Research Article

A Multimodality Information Synchronization Scheme for a Multisource Information System in the Electric Grid

Pan Zhang,¹ Fenggang Lai,¹ Jing Du,¹ Wenda Lu,² and Xiao Yu³

¹State Grid Information & Telecommunication Co., Ltd, Beijing 100761, China

²Information & Telecommunication Branch, State Grid Zhejiang Electric Power Corporation, Hangzhou 310000, China

³Department of Computer Science and Technology, Shandong University of Technology, Zibo 255022, China

Correspondence should be addressed to Xiao Yu; yuxiao8907118@163.com

Received 5 February 2021; Accepted 23 April 2021; Published 6 May 2021

Academic Editor: Zhe-Li Liu

Copyright © 2021 Pan Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A multisource and heterogeneous database is an important problem that disturbs the use of the electric power information system. The existing database synchronization scheme has some problems in practical applications, such as high resource loss and poor portability. This paper presents a high-efficiency database synchronization scheme for the electric power information system. The database is monitored, and its changes are captured by the shadow table and trigger method. Thus, data could be exchanged in trusted networks and nontrusted networks. In addition, a predetermined strategy is used to avoid data conflicts and ensure consistency and reliability of data synchronization. The above method is applied in the protection system of power networks. The results show that the synchronization scheme can effectively ensure the security of the system and has higher synchronization efficiency.

1. Introduction

The scale of the data network of the electric power information system increases rapidly with the development of the smart grid. Data scheduling is filled with dispatch centers, power plants, users, and so on. The security problems need to be researched and solved for services, especially for data security issues that are important and restrict the application in the electric power information system.

The telemechanical system controls the power station and substation through the telemetry data of the system [1–7]. Its security and reliability directly affect the safety and reliability of the whole power control system. Because the telemechanical system is interconnected with the local MIS (management information system) or connected to the internet, the power system design and construction, and lack of data network protection, the telemechanical system is easy to be attacked.

Network attacks occur frequently. Trojan horses, worms, and ransomware emerge one after another on the internet, which pose a serious threat to the safety in the production of the smart grid. Therefore, just like government networks,

military networks, and other classified networks, the security of the power information system, especially the production-related real-time data networks and dispatching data networks, should carefully formulate targeted strategies and set up a strong security guarantee system [3].

The traditional data backup method is to ensure the data security by the database maintainers under the premise of a continuously open network. However, in some networks where the degree of secrecy is relatively high, database synchronization should be implemented in a secure and isolated environment. Therefore, the new technique of database synchronization between isolated networks is worth further study.

The general network isolation scheme is implemented by installing a network isolation device between trusted intranet and nontrusted internet [8–19]. The principle is based on the idea of access control and physical isolation and defines relevant constraints and rules to guarantee the security strength of the network.

Our inspiration comes from the object-oriented database system and distributed client-server mechanism. On the

basis of the isolated network, we put forward a new efficient data synchronization scheme. In the scheme, the shadow table and trigger method are used to capture the data changes in the database. In this case, data between intranet and internet are synchronized without any disruption to system security. We implement two synchronous techniques to evaluate the scheme in the protection system of the information system of the smart grid. Experiment results show that our method could ensure system safety and achieve excellent performance.

The rest of the paper is organized as follows. We describe the overall framework of the isolation system in Section 2. Section 3 describes database synchronization in network isolation. We discuss our implementation in Section 4 and analyze numerical results in Section 5. Finally, we conclude our paper in Section 6.

2. Overall Structure of Network Isolation

Network isolation is a physical isolation method by a special isolation device between inner intranet and outer internet [19–28]. Network isolation technology ensures that the internal information of the trusted network is not leaked and uses shared storage to complete the safe exchange of data between networks [20–23].

In general, isolation devices are not connected to either trusted intranet or nontrusted internet. If there are requests for information exchange between them, the isolation device tries to connect to one of the two networks. Figure 1 shows the isolation structure used in the electric power information system.

The network isolation system abandons common network protocols such as TCP/IP and adopts a new type of proprietary security protocol to exchange data. The system blocks the TCP/IP connection, makes the internal network and external network completely lose the connection, and completely eliminates the hidden danger of the TCP/IP network attack. In addition, the system can effectively reduce the attack threat by using the vulnerability of the OS with the support of isolated hardware and software.

3. Principal of Database Synchronization

3.1. Description of Database Synchronization. Due to the physical isolation characteristics of the network isolation environment, the database distribution in two isolated networks includes the following three synchronization steps: capture changes of the source data, distribute the data, and update the data to the target database [23–28]. Figure 2 shows the structure of data synchronization. The process of data synchronization can be divided into several processes as follows: change capture, data distribution, network transmission, synchronization monitoring, and data update [29–32].

In data synchronization, the first step is to capture the data changes in the database. Thus, a combination of triggers and shadow tables is used. This approach uses an XML file as an intermediary to log changes' information such that every table of the source database corresponds to an XML file.

Data distribution adopts the server-client mode, in which each network device acts as a client to actively connect to the server. A monitor module that is continuously running is used to record the update of the source database and target database, respectively. As is mentioned in the graph, the data update module is used to update the data to the target database [33]. When the update operation is executed, the communication module starts to send (or receive) data according to the synchronous configuration policy. Thus, an entire synchronization procedure is completed.

3.2. Process of Database Synchronization. Data synchronization is the process of synchronizing data from the source database to the target database. Maintaining consistency of multiple copies of replicated objects and considering the synchronization efficiency can effectively reduce network overhead and shorten response time, thus improving the availability and reliability of the whole system. The synchronization approach takes the following three steps.

3.2.1. Change Capture. Change capture is the basis of synchronization, and it directly determines how the database is updated and how time is selected for synchronization. Changing sequence information is essential when synchronizing the target database. In addition, a large amount of control information needs to be synchronized as well [1, 34]. Due to different characteristics of the trigger method and shadow table method, the combined method of trigger and shadow table is adopted to ensure the performance of the source database and minimize the impact on the system operation.

- (i) Trigger method: this method is especially suitable for a large amount of data and often needs incremental synchronization. Trace triggers are created for data operations such as Insert, Delete, Modify, and Update in the source data table. When one of the above operations occurs on the source table, newer field data, action type, and action sequence number are stored in the log table, which provides synchronous updates to the source table.
- (ii) Shadow table method: this method is generally used in scenes with a small amount of data and low requirement of the real-time performance. The advantage of this method is that it has little impact on the business system and easy to deploy. When the source table needs to synchronize, a supporting shadow table is created to record the change tracking table. After that, the shadow table and source table are compared to extract the changing information. In this case, the shadow table is synchronized.

In the above combined synchronization method, data distribution could obtain synchronization information by the source table and change tracking table. Thus, the work of change capture is integrated in distributed modules and then encapsulated into the database layer [2, 23].

FIGURE 1: Structure of the network isolation system.

FIGURE 2: Overall structure of database synchronization.

Change capture is the process of capturing the sequence of changes in the source table. Based on the configuring parameters of the synchronization mode, the system automatically creates the tracking table and shadow table for the source table's change capture. The main idea of data synchronization technology to capture changes is to create a change tracking table for multiple related source tables, perhaps a single source table, or all tables of an entire database. Source table's field information is recorded, such as the type of operation, the sequence number of operations, the operation time, and the change's key field information.

3.2.2. Monitoring the Synchronization. Because of the independence of the JDBC platform, it has become one of the most popular methods to access the database. In this case, JDBC is used to connect to the database in the synchronous monitoring module. If there are changes in the source database, the synchronization system will start the communication module according to the synchronization mode.

3.2.3. Data Distribution. The primary purpose of data distribution is to implement change information from the source table to the corresponding target table. Based on the above change capture method in synchronization, the data distribution module obtains the corresponding SQL statements in sequence according to the sequence number in the change tracking table. The dispatcher then executes the SQL statement on the target server and applies the changes in the source table to the target table. After successful execution, the record corresponding to the sequence number in the change tracking table is deleted. The dispatcher is a coordinator and is also responsible for passing control information and mediation if replication conflicts are found.

3.2.4. Data Update. Data updates occur on network target nodes. When the XML file is sent to the network node, the SQL statement is immediately extracted. If the instruction contains one of the Insert, Delete, or Update operations, the SQL statement is executed directly, and the target table is updated accordingly. In particular, if it is a Create operation, the target database will create a new synchronization table and initialize it.

3.3. Conflict Detection and Resolution. Replication operations incurred by data synchronization can cause inconsistencies and conflicts between different copies [35, 36]. It is necessary to determine the cause and location of the conflict and resolve the conflict in accordance with the pre-determined strategy. In addition, the granularity of conflict detection, record level, field level, and so on, will affect the performance of the system. Therefore, a control information table was set up in the prototype to help resolve the conflict. The structure of the control information list is shown in Figure 3.

As shown in Figure 3, all pieces of control information are organized into a list, which is indexed by metadata. Each information item consists of the conflicting data item, conflict position, operator, and timestamp. The main function of each of these information items is as follows:

- (1) Conflicting data item: it provides a change copy of the conflicting data items, which can reflect relationships between the changes of data items
- (2) Conflict position: it identifies the subject causing the conflict
- (3) Operator: it indicates the location that caused the conflict, which may contain multiple server nodes
- (4) Timestamp: it provides the exact time when the conflict occurred

FIGURE 3: Structure of the control information list.

4. Network Isolation Deployment in the Electric Power Information System

4.1. Network Isolation Device. In addition to using basic firewalls and proxy servers for security, the power information system also uses network isolators as gateways. The isolation device adopts isolation technology that the intranet and internet disconnect to the isolation device.

The isolated transmission mechanism uses dedicated hardware and security protocols to achieve data exchange between internal and external networks. The data monitoring module has powerful control and management function by security mechanisms, such as access control, identity authentication, and encryption.

4.2. Isolation Scheme in the Electric Power Information System. The whole topology of the telemechanical network and MIS network of the smart grid is shown in Figure 4. Figure 4 shows that the telemechanical system could connect directly to the dispatch center and production site without network isolation. Furthermore, as the middle system of the MIS and internet connection, once the telemechanical system is deliberately destroyed by a virus or hacker, the whole power network will face serious risks.

To guarantee the reliability of the telemechanical system and intranet, network isolation should be implemented between the telemechanical system and MIS networks [29], and the original industry network will be protected. The connection between internet and remote database will also be blocked. In the MIS network, an image database of the remote database is created to replace the original remote database to provide services to the system.

In the opposite case, the image database should be synchronously updated to the original telemechanical database. The telemechanical web system is deployed on intranet so that the telemechanical engineer can browse the telemechanical data. Figure 4 shows the working scene that the original telemechanical database is the source database, and the mirror database is the target database.

4.3. Deployment and Configuration of Synchronization. The synchronization system consists of the sending end and receiving end. Typically, the sending end is deployed on a remote database server, and the receiving end is deployed on a mirrored database server to save resources [37, 38].

In the initial state of the synchronization, most of the tables need to be synchronized first, and then the tables need to be synchronized selectively. Because there are too many

tables in the remote database to synchronize every table in real time, tables are categorized as follows:

- (1) Nonsynchronous tables: these tables are always used in the external network, such as parameter tables for data acquisition, evaluation parameter settings, and intermediate evaluation calculation tables.
- (2) Synchronous tables: these tables are always involved in synchronous operations. Depending on the size of these tables and how often they are updated, appropriate synchronization method needs to be selected.

5. Experiment Results and Analysis

5.1. Experiment Environment. For our experiments, we used Dell PowerEdge Server which has Intel Xeon processors running at 3.20 GHz and 32 GB RAM. The OS is Windows Server 2016, and the test DBMS includes MySQL, Oracle, and SQL Server. The inner and outer networks are separated by special electric power information system isolation devices. During the prototype system operation, unnecessary programs and processes are closed as much as possible to minimize resource usage.

5.2. Operational Approach. After the prototype was implemented, data replication function was tested. On the image database server, an image database is created to test the performance of synchronization. Next, a synchronous receiver program is run to monitor the network at port 9098 connecting the database successfully.

On the source database server, the receiver is always running. After synchronization table configuration is completed, a series of actions including change tracking table, change tracking trigger, shadow table, and incremental change analysis are performed over and over again.

In general, system synchronization will read update data from source tables incrementally on the basis of the pre-configuration and then update to external internet. All of the above processes need no manual interventions.

5.3. Performance Analysis. There are approximately 3.5 million records in the test database, each of which is about 100 bytes long. We test the time efficiency and space efficiency of the system by randomly selecting some of record items. Figures 5 and 6 show the experiment results, respectively.

Figure 5 shows that time efficiency of the data synchronization technique increases linearly with the increase of the record number. Generally, time consumption of every record is smaller when there are more records. On the contrary, when few records are recorded, more time is spent. A reason for this could be that data synchronization adopts the method of copying by reading control information afterwards, and it brings limitations on synchronization efficiency. In terms of space efficiency, as shown in Figure 6, it remains a significant constant that indicates the system has been stable and reliable.

FIGURE 4: Network isolation topology.

FIGURE 5: Test result of time efficiency.

FIGURE 6: Test result of space efficiency.

TABLE 1: Time efficiency and resource efficiency of initialization/real-time synchronization.

Work type	CPU utilization (%)	Memory usage (MB)	Response time (ms)
Initialization synchronization	50	170	0.75
Real-time synchronization	≤ 10	110	6.0

We suppose the isolation network environment is an insecure and vulnerable network that is easy to be attacked by viruses and hackers. When the external database system is damaged, the restoration of the external server can quickly complete the reconstruction of the mirror database and normal operation of synchronization.

When the system performs operations of initializing the target table, the speed of synchronization can reach 80,000 per minute. So, the target database could complete initialization in 1 hour. In the initialization process, the receiver also needs to update the target database constantly, so the CPU utilization rate is about 50% that is still relatively higher. Memory consumption is about 170 MB.

When the target table of the database has been initialized, the remaining operations are incremental updates. Since then, the speed of synchronization is about 10,000 records per minute. In this case, the CPU utilization is at a relatively low level, less than 10%, and the memory consumption is about 110 MB. The main working parameters of initialization synchronization and real-time synchronization are given in Table 1.

It is worth mentioning that the system has been stable after entering the working state. Data in the mirror database could be updated immediately, and end users can barely feel the corruption of the system. With the same method, the system can also easily recover the corrupted data of external internet.

6. Conclusion

The hybrid isolation synchronization scheme for the electric power information system realizes secure and high-performance data exchange between trust and nontrust networks. This scheme is not only suitable for the isomorphic databases but also suitable for heterogeneous databases. With the data exchange captured by the hybrid method, the synchronization scheme could record and query all of the synchronization data.

Data Availability

The data are not available because of the supported funding privacy policy.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was sponsored by State Grid Technology Project "Research and Application of Key Technologies in Enterprise-Level Data Center Platform based on Full Service Unified Data Center" (Grant no. 5211XT190033) and

Zhejiang Lab (2020LE0AB02). This work was supported in part by the National Natural Science Foundation of China (nos. 61876019 and 62072037).

References

- [1] H. Cui, J. Feng, W. Ma, Q. Zhang, K. Pang, and K. Yu, "Research and implementation of heterogeneous database synchronization technology based on isolation gateway," *Software Engineering*, vol. 19, no. 2, pp. 10–13, 2016.
- [2] S. P. Kumar, "Adaptive consistency protocols for replicated data in modern storage systems with a high degree of elasticity," in *Document and Text Processing*, Conservatoire National Desarts et Metiers (CNAM), Paris, France, 2016.
- [3] B. Kemme and G. Alonso, "A suite of database replication protocols based on group communication," in *Proceedings of the 18th International Conference on Distributed Computing Systems (ICDCS)*, pp. 156–163, Amsterdam, Netherlands, May 1998.
- [4] S. K. Wong and M. Y. Siu, "Location spoofing attack detection with pre-installed sensors in mobile devices," *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, vol. 11, no. 4, pp. 16–30, 2020.
- [5] Z. Guan, X. Liu, L. Wu et al., "Cross-lingual multi-keyword rank search with semantic extension over encrypted data," *Information Sciences*, vol. 514, pp. 523–540, 2020.
- [6] F. Yan, X. Yang, J. Liu, H. L. Tang, Y.-A. Tan, and Y. Z. Li, "Optimizing the restoration performance of deduplication systems through an energy-saving data layout," *Annals of Telecommunications*, vol. 74, no. 7–8, pp. 461–471, 2019.
- [7] Y. Li, S. Yao, K. Yang, Y.-A. Tan, and Q. Zhang, "A high-imperceptibility and histogram-shifting data hiding scheme for JPEG images," *IEEE Access*, vol. 7, pp. 73573–73582, 2019.
- [8] Y. Li, X. Zhang, X. Xu, and Y.-A. Tan, "A robust packet-dropout covert channel over wireless networks," *IEEE Wireless Communications*, vol. 27, no. 3, pp. 60–65, 2020.
- [9] M. Park, S. Kim, and J. Kim, "Research on note-taking apps with security features," *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, vol. 11, pp. 63–76, 2020.
- [10] L. Zhang, S. Hao, and Q. Zhang, "Recovering SQLite data from fragmented flash pages," *Annals of Telecommunications*, vol. 74, no. 7–8, pp. 451–460, 2019.
- [11] Y.-A. Tan, X. Xu, C. Liang, X. Zhang, Q. Zhang, and Y. Li, "An end-to-end covert channel via packet dropout for mobile networks," *International Journal of Distributed Sensor Networks*, vol. 14, no. 5, 2018.
- [12] R. Zhu, B. Zhang, J. Mao, Q. Zhang, and Y.-A. Tan, "A methodology for determining the image base of ARM-based industrial control system firmware," *International Journal of Critical Infrastructure Protection*, vol. 16, pp. 26–35, 2017.
- [13] W. Wang, Y. Shang, Y. He, Y. Li, and J. Liu, "BotMark: automated botnet detection with hybrid analysis of flow-based and graph-based traffic behaviors," *Information Sciences*, vol. 511, pp. 284–296, 2020.
- [14] W. Wang, X. Wang, D. Feng, J. Liu, Z. Han, and X. Zhang, "Exploring permission-induced risk in android applications

- for malicious application detection,” *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 11, pp. 1869–1882, 2014.
- [15] W. Meng, W. Li, and L.-F. Kwok, “EFM: enhancing the performance of signature-based network intrusion detection systems using enhanced filter mechanism,” *Computers & Security*, vol. 43, pp. 189–204, 2014.
- [16] W. Li, W. Meng, Z. Tan, and Y. Xiang, “Design of multi-view based email classification for IoT systems via semi-supervised learning,” *Journal of Network and Computer Applications*, vol. 128, pp. 56–63, 2019.
- [17] Y. Yuan, L. Huo, Z. Wang, and D. Hogrefe, “Secure apit localization scheme against sybil attacks in distributed wireless sensor networks,” *IEEE Access*, vol. 6, pp. 27629–27636, 2018.
- [18] D. Berbecaru, A. Liroy, and C. Cameroni, “Supporting authorize-then-authenticate for wi-fi access based on an electronic identity infrastructure,” *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, vol. 11, pp. 34–54, 2020.
- [19] D. M.-E. Francesc, E. A.-I. José, and S. Melissa, “Evaluation of database replication techniques for cloud systems,” *Computing and Informatics*, vol. 34, no. 5, pp. 973–995, 2015.
- [20] Q. Fan, M. Tan, and Y. Luo, “Security setting and security hardening of computer equipment in LAN,” *Information Security and Communication Confidentiality*, vol. 17, no. 7, p. 31, 2008.
- [21] B. Dong, “Physical isolation technology for network security,” *Computer and Automation*, vol. 10, no. 2, pp. 108–110, 2004.
- [22] I. Kholod, A. Shorov, and S. Gorlatch, “Efficient distribution and processing of data for parallelizing data mining in mobile clouds,” *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, vol. 11, pp. 2–17, 2020.
- [23] W. Liu and L. Tong, “Design of change capture scheme in integration of heterogeneous databases,” *Workflow of Database Synchronization Sending End Computer Application Research*, vol. 23, no. 7, pp. 213–215, 2005.
- [24] R. Guerraoui and A. Schiper, “Software-based replication for fault tolerance,” *Computer*, vol. 30, no. 4, pp. 68–74, 1997.
- [25] H. Anada, “Decentralized multi-authority anonymous authentication for global identities with non-interactive proofs,” in *Proceedings of the 2019 IEEE International Conference on Smart Computing (SMARTCOMP)*, vol. 10, no. 4, pp. 23–37, IEEE, Washington, DC, USA, June 2019.
- [26] F. Guo, Q. Zhao, X. Li et al., “Detecting adversarial examples via prediction difference for deep neural networks,” *Information Sciences*, vol. 501, pp. 182–192, 2019.
- [27] Q. Zhang, Y. Zhao, Y. Wang, T. Baker, J. Zhang, and J. Hu, “Towards cross-task universal perturbation against black-box object detectors in autonomous driving,” *Computer Networks*, vol. 180, Article ID 107388, 2020.
- [28] A. Kumar and A. Segev, “Cost and availability tradeoffs in replicated data concurrency control,” *ACM Transactions on Database Systems*, vol. 18, no. 1, pp. 102–131, 1993.
- [29] J. B. Lim and A. R. Hurson, “Data duplication and consistency in a mobile, multidatabase environment,” in *Proceedings of the International Conference on Parallel and Distributed Systems (ICPADS)*, pp. 50–58, IEEE, Tainan, Taiwan, December 1998.
- [30] A. Paul, S. Lihua Wang, D. S. Sharmila, and C. P. Rangan, “Non-transferability in proxy re-encryption revisited,” *Journal of Internet Services and Information Security*, vol. 10, no. 3, pp. 1–30, 2020.
- [31] M. Nicola and M. Jarke, “Performance modeling of distributed and replicated databases,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 12, no. 4, pp. 645–672, 2000.
- [32] J. Wu and N. Li, “Design and application of safety isolation software for intelligent hydropower plant,” *Hydropower Plant Automation*, vol. 37, no. 4, pp. 67–69, 2016.
- [33] H. Yu, Y. Ding, X. Gao, N. Geng, X. Huang, and Y. Lu, “Data security transmission technology of water conservancy project standardization management based on one-way isolation gate,” *Zhejiang Water Science and Technology*, vol. 44, no. 5, pp. 48–50, 2016.
- [34] D. Filipović, D. Sokač, and R. Picek, “Bidirectional database synchronization to the cloud computing platform,” in *Proceedings of the ICISS 2020: 2020 The 3rd International Conference on Information Science and System*, London, UK, March 2020.
- [35] S. H. Phatak and B. R. Badrinath, “Conflict resolution and reconciliation in disconnected databases,” in *Proceedings of the 10th International Workshop on Database & Expert Systems Applications (DEXA)*, pp. 76–81, IEEE, Florence, Italy, September 1999.
- [36] R. K. Cassar, J. Vella, and J. Ellul, “A conflict resolution abstraction layer for eventually consistent databases,” in *Proceedings of the 2016 International Conference on Engineering & MIS (ICEMIS)*, pp. 1–5, IEEE, Agadir, Morocco, September 2016.
- [37] A. Marotta, D. Cassioli, M. Tornatore, Y. Hirota, Y. Awaji, and B. Mukherjee, “Reliable slicing with isolation in optical metro-aggregation networks,” in *Proceedings of the 2020 Optical Fiber Communication Conference and Exhibition*, IEEE, San Diego, CA, USA, March 2020.
- [38] J. Gray and A. Reuter, *Transaction Processing: Concepts and Techniques*, Morgan Kaufmann, San Mateo, CA, USA, 1993.