Indonesian Journal of Electrical Engineering and Computer Science
Vol. 28, No. 2, November 2022, pp. 973~979
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v28.i2.pp973-979

An automate failure recovery for synchronous distributed database system

Ahmad Shukri Mohd Noor[1], Auni Fauzi Che Fauzi[1], Ainul Azila Che Fauzi[2], Noraziah Ahmad[3,4], Mohamad Syauqi Mohamad Arifin[2]
[1]Faculty of Ocean Engineering Technology and Informatics, Universiti Malaysia Terengganu, Kuala Terengganu, Malaysia
[2]Faculty of Computer Science and Mathematics, Universiti Teknologi MARA (UiTM) Cawangan Kelantan, Machang, Malaysia
[3]Faculty of Computing, Universiti Malaysia Pahang, Pekan, Malaysia
[4]Centre of Software Development, and Integrated Computing, Universiti Malaysia Pahang, Pekan, Malaysia

Article Info

Article history:

Received Apr 4, 2022
Revised Jul 14, 2022
Accepted Aug 18, 2022

Keywords:

Database replication
Distributed databse
Fault recovery
Fault tolerance
Synchronous replication

ABSTRACT

Periodically, researchers have been sharing their constant attempts to improve the existing methods for data replication in distributed database system. The main goal is to work for an efficient distributed environment. An efficient environment may handle huge amount of data and preserve data availability. The occasionally failures in distributed systems will affect the end results, such as data loss, income loss etc. Thus, to prevent the data loss and guarantee the continuity of the business, many organizations have applied disaster recovery solutions in their system. One of the widely used is database replication, because it guarantees data safety and availability. However, disaster still can occur in database replication. Hence, an automatic failure recovery technique called distributed database replication with fault tolerance (DDR-FT) has been proposed in this research. DDR-FT uses heartbeat message for node monitoring. Subsequently, a foundation of binary vote assignment for fragmented database (BVAFD) replication technique has been used. In DDR-FT, the data nodes are continuously monitored while auto reconfiguring for automatic failure recovery. From the conducted experiments, it is proved that DDR-FT can preserve system availability. It shows that DDR-FT technique provides a convenient approach to system availability for distributed database replication in real time environment.

Corresponding Author:

Ahmad Shukri Mohd Noor
Faculty of Ocean Engineering Technology and Informatics, Universiti Malaysia Terengganu
21030 Kuala Terengganu, Terengganu, Malaysia
Email: ashukri@umt.edu.my

1. INTRODUCTION

Nowadays, getting access to high quality data can help to enhance the quality of life both for those working in the organization and for the people. Hence, it is crucial to ensure the database system is well managed and secure. There are two types of database systems which are centralized database system and distributed database system. Centralized database system is a database that is located, stored and maintained in a single location. When the single database server is crashed or mishap happens, everything will be lost. Thus, to increase and preserve the data availability and reliability, it is better to store the data in multiple servers rather than in one server [1]–[3]. This method is called distributed database system [1], [4], [5]. Hence, in the event of disaster, data availability is ensured because distributed systems have higher reliability and incremental growth [6].

Distributed databases systems (DDBS) are a set of logically networked computer databases, managed by different sites, locations and accessible to the user locally or via the Internet for different tasks as a single database by running transactions in parallel [4], [7]. The design of DDBS is a very demanding task since an optimum level of performance must be continuously satisfied [8], [9]. Data availability plays a major role in the success of information systems because data must be always available in order to meet the users' requirements.

Data replication is one of the widely used solutions to ensure data safety and availability in distributed database system [10], [11]. It plays a critical role in promoting business for any organisations, refining the quality of the data, improving data sharing as well as the process for big data analysis [12]. Therefore, the implementation of the data replication method itself can be vital when it involves failure interruption. Particularly, failures occur regularly on the internet, clouds and in scale-out data center networks [13]–[16]. Hence, a fault tolerance method is needed in data replication transaction [17]–[19]. A system using such services are required to be equipped with resources so that the system can guarantee to operate even in the presence of faults [17], [20]–[22]. Fault is supposed to be detected by using a reliable fault detector followed by a recovery technique.

In this research, a data replication technique called binary vote assignment for fragmented database (BVAFD) is combined with a proposed fault tolerance technique called distributed database system with fault tolerance (DDR-FT) with the aimed of assessing the efficiency of synthesizing data replication technique with fault tolerance technique for a better performance of a single database replication transaction in the event of failure. The paper is arranged as follows. Section 2 is the literature review which detailed out about BVAFD data replication technique. In section 3, the methodology describes the procedure of DDR-FT algorithm which is employed within BVAFD framework. The result and discussion is presented in section 4 where the outcomes obtained from a series of experiments that has been tested are discussed in this section. Finally, the conclusion of this research is provided in section 5, conclusion.

2. LITERATURE REVIEW

Figure 1 shows the concept of binary vote assignment for fragmented database (BVAFD) is copying some data from the primary node to some of the adjacent's nodes [23]. Full replication mechanism may waste a lot of of storage space and consume a lot of bandwidth [24] because it copies all data to all replication nodes. By using BVAFD, the replication execution time is reduced since it only copies some data to some nodes [24], [25]. BVAFD is pacing a new path in distributed database replication as it helps to maximize the write availability with low communication cost [26].

In BVAFD, each node has a primary database table, PDT. PDT will be copied to the neighbours' nodes, NS from the primary node, PS [23]. PS of any PDT and NS are assigned with vote one (1) or vote zero (0). This assignment is treated as an allocation of replicated copies and a vote assigned to the site results in a copy allocated at the neighbour [27] PS of any PDT and NS are assigned with different status depends on their condition. Status = 0 is shown when PS is accessible. Meanwhile, status = 1 is shown when PS is inaccessible or busy.

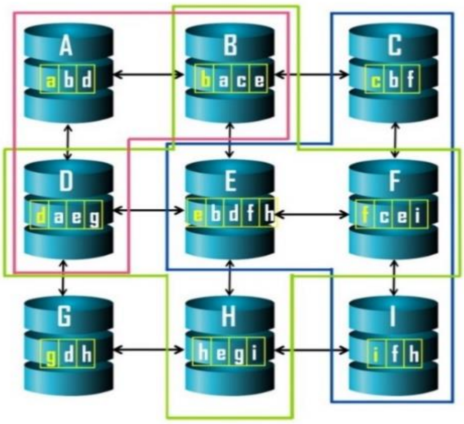

Figure 1. Examples of data replication in BVAFD

3. RESEARCH METHODS

In this section, distributed database system with fault tolerance (DDR-FT) is proposed by considering fault detection and fault recovery in BVAFD. The monitor node will always monitor all the heartbeat message from all the data nodes. Recovery process will only begin automatically once a fault has been detected. The following notations are defined:

- N_d is a node with data
- N_m is a monitor node
- N'_d is neighbour node, where d = 1, 2, 3, 4, or 5
- N_{d_f} is failure node transformed from N_d.
- F_∇ is when no failure occurs
- F_Δ is when failure occurs
- IP_p is a primary IP address
- IP_v is a virtual IP address

All N_d where d = {1, 2, 3, 4, 5} continuously send a heartbeat message, to N_m. Hence, N_m always monitor all heartbeat messages received from N_d. If N_m does not receives any heartbeat messages from N_d, it will assume $N_d = F_\Delta$ and transform N_d to N_{d_f}. Next, N_m will search the available N'_d of N_{d_f} . Once N_m get the N'_d, it will assign the N'_d as the backup node for the N_{d_f}. Then, N'_d will create IP_v contains an IP of N_{d_f}. Now N'_d will have two IP addresses which are its own IP_p and IP_v of N_{d_f}. Once N_m detect h=1 from N_{df}, then $N_{df} = F_\nabla \rightarrow N_d$, it will delete the IP_v from N_{df} and continue monitoring the nodes.

4. RESULT AND DISCUSSIONS

In this section, experiments of cases that occur during real time transactions are presented. It involved three replication nodes called node 1, node 2 and node 4 in one distributed database systems. All data in these three servers are supposed to be the same replicated data. Therefore, all the data will be updated synchronously.

4.1. Experiment 1: no failure occurs in any nodes

Figure 2 shows $N_1 = F_\nabla$. N_m receives all the heartbeat messages sent from N_d where d = 1, 2 and 4. Thus, it will assume no failure occurred. Figure 3 shows monitor node, N_m receive heartbeat from node 1, 2 and 4. During this time, all N_d are sending their heartbeat message to the N_m without any failure.

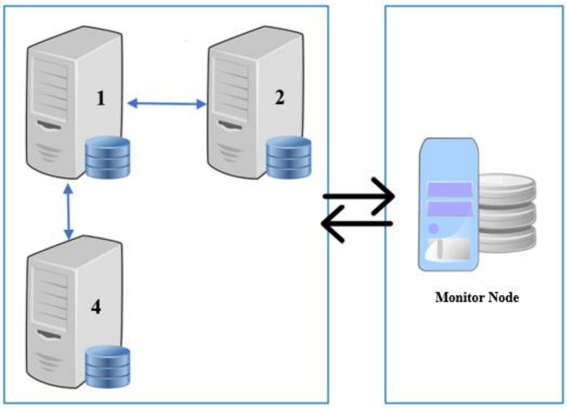

Figure 2. Data nodes without any failure occurrence

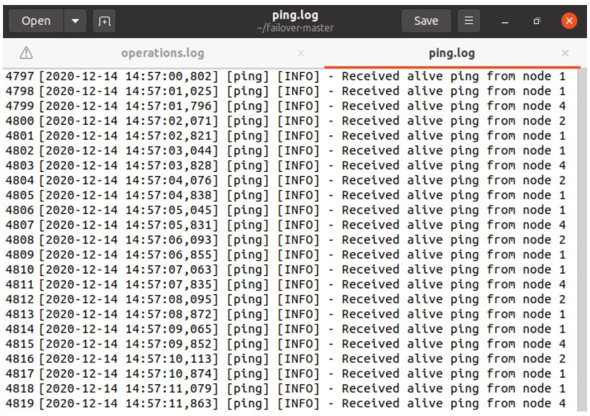

Figure 3. Monitor node receive heartbeat

4.2. Data node 1 held failure occurrence without DDR-FT

Figure 4 shows $N_1 = F_\Delta \rightarrow N_{1_f}$ which means a failure occurs during the transaction. Node 1 is not connected through the network. When this happens, N_{1_f} will try to reconnect to the N_m every 5 seconds. However, since there is no fault tolerance in this system, this problem will not be solved until it is fixed manually. Any transaction will not be able to procced until the problem is solved.

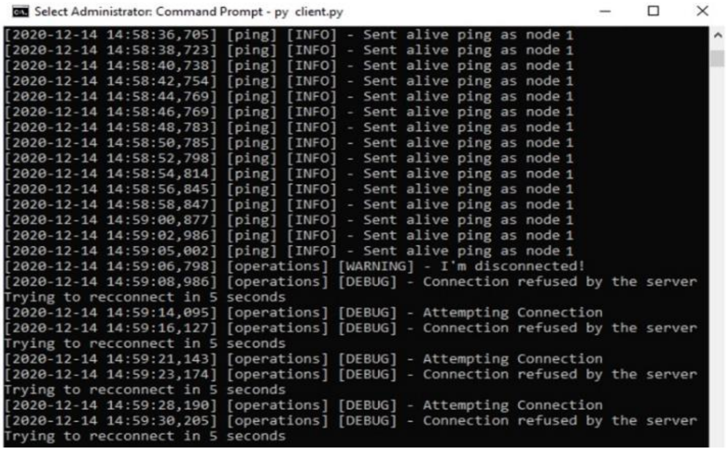

Figure 4. Node 1 attempting to send it heartbeat

4.3. Data node 1 held failure occurrence with DDR-FT

Figure 5 shows $N_1 = F_\Delta \longrightarrow N_{1_f}$ which means a failure occurs during the transaction. In this experiment, DDR-FT has been applied in the BVAFD for failure detection and recovery purposes. Node 1 is not connected through the network which will be detected by monitoring node. The monitor node then will begin the recovery process. In this experiment, N_1 do not send any heartbeat messages to the N_m. As shown in Figure 6, N_m does not receive any heartbeat messages from N_1. Hence, it will assume $N_1 = F_\Delta$ and transform N_1 to N_{1_f}. N_m will begin the recovery process. N_m will search for any N'_{d} where d = 2 or 4 to backup N_1.

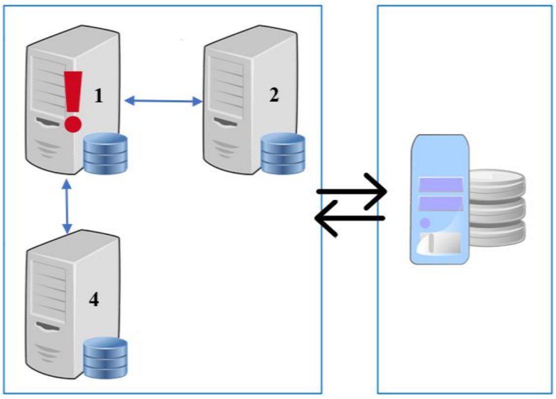

Figure 5. Failure nodes without fault tolerance

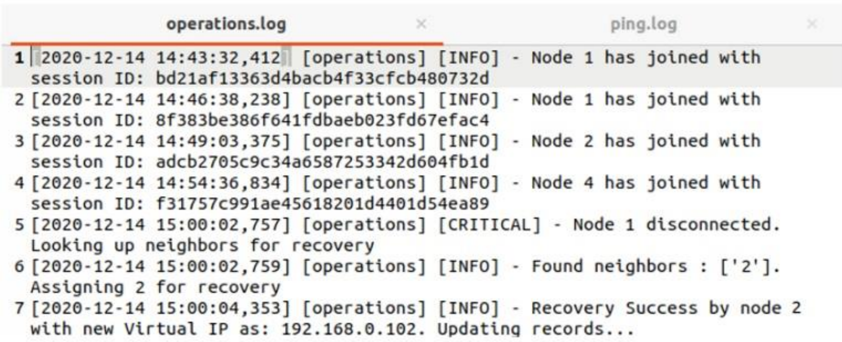

Figure 6. Failure recovery process from monitor node operations log

After N_m found the available neighbor node, for example in this experiment, N'_2, N'_2 then will create an IPv of N_{1_f} which is IP address for N1 as shown in Figure 7. Figure 8 shows N'_2 now have two IP addresses, which are IP_p = 192.168.0.101 and the IP_v = 192.168.0.102. From figure 8, N'_2 have two IP addresses which are its own IP address and virtual IP address from the failure node, N_{1_f}.

```
06:17,456] [operations] [INFO] - Received recovery request. Recovering node 1
06:17,456] [operations] [INFO] - creating virtual IP address: 192.168.0.102
06:17,456] [operations] [CRITICAL] - netsh interface ipv4 add address "Ethernet0" 192.168.0.102 255.255.255.0
06:17,782] [operations] [CRITICAL] - Success
06:17,797] [operations] [INFO] - Successfully created secondary IP as 192.168.0.102 Notifying Server...
```

Figure 7. Node 2 receive recovery request

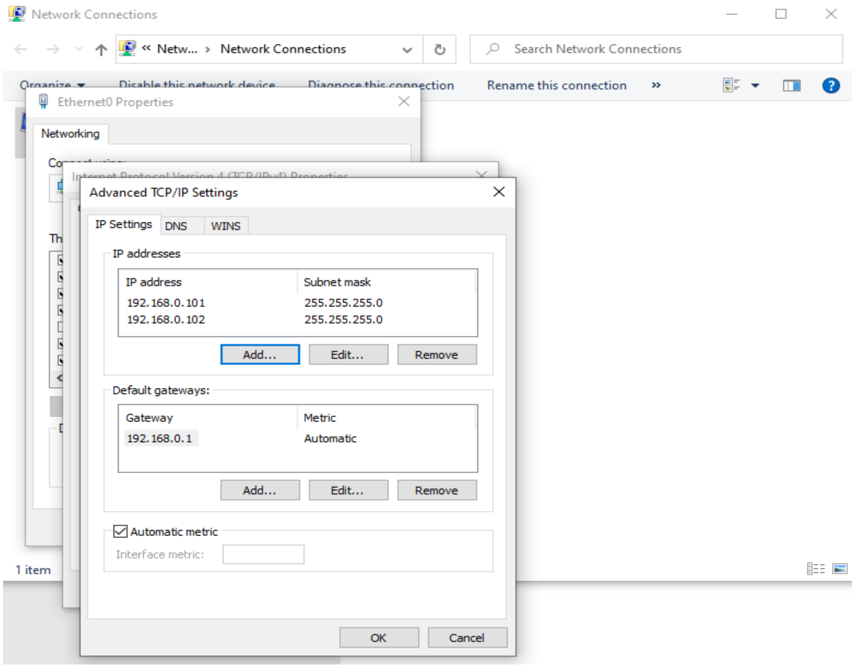

Figure 8. IP address for node 2

5. CONCLUSION

This study proposed an automatic failure recovery technique called DDR-FT. DDR-FT focuses on the detection of fault and failure recovery. It is run in a replication technique called BVAFD. From the series of experiments that have been conducted, it is proved that DDR-FT can preserve system availability in BVAFD during the event of a failure. It also shows that DDR-FT provides a convenient approach to data availability for distributed database replication in real time environment. However, DDR-FT can be improved in many ways. As we know, data consistency is very important in distributed database. Currently DDR-FT does not support post recovery process. In future, DDR-FT will handle the data consistency after the failure node is operating again. In addition, this method is aimed to be fully implemented in various systems and compare the performances against other existing fault tolerance methods. The framework will be expanded to cover not only the ability to detect and recover failure, but also allow the application to automatically access various recovery and masking mechanisms.

ACKNOWLEDGEMENTS

This research is funded by Fundamental Research Grant Scheme (FRGS) with the Ref: FRGS/1/2018/ICT04/UMT/02/2. FRGS is a research grant from the Ministry of Higher Education (MOHE) Malaysia.

REFERENCES

[1] A. Noraziah, A. A. C. Fauzi, S. H. S. A. Ubaidillah, B. Alkazemi, and J. B. Odili, "BVAGQ-AR for fragmented database replication management," *IEEE Access*, vol. 9, pp. 56168–56177, 2021, doi: 10.1109/ACCESS.2021.3065944.

[2] B. A. Milani and N. J. Navimipour, "A comprehensive review of the data replication techniques in the cloud environments: Major trends and future directions," *Journal of Network and Computer Applications*, vol. 64, pp. 229–238, Apr. 2016, doi: 10.1016/j.jnca.2016.02.005.

[3] J. Wang, H. Wu, and R. Wang, "A new reliability model in replication-based big data storage systems," *Journal of Parallel and Distributed Computing*, vol. 108, pp. 14–27, Oct. 2017, doi: 10.1016/j.jpdc.2017.02.001.

[4] B. S. M. Sawiris and M. A. Abdel-Fattah, "A novel solution for distributed database problems," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 10, pp. 468–474, 2020, doi: 10.14569/IJACSA.2020.0111059.

[5] K. Maabreh, "Adaptive lockable units to improve data availability in a distributed database system," *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 1, 2016, doi: 10.14569/ijacsa.2016.070168.

[6] F. N. Al-Wesabi, "Improving availability in component-based distributed systems," *Intelligent Automation and Soft Computing*, vol. 26, no. 6, pp. 1345–1357, 2020, doi: 10.32604/iasc.2020.013835.

[7] H. Guo, X. Zhou, and L. Cai, "Lock violation for fault-tolerant distributed database system," in *Proceedings - International Conference on Data Engineering*, Apr. 2021, vol. 2021-April, pp. 1416–1427. doi: 10.1109/ICDE51399.2021.00126.

[8] F. Castro-Medina, L. Rodriguez-Mazahua, A. Lopez-Chau, I. Machorro-Cano, and M. A. Abud-Figueroa, "Design of a horizontal data fragmentation, allocation and replication method in the cloud," in *IEEE International Conference on Automation Science and Engineering*, Aug. 2019, vol. 2019-August, pp. 614–621. doi: 10.1109/COASE.2019.8842934.

[9] P. Y. Zhang and M. C. Zhou, "Dynamic cloud task scheduling based on a two-stage strategy," *IEEE Transactions on Automation Science and Engineering*, vol. 15, no. 2, pp. 772–783, Apr. 2018, doi: 10.1109/TASE.2017.2693688.

[10] J. Mendonca, W. Medeiros, E. Andrade, R. MacIel, P. MacIel, and R. Lima, "Evaluating database replication mechanisms for disaster recovery in cloud environments," in *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, Oct. 2019, vol. 2019-October, pp. 2358–2363. doi: 10.1109/SMC.2019.8914069.

[11] Y. Yan, Y. Song, and B. Wang, "DRF-FTS: A dynamic replication factor replication scheme based on fault-tolerant set," in *Proceedings of 2021 IEEE 3rd International Conference on Civil Aviation Safety and Information Technology, ICCASIT 2021*, Oct. 2021, pp. 974–979. doi: 10.1109/ICCASIT53235.2021.9633522.

[12] Y. Jiang, N. Zhang, and Y. Fang, "The analysis and design of ship monitoring system based on hybrid replication technology," in *Proceedings - 2019 International Conference on Intelligent Transportation, Big Data and Smart City, ICITBS 2019*, Jan. 2019, pp. 456–459. doi: 10.1109/ICITBS.2019.00118.

[13] A. Gorbenko, A. Karpenko, and O. Tarasyuk, "Analysis of trade-offs in fault-tolerant distributed computing and replicated databases," in *Proceedings - 2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies, DESSERT 2020*, May 2020, pp. 1–6. doi: 10.1109/DESSERT50317.2020.9125078.

[14] R. Potharaju and N. Jain, "When the network crumbles: An empirical study of cloud network failures and their impact on services," Oct. 2013. doi: 10.1145/2523616.2523638.

[15] E. Katz-Bassett *et al.*, "LIFEGUARD: Practical repair of persistent route failures," *Computer Communication Review*, vol. 42, no. 4, pp. 395–406, Sep. 2012, doi: 10.1145/2377677.2377756.

[16] A. Gorbenko, V. Kharchenko, O. Tarasyuk, Y. Chen, and A. Romanovsky, "The threat of uncertainty in Service-Oriented Architecture," in *Proceedings of the 2008 RISE/EFTS Joint International Workshop on Software Engineering for Resilient Systems, SERENE'08*, 2008, pp. 49–54. doi: 10.1145/1479772.1479781.

[17] E. Abdelfattah, M. Elkawkagy, and A. El-Sisi, "Hybrid fault tolerance approach based on replication impact heuristic for cloud computing," in *29th International Conference on Computer Theory and Applications, ICCTA 2019 - Proceedings*, Oct. 2019, pp. 60–67. doi: 10.1109/ICCTA48790.2019.9478819.

[18] M. A. Mukwevho and T. Celik, "Toward a smart cloud: a review of fault-tolerance methods in cloud systems," *IEEE Transactions on Services Computing*, vol. 14, no. 2, pp. 589–605, Mar. 2021, doi: 10.1109/TSC.2018.2816644.

[19] P. Kumari and P. Kaur, "A survey of fault tolerance in cloud computing," *Journal of King Saud University - Computer and Information Sciences*, vol. 33, no. 10, pp. 1159–1176, Dec. 2021, doi: 10.1016/j.jksuci.2018.09.021.

[20] N. Malhotra and M. Bala, "Fault-tolerant communication induced checkpointing and recovery protocol using iot," *Intelligent Automation and Soft Computing*, vol. 30, no. 3, pp. 945–960, 2021, doi: 10.32604/iasc.2021.019082.

[21] M. N. Cheraghlou, A. Khadem-Zadeh, and M. Haghparast, "A survey of fault tolerance architecture in cloud computing," *Journal of Network and Computer Applications*, vol. 61, pp. 81–92, Feb. 2016, doi: 10.1016/j.jnca.2015.10.004.

[22] A. Kumar and D. Malhotra, "Study of various proactive fault tolerance techniques in cloud computing," *International Journal of Computer Sciences and Engineering*, vol. 06, no. 03, pp. 81–87, Apr. 2018, doi: 10.26438/ijcse/v6si3.8187.

[23] A. Noraziah, A. A. C. Fauzi, S. H. S. A. Ubaidillah, Z. Abdullah, R. M. Sidek, and M. A. I. Fakhreldin, "Managing database replication using binary vote assignment on grid quorum with association rule," *Advanced Science Letters*, vol. 24, no. 10, pp. 7834–7837, Oct. 2018, doi: 10.1166/asl.2018.13027.

[24] J. P. Yang, "Elastic load balancing using self-adaptive replication management," *IEEE Access*, vol. 5, pp. 7495–7504, 2017, doi: 10.1109/ACCESS.2016.2631490.

[25] A. A. C. Fauzi, A. Noraziah, T. Herawan, Z. Abdullah, and R. Gupta, "Managing fragmented database using BVAGQ-AR replication model," *Advanced Science Letters*, vol. 23, no. 11, pp. 11088–11091, Nov. 2017, doi: 10.1166/asl.2017.10226.

[26] S. H. S. A. Ubaidillah, B. Alkazemi, and A. Noraziah, "An efficient data replication technique with fault tolerance approach using BVAG with checkpoint and rollback-recovery," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 1, pp. 473–480, 2021, doi: 10.14569/IJACSA.2021.0120155.

[27] S. H. S. A. Ubaidillah, N. Ahmad, and J. Beneoluchi Odili, "Fragmentation techniques for ideal performance in distributed database–a survey," *International Journal of Software Engineering and Computer Systems*, vol. 6, no. 1, pp. 18–24, May 2020, doi: 10.15282/ijsecs.6.1.2020.3.0066.

BIOGRAPHIES OF AUTHORS

Ahmad Shukri Mohd Noor is an Associate professor of computer science at Faculty of Ocean Engineering Technology and Informatics, Universiti Malaysia Terengganu (UMT). He received BSc in Computer Science from Coventry University, UK. MSc and PhD in distributed computing from UTHM, Malaysia and KUSTEM, Malaysia respectively. He was a post-doctoral research fellow with Distributed, Reliable, Intelligent Control and Cognitive Systems Group (DRIS) at the Department of Computer Science, University of Hull, UK. He has published more than 50 research papers in various referred journals, conferences, seminars and symposiums. His research interests include distributed databases, data grid, data mining and soft computing and various other referred journals. Currently, he supervises 11 post graduate students. He is also a Google Online Professional Certified as well as Cloud Professional Certified. He can be contacted at email: ashukri@umt.edu.my.

Ainul Auni Che Fauzi is a master student in the Department of Ocean Engineering Technology and Informatics, University Malaysia Terengganu, Kuala Nerus, Terengganu Darul Iman, Malaysia. Received the B.Sc. degree in software engineering and Diploma in computer science from Universiti Malaysia Pahang. During her studies she received 7 of exhibition awards from her bachelor project and currently 1 exhibition award from her master project. She continues her bachelor project by adding fault tolerance value for her master project. She can be contacted at email: p4155@pps.umt.edu.my.

Ainul Azila Che Fauzi received the B.Sc. Degree and M.Sc. Degree in Software Engineering from Universiti Malaysia Pahang, and Ph.D. Degree also from Universiti Malaysia Pahang. She is currently working as a senior lecturer in Universiti Teknologi MARA Cawangan Kelantan (UiTMCK), Malaysia. She is also the publication coordinator in UiTMCK since 2020. She has supervised and co-supervised more than 5 bachelor’s degree students. She has authored or coauthored more than 30 publications: 12 proceedings and 22 journals, with 7 H-index and more than 100 citations. Her research interests include distributed database system, cloud computing, grid computing, data mining and association rules. She can be contacted at email: ainulazila@uitm.edu.my.

Noraziah Ahmad received the Ph.D. degree in database from University Malaysia Terengganu, Malaysia. She is currently an Associate Professor with the Faculty of Computing, and a Research Fellow with the Centre for Software Development and Integrated Computing, University Malaysia Pahang, Malaysia. She had published 280 scientic research articles. She also supervised more than 20 postgraduate students, and obtained many grants and awards related to her research expertise. Her research interests include distributed database, data grid, data mining, big data, and computational intelligence. She has a professional membership in IEEE Computer Society, IACSIT, MNCC, and IAENG. In addition, she also received several international and national awards and recognitions. She served as an International Program Committees and reviewers for many numerous international journals and conferences. She can be contacted at email: noraziah@ump.edu.my.

Mohamad Syauqi Mohamad Arifin received the B.Sc. Degree and M.Sc. Degree in Information System Management from Universiti Teknologi MARA. He is currently working as a lecturer in Universiti Teknologi MARA Cawangan Kelantan (UiTMCK). His research interest included E-commerce and administration of database security. He can be contacted at email: mohdsyauqi@uitm.edu.my.