Performance Verification of the Digital Archive Subsystem for Digital Twin-based Disaster Management Platform

Ki-Sook Chung
Disaster & Safety AI Convergence Center
Electronics and Telecommunications Research Institute
Daejeon, Korea
kschung@etri.re.kr

Woosug Jung
Disaster & Safety AI Convergence Center
Electronics and Telecommunications Research Institute
Daejeon, Korea
wsjung@etri.re.kr

***Abstract*— Underground Utility Tunnel (UUT) is responsible for the lifeline of a city area and would threat the safety of daily life of citizens when the disaster like fire, earthquake, and flood happens. We are developing the digital twin platform to monitor UUT and prepare for disaster situation. The digital twin-based disaster and safety management platform generates a digital twin model by hierarchically managing and integrating various types of information such as data collected on-site, spatial information, modeling, and simulation. The digital archive subsystem plays important roles in the platform by managing data collected, generated, and used by the various h/w or s/w components of the platform. This paper shows the result of performance test of the digital archive assuming both normal case scenarios and disaster case scenarios.**

***Keywords*— Digital Twin, Digital Archive, Sensor Data, Twin model)**

I. INTRODUCTION

According to Gartner's definition, a digital twin is a digital representation of a real-world system or object[3]. Digital twin technology, which was first proposed by Dr. Michael Grieves in the United States in 2002[2], has recently emerged as a key technology that can be used in various fields, including manufacturing, healthcare, cities, defense, and disaster safety, to identify problems and derive solutions.

In this study, we focus on the performance evaluation of the digital archive, which plays a crucial role in a disaster safety platform designed to minimize damage to urban infrastructure and human and material resources in the event of disasters such as fire, earthquakes, and floods. The digital archive subsystem is responsible for storing and managing the key data collected and generated by the platform, and thus, we conducted a load test based on both normal and disaster scenarios to evaluate its performance under different conditions such as data volume and frequency.

Chapter 1 introduces digital twin-based disaster and safety management platform and the digital archive subsystem currently under development, while Chapter 2 presents the test environment and scenarios, test items, and test results of the digital archive subsystem. Finally, we analyze the test results and draw conclusions.

II. PERFORMANCE TEST OF DIGITAL ARCHIVE SUBSYSTEM

A. Digital Twin-based disaster management platform

Underground Utility Tunnel (UUT) is one of the national infrastructure facilities that accommodates lifelines such as electricity, communication, water, and heating, providing convenient urban living to citizens. Safety accidents in underground complexes caused by fires or earthquakes are directly related to the daily lives of citizens and can even result in the paralysis of urban functions. In order to ensure the safe management of underground complexes, the "Digital Twin-based Integrated Platform for Fire and Disaster Support in Underground Utility Tunnel (UUT)" project has been underway since 2020 as a multi-ministry collaboration between the Ministry of Science, ICT and Future Planning, the Ministry of the Interior and Safety, the Ministry of Land, Infrastructure and Transport, and the Ministry of Trade, Industry and Energy to develop efficient disaster prevention and rapid response systems by introducing digital twin technology[1].

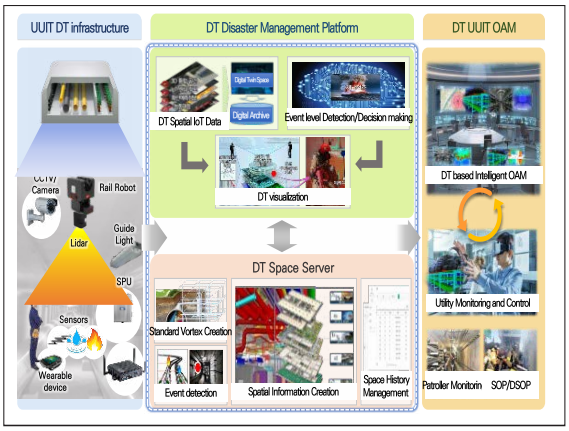

Fig 1. Digital Twin based UUT Disaster Management Platform

Figure 1 illustrates the digital twin-based disaster safety management

platform for UUT. By modeling the spatial information of the underground complex, the platform processes and analyzes real-time data collected on site, and generates various layers of information to create a digital twin model. Based on the digitized model, the platform enables disaster diffusion prediction through situation judgment and simulation, and supports immediate response in case of actual disasters. To create the digital twin model of the underground utility tunnel, various forms of data such as real-time sensor information, underground complex patrol and operational management information are required starting with spatial information. The digital archive subsystem within the platform provides a crucial role in storing and synchronizing this data.

A.1 Digital archive dataset

The types of datasets managed in the digital archive are categorized into 5 datasets such as spatial data, sensing data, modeling and simulation data, dynamic data, and management data. The digital twin model data generated from these datasets is also stored and managed.

1. Spatial data (D1): 3D spatial information data to represent the underground utility tunnel
2. Sensing data (D2): Sensor data collected from the underground utility tunnel environment and situations
3. Simulation data (D3): Data used to simulate and predict risks and their spread
4. Dynamic data (D4): Dynamic service data such as provided to workers or operators in the underground utility tunnel site
5. Management data (D5): Relevant and linked agency and management data for responding to risk situations

Each of the datasets defined above is used as the fundamental data to generate various digital twin models, depending on the combination method as shown in Figure 2.

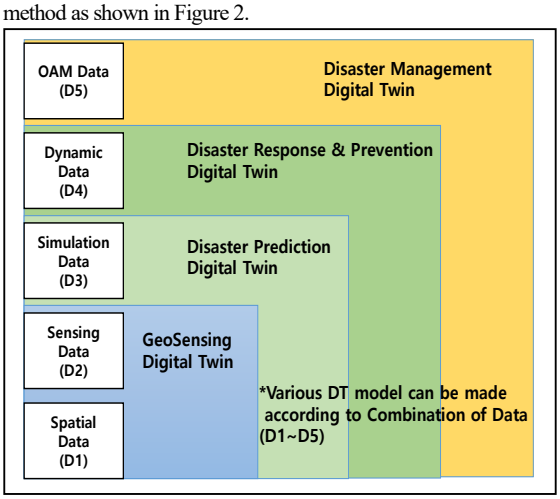

Fig. 2. Digital Archive Dataset and Digital Twin

A.2 Digital Archive Subsystem

The digital archive subsystem receives real-time on-site data of the underground utility tunnel, stores it, and processes and provides it to create models on the platform. It stores all data that constitutes the digital twin model for a certain period and synchronizes it with an external digital archive in the urban integrated center, including operational management as shown in Figure 3.

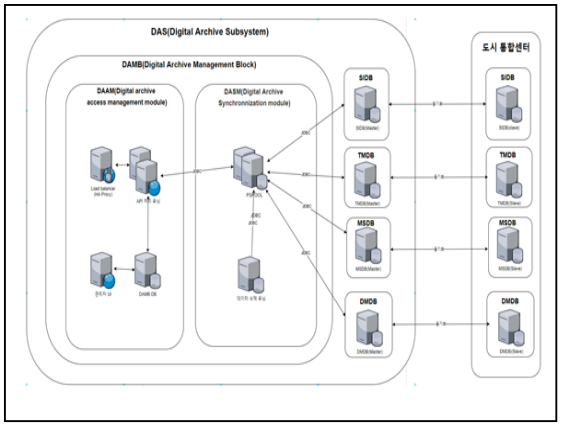

Fig3. Digital Archive Subsystem architecture

The digital archive consists of four DB storage blocks and a DAMB (Digital Archive Management Block) that provides access control and synchronization functions, as shown in Figure 3. The four types of DB storage blocks are used to categorize and store the datasets (D1~D5) managed by the digital archive, as well as to store and manage the digital twin models generated by the platform.

- SIDB (3d Space Information Database Block): Manages D1 and D2 datasets, including spatial shape and spatial sensing information.
- MSDB (Modeling and Simulation Database Block): Stores and manages D3 datasets, which simulate and predict the direction and speed of the spread of disasters such as fires, earthquakes, and floods, to enable immediate response.
- DMDB (Disaster Management Database Block): Stores and manages the overall management data of the underground facility, including operational management and response information for D4 and D5 datasets.
- TMDB (Twin model Management Database Block): Stores and manages the digital twin models generated through the analysis and processing procedures of the disaster safety management platform.

In this paper, we describe the process and results of a load test conducted to verify the performance of the digital archive, which stores and manages various types of data collected and delivered in real-time, as well as the digital twin models generated for disaster safety management.

B. Load tests of Digital Archive

B.1 Test Environment of system

The physical environment setup for the load test is shown in Figure 4, and the specifications for each system are listed in Table 1.

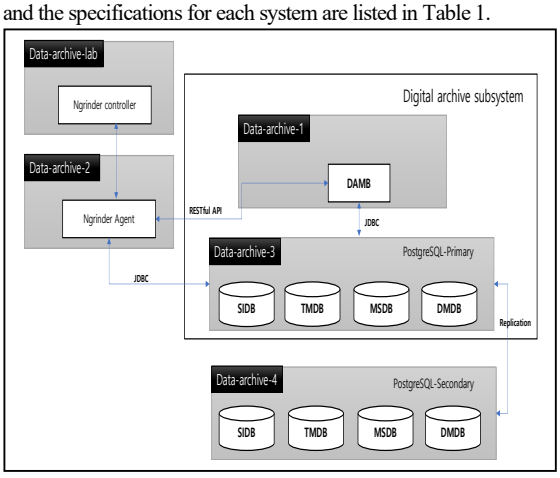

Fig 4. Architecture of the System Under Test

Table 1 Specification of the System under test

Host	OS	CPU	Mem	Disk	Network
data-archive-lab	Ubuntu 20.04.2	Intel(R) Core(TM) i9-10920X CPU @ 3.50GHz, 1 * 12 Core	128 GB	1TB	1Gbit/s
data-archive-1	Ubuntu 20.04.2	Intel(R) Xeon(R) Gold 6230R CPU @ 2.10GHz, 2 * 26 core	128 GB	1TB, 10TB	10Gbit/s
data-archive-2	Ubuntu 18.04.6	Intel(R) Xeon(R) Gold 6230R CPU @ 2.10GHz, 2 * 26 core	192 GB	1TB	10Gbit/s
data-archive-3	Ubuntu 18.04.6	Intel(R) Xeon(R) Gold 6230R CPU @ 2.10GHz, 2 * 26 core	192 GB	1TB	10Gbit/s
data-archive-4	Ubuntu 18.04.6	Intel(R) Xeon(R) Silver 4210R CPU @ 2.40GHz, 2 * 10 core	128 GB	13TB, 388GB	10Gbit/s

The Digital Archive subsystem stores and manages basic data and digital model data that make up digital twin models, including space-based sensing data collected from external sources at second-level intervals and operational management data such as patrols. To perform the load test, input data was constructed to mimic the collection and creation environment, generating data that matched the conditions of each test item.

The software installed for the test is as shown in Table 2. For generating the load, an open-source-based "ngrinder-3.5.5-p1" was used, and "PostgreSQL 14.3" was used as the DBMS server software. Additionally, "repmgr" was used to provide DB synchronization functionality, and "pg-scheduler" was used to periodically delete records from the DB table to maintain a certain capacity. "MINIO" was used to manage and synchronize file-based data in order to manage file-based data.

Table 2. Software list for testing

Software	Role
PostgreSQL 14.3	DBMS
pgAdmin 4.27	PostgreSQL administration tool
minio version RELEASE.2022-06-25T15-50-16Z	File replication
DAMB	To Process Restful AP Requests and responses To Control the access to the digital archive
ngrinder-controller-3.5.5-p1	Web Interface for Performance measurement. To Create/Control Test processes/Scripts To Show the Statistics for Test Results
ngrinder-agent-3.5.5-p1	To execute processes and threads to load test
ngrinder-monitor-3.5.5-p1	To monitor the usage of resources of under test systems

The scope of performance measurement for this test is as follows.

- Performance range is limited to the performance of inputting or searching data into/from the digital archive after it is acquired from the data collection server. Additional performance measurements for generating virtual data to be input or for ancillary operations such as creating REST/JDBC requests are excluded.
- For operations using REST APIs, the performance is measured from the time of generating the REST request message to sending it to the DAMB server and receiving the result.
- For REST requests that search for data, the time to parse the received message and bind it to a local variable is included.
- For operations using JDBC connections to make requests, the performance is measured from the time of creating the request query and sending it to the DB server to receiving the result.
- For JDBC queries that search for data, the time to parse the received data and bind it to a local variable is included.

During the test, only one PostgreSQL instance was created, and it was configured to operate with four DBs (SIDB, TMDB, DMDB, and MSDB). For the basic function of the digital archive subsystem, which is synchronization, the DB and files were configured to be synchronized in real-time, and the module for deleting data at regular time intervals was

configured to operate continuously.

B.2 Test cases and results

The test items were conducted by adjusting the values of data input size and interval for general and disaster situations, as shown in Table 3.

Table 3. Test cases

TEST CASE	Description
TC-01	# of values in the INSERT command
TC-02	File upload time
TC-03	Parsing time based on Interface type (REST, jdbc)
TC-04	Size of Sensing Data Input on Normal case
TC-05	Frequency of Sensing data Input on Normal case
TC-06	Size of Sensing Data Input on Disaster case
TC-07	Frequency of Sensing data Input on Disaster case
TC-08	Duration of Disaster

During the test, load testing was conducted by adjusting the values of data input size and frequency for general and disaster situations as shown in Table 3. In the event of a disaster, simulation data is used to predict the spread of the disaster, and as the digital twin reflects the disaster situation over time, the flow or capacity of the data increases. Therefore, the same conditions were tested for both situations. This report summarizes the major test items and results.

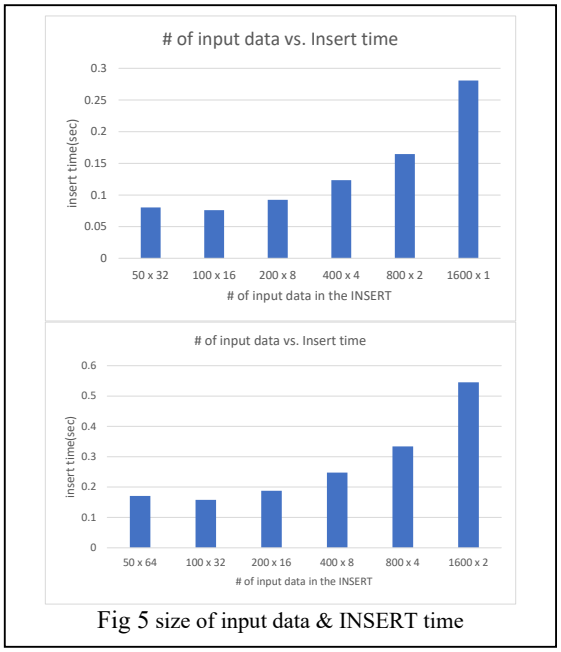

Fig 5 size of input data & INSERT time

TC-01) Performance evaluation of INSERT based on the size of the VALUE in the INSERT query.

- Multiple sensor data are input simultaneously by including multiple VALUEs in the INSERT statement for the INSERT query, and the test determines whether there is a difference in performance depending on the size of the VALUE.
- The performance of the INSERT query is measured by changing the number of VALUEs for a given number of sensor data.

As shown in Figure 5, for all experimental results, it was found that the performance was best when the number of VALUES included in one INSERT query was limited to 100 and multiple INSERT queries were performed.

TC-08) Performance test according to the duration of a disaster situation

When a disaster occurs, data traffic for predicting and responding to the spread of disasters increases, and digital twin model data continues to evolve and accumulate. At this time, it is expected to be the most heavily loaded, and the test was conducted to see how the system performance is affected as the disaster situation persists.

- Test the performance of data input and digital twin data management according to the duration after a disaster occurs.

Fig. 7. The Usage of Resource on the DB replication server during Disaster

- The performance results according to the duration of the disaster situation (30 minutes, 1 hour 30 minutes, 2 hours, 2 hours 30 minutes)

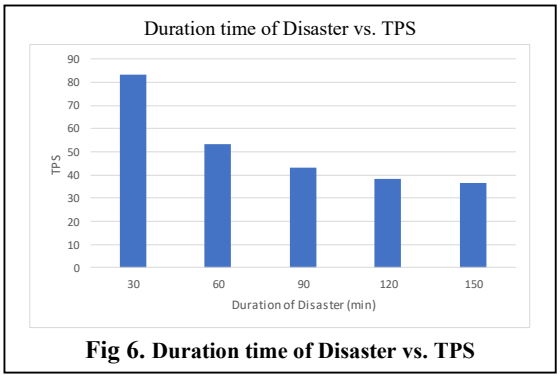

Fig 6. Duration time of Disaster vs. TPS

It can be seen that the performance to handle the increasing size of the digital twin model is degraded, and once the increase rate of the digital twin model size is determined based on the initial twin model size and the sensing data input frequency, it can be utilized as a benchmark for the DB I/O performance in the future. Figure 7 shows the usage of resources such as memory and CPU used in such situations, and it is judged that it does not impose a significant burden on the performance of the server hardware configuration.

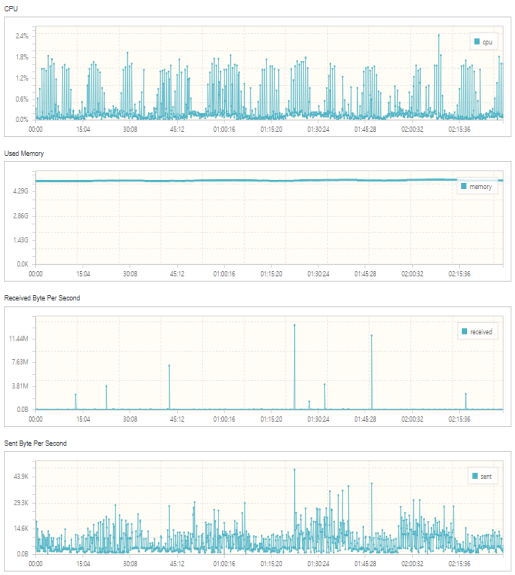

Fig 7 The Usage of Resource on the DB replication server during Disaster

Summarizing the results of the load tests on the digital archive subsystem, we have:

- In this test environment, the performance of processing INSERT queries is the best when the number of VALUES included in the INSERT query is limited to 100.
- Uploading a 3MB file cannot be processed within 1 second, and uploading more than 6 files cannot be processed within 1 second.
- Searching for JSON string data results in more than 5 times performance improvement compared to applying JDBC interface or utilizing REST API.
- When inputting sensor data using multiple threads and 30 threads simultaneously input, up to 147,893 pieces of data can be input per second.
- The performance of sensor data input is determined by the number of concurrently inputted sensor data, with negligible impact from the input interval.
- In a disaster situation, if sensor data is inputted at one-second intervals for one hour, it cannot be processed within 1 second if the number of sensor sets is over 1,700 (sensor data: 13,600 pieces).
- If sensor data for 1,600 sensors (12,800 pieces of sensor data) is inputted every one second for one hour, it can be processed within 1 second, but it cannot be processed within 1 second if it is inputted for 7 hours 59 minutes and 59 seconds.
- In the test environment, there are no significant constraints on resource usage overall because of the high performance of the test equipment.

However, these performance test results are specific to the test environment and may vary in different environments. Therefore, to improve performance in a specific system, you need to find optimized methods based on the system's characteristics and environment.

III. Conclusion

This article introduces the digital archive of a disaster safety management platform based on digital twin technology, and describes the dataset required for the digital twin model of the underground shared space and the load test to verify its performance. Generally, it has been confirmed that the digital archive synchronization and data storage functions will work well in normal and disaster situations based on hardware with good performance. However, it is expected that re-validation will be necessary in the event of exceptional situations or additional requirements in actual environments.

ACKNOWLEDGMENT

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT, MOIS, MOLIT, MOTIE) (No. 2020-0-00061, Digital Twin based Disaster Lifecycle Management Technology for Underground Infrastructure)}

References

[1] Misuk Lee, Woo-Sug Jung, Eunsol Kim, “A Study on the Disaster Safety Management Method of Underground Lifelines based on Digital Twin Technology”, Journal of Computing Science and Engineering, Feb., 2021

[2] Michael Grieves, “Origins of the Digital Twin Concept,” August, 2016

[3] Gartner Glo, https://www.gartner.com/

[4] ETRI, “Digital Twin Technology Report”, January, 2021

[5] S. Rautmare and D. M. Bhalerao, "MySQL and NoSQL database comparison for IoT application," 2016 IEEE International Conference on Advances in Computer Applications (ICACA), 2016, pp. 235-238, doi: 10.1109/ICACA.2016.7887957.

[6] M. Muniswamaiah, T. Agerwala and C. C. Tappert, "Performance of databases in IoT applications," 2020 7th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2020 6th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom), 2020, pp. 190-192, doi: 10.1109/CSCloud-EdgeCom49738.2020.00041.

[7] Asiminidis, Christodoulos & Kokkonis, George & Kontogiannis, S..

(2018). Database Systems Performance Evaluation for IoT Applications. SSRN Electronic Journal. 10. 10.2139/ssrn.3360886.

[8] PostgreSQL, https://www.postgresql.org/