

移动数据管理的研究与发展^{*}

Development and Research of Mobile Data Management

王意洁 卢锡城

(国防科技大学计算机学院并行与分布处理国家重点实验室 长沙410073)

Abstract The features of mobile computing include low bandwidth, frequent disconnections, predictable disconnections, high bandwidth variability, low security, fast changing locations, limited resources, limited battery power, expensive, etc., by which the higher demands of mobile data management are proposed. In this paper, key techniques in mobile data management are analyzed, which include location data management, cache consistency, data replication, query processing, and transaction processing. Moreover, several new upcoming research directions in location-dependent query processing, mobile view maintenance, mobile workflow, mobile digital library, mobile Web and e-commerce, and mobile data security are highlighted.

Keywords Mobile computing, Mobile data management, Location, Cache consistency, Data replication, Query, Transaction

一、引言

随着社会信息的不断发展和普及,Internet 和无线移动通信技术的成熟,以及计算机处理能力的不断提高,新的业务和应用不断涌现,人们要求不仅能够在办公室和家庭等固定场所中获得和处理信息,而且能够在移动环境中进行相应的工作。于是,移动计算,这一满足人们移动信息处理要求的新兴技术应运而生。

简单地讲,移动计算就是动态的分布式计算,其网络结点之间的连接经常发生动态改变,而且网络结构也不是固定不变的。另外,固定计算机与移动计算机在计算能力和内存容量等诸多方面都存在着差异。因此,传统分布式计算中的许多技术都不能直接用于移动计算中。相对于传统的分布式计算而言,移动计算的特点主要包括:低带宽、可预知的频繁的断接、带宽的易变性、安全性较差、位置的快速变动性、有限的资源、电池能量有限、成本高等。这些都为移动计算中的数据管理提出了更高的要求。

1.1 有限的带宽和能量

由于无线网络的带宽有限,而且经常发生变化,所以移动计算特别关注网络带宽的使用和变化。通常,带宽被同一蜂窝中的不同移动用户分割使用,所以,可供每个移动用户使用的传输带宽更是非常有限。对于移动式设备而言,节约能量至关重要。现在,对能效(Energy Efficient)CPU、内存和系统软件的需求越来越迫切,特别是能效的数据访问协议和算法。

目前,各国学者针对有限的带宽和能量进行了广泛的研究,提出了一系列的解决方案。

·定期进行数据广播,而不是在有需求时才发送数据,诸如路况信息、股票行情、各种新闻等信息都适于进行定期广播。这样,移动用户可以减少发送请求,只在需要时才从睡眠状态中醒来,从而节约能量。获取广播数据只需要监听,不需要上行链路,许多移动用户都可以同时监听广播数据,因此具有良好的可扩展性。

·利用数据压缩、日志等技术来提高带宽的利用率,节约能量。数据压缩可以有效节约内存和带宽,但需要占用 CPU 资源来进行解压缩。日志有助于将众多小请求的应答组织成一个较大的应答,从而提高带宽的利用率。

·尽量在固定服务器上进行处理,减少移动用户的工作量。这样,由于移动用户的能量有限,需要在局部可访问数据量和服务器处理数据量之间进行权衡,也就是说,需要在服务器和移动用户之间进行数据划分。另外,还应该考虑到处理速度,允许的数据处理的等待时间越长,移动用户消耗的能量就越少。

由此可见,处理速度、存储开销、传输数据量和等待时间是影响数据访问和数据组织的重要因素。

1.2 可预知的频繁断接

由于低带宽、出错率较高和较为频繁的断接等因素,无线连接的质量较低。重传、超时、差错控制和断接等因素容易延长通信时间,提高通信开销。由于移动性的原因,无线连接也容易丢失,特别是当用户进入会场等高密度、高干扰度的地方,这些地方经常发生网络超载。由此可见,移动环境比传统的分布式计算环境更加容易出现故障,但是,其中某些故障是可以预知的。对于可预知的断接,系统可以采取某些特殊的措施,比如:

·如果不再需要与用户交互,那么可以将事务处理转移到固定的计算机上。

·在发生可预知的断接之前将远程数据下载下来,以便发生断接后交互事务在移动计算机上继续执行。

·由于移动计算的不稳定性,移动计算机在灾难性的故障中最易受到攻击,因此,在断接发生之前应该将移动计算机上的日志记录传送到固定计算机上,形成日志的复本。

二、移动数据管理

移动数据管理可以分为全局数据管理和局部数据管理。全局数据管理主要处理网络层的问题,具体包括:定位、寻址、

^{*} 本文得到国家自然科学基金项目(69903011,69933030)的资助。王意洁 博士,副教授,研究领域为移动计算,数据库和并行分布处理技术。卢锡城 中国工程院院士,教授,博士生导师,研究领域为大规模并行处理技术,移动计算和网络技术。

复制和广播等。局部数据管理涉及用户层的问题,具体包括能有效的数据访问、断接管理和查询处理等。

2.1 位置数据管理

在移动计算中,移动用户的定位至关重要,移动用户的位置可以视为一个随用户的移动而改变的数据项,移动计算中的位置管理实际上就是一个数据管理问题。位置管理的研究内容主要包括:1)如何确定移动用户的当前位置;2)如何存储、管理和更新位置信息。为了对移动用户进行定位,人们利用分布式位置数据库来维护移动用户的当前位置信息。移动用户位置的改变将直接引起位置数据库及其副本数据库中相应位置信息的更新。目前,关于位置管理的研究比较深入,有代表性的研究工作^[1]主要包括:基于本地位置寄存器(Home Location Register)的局部位置管理策略,基于层次分布式位置数据库的全局位置管理策略等。

2.2 缓存一致性

在移动计算中,对经常访问的数据进行缓存,可以缓解弱连接和断接对系统性能和可用性的限制,还可以减轻数据访问对低带宽的激烈争夺,从而提高查询处理效率,并对弱连接操作和断接操作提供支持。由于移动环境中的连接可以分为强连接、弱连接和断接,它们对缓存一致性的需求也不同,连接性越弱,对缓存一致性的要求越低,因此,缓存一致性也分为若干等级。缓存一致性主要受断接和用户移动性的影响。通常,通过服务器定期广播各种信息来实现缓存一致性,广播的信息包括实际数据、数据的更新信息、锁表和日志等控制信息等。广播不需要服务器知道移动用户的当前位置和连接状态,移动用户也无需建立与服务器的上行链路,移动用户无需发送数据请求,从而节约了能量,而且无需额外开销,广播的信息就可以被许多移动用户接收。缓存的一致性直接取决于广播的频率,广播越频繁,缓存更新就越及时,缓存一致性就越高。美国卡内基梅隆大学开发的 Coda^[2]是一个支持移动计算的分布式文件系统。在 Coda 中,当发生断接时,移动用户就直接对自己的缓存进行数据访问;当恢复连接后,系统将发布更新信息,更新用户的缓存。

2.3 数据复制

数据复制对提高移动计算的系统性能和可用性也是至关重要的。共享数据的同步约束因数据语义和用法的不同而不同。复制系统应该对断接模式、数据分歧、应用定义的协调过程、乐观并发控制等提供支持。用户和服务的移动性对数据复制和数据迁移带来的影响是当前需要解决的关键技术问题。目前,针对数据重分配、用户和服务的移动性等研究主要包括:1)在保证一定的一致性、持久性和可用性的前提下,如何管理数据复制;2)如何定位感兴趣的数据对象,是否需要复制位置信息;3)如何确定在移动用户上复制数据的条件;4)用户移动如何影响复制模式,是否需要动态复制模式;5)根据移动计算的需求,如何修改传统的分布式计算中的复制算法,是否需要设计新的复制算法。现在,关于上述问题的研究^[3]已经广泛开展,并提出了多层复制算法和动态复制模式等有效的数据复制管理策略。

2.4 查询处理

移动计算中的查询可以分为两类:第一类查询只涉及数据库的内容,第二类查询涉及到位置相关数据(Location-Dependent Data)。对第二类查询进行查询优化需要引入新的参数。在查询执行过程中,位置信息会发生变化。由于位置信息变化较快,查询结果多为近似结果。位置相关查询遇到的主要

问题包括:1)如何在变化的环境中实现对广播数据的查询;2)如何查询多个通道上的广播数据;3)如何组织广播数据,从而使移动用户的开销最小;4)哪些信息适于广播,哪些信息在需要时才提供。针对这些问题,目前已经提出了几种适于移动计算的查询处理模型^[4]。

2.5 事务处理

移动环境中的事务与传统事务的区别可以归纳为以下几点:

1)移动事务将其计算划分为若干个操作集合,一部分在固定计算机上执行,一部分在移动计算机上执行。由于移动性和断接等因素,移动事务之间共享部分结果和某些状态信息。

2)移动事务需要固定计算机提供对计算和通信的支持。

3)移动事务执行过程中,当用户移动时,移动事务继续执行,执行地点由移动用户确定。

4)当移动用户跨蜂窝移动时,事务的状态、所访问数据对象的状态和位置信息也需要移动。

5)由于数据和用户的移动性,以及频繁的断接,移动事务容易形成长事务。

6)移动事务支持并发、恢复、断接和复制数据的一致性。

鉴于此,移动事务模型应适应通信可靠性低、能量有限、低带宽、存储容量小等移动计算的特点。移动事务应尽量减少卷回,事务执行过程尽量减少固定计算机与移动计算机之间的通信开销,提高并发度。移动事务模型应提供局部自治性,允许在临时断接时,事务可以在移动计算机上执行并提交。目前提出的移动事务模型主要包括:基于语义的事务处理模型、乐观并发控制模型、事务代理机制、开放嵌套事务模型、Kangaroo 事务模型^[5]等。

三、发展趋势

3.1 位置相关查询处理

对位置相关数据的查询称为位置相关查询。许多位置相关查询都是聚合查询(Aggregate Query),例如,找出所在地区的所有五星级宾馆,找出离现在位置最近的医院,等等。由于位置信息是在不断变化的,因此要求找出给定查询针对不同位置信息的正确结果,查询结果需要满足关于位置信息的不同约束条件,比如,查询的发起位置,查询结果的接收位置等。查询涉及到的数据可以简单地分为三种:1)与位置无关的数据,如,人的姓名、身份证号码等;2)与位置有关的数据,如,人的房租,因为不同地区的房租标准是有差别的;3)与查询发起位置或查询结果接收位置有关的数据,如,驾车外出旅游者在驾车过程中寻找最近的旅馆,其查询结果要么是查询发起位置最近的旅馆,要么是查询结果接收位置最近的旅馆。目前,针对移动对象的描述问题,研究者们提出了动态属性的概念^[6],专门描述不断变化的信息,对动态属性建立索引,设计时空查询语言。

3.2 移动视图维护

由于移动计算机到固定计算机的上行链路有限,发送信息会消耗移动计算机有限的能量,因此移动计算机在线数据访问的成本较高。为了解决这个问题,人们在移动计算机上维护一个视图,用以存储固定计算机上的数据,固定计算机通过发送数据更新信息来更新移动计算机上的视图。在移动计算机上维护视图还可以有效提高数据访问的效率。维护视图主要涉及位置相关数据、时间相关数据、视图的动态分配^[7]等关键问题。视图与原始数据的数据分歧问题是视图动态分配的核心

问题。

3.3 移动工作流

工作流(Workflow)管理系统能够有效协调一个机构的运营过程,提高整个机构的工作效率,因此得到了长足的发展。工作流系统与移动计算相结合能够更加有效地提高系统对各种故障的恢复能力^[8]。移动环境中的工作流系统的关键技术包括:1)当移动用户处于断接状态或者跨蜂窝移动时,如何协调移动用户正在执行的任务;2)当移动用户处于断接状态时,如何保证用户任务的正确执行,如何协调移动用户之间的交互;3)当移动用户跨蜂窝移动时,如何解决一致性问题。目前,针对这些问题,已经提出了初步的解决方案。在可预知的断接发生之前,将与任务相关的信息提前下载到移动计算机上;任务在断接状态执行结束后,当连接恢复时,将执行结果上传;当移动用户跨蜂窝移动时,需要在新旧工作流服务器之间进行交接,即将移动用户与旧工作流服务器的连接迁移到新的工作流服务器上,并保证状态一致性。

3.4 移动数字图书馆

数字图书馆能够有效管理分布式环境中的GB级多媒体数据,具体包括文本、图片、图像、声音等。目前,对数字图书馆进行信息访问的用户都是静态的。在不久的将来,用户将通过无线连接访问数字图书馆。数字图书馆需要对大量的多媒体数据进行复杂处理,因此在移动环境下提供优质的数字图书馆服务是一个具有挑战性的研究课题。

3.5 移动Web和电子商务

随着Internet的发展和日益普及,人们越来越需要通过无线方式上网浏览,因此,建立移动Web,并进一步开展电子商务,已经不是遥不可及的幻想了。然而,移动环境的断接、可靠性低、带宽有限、安全性低等特点难以直接满足电子商务对可靠性和安全性的要求。这也是移动Web和电子商务需要解决的核心问题^[9]。

3.6 移动数据安全性

随着移动计算的不断发展,数据安全性已经成为日益紧迫的研究课题。移动计算中的数据安全性主要包括以下几个方面:1)保证原始数据的安全性,文^[10]提出对原始数据进行信息提取,形成元数据,移动用户只保存元数据,原始数据放在移动业务台(Mobile Service Station);2)针对频繁的断接,建立高效的恢复机制;3)提供安全认证服务;4)维护移动用户

的隐私权,即,保护移动用户的身份信息、位置信息等;5)建立完善的访问控制机制。

结论 移动计算中的数据管理面临着许多新的挑战,其研究内容涉及到网络层和用户层两个层次。为了适应移动计算的低带宽、可预知的频繁的断接、带宽的易变性、安全性较差、位置的快速变动性、有限的资源、电池能量有限、成本高等特点,在现有分布式数据管理的基础上,必须引入许多新的概念和思想。本文在分析移动计算主要特点的基础上,探讨了移动数据管理的关键技术,并展望了未来发展趋势。

参考文献

- 1 Jain R. Reducing Traffic Impacts of PCS using Hierarchical User Location Databases. In: Proc. of the IEEE Intl. Conf. on Communications, 1996
- 2 Satyanarayanan M. Mobile Information Access. IEEE Personal Communications, 1996, 3(1)
- 3 Wu Shiow-yang, Change Y. An Active Replication Scheme for Mobile Data Management. In: IEEE Proc. of 6th DASFAA, Taiwan, 1999
- 4 Massari A, Weissman S, et al. Supporting Mobile Database Access Through Query by Icons. Distributed and Parallel Databases: An International Journal, 1996
- 5 Eich M H, Helal A. A Mobile Transaction Model That Captures Both Data and Movement Behaviour. ACM/Baltzer Journal on Special Topics on Mobile Networks and Applications, 1997
- 6 Wolfson O, Sistla A P, et al. Updating and Querying Databases that Track Mobile Units. Distributed and Parallel Databases, 1997, 7(3)
- 7 Dong G, Mohania M. Algorithms for View Maintenance in Mobile Databases. In: Proc. of the First Australian Workshop on Mobile Computing and Databases and Applications (MCDA'96), Melbourne Australia, 1996
- 8 Alonso G, et al. Exotica/FMDC: Handling Disconnected Clients in a Workflow Management Systems. In: 3rd Intl. Conf. on Co-operative Information Systems, May 1995
- 9 Voelker G M, Bershad B N. Mobisaic: An Information Management System for Mobile Wireless Computing Environment. In: T. Imienlinski, H. Korth, eds. Mobile Computing, Kluwer Academic Publishers, 1996. 375~395
- 10 Madria S K, Mohania M, Roddick J. A Query Processing Model for Mobile Computing Using Concept Hierarchies and Summary Databases. In: proc. of 5th Intl. Conf. on Foundation for Data Organization (FODO'98), Japan, Nov. 1998

(上接第33页)

参考文献

- 1 Sbim S S Y, Pendyala V S, et al. Business-to-Business E-Commerce Framework. Computer, 001-9162, October 2000
- 2 Business Process Management Initiative. Business Process Modeling Language Specification, March 2001. <http://www.bpml.org>
- 3 Korok A. Introduction to XML and EDI. Aug. 1999. <http://www.xml.com>
- 4 Martin R. Building E-Commerce with XML. Beijing Hope Electronic Press, 2001/4/12
- 5 World Wide Web Consortium. XML Schema Specification, 2001. <http://www.w3c.org>
- 6 Open Buying On the Internet. OBI Technical Specifications Release V2. 0. <http://www.openbuy.org>
- 7 CommerceNet. The eCo Specification. Sept. 1999. <http://eco.commerce.net/specs/index.chm>
- 8 RosettaNet. RosettaNet Specifications, Dictionaries, Implement Framework, Partner Interface Protocol. 2001. [tanet.org](http://www.roset-

</div>
<div data-bbox=)

- 9 Aruba Inc. cXML Language Specification. Jun. 2000. <http://www.cxml.org>
- 10 Microsoft Inc. BizTalk Framework Specification. 2001. <http://www.microsoft.com/biztalk/production/doc/framework20.htm>
- 11 UN/CEFACT and OASIS. ebXML Technical Architecture Specification v1. 04. Feb. 2001. <http://www.ebxml.org/specdrafts/ebxml-ta-v1.0.1.pdf>
- 12 UDDI Org. UDDI Specifications. 2001. <http://www.uddi.org/specification.htm>
- 13 Microsoft Inc. Simple Object Access Protocol Specification. 2001. <http://msdn.microsoft.com/general/soapspec.asp>
- 14 Commerce One Inc. xCBL 3. 0 Release 2. 2001. <http://www.xcbl.org>
- 15 Koushik S. Understanding B2B e-Business Solutions. 2001. <http://www-3.ibm.com/e-business/index.jsp>
- 16 cnXML Org. Chinese B2B transaction Language Specification V0. 83. 2001. <http://www.cnxml.org.cn/publish/content.asp?id=19&cat=A23>