

**本科生毕业论文(设计)**

**外文翻译及原稿**

|  |  |
| --- | --- |
| 译文题目： | 使用Web服务数据实现与数据标记的数据同步 |
| 原稿出处： | Implementation of Data Synchronization with Data Marker using Web Service Data |

**学 院： 计算机科学与技术学院**

**专 业： 软件工程（中外合作办学）**

**班 级： 2019软件工程（中外合作办学）03**

**学 号： 201906150319**

**学生姓名： 卫晨超**

**指导老师： 李小薪**

**提交日期： 2023年3月**

**使用 Web 服务数据实现与数据标记的数据同步 (外文翻译之译文题目)**

**摘要：**随着越来越多的人拥有不止一台设备或智能手机，如果他们使用不同的设备，必须一遍又一遍地输入相同的数据，每个用户都会感到不方便。本文的研究目标是实现一种同步算法，能够通过 web service 维护不同数据库拥有的数据的一致性，使数据始终是最新的。在本研究的实现中，将描述通过 web service 在不同设备上单向同步和双向同步的数据同步算法的区别，使用户在不同设备上仍然可以使用相同的数据。该实现还添加了额外的信息，称为标记，以决定数据是需要忽略还是同步到另一个数据库。

**关键词：** 数据同步; 数据库;web 服务; 移动

# 1背景

在每个设备上存储数据有几种方法，其中之一是使用云计算技术，通过使用这种技术，用户可以通过使用互联网连接来访问数据，从而将所有数据存储在云中。但如果用户没有互联网连接或处于离线状态，这将是一个问题。用户不能访问云中的数据库，也不能操作云中的数据。

Isak Shabani, Betim Çiço 和 Agni[1]提出了一种数据同步算法，该算法允许软件应用程序在离线或在线配置下正常工作。这是在普里什蒂纳大学 (UP) 的电子学生管理系统 (ESMS) 中通过 web 服务与适当的模块实现的。

此外，在Sudha、Brindha、Sai Vamsy Krishna、Gokul和Sanath Kumar[2]进行的研究中，他们提到云计算通常由前端用户设备和后端云服务器组成，这为用户提供了对云上大量存储的访问。文件可以从移动设备或PC上传到云存储，当用户连接到互联网时，这些文件将自动同步到用户的设备上，这一范例在Windows平台上实现。

Naveen Malhotra和Anjali Chaudhary[3]进行了进一步研究，作者对数据库同步技术进行了说明。该技术分为单向同步和双向同步两种方式。单向同步用于将驻留在目标数据库中的所有数据替换为驻留在数据库服务器中的数据。而双向同步用于将驻留在目标数据库中的数据与数据库服务器结合起来，以便保持最新。

在这项研究中，我们开发了Naveen Malhotra和Anjali Chaudhary通过web服务实现的同步概念，并添加了额外的信息作为标记，以防止从未被操作的数据被递归同步，并能够将数据保存在所有设备中。因此，当设备未连接到服务器时，用户始终可以查看数据。

下面的场景将说明数据同步的重要性。旅行者有智能手机;在路上，他使用该应用程序在智能手机中存储他去过的地方的数据。该应用程序使用了一个只能通过Internet访问的云数据库。有一天，旅行者没有网络连接，导致他无法上传近照。因为应用程序使用的旅行者不使用本地数据库，只使用云数据库，因此旅行者不能使用驻留在智能手机上的数据。

该场景说明，如果用户设备处于离线状态或没有使用Internet连接，则用户无法将数据发送到云中的数据库服务器。因为应该发生的情况是，当用户处于脱机状态或没有Internet连接时，应用程序应该使用本地数据库以及支持数据存储。能够同时使用本地数据库和服务器云数据库是非常重要的，因为使用这些数据库可以帮助用户在数据存储方面更加高效。

为了克服这一问题，采用客户端-服务器数据同步算法是解决这一问题的方法之一。数据同步是使每个数据库中的数据保持最新的过程。通过使用这种算法，尽管用户处于脱机状态，但仍然可以访问数据，因为所有用户数据都存储在与云中的数据库服务器集成的本地数据库中。本研究的目的是通过web服务为不同平台的数据同步提供新的结果，它允许软件应用程序在线和离线执行。

# 2数据同步

Aashima和AnitKaur[4]认为，在计算机科学中，数据同步是指从源数据到目标数据存储之间建立数据一致性的过程，反之亦然，包括随着时间的推移数据的持续协调。同步指的是两个截然不同但相关的概念之一，即数据同步和进程同步。为了达成协议或承诺一定的行动顺序，流程同步指的是多个流程在某一点上合并或握手。同时，数据同步是指保持数据完整性或保持数据集的多个副本彼此一致的思想。数据同步通常使用进程同步原语实现。

Naveen Malhotra和Anjali Chaudhary进行了一项研究，旨在提供一种算法来解决当所有客户端依赖于单个服务器时发生的问题。当服务器不可用时，由于服务器故障或计划的服务器停机，所有用户(远程工作者)将与他们的数据断开连接，远程工作者使用的所有数据将存储在他们的本地系统中。当用户连接到互联网时，数据将自动从他们的本地系统(客户端)传输到服务器。文件处理系统也采用相同的行为，当网络不可用时，用户上传的所有文件都会保存在客户端机器的文件夹中，当网络可用时，所有文件都会自动从客户端传输到服务器。自动传输过程完成后，本地存储的数据和文件都会被删除。该研究留下了一些问题，其中之一是如果由于多个实体中的数据或文件更改而导致数据冲突，则必须手动进行协调。

根据Naveen Malhotra和Anjali Chaudhary的说法，数据同步的一个明显不同(但相关)的概念是需要保持一组数据的多个副本，这些副本彼此一致。数据同步是在数据之间建立一致性的过程，从源数据存储到目标数据存储，反之亦然，并且随着时间的推移不断协调数据。例如，用户在一台移动设备上的联系人列表可以与其他移动设备或计算机同步。该技术旨在同步两个或多个设备之间的一组数据，并自动将更改从一个复制到另一个，反之亦然。

当一个团队的人在偏远的地方工作，不能访问中央数据库时，这些工作人员通常需要彼此共享信息。远程数据库可以通过某种点对点网络自由地与其他数据库交换信息，这种解决方案是有用的。根据需要，缓存计算机可以修改原始数据版本，该过程在分布式系统中实现。数据同步是通过专门的软件来实现的，该软件跟踪在多台计算机或系统之间创建和使用数据单元的数据版本。

# 3 Web服务

根据Anil Dudhe和S.S.Sherekar博士[5]的研究，web服务是一种使用不同技术开发的计算机程序，如XML、RESTfulAPI存储在一些服务器上，这些服务器可以使用不同的协议访问，如。Web服务与平台无关，通常基于文本，可以在异构技术上开发、运行和访问。所以它们是可互操作的。

过HTTP协议实现分布式超媒体系统的一种新的web服务架构。为了关注组件的角色，它们与其他组件交互的约束，以及它们对重要数据元素的解释[6];REST忽略了组件实现和协议语法的细节。它包含了对组件、连接器和数据的基本约束，这些组件、连接器和数据定义了Web体系结构的基础，从而包含了其作为基于网络的应用程序的本质行为。

基于Anil Dudhe和S.S.Sherekar博士的研究，在面向服务架构的帮助下，应用程序可以在万维网上使用web服务。服务提供特定的功能，就像任何业务功能一样，例如分析业务历史、销售和采购或天气预报信息等。SOA并不是一个新概念，已经使用了很多年。因为它在应用程序中提供了松耦合，这是它的美丽之处。客户端不需要知道服务是如何编码的，也不需要知道使用什么平台来运行服务。

# 4 实现

根据前面的解释，实现同步的一种可能方法是使用Web服务。在同步过程中，Web服务被用作客户端和服务器之间的辅助器。我演示了一个系统设计的框架，web服务控件作为智能手机、个人计算机和数据库服务器之间的中介。

单向同步过程阶段示例: 1. 管理员在服务器中插入’foo bar’ 2. 用户同步本地数据库。 3.然后，当同步过程执行时，' foo bar’数据将进入本地数据库。

双向同步过程阶段示例:1.用户在凌晨02:00在设备A中更新'foo bar'。2.然后，用户在凌晨03:00在设备B中再次更新'foo bar'。3.当同步过程执行时，输入到服务器数据库的数据是'foobar'的设备B为设备B中最近更新的数据 4. 所以，用户将从设备B获得' foo bar '设备A中最近更新的数据。

前面的同步过程示例可能存在问题，因为输入到数据库服务器的最新数据不一定是最新的数据。为了解决这些问题，它需要一个标记，标记作为新数据的基准或标志。该标记将在创建新数据或操作数据后同时写入。本研究中使用的标记是时间戳。因此，它使用“时间”作为最新数据的行列式。时间戳用于保存输入数据的时间。所以，时间戳最大的数据就是最新的数据。此外，还使用时间戳来避免重复的数据请求。系统将只请求上次同步后更新过的数据。

时间戳将分为3种类型: 1. 上次同步:用户最近一次同步的时间 2. 创建时间:获取数据的创建时间 3.更新时间:获取到的数据更新的时间。

# 5 总结

在本文中，我们研究并提出了使用数据同步的重要性。当数据只传输发生变化的数据时，数据同步将有效运行。因此，每个同步进程都需要一个标记来确定最新的数据。我们使用时间戳作为标记，表示具有最大时间戳的数据是最新的数据。在我们的研究中，我们研究了单向数据同步和双向数据同步的区别。单向数据同步用于仅在一个源或服务器中发生数据更改。以及在每个设备中的每个数据库中发生数据更改时执行的双向数据同步。我们提供了15分钟的时差容差来解决数据差异。同时，我们使用GMT+0作为基准时间，忽略了每个区域的时差。我们使用同步来消除用户拥有的每个设备上的数据差异。通过使用数据同步，用户无论使用不同的设备都可以使用相同的数据。

References

1. Shabani, I., Cico, B., Chaudhary, A. Solving Problems in Software Applications through Data Synchronization in Case of Absence of the Network.International Journal of Computer Science Issues; 2014.
2. Senthilkumar,S., Karthikeyan,B., Krishna, S. V., Gokul K, Kumar, S.Data Synchronization Using Cloud Storage.International Journal of Advanced Research inComputer Science and Software Engineering;2012.
3. Malhotra, N., Chaudhary, A. Implementation of Database Synchronization Technique between Client and Server. International Journal of Engineering Science and Innovative Technology; 2014.
4. Kaur, A., Aashima. Synchronized Algorithm For Database And Image Processing Between Client And Server. International Journal of Computer Science and Information Technologies; 2014.
5. VDudhe, A., & S.S. Sherekar, P. D. Performance Analysis of SOAP and RESTful Mobile Web. International Journal of Computer Applications, 1; 2014.
6. R. Fielding. Architectural Styles and the Design of Network-based Software Architectures, PhD thesis, University of California, Irvine, USA; 2000