A hybrid cloud-based distributed data management infrastructure for bridge monitoring

*Seongwoon Jeong[1)], Rui Hou[2)], Jerome P. Lynch[3)], Hoon Sohn[4)] and Kincho H. Law[5)]

[1),5)] *Dept. of Civil and Environ. Eng., Stanford University, Stanford, CA 94305, USA*
[2), 3)] *Dept. of Civil and Environ. Eng., University of Michigan, Ann Arbor, MI 48109, USA*
[4)] *Dept. of Civil and Environ. Eng., KAIST, Daejeon 305-600, Korea*
[1)] *swjeong3@stanford.edu*

ABSTRACT

This paper describes a hybrid cloud-based distributed data management infrastructure platform for bridge monitoring applications. As the deployment of sensors and the collection of monitoring data continue to grow, proper management of the data becomes a paramount issue. Cloud computing is one viable approach that is popular among IoT and big data vendors. Cloud service provides many useful features, including reliability, flexibility and scalability and allows efficient and cost-effective use of resources. In practice, however, there is also a desire to maintain a local (private) server, for example, to handle sensitive data that bridge managers hesitate to upload on a public cloud. In this study, we design a hybrid framework in which the bridge monitoring data is distributed across public cloud platforms and private servers. Data distributed over the hybrid cloud environment can be accessed through unified web services with appropriate authority and access controls. The conceptual framework has been prototyped and demonstrated using Microsoft Azure cloud service and a private server.

1. INTRODUCTION

As the deployment of sensors and the collection of monitoring data continue to grow, proper management of the data becomes a paramount issue for bridge monitoring applications. Current practice of bridge monitoring usually adopts a

---

[1)] Graduate student
[2)] Graduate student
[3)] Professor
[4)] Professor
[5)] Professor

traditional approach in managing and storing monitoring data on a server. The private server-based approach, however, requires significant efforts in operating and maintaining the computer system. Furthermore, the lack of scalability and flexibility makes the traditional server-based approach unable to meet today's computing requirements (Marston et al. 2011). Cloud computing is one viable approach that is popular among IoT and big data vendors. Cloud service provides many useful features, including reliability, flexibility, scalability and the pay-as-you-go pricing allows efficient and cost-effective use of resources (Zhang *et al*. 2010). A few research efforts have been reported on the adoption of cloud computing for bridge monitoring applications (Zhang *et al*. 2016, Alampalli *et al*. 2016, Jeong *et al*. 2017). In practice, however, there is a desire to maintain a private server. For example, bridge managers maybe hesitate to upload sensitive data to a public cloud because of security. Hybrid cloud, which combines public cloud and a private server, can be a desirable alternative in that data can be stored in either public cloud or private server depending on the sensitivity or importance of the data (Huang and Du 2013).

In this study, we propose a cyberinfrastructure framework for bridge monitoring where the underlying computing infrastructure is built upon a hybrid cloud paradigm. In this framework, bridge monitoring data is distributed across public cloud and a private server. Data distributed over the hybrid cloud environment can be accessed through unified web interfaces with appropriate authority and access controls. The proposed framework is demonstrated using the bridge monitoring data collected from the bridge along the I-275 corridor located in Michigan.

2. CYBERINFRASTRUCTURE FRAMEWORK BASED ON HYBRID CLOUD

Fig 1 depicts the conceptual framework of the hybrid cloud-based cyberinfrastructure. The framework is mainly composed of two autonomous systems - public cloud-based system and a private server-based system - and a middleware that bridges the autonomous systems. The public cloud-based system offers scalable, flexible and cost-effective computing resources, while the private server enables full management control over the data as well as the physical computer system. The middleware bridges the two heterogeneous systems and provides the users with unified interfaces that resolves some of the underlying complexities of the hybrid cloud infrastructure.

In the cyberinfrastructure framework, the public cloud-based system plays a role in managing large volume of data (e.g., sensor data and video image data). A public cloud-based system is composed of virtual machines, distributed database and web server (Jeong *et al*. 2017). A virtual machine (VM) is a computing infrastructure service offered by cloud service vendors. In cloud computing environment, VMs can be rapidly provisioned through cloud service interfaces. Furthermore, the computational capabilities, storage size and the number of VMs can be easily configured and dynamically modified as needed, thereby facilitating cost-effectiveness and scalability. Distributed database system, which runs on multiple VMs in a decentralized manner, serves as a permanent data store. For the scalable data management, the framework employs Apache Cassandra (http://cassandra.apache.org/), one of the most widely

used distributed NoSQL database. Built upon P2P architecture, Cassandra prevents single point of failure and enables linear scaling of performance and capacity as additional VMs are deployed. Web server serves as a wrapper that provides standardized web-based interface through which the middleware platform and authorized clients can access the database on the public cloud. To provide a standardized and light-weight interface, the web server hosts web services that adhere to the *de facto* RESTful web service design style.

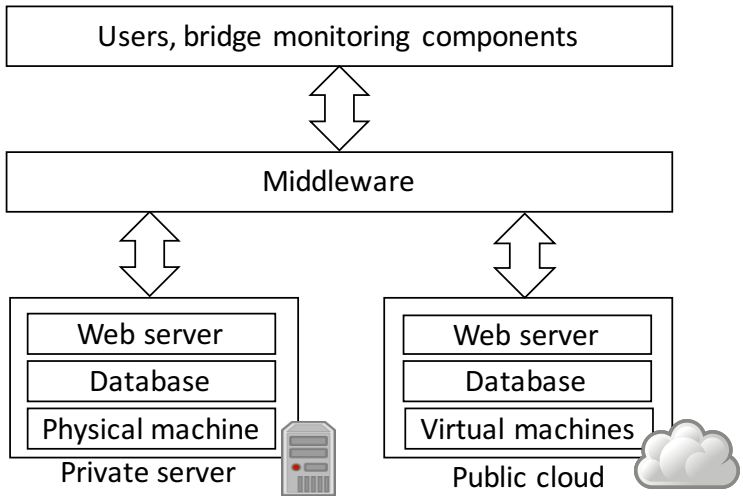

Fig. 1 Conceptual framework of hybrid cloud-based cyberinfrastructure

The private server, on the other hand, manages sensitive and less-voluminous data, such as authorized user list, sensor information and bridge models. Similar to the public cloud-based system, the server employs a database and web server for data management and interface, respectively. In the current design of the framework, we employ Cassandra database system for the server. However, it should be noted that the server does not need to employ the same database system employed on the public cloud-based system. The web server of the private server also serves as wrapper providing standardized web interface for accessing the database.

The middleware platform in the proposed framework is another web server that receives requests from the client users and systems, and then forwards the request to the appropriate recipient (i.e., either public cloud-based system or on premise server). Once the recipient processes the request and returns a response, the middleware delivers the response to the client.

3. EXAMPLE SCENARIO

For demonstration, the bridge monitoring data collected from the Telegraph Road Bridge (as shown in Fig 2), located in Monroe, Michigan, is employed to illustrate the hybrid infrastructure platfrom. The conceptual framework has been prototyped and demonstrated using Microsoft Azure VM service (https://azure.microsoft.com/) and a

proprietary Linux server computer. A web server for the middleware of the framework is deployed on another VM on the Microsoft Azure cloud.

Fig. 2 Testbed bridges: Telegraph road bridge (Monroe, Michigan)

In the example scenario, as depicted in Fig 3, a client retrieves sensor information and sensor data, which are managed by the public cloud-based system and the private server-based system, respectively. In the first step, the client sends a request for retrieving the list of accelerometers to the middleware platform, which then forwards the request to the private server by invoking the sensor information retrieval service on the server. Once processed, the service response, including the accelerometer list, is delivered from the server to the middleware, and to the client. Given the list of accelerometers, which includes sensor IDs, the client sends another request for retrieving acceleration data collected from one of the accelerometers to the middleware platform. The middleware then forwards the request to the public cloud-based system by invoking sensor data retrieval service on cloud platform. After processing the query request, the retrieved acceleration data is delivered from the public cloud to the client via the middleware. Fig 4 shows the actual log record at the middleware and the data retrieval results corresponding to the example scenario. The result shows that the client can retrieve data from the hybrid cloud platform through the unified interface without knowing whether the data resides on the public cloud or on the private server.

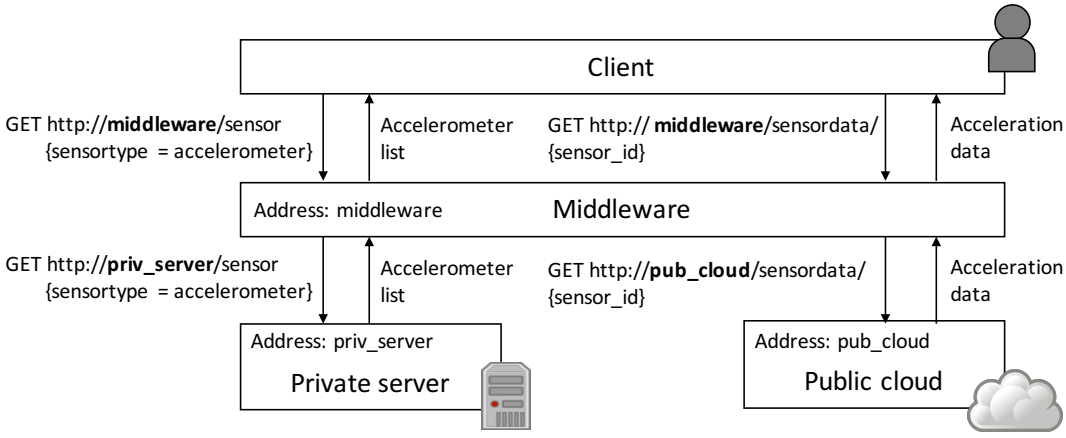

Fig. 3 Data retrieval scenario

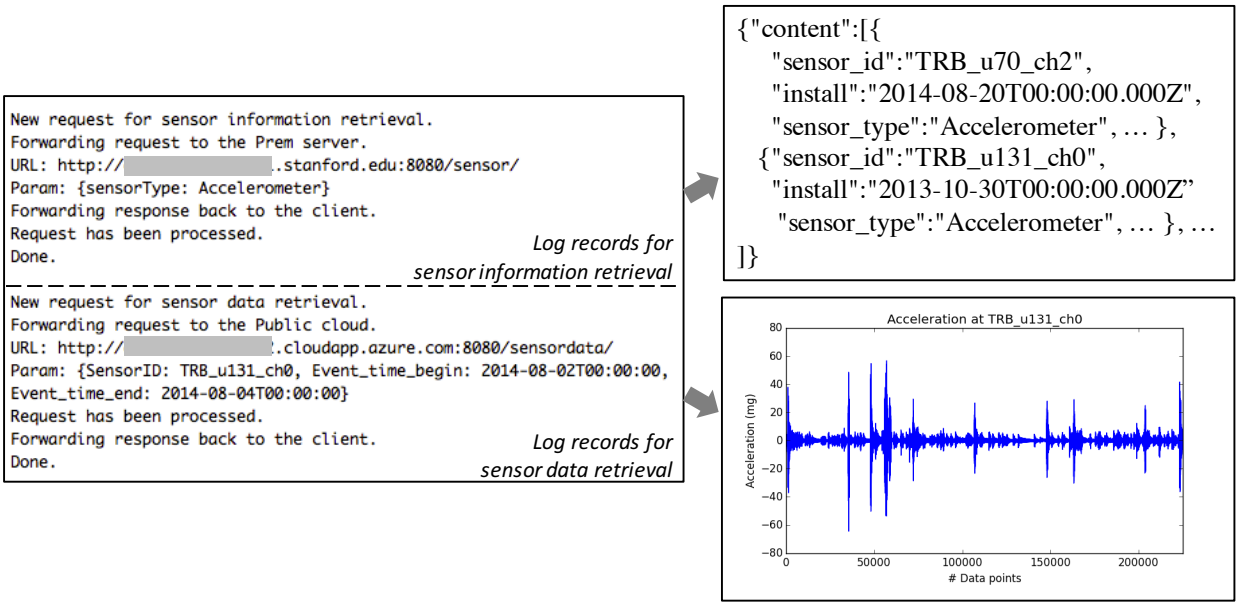

Fig. 4 Log records and the query results corresponding to the example scenario

4. CONCLUSIONS

In this paper, we discuss a hybrid cloud-based data management framework for bridge monitoring applications. The framework consists of two autonomous systems, namely, a system deployed on public cloud and a system deployed on a private server. A middleware is designed as a means to seamless integrate the two systems. The public cloud system is used to manage the large volume of data, while the private server is used to manage sensitive data that bridge managers hesitate to upload on the public cloud. The results show that the hybrid distributed data management infrastructure can take full advantages of the public cloud services as well as enable secure local storage and access to privately stored data.

ACKNOWLEDGMENTS

This research is supported by a Grant No. 13SCIPA01 from Smart Civil Infrastructure Research Program funded by Ministry of Land, Infrastructure and Transport (MOLIT) of Korea government and Korea Agency for Infrastructure Technology Advancement (KAIA). The research is also partially supported by a collaborative project funded by the US National Science Foundation (Grant No. ECCS-1446330 to Stanford University and Grant No. ECCS-1446521 to the University of Michigan). The authors thank the Michigan Department of Transportation (MDOT) for access to the Telegraph Road Bridge and the Newburg Road Bridge and for offering support during installation of the wireless monitoring system. Any opinions, findings, conclusions or recommendations expressed in this paper are solely those of the authors and do not necessarily reflect the views of NSF, MOLIT, MDOT, KAIA or any

other organizations and collaborators.

REFERENCES

Alampalli, S., Alampalli, S. and Ettouney, M. (2016), “Big data and high-performance analytics in structural health monitoring for bridge management,” in *2016 SPIE Smart Structures/NDE Conference*, art no. 980315.

Huang, X. and Du, X. (2013), “Efficiently secure data privacy on hybrid cloud,” *IEEE ICC 2013 - Communication and Information Systems Security Symposium*, pp. 1936-1940.

Jeong, S., Hou, R., Lynch, J.P., Sohn, H. and Law, K.H. (2017), “A Distributed Cloud-based Cyberinfrastructure Framework for Integrated Bridge Monitoring,” in *2017 SPIE Smart Structures/NDE Conference*, art no. 101682W-101682W.

Marston, S., Li, Z., Bandyopadhyay, S., Zhang, J. and Ghalsasi, A. (2011), “Cloud computing—The business perspective,” *Decision support systems*, **51**(1), pp.176-189.

Zhang, Q., Cheng, L. and Boutaba, R. (2010). “Cloud computing: state-of-the-art and research challenges*,” Journal of Internet Services and Applications*, **1**(1), 7–18.

Zhang, Y., O’Connor, S.M., van der Linden, G., Prakash, A. and Lynch, J.P. (2016), “SenStore: A Scalable Cyberinfrastructure Platform for Implementation of Data-to-Decision Frameworks for Infrastructure Health Management,” *Journal of Computing in Civil Engineering*, 30(5), 04016012.