

【软考达人】

软考资料免费获取

- 1、最新软考题库
- 2、软考备考资料
- 3、考前压轴题



微信扫一扫，立马获取



6W+免费题库



免费备考资料

PC版题库：ruankaodaren.com

论分布式数据库的设计与实现

摘要：

分布式数据库系统把应用所需的数据存放在多个数据库服务器上，完成某个数据操作要涉及到访问多个服务器，这适用于某种特定需要的应用。本文描述了我在主持设计开发的一个MIS系统中，为了达到了在低速网络通道下有效提高应用程序性能的目的，使用Sybase分布式数据库技术的过程。系统采用典型的C/S结构，但许多客户端连接服务器的网络采用电话线拨号，速度有限，传统Windows界面的客户端应用程序相应速度比较慢。考虑到B/S结构也避免不了大量数据从服务器端传输到客户端，我认为WEB界面并不能有效解决这个问题，所以采用了优化数据库结构的方法，把数据分两部分存放，基础数据放客户机，会员资料主要采用键码放服务器，应用程序再现数据时从服务器取键码，到客户机取对应的解释，由于键码的数据量少，网络传输便快。在构建这个分布式数据库系统的过程中，我着重研究并解决了数据同步和事务协调的问题，取得了良好的应用效果。我认为，分布式数据库系统的技术在Internet时代正当其道，大有发展前景。

正文：

分布式数据库系统把数据存放在多个数据库服务器上，当应用提取所需数据时，要访问多个服务器，综合多点数据才能完成。分布式数据库技术在很多场合得到了应用。譬如某企业随着业务量的扩大，原有数据库服务器已经达到了容量和性能极限，如果不希望丢弃原有投资，可以建立另外一套新的数据库，跟原有的系统组成一个分布式数据库系统，给应用提供透明统一的数据访问。还有，如果某企业分成多个业务部门，而且地域分散，可以在某个部门放置单独的数据库服务器，用于存放该部门最常用的数据，而部门和部门之间相互引用的数据可以通过分布式数据库技术来方便地完成。分布式数据库不是简单地把集中数据库分散实现，而是针对某种特定应用需要而诞生，它必然具有自己特有的性质和特征，需要在上面做许多的工作，来满足应用的要求。我在设计、开发一个MIS系统时，针对应用的需要而引入分布式数据库技术，取得了良好的效果。该系统针对会员资料的管理而设计，用于管理会员入会、缴纳会费、申请资助、办理资助审批、关系转移、退会和注销手续等等业务流程。分三个级别的应用权限——基层单位级、总公司级和集团公司级，各个级别只能操作各自范围内的业务数据。该系统采用典型的C/S结构，后台数据库采用Sybase，前端应用采用PB开发工具来设计标准的Windows操作界面。我在其中任系统分析和数据库设计的角色，担任了调查业务需求、业务建模和数据库建模、数据库设计以及指导应用程序测试、优化系统和应用的性能等一系列工作。由于客户端地域的分散，遍及多个省境内，许多使用该系统的基层单位连接服务器数据库的网络采用电话线拨号方式，速度有限，在使用客户端应用程序时感觉界面速度很慢。经过分析，认识到许多操作都要从服务器中取数据，速度慢就慢在数据访问上。服务器是没有性能瓶颈的，问题出在网络速度上。不可能要求众多使用客户改善和升级他们的网络，只能充分挖掘软件的潜力，来适应这种低速网络的使用模式。经探讨，结合关系数据库的知识，认识到，应用程序的每一次数据库操作，都要访问多个相关联的表，其中，有会员资料表和基础数据表，会员资料表中存放许多的键码值，在基础数据表中有键码相应的解释。键码值的数据量比较少，而基础数据是静态的，几乎不会更改。如果考虑把会员资料放在服务器上，基础数据放在客户端，当应用程序中访问数据时，总是从服务器上存取会员资料，从客户端提取会员资料中键码的相应解释。由于键码的数据量少，

便减少了网络上传递的数据量，从而提高了界面的响应速度。同时考虑到基层单位总是操作自己所属的部分会员，增删转移操作少，会员列表比较固定，而每一项业务操作都涉及到要从会员列表中查找定位到某个会员，所以会员列表是最常访问的数据项。把会员列表从会员资料数据中抽取出来，也放置在客户端，这样，便进一步改善了性能。把数据分散存放只是工作的第一步，接下来要考虑应用程序怎样访问这种分布式数据。开发应用时，如果每一功能都针对两个数据库进行，就带来了很多麻烦。所以，我们研究了 Sybase 的分布式数据库技术，决定采用了 CIS (组件集成服务) 部件，来合并两个数据库成一个统一的分布式数据库。应用程序只要连接一个数据库，就可以透明统一访问到两个数据库中的数据。该技术具体实施方法是，在客户端数据库中建立一个对服务器数据库的远程访问服务名，包含访问地址、登录用户名、登录密码等等关键的连接信息；并且对服务器中会员资料数据表建立一个本地代理表，结构和服务器中远程表完全一样，它是访问服务器中会员资料的中转和代理。客户端应用程序访问本地代理会员资料表时，实际上是通过预定义的远程访问服务名中包含的连接信息到服务器中对应的实际会员资料表中访问数据。这种访问对于客户端完全透明，感觉不到是从物理上独立的两个服务器中存取数据。所以，这种数据库结构是典型的分布式数据库。部署这种分布式数据库不是坏事，只要在客户端和服务器上安装 12.0 版以上的数据库服务器，在客户端服务器上建立远程服务名和代理表即可。由于 Sybase 数据库的安装支持脚本方式，在客户端应用程序的标准安装过程中，嵌入 Sybase 数据库的安装和配置脚本，就自动化地完成了所有工作。在实际使用该分布式数据库系统的过程中，遇到了几个问题。第一，数据同步。客户端基础数据不是绝对静态的，也有变化，因此在服务器端要设置一个统一的基准，称为主点数据。客户端总是复制使用，称为复制点数据。如何及时感知到服务器端主点数据的变化，有效率地复制到客户端，是个难题。Sybase 针对这种应用场合，提供了复制服务器技术，但为了避免过于复杂，我们实际采用应用程序来管理同步。当服务器端主点数据有了更改时，保存一个相应的标识和时间戳，客户端应用在登录服务器时，检查这种标识，一检测到了数据有更新，就首先下载，然后再进入系统正常使用。这种方法实现起来，增加了额外的开发量，且不能判别绕过应用程序对数据的直接修改，但是，是最简单和有效的方法。第二个问题是事务协调问题。物理上独立的两个数据库，在协同操作时，如果服务器正好停机或者网络故障，完整的一个事务没能完成，就会“事务崩溃”。虽然 Sybase CIS 内嵌了两阶段提交技术，能够自动恢复。但是应用程序在这种情况下，敏感性不够，操作界面会无端凝固，影响了使用的方便性。我们针对 PB 对于连接的判断和感知，用了一个小小编程技巧，使应用程序能够及时感知到数据库连接故障，及时停止和恢复事务，使操作界面表现友好灵活。以上遇到的这些问题，都找到了解决办法。分布式数据库技术的应用并不是非常复杂，它往往为解决特定问题、满足特定需要而被采纳，使用得当，会给应用带来了许多便捷。在当今信息社会里，互联网络带来了相互连通的便捷，而且知识爆炸，数据的分布式访问是个必然趋势。潮流兴起的 XML 技术，提供了各种平台数据库之间的一个公共数据访问标准，可能会用来构建更加灵活、适应性更强的分布式数据库技术。