Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 59 (2015) 366 – 372

Procedia Computer Science

International Conference on Computer Science and Computational Intelligence (ICCSCI 2015)

Implementation of Data Synchronization with Data Marker using Web Service Data

Muhsin Shodiq*, Rini Wongso, Rendy Setya Pratama, Eko Rhenardo, Kevin

Bina Nusantara University, Jl. K.H. Syahdan No.9 Kemanggisan, Palmerah Jakarta Barat 11480, Indonesia

Abstract

With the increasing number of people who have more than one device or smartphone, every user feel inconvenience if they have to enter the same data over and over again if they use a different device. The research objective is to implement an synchronization algorithm which is able to maintain consistency of data owned by different databases through web service, so that data is always up to date. In this study, the implementation will describe the difference of data synchronization algorithm between unidirectional synchronization with bidirectional synchronization in different device through Web Service so users can still use the same data although using a different device. The implementation also adding additional information called as a marker to decide whether the data is need to ignore or synchronized to another database.

Peer-review under responsibility of organizing committee of the International Conference on Computer Science and Computational Intelligence (ICCSCI 2015)

*Keywords:*data synchronization; database; web service; mobile

1. Background

There are several ways to store data residing on each device, one of them is the use of cloud computing technology, and by using this technology the user can store all data in the cloud by using an internet connection to access it. But it will be a problem if the user does not have an Internet connection in it or are in the offline condition. Users cannot access the database to the cloud and cannot manipulate the data that is in the cloud.

* Corresponding author. Tel.: +62817 678 4564; fax: +6221 534 5830.
E-mail address: muhsinshodiq@gmail.com

Peer-review under responsibility of organizing committee of the International Conference on Computer Science and Computational Intelligence (ICCSCI 2015)
doi:10.1016/j.procs.2015.07.538

Isak Shabani, Betim Çiço and Agni Dika[1]had presented an algorithm for data synchronization, which allows software applications to work well on offline or online configurations. This is implemented in Electronic Student Management System (ESMS) at the University of Prishtina (UP) through web service with the appropriate module.

Furthermore, in the research conducted by Sudha S, Brindha K, Sai Vamsy Krishna S, Gokul K and Sanath Kumar M[2], they mentioned that Cloud computing usually consists of front-end user devices and back-end cloud servers, which gives access to a large volume of storage on the cloud for users. Files can be uploaded from mobile devices or PC to the cloud storage, which will be automatically synchronized to the user's devices when they are connected to the Internetand this paradigm is implemented in Windows Platform.

Further research done by Naveen Malhotra and Anjali Chaudhary[3], authors had explained the technique of database synchronization. This technique is divided into two ways, unidirectional synchronization and Bidirectional synchronization. Unidirectional synchronization is used to replace all the data residing in the destination database with the data residing on the database server. While the Bidirectional synchronization is used to combine data residing on the destination database with the database server in order to stay up to date.

In this study, we developed the concept of synchronization has been done by Naveen Malhotra and Anjali Chaudhary[3] through web service with additional information called as marker to prevent the data that never manipulated is sync recursively with capability to keep data in all devices.Therefore, the user always can view the data when the device is not connected to the server.

The following scenario will illustrate how important the data synchronization. A traveler has a smartphone; on his way he uses the application to store data where he had visited in the smartphone. The application uses a cloud database that can only be accessed via the Internet. One day, the traveler does not have Internet access which resulted in him not be able to upload a recent photo. Because the application used a traveler that does not use a local database but only uses cloud database, so that the traveler cannot use the data residing on the smartphone.

The scenario illustrates that if a user device in the offline condition or not using an Internet connection, the user cannot send data to the database server that is in the cloud. Because the conditions that should happen is that the application should use the local database as well as to support data storage when the user is in the offline or do not have an Internet connection. Its' very important to be able to use local databases and server cloud databases at the same time, due to the use of these databases can assist the user in data storage to be more efficient.

To overcome that issue, one of solution of this problem is to use data synchronization algorithm client - server. Data synchronization is process to make data up to date in each database. By using this algorithm, the user can still access the data in spite of being in the offline because all user data is stored in a local database that is integrated with a database server that is in the cloud. The aim of this research is to provide new results for data synchronization in different platforms through web services, which allow software applications to be executed bothonline and offline.

1.1. Data Synchronization

According to Aashima and Anit Kaur[4], data synchronization in computer science, is the process of establishing consistency among data from a source to target data storage and vice versa, including the continuous harmonization of the data over time. Synchronization refers to one of two distinct but related concepts, which are synchronization of data and synchronization of processes. In order to reach an agreement or commit to a certain sequence of action, process synchronization refers to the idea that multiple processes are to join up or handshake at a certain point. Meanwhile, data synchronization refers to the idea of maintaining data integrity or keeping multiple copies of a dataset in coherence with one another. Data synchronization is commonly implemented using process synchronization primitives.

Naveen Malhotra and Anjali Chaudhary[3]conducted a research, which aims to provide an algorithm in order to solve the problem, which occurs when all clients rely to a single server. When the server is unavailable, due to server failure or planned server downtime, all users – remote workers -- will be disconnected from their data, and all the data used by the remote workers will be stored in their local system. Data will then be automatically transferred from their local system (client) to the server when user connects to the Internet. The file handling system use the same behavior as well, all the files uploaded by users will be saved on the client machine folder when Internet is not available and all files will be automatically transferred from client to server when Internet connection is available. After the automatic transfer process is completed, both of data and file in the local storage are always deleted. The

research left some issues, one of them is if there is a data conflict due to data or file changes in more than one entity, it must be reconciled manually.

According to Naveen Malhotra and Anjali Chaudhary[3], a distinctly different (but related) concept of data synchronization is the need to keep multiple copies of a set of data which coherent with one another. Data synchronization is the process of establishing consistency among data from a source to target data storage and vice versa, and the continuous harmonization of the data over time. For example, a user's contact list on one mobile device can be synchronized with other mobile devices or computers. The technology is designed to synchronize a single set of data between two or more devices and automatically copying changes one to another and vice versa.

When a team of people working in remote locations and do not have access to a central database, these workers often need to share information amongst each other. The solution in which a remote database is free to exchange information with other databases through some sort of peer-to-peer network can be useful. Depending on requirements, each computer may modify original data versions and the process is implemented in distributed system. Data synchronization is enabled through specialized software that tracks data version where data elements are created and utilized between several computers or systems.

1.2. Web Service

Based on research by Anil Dudhe and S.S. Sherekar, Ph. D[5], web services is a computer program developed using difference technologies like, XML, RESTful API stored on some server which can accessed using different protocols like. Web services are platform neutral and generally text based which can developed, run and accessed on heterogeneous technologies. So they are interoperable.

Based on Roy Fielding dissertation about the Representational State Transfer (REST)[6], its is an new version of the web service architecture which can be distributed hypermedia system via HTTP protocol. In order to focus on the roles of components, the constraints upon their interaction with other components, and their interpretation of significant data elements; REST ignores the details of component implementation and protocol syntax. It encompasses the fundamental constraints upon components, connectors, and data that define the basis of the Web architecture, and thus the essence of its behavior as a network-based application.

Based on research by Anil Dudhe and S.S. Sherekar, Ph. D[5], with the help of service-oriented architecture applications makes use of web services available on the World Wide Web. A service provides a particular function, like any business function, such as analyzing a business history, sales and purchase, or weather forecast information etc. SOA is not a new concept, have been used for years. Because it provides loose coupling in the application is its beauty. Clients do not need to know how the service is encoded and what is the platform is used to run the service.

2. Implementation

Based on the explanation previously, one possible method to implementing synchronization is using Web Service. Web service is used as a facilitator between the client and server in the synchronization process.Fig. 1 illustrates the framework of a system designed, web service control as an intermediary between smartphones, personal computers, and the database server.

Fig. 1. System design

Examples of the unidirectional synchronization process stages:

1. Administrator insert ‘foo bar’ in server.
2. Users synchronizethe local database.
3. And then, when the synchronization process is executed, the ‘foo bar’ data will entered into local database.

Examples of the bidirectional synchronization process stages:

1. User update ‘foo bar’ in device A at 02:00 AM.
2. And then, user update ‘foo bar’ again but in device B at 03:00 AM.
3. When the synchronization process is executed, the data are entered into server database is ‘foo bar’ of device B as the data last updated in device B
4. So, user will get ‘foo bar’ last updated data in device A from device B.

Example of the synchronization process previously can be problematic, because the latest data entered into the database server is not necessarily the most recent data. To solve these problems, it takes a marker, marker serves as a benchmark or a sign of a new data. The marker will be written simultaneously when new data is created or after the data is manipulated. Marker used in this research is the timestamp. So it is used 'time' as a determinant of the most up to date data. Timestamp is used to save the time of the input data. So, the data that have the greatest timestamp is the latest data.In addition, the timestamp is also used to avoid repetitive data requests. Where the system will only ask for data that has been updated after the last synchronization.

Timestamp will be divided into 3 types:

1. Last time sync: Last time user did synchronize
2. Created At: The obtained time when the data was created
3. Updated At: The obtained time when the data was updated

Fig. 2 will show the bidirectional synchronization flow and Fig. 3 will show the unidirectional synchronization flow with timestamp as a marker.

4. Conclusion

In this paper we have studied and presented the importance of using data synchronization. Data synchronization will run efficiently when data is transmitted only data that has changed. Therefore, each synchronization process takes a marker to determine the latest data. We use the timestamp, as a marker indicating that the data has the greatest timestamp is the most recent data. In our study, we have investigated the difference between unidirectional data synchronization and bidirectional data synchronization. Unidirectional data synchronization used when the data changes only occurred in one source or in server. And the bidirectional data synchronization executed when the data changes occur in each database that resides in each device. We provide tolerance of the time difference of 15 minutes to resolve the data discrepancy. And also we ignore the time difference on each area by using GMT+0 as a benchmark time. We used synchronization to eliminating differences in data on each device owned by the user. By using data synchronization, the user can always use the same data despite using different devices.

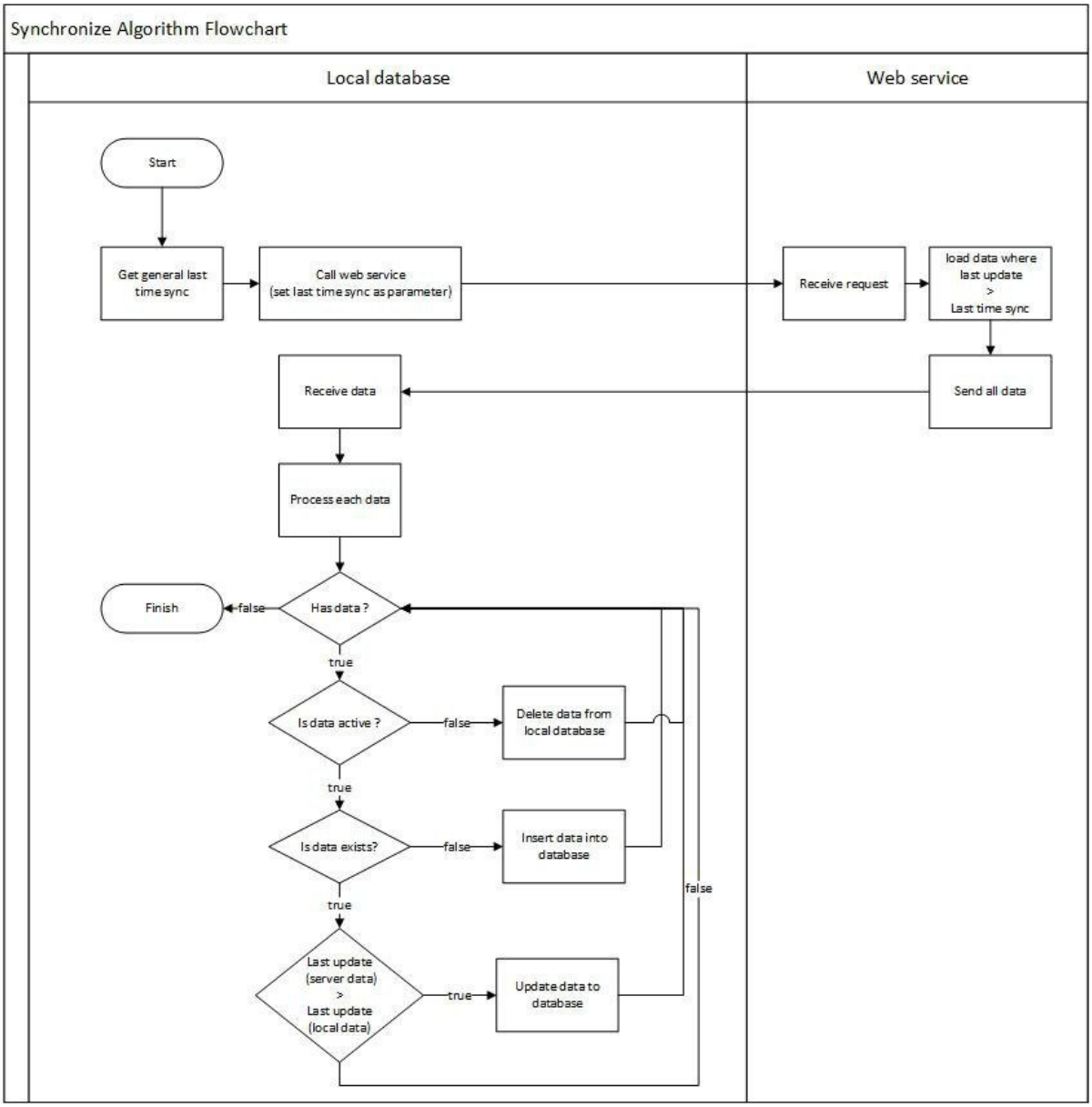

Fig. 2. Bidirectional synchronization

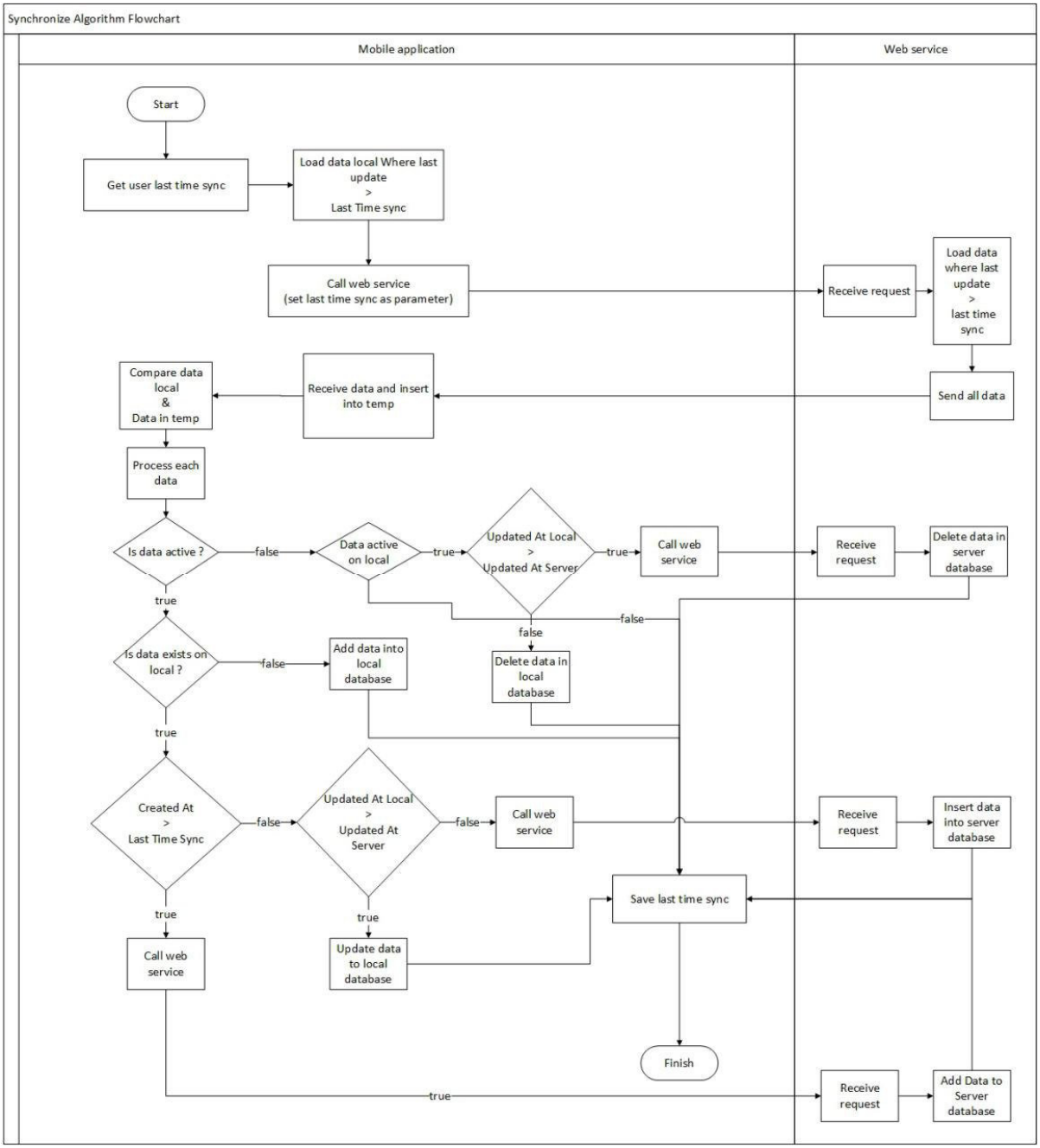

Fig. 3. Unidirectional synchronization

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