这个函数调用之间的过程，很是疑惑，请问从这个调用是在解析token的过程中进行的，还是已经解析完token了在构建DOM树的过程中调用的呢？这中间的调用流程是怎么样的？希望指点一下。谢谢！

Re: 红心地瓜 2011-12-04 20:05发表



回复lwjtutu_1985 : 边解析边构建DOM tree , 在构建dom tree的过程中
WebCore::ImageLoader::updateFromElementIgnoringPreviousError
WebCore::HTMLImageElement::parseMappedAttribute
WebCore::StyledElement::attributeChanged
WebCore::Element::setAttributeMap WebCore::HTMLConstructionSite::createHTMLElement
WebCore::HTMLConstructionSite::insertSelfClosingHTMLElement
WebCore::HTMLTreeBuilder::processStartTagForInBody WebCore::HTMLTreeBuilder::processStartTag
WebCore::HTMLTreeBuilder::processToken WebCore::HTMLTreeBuilder::constructTreeFromAtomicToken
WebCore::HTMLTreeBuilder::constructTreeFromToken WebCore::HTMLDocumentParser::pumpTokenizer

Re: 红心地瓜 2011-12-04 20:06发表



回复dlmu2001 : 边解析token边构建DOM tree

13楼 [liguang83942488](#) 2011-11-09 10:14发表



膜拜

12楼 [lucifer_z7](#) 2011-11-01 12:19发表



请问,如果要传出HTTP状态码比如404,怎么处理?在哪传出?

Re: 红心地瓜 2011-11-02 14:04发表



回复lucifer_z7 : ResourceLoader::didReceiveResponse已经受到http header , 可以沿这条线下去, 比如
FrameLoaderClient***.cpp的dispatchDidReceiveResponse

11楼 [boundis](#) 2011-09-09 10:36发表



问下楼主, webkit 负责对页面内容中类似url的地址如: a.d.com 高亮加下划线等操作 是在哪部分完成的呢?

10楼 [zfreeze](#) 2011-08-11 17:28发表



楼主写的很好!支持一下!

请问一下,插件资源的加载流程是怎样的?比如网页中的Flash对象.

9楼 [fa407](#) 2011-06-15 14:14发表



请问,如果要添加网页打开失败的提示,怎么处理?

是在MainResourceLoader类中的receiveError中添加消息处理么?
ResourceNotifier 类主要起什么作用?

Re: 红心地瓜 2011-06-16 13:34发表



回复 fa407 : FrameLoaderClientQt::callErrorPageExtension.ResourceNotifier实现了将相应的内核事件通知使用者的作用。

8楼 [sxzbsnow](#) 2011-06-03 16:52发表



楼主可否写一个调试的方法。。[e01][e01]

Re: 红心地瓜 2011-06-16 13:31发表



回复 sxzbsnow : 就是gdb调试。网上很多相关的文章。

7楼 [redarclive](#) 2011-05-03 16:58发表



不知道楼主看的是 哪个webkit, android的 chrom的还是safari的, android版的webkit调试相对容易, 打log即可, resourceload是载入页面的第一步

Re: 红心地瓜 2011-05-03 17:25发表



回复 redarclive : qt. gdb+vim调试

6楼 [dongguac10](#) 2011-05-03 10:58发表



[e03]顶顶·-

5楼 红心地瓜 2011-04-29 11:07发表



时间就像**，挤挤总会有的,[e04]

4楼 blazeman 2011-04-28 16:32发表



你们公司一直可以学习吗? 工作不忙?

3楼 sunshine009 2011-04-28 13:44发表



好好学习下！

2楼 Moonkiller1026 2011-04-28 10:18发表



[e01][e02][e03][e04][e05][e06][e07][e08][e09]

1楼 imshhui 2011-04-27 13:53发表



看不懂，学习的路还很长[e06]

您还没有登录,请[登录](#)或[注册](#)

* 以上用户言论只代表其个人观点，不代表CSDN网站的观点或立场

专区推荐内容

Android 游戏教程:让人物...
Linux 下的 UltraEd...
关于HTML怎样用图片做背景
Android 游戏教程:让人物...
HTML5应用性能调优工具WAP...
Android应用开发



更多招聘职位

我们公司职位也要出现在这里

【上海全成通信科技有限公司】高级java工程师
【上海香象信息技术有限公司】软件UI设计师
【深圳爱展销网络有限公司】高级（.net）开发工程师
【上海锐斯康宁】android开发工程师
【数赢网络科技有限公司】网络营销项目副总监
【广州河谷医疗科技有限公司】系统开发人员

公司简介 | 招贤纳士 | 广告服务 | 银行汇款帐号 | 联系方式 | 版权声明 | 法律顾问 | 问题报告

QQ客服 微博客服 论坛反馈 联系邮箱：webmaster@csdn.net 服务热线：400-600-2320

京 ICP 证 070598 号

北京创新乐知信息技术有限公司 版权所有

世纪乐知(北京)网络技术有限公司 提供技术支持

江苏乐知网络技术有限公司 提供商务支持

Copyright © 1999-2012, CSDN.NET, All Rights Reserved 

第9页 共9页

2013年10月15日 11:30