

WEBGIS 混合缓存的应用与研究^①

Application and Research of WebGIS Hybrid Cache

魏祖宽¹ 胡娟¹ 金在弘² (1.电子科技大学 计算机科学与工程学院 四川 成都 610054;

2.永同大学 计算机工学科 韩国 忠清北道 永同郡 370701)

摘要: 大数据量的网络传输速度和空间信息的分布式处理效率成为影响 WebGIS 系统性能的关键因素。缓存技术是改进系统性能的关键技术。针对胖客户模式、瘦客户模式存在的不足,提出一种适用于电网可视化管理系统的混合缓存模式,即通过运用服务器缓存和客户端缓存技术,在服务器端基于地图缓存和无状态的池式服务,在客户端采用内存缓存,提高用户请求响应应用速度。实验表明,这种模型可以减少数据的网络传输量,提高空间数据的存取效率和并发访问能力。

关键字: WebGIS 地图切片 地图缓存 混合缓存 电网分析 缓存分析

在分布式 WebGIS 中, WebGIS 在执行区域查询时需要从不同的远程结点上获取大量的空间数据进行集成处理,最后生成结果数据。大量数据的传输将带来较大的网络通信流量,并降低查询的响应速度。为此,一些 WebGIS 系统采用缓存机制来减少网络通信流量。缓存是分布式系统中提高性能的关键技术,其基本思想是将用户经常访问的内容保存在 Web 上距用户较近的结点,以便能以较快的速度查询到这些内容,从而减轻数据源服务器的负担,避免频繁访问远程数据库,降低通信代价和存储代价。

缓存的缺点:一、缓存中数据项变化太快、缓存时间过长或过短等因素,会降低 Web 应用程序的响应速度;二、使用缓存时要消耗一定的内存,所以在使用缓存时一定要注意内存的问题。

1 WEBGIS缓存模式

WebGIS 采用由数据库、服务器和客户端组成的三层体系结构,客户端一般为 Web 浏览器。WebGIS 系统具有空间数据量大和空间处理复杂的特点,因此产生了计算模式的概念。WebGIS 的计算模式主要是指 GIS 功能在客户端和服务器端的分配,WebGIS 计

算模式的选择决定了整个 WebGIS 系统的实现。WebGIS 的计算模式主要包括以下三种:胖客户模式、瘦客户模式和混合模式。

胖客户模式适合于客户端处理能力较强,用户需要对数据处理过程进行控制的环境,这种模式的客户端的计算能力难以保证,并且各种网络安全漏洞比较多;瘦客户模式则适用于广域网环境或对 GIS 分析功能要求较高的应用,客户的每一次操作都要与服务器端交互一次,客户与服务器之间的会话及数据传输频繁;而混合模式结合了胖客户模式和瘦客户模式的优点,它既不是把全部的空间处理功能模块和数据下载到本地,在客户端进行所有的空间操作;也不是把全部的空间处理功能放置在服务器端,在服务器进行所有的空间操作;而是根据 Web 应用的特点和网络的情况,在客户端和服务器端进行空间处理功能的分配。

从以上分析来看,混合模式是一种符合 WebGIS 应用需求的系统开发模式。但与其它的信息系统一样,不存在一种万能的计算模式,因此需要根据具体的应用需求和运行环境,对计算模式进行选择,以使开发的 WebGIS 应用系统能最大可能地满足应用的需求。

^① 基金项目:成都科技局攻关项目计划(O6GGYB801GX-032)

收稿时间:2008-11-23

2 混合缓存技术及其应用

对于目前的电力部门来说,正确、全面、及时地获取电力企业的各种资源信息,并加以提炼、分析,为电力企业的管理者和决策者提供辅助决策的依据,从而保障电力网络安全、高效地运作,以及为更多用户提供更加完善周到的服务,这是在电力企业内建设信息系统的根本出发点。

凭借对电网企业业务流程和电力网络拓扑的深入分析和提炼,结合当前最新的 3S 技术和 IT 技术,以面向电网为核心,形成了端到端的配电管理系统整体解决方案—电网可视化管理系统 GVMS (Grid Visual Management System)。

GVMS 系统是一套基于多层架构的软件系统,它由数据层、业务逻辑层和展现层组成(如图 1 所示)。在 GVMS 系统 B/S 应用中,基于 ArcGIS Server ADF for.NET 开发平台,为了提高空间数据的访问效率和并发性能,减少空间数据的传输,以及考虑到应用程序的可扩展性,我们必须使用混合缓存技术。

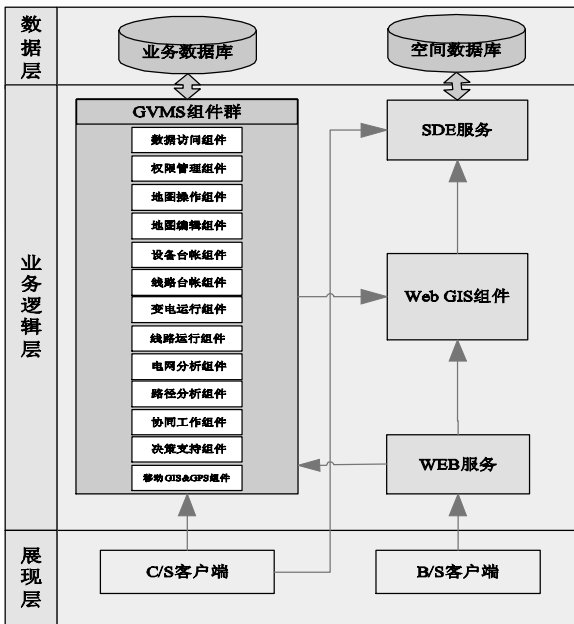


图 1 GVMS 系统部署图

2.1 混合缓存原理

混合缓存模型包括两个层次:服务器缓存和客户端缓存。它采用基于地图切片的地图缓存原理,地图切片之后将其放置于服务器的虚拟目录中,在需要显示某个范围的地图时,借助客户端技术将这些图片无

缝地拼接在一起,即可得到用户所需要的地图。

2.2 服务器缓存

2.2.1 地图切片原理

地图切图基于预生成技术的 Web 地图显示系统,地图的预生成一般将指定范围的地图按照指定尺寸和指定格式切成若干行及列的正方形图片,切图只能得到二维矢量数据对应的栅格图。

2.2.2 地图切片算法

(1) 参数定义

- ① 地图等级为从 1 开始、行列数为从 0 开始的整数;
- ② 地图范围的极值用 $XMin, XMax, YMin, YMax$ 表示;
- ③ 用 Δx 表示横坐标差, Δy 表示纵坐标差;
- ④ 地图瓦片的命名一般采用包含地图等级、行列数的方式如 "zoom_x_row_column", 其中 x 表示地图等级, row 和 column 分析表示当前切片所处的行数和列数。

(2) 地图数量计算

设第一个等级地图为 m 行 n 列,则第 level 级的地图行列数如下:

$$\text{行数: } m \times 2^{level-1}, level \geq 1 \quad \text{列数: } n \times 2^{level-1}, level \geq 1$$

(3) 算法流程 (如图 2 所示)

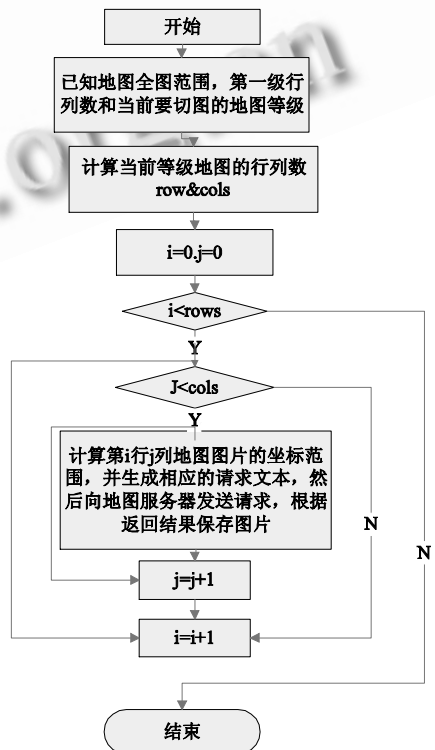


图 2 地图切片的算法流程

(4) 地理坐标与图片位置正反算

①正算

输入：行列对 (i, j)

输出：范围(Extent) $curXMin, curXMax, curYMin, curYMax$ 第 level 级下第 i 行 j 列位置切片的坐标范围计算公式如下：

$$curXMin = XMin + j \times \frac{\Delta x}{m \times 2^{level-1}}$$

$$curXMax = XMin + (j + 1) \times \frac{\Delta x}{m \times 2^{level-1}}$$

$$curYMin = YMax - (i + 1) \times \frac{\Delta y}{n \times 2^{level-1}}$$

$$curYMax = YMax - i \times \frac{\Delta y}{n \times 2^{level-1}}$$

②反算

输入： x, y (坐标值)

输出： i, j (图片的行列数)

$$i = \text{Math.Floor} \left[\frac{YMax - y}{\Delta y} \times n \times 2^{level-1} \right]$$

$$j = \text{Math.Floor} \left[\frac{x - XMin}{\Delta x} \times m \times 2^{level-1} \right]$$

2.2.3 地图缓存

地图切片之后可将其放置于本地或服务器的虚拟目录中，在需要显示某个范围的地图时，根据上面坐标与图片位置正反算公式即可计算出需要哪些图片，然后借助客户端技术将这些图片无缝地拼接在一起，即可得到用户所需要的地图。视觉上感觉是连续的地图在后台都是一张张命名规则的尺寸相同的预先切好的图片，但都借助预生成技术和界面友好的客户端提高了地图浏览速度，增加了用户体验。一个完整的地图缓存包括：

缓存目录：包含在特定比例尺级别下，地图范围内的图片块，以及一个描述缓存块结构的文件(Conf.xml)。如图 3 所示。

Web 服务器：Web 的虚拟目录参考实际的缓存目录。

GIS 服务器：提供地图缓存信息，支持查询和数据操作的地图服务，当缓存不可用时，就动态生成地图。

2.2.4 无状态的池式模式

服务器对象是否使用状态与服务器对象是否是共享池模式有密切关系。一个池式的服务器对象是默认为无状态使用的，因为这个服务器对象可能会给多个

用户访问，而且它用完后是会放回共享池中的，如果改变了它的状态，则意味着不同的访问者访问它时获得的信息不一致，因此，保证本身具有“共享”特征的池式服务器对象的无状态使用是有效满足多用户访问的方式。

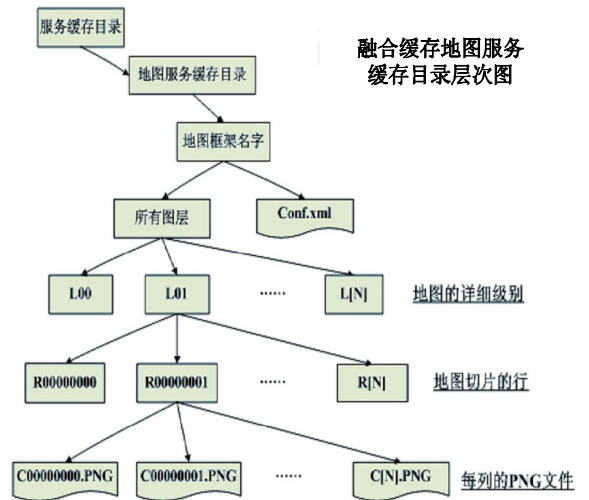


图 3 单一融合缓存地图服务缓存目录层次图

2.3 客户端缓存

服务器的缓存让你的页面访问起来非常地快，但它依然需要依赖浏览器下载并输出，这种方式带来了系统配置和维护上的难题，并且不能发布海量空间数据，而当你加入客户端缓存时，即服务器将它最常访问的文件复制到本地存储器中，每当用户请求该页面时，服务器就从存储的副本中验证文件是否尚未更改。服务器从该页面的源中发布一个已更改的页面，而从本地高速缓存中发布一个未更改的页面，以减少网络通信流量和缩短访问所需的时间。

2.4 混合缓存机制的工作流程

GVMS 是一个三层的分布式系统，第一层是客户端、第二层是 WEB 服务器和 GIS 服务器，第三层是数据库服务器。服务端缓存位于 WEB 服务器和 GIS 服务器上，Web 服务器用于接收客户端请求并运行 WEB 应用程序和服务；GIS 服务器负责管理和运行服务器对象。(如图 4 所示)

服务端缓存存储从空间数据库中获取的反馈给客户端的数据。根据客户端请求类型，Web 服务器会将请求转发给对应的应用服务器，如果请求的空间数据存在于服务端缓存目录中(如图 4 橙色线所示)，则应

用服务器直接从服务端缓存中取出数据,从而减少访问次数,提高了数据获取效率。

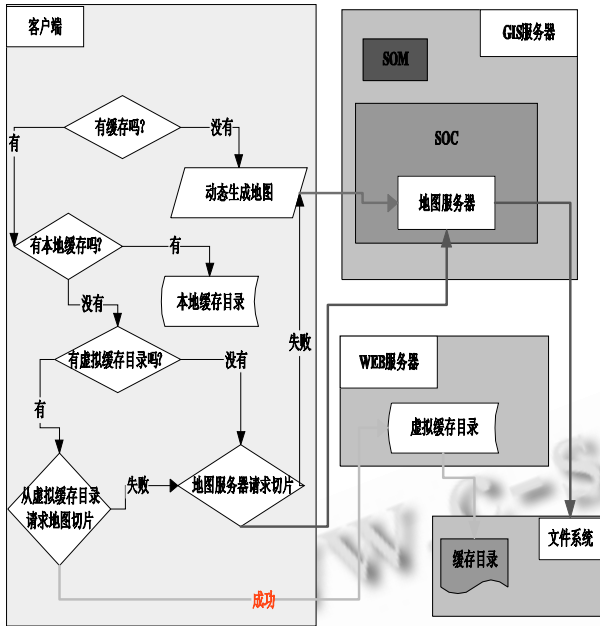


图 4 客户端访问缓存块流程图

3 WEBGIS缓存分析

3.1 缓存类型

3.1.1 融合缓存

融合缓存就是预先将要显示的地图的所有图层,合在一起生成一序列静态的图片放在服务器上,客户端在浏览地图时就可以直接读取这些图片。在融合缓存中,负担比较重的是WEB服务器和网络,它们负责传输图片;而SOC机器的负担相对较小,它主要计算地图的显示比例。缺点:对数据进行编辑后,需要重新生成缓存。采用这种缓存方式,静态地图需要1.5秒,而动态生成地图需要4秒。

3.1.2 分层缓存

分层缓存包含了每个级别的每个单独图层创建的图片块。客户端浏览时需要先解压每个图层的图片,然后将这些图片再动态合成在一起,再将合成后的图片压缩后传输到客户端。这种方式下,WEB服务器的负担比较重,SOC机器和数据库服务器的负担很小或者几乎没有。在图层较多的时候,对性能的改善没有明显提高。多应用于三维地图。

3.1.3 按需缓存

按需缓存:不需要缓存整个地图,只需要为经常访问的区域创建缓存,其它的区域可以设置为按需缓

存。一旦访问设置了按需缓存的地图就会促使GIS服务器创建该区域的缓存,并且保存用于下次访问。这样可以节省缓存创建时间和磁盘空间。

3.2 性能测试

WebGIS系统的性能评价从以下几个方面来分析:用户界面的响应速度、多用户并发操作时的效率、系统处理大数据的能力、数据库的执行效率、系统的可移植性、可扩展性、安全性等方面。其中,数据响应速度和并发性能是用户所关注的首要性能,由于WebGIS系统中数据处理量传输量都很大这个特殊的限制,WebGIS系统大部分技术都是以提高用户响应速度为主要目标,上述的其他诸多性能也以提高速度为最终目的。

在电网可视化管理系统中,基于混合缓存模型,客户端使用Act软件+Fiddler软件工具测试性能。由Web服务器分配客户端请求,并向GIS服务器请求数据,服务器端缓存由Web服务器实现。网络传输速率为10M/s,地图数据为成都市地图,并分别在8级比例尺下进行图片栅格化及地图分割,将地图分割为大小是512*512的栅格图,最后将图片压缩为PNG格式保存在GIS服务器端,大小约87MB。客户端测试的基本操作包括:

- ① 起始页、地图放大、地图再次放大、地图移屏、地图再次移屏,全图六次操作作为一个操作流程。
- ② 数据查询,将查询放在循环体中执行(将查询结果放到一个临时表中)将一行数据处理10次、100次、1000次。将记录到的100次最慢的查询时间记录和100次最快的查询时间记录舍弃,然后根据剩余的执行时间记录计算查询的平均执行时间。

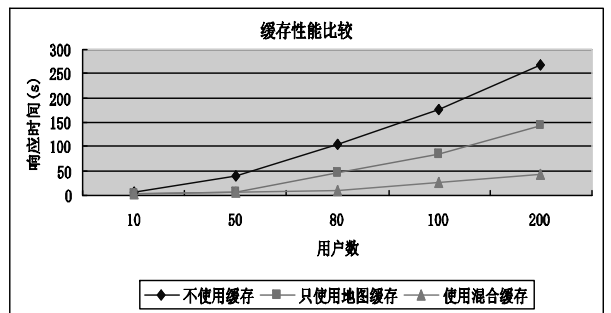


图 5 缓存性能的比较

实验结果如图5,分别在不使用缓存、只使用服务器缓存和使用混合缓存三种情况下对系统响应时间

进行测试。通过对实验结果进行分析,发现在相同用户数访问的情况下,使用混合缓存比其它两种情况的系统响应速度明显加快,且在用户数加大的情况下,时间响应曲线相对较为平缓,说明在面向多用户并发访问的情况下,采用客户端缓存对系统的性能提升有着较为明显的作用。

4 总结

针对如何提高分布式 WebGIS 快速响应用户请求的问题,本文提出了一种混合缓存机制,基于地图缓存和池式无状态服务的服务器缓存加客户端缓存的混合缓存原理,讨论了在客户端访问地图缓存的工作流程,并对其性能进行了分析和比较。经过实验可以得出,基于地图缓存的混合缓存模型能够明显提高用户请求响应速度和并发性能。今后的研究工作主要:通过对二维地图预存取技术的研究,扩展三维地图预存取技术;在数据缓存方面,基于兴趣区域和按需缓存技术,将重点研究三维地图图层之间的关系,通过建立分层缓存,提高系统的性能。

参考文献

- 1 王强.分布式 WebGIS 地图缓存策略的应用研究.软件导刊, 2008,7(5):16-18.
- 2 祁羽,陈莹,张瑞雪,等.基于双缓存机制的分布式 WebGIS 数据集成访问策略.计算机工程与科学, 2007,29(5):41-44.
- 3 李浩松,朱欣焰,李京伟,等.WebGIS 空间数据分布式缓存技术研究.武汉大学学报·信息科学版, 2005, 30(12):1092-1095.
- 4 谢强,于雅丽,丁秋林,等.基于对象池和数据缓存技术的 Web_OLAP 系统.应用科学学报, 2007,25(2):161-165.
- 5 石磊,卫琳,古志民,等.Web 对象可缓存性研究及加速方案.计算机工程, 2005,31(18):74-75.
- 6 涂小鹏,汪林林.分布式空间数据库中基于事务的客户端高速缓存技术研究.计算机科学, 2004,31(6):76-78.
- 7 罗英伟,汪小林,许卓群.构件化 WebGIS 及其缓存框架.计算机辅助设计及图形学学报, 2005,17(2):320-326.
- 8 商新娜.Web 应用中的海量数据访问缓存技术.北京联合大学学报(自然科学版), 2007,21(3):45-49.
- 9 罗英伟,汪小林,许卓群.层次化 WebGIS 构件系统的设计与实现.计算机学报, 2004,27(2):177-185.
- 10 肖心智,苏奋振,杜云艳,等.WebGIS 性能分析与优化.测绘与空间地理信息, 2005,28(4):1-3.